INTERNATIONAL JOURNAL
of
CONTEMPORARY
EDUCATIONAL RESEARCH

IJCER

INTERNATIONAL JOURNAL
of
CONTEMPORARY
EDUCATIONAL RESEARCH

IJCER

Table of Contents

Exploring the Effect of Collaborative Learning on Teacher Candidates' Intentions to Use Web 2.0 Technologies

Erhan ÜNAL[1]
[1]Afyon Kocatepe University

To cite this article:

Exploring the Effect of Collaborative Learning on Teacher Candidates' Intentions to Use Web 2.0 Technologies

Erhan ÜNAL[1*]

[1]Afyon Kocatepe University

Abstract

The purpose of this study was to examine teacher candidates' pre and post intentions to use web 2.0 technologies for teaching before and after a collaborative learning process. An extended TAM was used in the current study. This study was designed as a one-group pretest-posttest design with 56 teacher candidates. In the experiment process, the instructor taught the design principles and how to use and design instructional materials with selected web 2.0 technologies. Then the teacher candidates worked in small groups to design the instructional materials with web 2.0 technologies. The data collection tool was administered before and after the experiment. Data were analyzed through partial least squares structural equation modeling. The results indicated that 4 of the 7 hypotheses were supported in the pre-acceptance model while all hypotheses were accepted in the post-acceptance model. The proposed model can be used as an appropriate framework for examining factors influencing teacher candidates' intentions to use web 2.0 technologies for teaching.

Keywords: Web 2.0 technology, structural equation modeling, teacher candidate, acceptance.

Introduction

New technologies can both provide numerous opportunities for learning environments and improve the 21st skills of learners in the digital age (OECD, 2019). One of these technologies is web 2.0 technologies. They are a kind of internet technology and can be used for enhancing learning and teaching environments. Web 2.0 technologies are web applications that allow collaboration, communication, and content share among users (Butler, 2012). There are many web 2.0 tools developed for different purposes in education one of which can be used to create different instructional materials such as textbooks, infographics, posters, quizzes, and mind maps.

Today, students can be described as digital natives being experts in using a different technology (Prensky, 2001). For example, they use several web 2.0 technologies such as social networks, blogs, social bookmarking (Pence, 2007). Furthermore, these technologies provide many opportunities during the learning process such as sharing information and collaboration among students (Ajjan & Hartshorne, 2008) In this regard integrating web 2.0 technologies into the learning environments of the digital age is important. Thus, teachers should not only use the web 2.0 technologies for personal aims but also use them to create learning activities to support students' learning (Jimoyiannis, Tsiotakis, Roussinos, & Siorenta, 2013). Moreover, it is emphasized in many reports such as the International Society for Technology in Education (ISTE, 2017) and UNESCO (2008) that teachers have to be equipped with some technological skills. In this regard, instructors should teach the opportunities that web 2.0 technologies provide and how to overcome problems when using them to the teacher candidates in the courses. However, training or courses may not ensure that teacher candidates will use web 2.0 technologies in the future. Therefore, as teachers of the future, teacher candidates' intentions to use web 2.0 technologies for teaching can be investigated with the help of the technology acceptance model (TAM) (Davis, 1989). Some studies investigated teacher candidates' intentions to use several technologies in the future teaching such as computers (Teo, 2009; Teo, Ursavaş, & Bahçekapili, 2012), interactive whiteboard (Baydaş & Yilmaz, 2017; Wong, Russo, & McDowall, 2013), and web 2.0 technologies (Cheon, Song, Jones, & Nam, 2010; Sadaf, Newby, & Ertmer, 2012). Accordingly, there are limited examining factors that influence teacher candidates' pre and post intentions to use web 2.0 technologies for teaching. However, it should be noted that instructors can design courses to teach these tools in a meaningful way to enhance their intentions to use them in

---

[*] Corresponding Author: *Erhan ÜNAL, eunal@aku.edu.tr*

their future teaching (Valtonen et al., 2015). One of the innovative ways is the collaborative learning approach where learners can work in a group and learn the content together (Dillenbourg, 1999). This approach was used for designing instructional materials with web 2.0 technologies in the current study. Therefore, the collaborative activities that teacher candidates do can be taken into consideration and added to the TAM to explain their acceptance behaviors. In this context, teacher candidates' intentions to use web 2.0 tools for teaching can be investigated before and after the collaborative learning process. As a consequence, the current study aimed to examine teacher candidates' pre and post intentions to use web 2.0 technologies for their future teaching before and after the collaborative learning process.

Theoretical Framework

Web 2.0 Technologies

Almost all students can be seen as an expert in using different information and communication technologies (ICT) for learning as they live and grow up with digital technologies doing different activities such as blogging, playing online games, and sharing interests (Redecker, 2009). Furthermore, students as digital natives prefer graphics when studying something (Prensky, 2001). Therefore, teachers that educate the new generation should have technological skills.

Web 2.0 technologies are services and tools that users can access any content, create knowledge, and share new knowledge with other users. Web 2.0 technologies have some benefits such as facilitating content creation and sharing it easily, accessing information easily, and collaborating with users (Mcloughlin & Lee, 2007). Due to the characteristics of web 2.0 tools, interactive learning environments can be designed where users can participate in the knowledge construction actively (Cakir, Yukselturk, & Top, 2015; Conole, 2010). There have been many attempts to integrate web 2.0 technologies in education. Hew and Cheung (2013) reviewed some of these studies. They expressed that there was a positive relationship between learning achievement and web 2.0 tools usage. In another review, Redecker (2009) reported the potential benefits of web 2.0 technologies as increasing accessibility of instructional materials, improving achievement, developing learning and higher-order skills, easing knowledge management, and providing advanced tools for a certain subject. Through these potentials of web 2.0 technologies, teacher candidates should be trained about how to use these technologies for teaching.

Collaborative Learning

To form teacher candidates' beliefs, a collaborative learning approach can be useful. The underlying premise of the collaborative learning approach is that students work in small groups for maximizing their learning together (Johnson &Johnson, 1996). Collaborative learning is grounded on social constructivism which is based on the studies of Vygotsky (Duffy & Cunningham, 1996). Collaborative learning assumes that teachers are seen as facilitators and two or more students form a group and engage in learning to solve a problem or achieve a goal doing activities such as participating in the process actively, taking responsibilities, discussing groupmates' findings or ideas, sharing experiences (Matthews, Cooper, Davidson, & Hawkes, 1995). On the other hand, Vygotsky stated the importance of social interaction when constructing knowledge (Duffy & Cunningham, 1996). Therefore, using ICT can be beneficial for the learning process.

Integration of ICT in the collaborative learning environment makes critical contributions to students through providing opportunities for social interaction (Kreijns, Kirschner, & Jochems, 2003). Use of ICT with collaborative learning approach has several benefits such as improvement of higher-order thinking skills, satisfaction with the learning experience, development of productivity (Resta & Laferrière, 2007), knowledge gain, skill acquisition, and positive perception development (Chen, Wang, Kirschner, & Tsai, 2018). Additionally, the use of ICT in the collaborative learning environment enhances group task performance and social interaction (Chen et al., 2018). In this regard, using ICT with the collaborative learning approach can be a better way for equipping teacher candidates with skills about web 2.0 technologies and forming their intentions because users' intentions to use any information system can change with experiences (Venkatesh & Davis, 2000).

Technology Acceptance Model

Some intention-based theories and models have been developed to investigate users' acceptance and adoption of technologies. For example, Davis (1986) introduced TAM for the explanation of computer usage behavior. Thus, the literature counts various applications of not only TAM but also other intention-based theories and models in the research.

There are several studies about teacher candidates', or university students' intentions to use web 2.0 technologies with the help of different intention-based models. For example, Cheon et al. (2010) explored teacher candidates' intentions to adopt web 2.0 technologies with the help of expectancy-value theory after a training session. Yueh, Huang, and Chang (2015) explored university students' adoption of the Wiki system with the Unified Theory of Acceptance and Use of Technology (UTAUT) in the course. Sadaf, Newby, and Ertmer (2016) investigated teacher candidates' intentions and actual uses of web 2.0 technologies in classrooms with the decomposed theory of planned behavior. In a similar study, Huang, Hood, and Yoo (2013) examined teacher candidates' acceptance of web 2.0 tools with the UTAUT model after they enrolled in an educational technology course. Yilmaz and Baydas (2016) examined the factors influencing teacher candidates' behavioral intention to make educational animated movies with a web 2.0 tool at the end of the course. Altanopoulou and Tselios (2017) explored university students' intentions to use wiki before and after the actual use of the wiki system with the help of TAM and added constructs as Big Five personality characteristics. Kul and Çelik (2018) designed a multiple case study and explored teacher candidates' intentions to use web 2.0 technologies for teaching mathematics by using the decomposed theory of planned behavior at the beginning, middle, and end of the course. Arslan (2019) aimed to determine teacher candidates' perceptions towards Web 2.0 tools and intentions at the end of the instructional technology and material development course using TAM. Teo, Sang, Mei, and Hoi (2019) designed a cross-sectional study aiming to explore teacher candidates' acceptance of web 2.0 technologies using extended TAM. Alkhayat, Ernest, and La Chenaye (2020) used the decomposed theory of planned behavior and designed a qualitative study to examine early childhood teacher candidates' intentions to use Web 2.0 technologies in their future.

To summarize the related studies, some of them were based on participants' perceptions toward web 2.0 technologies without any experience process while others were conducted to explore participants' perceptions after the usage of web 2.0 technologies with different intention models. There have been limited studies that examine teacher candidates' intentions to use web 2.0 technologies before and after a training program or a course. In other words, the effect of meaningful and collaborative learning on teacher candidates' intention to use web 2.0 technologies can be tested. Therefore, the difference can be observed thanks to this study as it is emphasized that experience over time can affect intention (Davis, Bagozzi, & Warshaw, 1989). Moreover, policymakers, university administrators, and instructors can benefit from the result of the study about what may influence teacher candidates' intentions to use web 2.0 technologies. Thus, TAM was selected in this research as it is a valid and robust model confirmed by many studies (King & He, 2006). The strengths of TAM are being simple, and applicability to the different technology-related fields. However, researchers attempt to refine the model. In this regard, they extend the model by adding external variables that influence the core variables of the model (Granić & Marangunić, 2019). Thus, TAM can be extended within this research aim by adding the collaboration construct to increase prediction.

The intention to use (IU) is influenced by three main constructs in TAM. These are perceived usefulness (PU), perceived ease of use (PEU), and attitude (AT). Davis (1989, p. 320) defined PU as "the degree to which a person believes that using a particular system will enhance his or her job performance". He explained PEU as "the degree to which a person believes that using a particular system will be free of effort (Davis, 1989, p. 320). Another construct in the model is AT that is "a learned predisposition to respond in a consistently favorable or unfavorable manner with respect to a given object" (Fishbein & Ajzen, 1975, p. 6). PU and AT have a direct effect on IU while PEU and PU are predictors of AT. Moreover, PEU has a direct effect on PU. Venkatesh and Davis (2000) reported that many studies using TAM explained 40% variance in behavioral intention. However, some of the researchers extended TAM by adding constructs to get a better explanatory power (Tarhini, Hone, Liu, & Tarhini, 2017; Yadegaridehkordi, Shuib, Nilashi, & Asadi, 2019). Therefore, the collaboration effort of teacher candidates was added to the model.

Research Model and Hypotheses

TAM was extended with a collaboration construct in the current study. Thus, the effect of the external variable on behavioral intention can be mediated by PU and PEU. The proposed model consists of the following variables: PU, PEU, AT, collaboration, and IU (see Fig. 1).
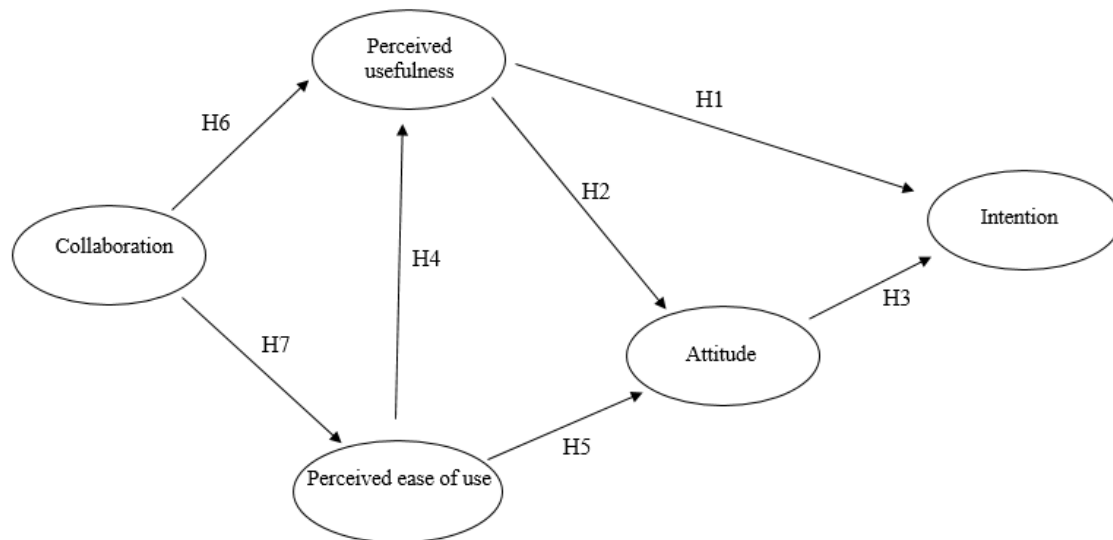
Figure 1. The research model

PU is described as using web 2.0 technologies will enhance teacher candidates' teaching performance. Therefore, if teacher candidates believe that incorporating web 2.0 technologies in the learning environments can be beneficial for their teaching performance, they will use these tools in future teaching. Moreover, this belief can influence their attitudes toward these tools. Many studies support the PU-BI (Esteban-Millat et al., 2018; Usoro, Echeng, & Majewski, 2014) and PU-AT (Ifinedo, 2017; Wang & Huang, 2016) relationships. Therefore, the following hypotheses are formulated:
H1: PU has a positive effect on IU.
H2: PU has a positive effect on AT toward web 2.0 tools.

AT can be explained as teacher candidates 'positive or negative feelings toward using web 2.0 technologies for teaching activities. These feelings of teacher candidates can influence their future intentions. In the literature, this relationship is supported by previous studies (Sadaf et al., 2012; Ursavaş, Şahin, & Mcilroy, 2014). Therefore, it can be hypothesized that:
H3: AT toward web 2.0 tools has a positive effect on IU.

PEU is described as the degree to which teacher candidates believe that using web 2.0 technologies for teaching activities will be free of effort. This perception can affect their feelings toward web 2.0 technologies and their perceptions of usefulness about web 2.0 technologies. Previous studies support PEU-PU (Teo, Lee, & Chai, 2008; Venkatesh, 2000) and PEU-AT (Hsu & Lin, 2008; Park, 2009) relationships. Accordingly, the following hypotheses are proposed:
H4: PEU has a positive effect on PU.
H5: PEU has a positive effect on AT toward web 2.0 technologies.

Collaboration (C) can be explained as supporting learning in a group with the help of web 2.0 technologies that facilitate teacher candidates' collaborative efforts such as collaboration, communication, and interaction to accomplish the goal (Yadegaridehkordi et al., 2019). As a general statement, web 2.0 technologies support collaborative learning by facilitating activities such as accessing, sharing, and co-creating the content (Jimoyiannis, Tsiotakis, Roussinos, & Siorenta, 2013). In the current study, when teacher candidates used some of the web 2.0 technologies to collaborate, communicate and interact with groupmates, instructors, and other people around them, they might believe that these technologies helped them to complete their work easily and increase their performance. Moreover, this belief can affect their perceived ease of use of web 2.0 technologies. Previous studies stated the importance of the collaboration feature of web 2.0 tools for learning (Cheung & Vogel, 2013; Cilliers, 2017). Thus, the following hypotheses are presented:
H6: C has a positive effect on PU.

H7: C has a positive effect on PEU.

Method

Research Model

This study was designed as a one-group pretest-posttest design which was a type of quasi-experimental design (Fraenkel, Wallen, & Hyun, 2012). At the beginning of this study, teacher candidates responded to a survey about extended TAM elements and demographics. In the treatment session of the course, the instructor taught the design and development principles of instructional material and how to design digital materials with web 2.0 technologies. Teacher candidates created groups with 3-4 teacher candidates and worked in small groups to prepare four different instructional digital materials. After the session finished, they took the survey again as a posttest. To examine changes in factors that influenced teacher candidates' intentions, the pretest and posttest data were used to establish structural models. Then pre-acceptance and post-acceptance models were compared according to the research hypotheses.

Participants

The participants of the study were 80 teacher candidates from the faculty of education at a public university in the Aegean Region of Turkey. All teacher candidates were in the second year at their departments. Although all participants participated in the study and fulfilled the survey, some participants did not complete the experiment. Therefore, the study was completed with 56 teacher candidates. Among them, 83.9% (n = 47) of the participants were female while 16.1% (n = 9) of them were male. The age of teacher candidates ranged from 19 to 28 and the average was 19.8. 42.9% (n = 24) of teacher candidates were from the department of primary school education, 37.5% (n = 21) of them were from the department of mathematics education, and 19.6% (n = 11) of them were from the department of Turkish Language Education. Additionally, 96.4% (n = 54) of the participants had no experience with web 2.0 usage in education while 3.6% (n = 2) of the participants had experiences with web 2.0 usage in education.

Experimental Procedure

This study was conducted in the Instructional Technology course. The course includes topics about basic concepts of instructional technology, instructional tools and materials, instructional material design and development, and evaluation of instructional materials. The experiment session of the study was started with the topic of instructional materials and material design and development. It continued for five weeks (see Fig. 2). Before the experiment was started, the survey was administered to teacher candidates and they created collaborative groups with 3-4 teacher candidates. The instructor gave lectures about the use of web 2.0 technologies in education, principles of designing visual materials, and which web 2.0 tool would be appropriate for designing each material weekly. Then the instructor introduced infographics, posters, mind maps, and quizzes as instructional materials. Therefore, Canva was selected for creating infographics and posters, Bubbl.us was preferred for designing mind maps, and Google Forms was selected for preparing online quizzes. For the collaborative design of instructional materials, each group created a group account in each web 2.0 tool. Therefore, each teacher candidate in groups made contributions to the digital material. Additionally, each group created a folder in Google Drive and shared it with groupmates and the instructor. Using Google Drive allows collaborative groups to work anywhere and anytime. These facilities enhanced interaction, collaboration, and communication among teacher candidates. As a consequence, each teacher candidate engaged in the learning process actively such as searching for design issues with web 2.0 tool and discussing and sharing the ideas. Finally, they developed the instructional digital materials collaboratively. At the end of this session, the survey was administered to teacher candidates as a posttest. According to the pretest data, the pre-acceptance model was established and analyzed while the post-acceptance model was established and analyzed based on the posttest data.
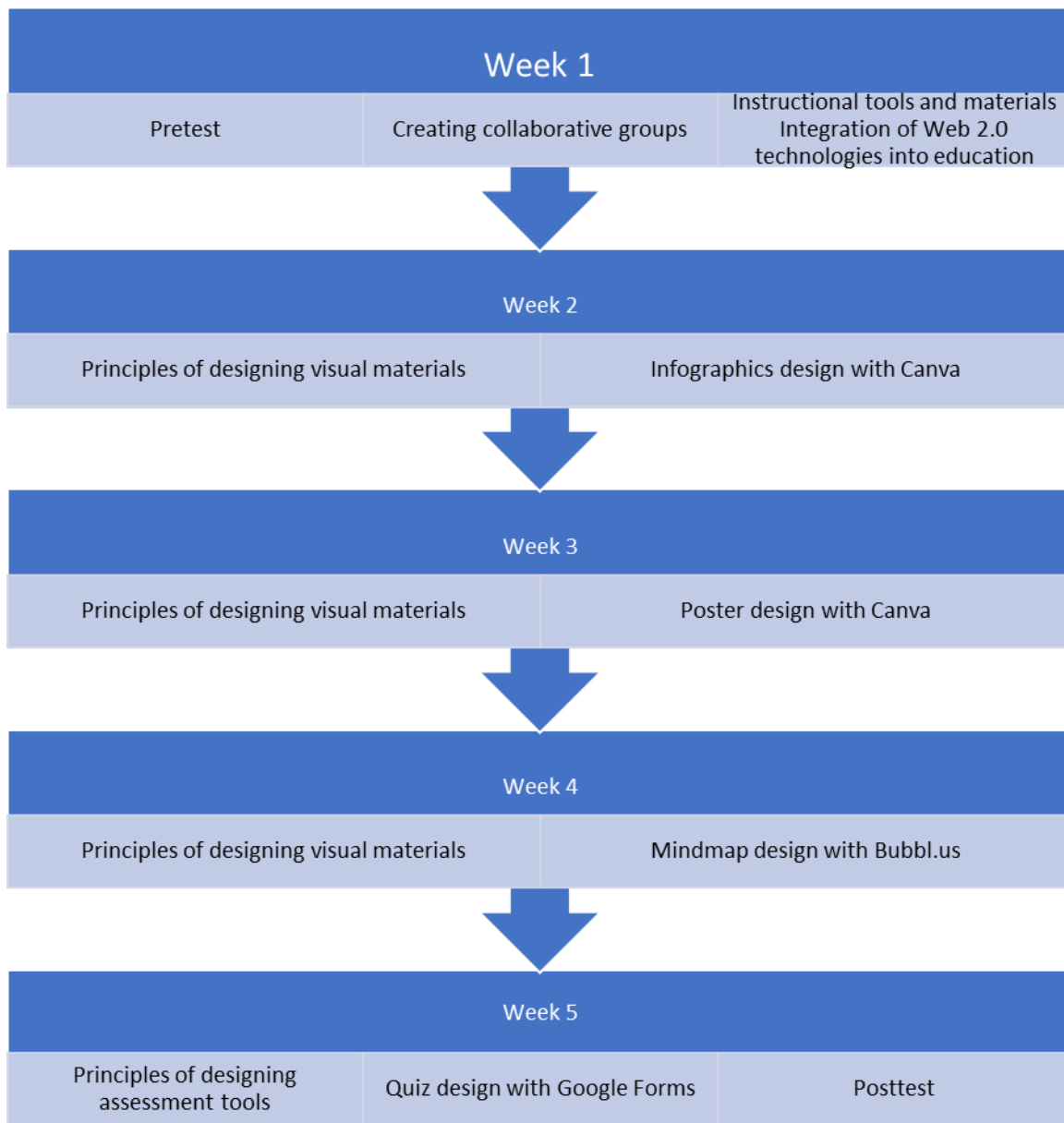
Figure 2. The implementation of the study

Data Collection Tools

The survey of this study consisted of two parts as demographics and scale items. Age, gender, department, experience with web 2.0 tools were used as demographic questions. The scale comprised of 19 items measured by 7 -point Likert scales ranging from "1=strongly disagree" to "7-strongly agree". The constructs of the scale were based on TAM. The items were in English. Therefore, the items were selected from previous studies to fit the purpose of this research. To do this, the selected items were translated into Turkish by two experts. Then two experts in both languages controlled the translated items. The final form of the scale was revised based on expert views (see Table 1).

Table 1. Items and constructs of the scale

Construct	Items	Reference
Perceived Usefulness	Pu1	Davis (1989)
	Pu2	
	Pu3	
	Pu4	
	Pu5	

Perceived Ease of Use	Peu1	Davis (1989)	
	Peu2		
	Peu3		
	Peu4		
Attitude	At1	Davis (1989)	
	At2		
	At3		
Intention to Use	Iu1	Bhattacherjee (2001a, 2001b)	
	Iu2		
	Iu3		
Collaboration	C1	(Baas, 2010)	
	C2		
	C3		
	C4		

Data Analysis

The Partial Least Square Structural Equation Modelling (PLS-SEM) was used to analyze the data. PLS-SEM is preferred than other SEM techniques because PLS-SEM is an appropriate technique where the sample size is small and the aim is to predict the key constructs on the dependent variable (Hair, Ringle, & Sarstedt, 2011). Therefore, the two-step analysis method was adopted with the SmartPLS 3.2.7 program. First, the measurement model was tested to find out the validity and reliability of measures for the pre-acceptance model and the post-acceptance model. Then the structural model was tested to reveal the relationship between constructs for both models.

Results

Results of the Measurement Models

Based on the data collected before the experiment, the pre-acceptance model was established while the post-acceptance model was established with the post-test data. First of all, the validity and reliability analysis was conducted to test the measurement models. Table 2 presents item loadings, internal consistency reliability values (Cronbach's Alpha and composite reliability (CR)), and convergent validity values. The item loadings in the pre and post-acceptance models were greater than 0.50 indicating that indicator reliability was ensured (Fornell & Larcker, 1981). Internal consistency was assessed with Cronbach's Alpha and CR values. The values of Cronbach's alpha and CR of all constructs in two models met the threshold of 0.7 (Hair et al., 2011). Convergent validity was assessed through the values of average variance extracted (AVE) which were exceeded 0.5 (Hair, Black, Babin, Anderson, & Tatham, 2006).

Table 2. Reliability and validity of measurement models

Construct	Item	Loading (Pre)	Loading (Post)	AVE (Pre)	AVE (Post)	CR (Pre)	CR (Post)	Cronbach's Alpha (Pre)	Cronbach's Alpha (Post)
Attitude				.838	.821	.939	.932	.903	.891
	AT1	0.914	0.869						
	AT2	0.913	0.919						
	AT3	0.919	0.929						
Intention to use				.860	.817	.949	.931	.919	.888
	IU1	0.921	0.905						
	IU2	0.940	0.870						
	IU3	0.921	0.935						
Perceived ease of use				.742	.706	.920	.905	.885	.860
	PEU1	0.784	0.816						
	PEU2	0.871	0.899						
	PEU3	0.870	0.801						

	PEU4	0.914	0.841					
Perceived usefulness				.802	.770	.953	.944	.938	.925
	PU1	0.891	0.882					
	PU2	0.858	0.913					
	PU3	0.925	0.922					
	PU4	0.929	0.844					
	PU5	0.872	0.825					
Collaboration				.802	.779	.953	.933	.927	.903
	C1	0.857	0.756					
	C2	0.955	0.917					
	C3	0.921	0.929					
	C4	0.890	0.917					

This study employed Fornell and Larcker's (1981) criteria to assess discriminant validity. The square root of AVE values of each construct should exceed the correlation than its correlation with other constructs in the models. Table 3 exhibits the Fornell-Larcker criteria values for the pre-acceptance model while Table 4 presents the Fornell-Larcker criteria values for the post-acceptance model. The values were acceptable for both cases.

Table 3. Results of discriminant validity for pre-acceptance model

	AT	IU	C	PEU	PU
AT	**0.915**				
IU	0.825	**0.928**			
C	0.753	0.754	**0.907**		
PEU	0.464	0.522	0.522	**0.861**	
PU	0.579	0.586	0.488	0.541	**0.895**

Note. Bold diagonal: square root of AVE, below diagonal

Table 4. Results of discriminant validity for post-acceptance model

	AT	IU	C	PEU	PU
AT	**0.906**				
IU	0.781	**0.904**			
C	0.789	0.719	**0.883**		
PEU	0.682	0.560	0.625	**0.840**	
PU	0.637	0.665	0.684	0.723	**0.878**

Note. Bold diagonal: square root of AVE, below diagonal

Results of the Structural Models

To test the causal relationships of the structural models, the bootstrapping approach was handled with 5,000 subsamples. The predictive power of the model with R^2 values (see Table 6), beta coefficients (β), significance of the path coefficients (p), and t values are presented in Table 5 and Fig. 3 and 4.

Table 5. Summary of hypotheses testing

		β		t		p		Supported	
Hypothesis	Relationship	pre	post	pre	post	pre	post	pre	post
H1	PU→IU	0.163	0.281	1.554	3.039	0.120	0.002	No	Yes
H2	PU→AT	0.464	0.302	2.931	2.072	0.003	0.038	Yes	Yes
H3	AT→IU	0.730	0.602	8.314	5.622	0.000	0.000	Yes	Yes
H4	PEU→PU	0.393	0.484	2.266	4.306	0.023	0.000	Yes	Yes
H5	PEU→AT	0.213	0.464	1.173	3.178	0.241	0.001	No	Yes
H6	C→PU	0.283	0.381	1.777	3.680	0.076	0.000	No	Yes
H7	C→PEU	0.522	0.625	4.740	7.240	0.000	0.000	Yes	Yes

Table 6. Predictive relevance and explained variance

	R^2	
Endogenous Variables	Pre	Post

AT	0.367	0.508
IU	0.698	0.657
PEU	0.273	0.391
PU	0.350	0.611

The analysis results of the pre-acceptance model showed that teacher candidates' intentions were predicted only by AT with a significant β value of 0.730 (p < 0.001). However, in the post-acceptance model, PU and AT were significant predictors of IU with significant β values of 0.281 ($p < 0.05$) and 0.602 ($p < 0.001$) respectively. Therefore, the H3 hypothesis was accepted for both models but the H1 hypothesis was supported for the post-acceptance model. The H2 hypothesis which stated the PU-AT relationship was confirmed with significant β values of 0.464 ($p < 0.01$) and 0.302 ($p < 0.05$) respectively for both models while the H5 hypothesis which stated the PEU-AT relationship was supported only for the post-acceptance model with significant β value of 0.464 ($p < 0.01$). The C and PEU had a significant effect on PU with significant β values of 0.381 ($p < 0.001$) and 0.484 ($p <0.001$) respectively for the post-acceptance model whereas PEU had a significant effect on PU only in the pre-acceptance model with significant β values of 0.393 ($p < 0.05$). Therefore, the H4 hypothesis was accepted in both models while the H6 hypothesis was supported only for the post-acceptance model. H7 hypothesis which was about C-PEU relationship was supported for both models with significant β values of 0.522 ($p < 0.001$) and 0.625 ($p < 0.001$) respectively.

The predictive power of the models was assessed by R^2 values. The pre-acceptance model explained 69.8% variance on IU with AT construct while the post-acceptance model revealed 65.7% variance on IU with PU and AT constructs. According to Hair et al. (2011), the values of 0.75, 0.50, and 0.25 R^2 indicate a substantial, moderate, and weak model. Therefore, the pre and post-acceptance models can be seen as moderate models. Other R^2 values for AT, PU, and PEU were 0.367, 0.350, and 0.273 respectively for the pre-acceptance model while these values increased to 0.508, 0.611, and 0.391 for the post-acceptance model.
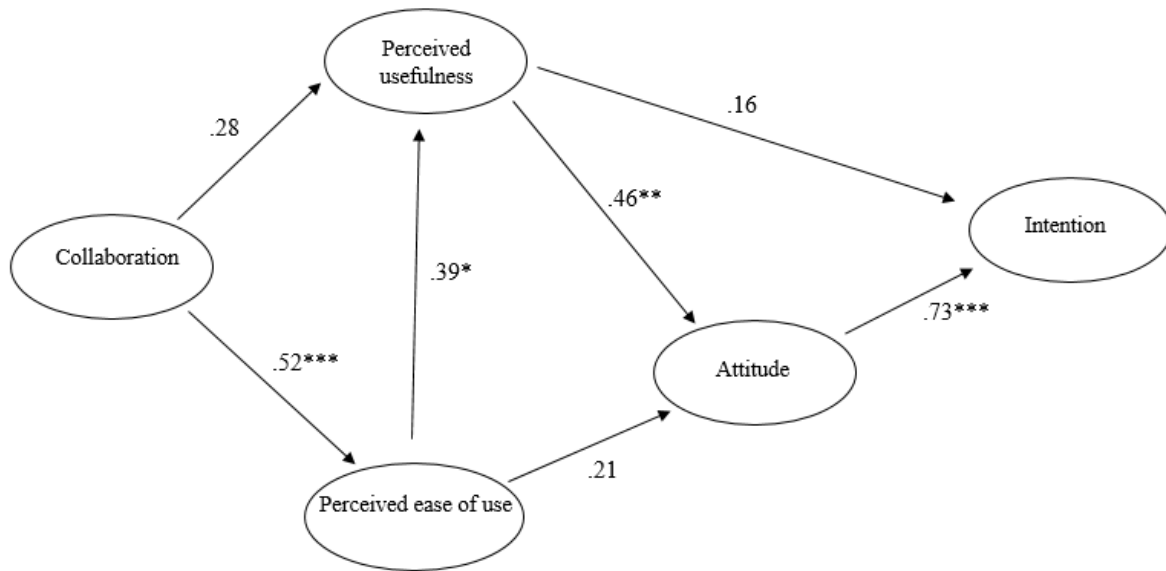
Note. *$p < .05$; **$p < .01$; ***$p < .001$
Figure 3. Results of structural model analysis for pre-acceptance
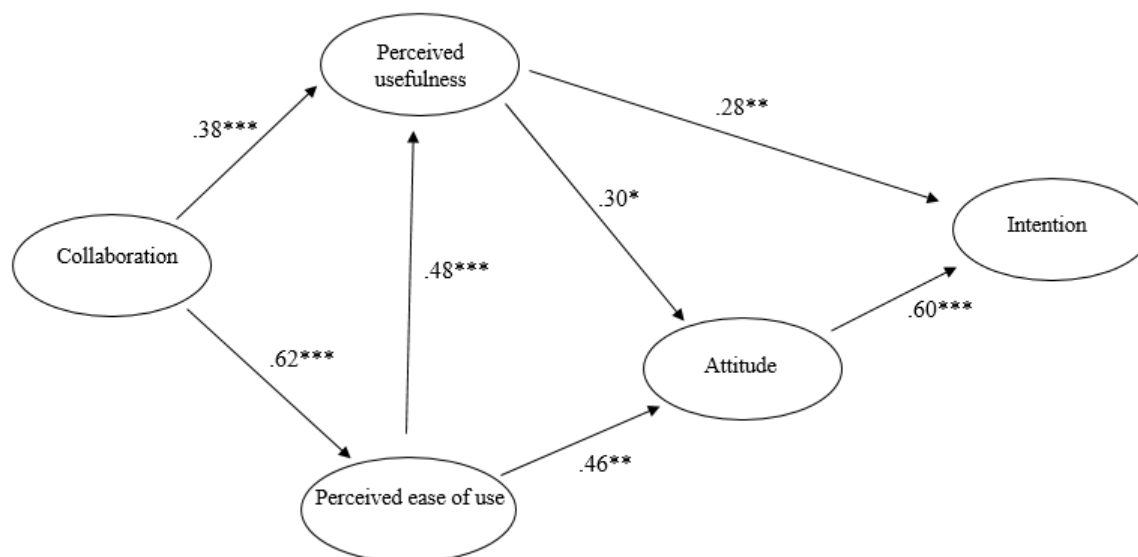
Note. *$p < .05$; **$p < .01$; ***$p < .001$
Figure 4. Results of structural model analysis for post-acceptance

Discussion

This study was designed to examine possible changes in teacher candidates' intentions to use web 2.0 technologies for teaching in the future by incorporating collaboration construct into TAM. Therefore, a pretest-posttest design was used to monitor changes of intentions over a course session. The results indicated that four of seven hypotheses were supported in the pre-acceptance model while all hypotheses were accepted in the post-acceptance model. Although the pre-acceptance model explained a little higher variance in intention to use than the post-acceptance model, both models were moderate. Moreover, other R^2 values tended to increase from pre-acceptance to post-acceptance.

AT was the strongest and significant predictor of IU in both cases while PU was the significant predictor for the post-acceptance model. This finding is in line with previous research findings (Moreno, Cavazotte, & Alves, 2017; Wang & Huang, 2016). As an expected result, teacher candidates could have some positive feelings toward ICT before the course. After the collaborative learning activities, their feelings did not change so much. However, PU did not affect IU in the pre-acceptance model. As most of them were less experienced, they could evaluate the system as heuristically (Castañeda, Muñoz-Leiva, & Luque, 2007). Therefore, they did not have any perception about the effect of web 2.0 technology on their teaching in the future. As they used web 2.0 technologies in the collaborative groups, they realized the capabilities of web 2.0 technologies such as collaboration, interaction, and communication. Through collaborative activities for designing digital materials, they believed that these tools would help them to achieve their teaching goals in the future.

The other important result was that AT was predicted by PU in the pre-acceptance model whilst PEU and PU were the significant factors that influenced AT in the post-acceptance model. This finding is consistent with some previous research results (Altanopoulou & Tselios, 2017; Huang, 2017) In other words, teacher candidates' beliefs about tangible benefits of the web 2.0 technologies had an impact on their attitudes toward the use of web 2.0 tools before and after the experiment. Accordingly, PU had an indirect effect on IU in the pre-acceptance model. The positive change of effect of PEU on AT could be explained that teacher candidates worked in small groups and helped each other to design digital materials with these tools. As most of them utilized and operated these tools for the first time, they did not have any opinion about their use. As a result, teacher candidates who found web 2.0 technologies easy to use tended to have favorable positive attitudes toward the use of web 2.0 technologies for teaching during the experiment process.

Based on the results, C had a significant impact on PEU for both cases while C had an impact on PU in the post-acceptance model. In a similar study, Yadegaridehkordi et al. (2019) found that students' perceptions of collaboration towards online collaborative learning tools had an impact on their PU and PEU. Therefore, collaboration was positively associated with the intention to use web 2.0 technologies in future teaching which

was mediated by PEU and PU at the end of the process. Moreover, its impact on PU and PEU was increased according to the coefficient values from the pre to post-acceptance models. Especially, the significant effect of C on PU after the experiment can be explained that as teacher candidates used the collaboration and interaction features of the web 2.0 technologies, their perception of usefulness would increase (Baas, 2010). For example, teacher candidates studied topics such as instructional materials, design principles of these materials, effective design with web 2.0 technologies through collaborative activities such as accessing information, sharing it with groupmates, discussing the topic with groupmates and instructor, and communicating with groupmates and instructor whenever they want. Therefore, they gained experience in the design of instructional materials with web 2.0 technologies during collaborative learning. In other words, their perceptions toward collaborative activities tended to enhance their productivity and become helpful in using the web 2.0 tool easily. On the other hand, the significant effect of C on PEU in both cases can be explained as teacher candidates could have experiences about collaborative learning with a technological tool. Thus, this belief may affect their perception that the use of web 2.0 technologies is free of effort. Additionally, the collaboration continued to affect the PEU with a more powerful coefficient value at the end of the process.

Regarding the research findings, this study has offered some implications for theory and practice. First, this study proposed a model for understanding the determinants of the teacher candidates' intentions to use web 2.0 technologies before and after the collaborative learning. In addition to the cross-sectional methods used widely in the technology acceptance studies, this study employed the experimental design to observe changes of intention to use. This method was used because any trainings related to the technology that users take might change their intention (Davis, Bagozzi, & Warshaw, 1989). In this regard, the proposed model has become helpful in explaining more factors that influence teacher candidates' intention to use web 2.0 technologies in the post-acceptance model. On the other hand, the effect of collaborative learning on teacher candidates' intention to use web 2.0 technologies was examined with extending TAM in the current study. Therefore, an application of extended TAM can offer practical value for examining teacher candidates' or teachers' intention to use technology in the collaborative learning environment.

For practice, to increase teacher candidates' intentions to use web 2.0 technologies, collaborative learning activities can be used in the learning environments. First, teacher candidates learn how to use them in education, participate in the learning process more actively, and interact with the environment. Then they may perceive that the web 2.0 tools improve their teaching performance. Second, hands-on projects and meaningful learning activities can affect their perceived usefulness and attitudes. Third, thanks to collaborative learning activities with technology, teacher candidates may learn how to operate a web 2.0 tool easily and in turn, this may affect their intentions.

This research has some limitations. First, this study was carried out with a small sample. Future research can be conducted with larger groups to generalize the research findings. Moreover, two groups can be selected and the effect of the traditional paradigm and active learning paradigm on users' intentions can be investigated. Second, while the participants of the study were in the second year at their departments, a similar study can be conducted with teacher candidates who are in 3rd or 4th year at their departments. Third, teacher candidates used Google Drive, Bubbl.us, Canva, and Google Forms. A future study can incorporate synchronous tools such as Google Hangouts and Skype, social learning management systems such as environment Edmodo, Facebook, and wiki for collaboration, and other instructional material design tools. Finally, TAM was extended with the collaboration construct. In further studies, factors that may influence intention such as subjective norms, self-efficacy, and perceived enjoyment can be added to TAM to get a holistic view.

Conclusion

This study described teacher candidates' intention to use web 2.0 technologies for future teaching over the collaborative learning process with the help of extended TAM. The study was one of the first studies about the related topic that monitor changes in intentions of teacher candidates. The proposed model indicated reliable and valid results. From this aspect, this study made a contribution to technology acceptance literature by the development of intention with the experimental design. Moreover, this study made a valuable contribution for the practice that ICT-supported collaborative learning environments can be used for improving teacher candidates' or teachers' intentions to use web 2.0 technologies for teaching in any course or training session.

References

Ajjan, H., & Hartshorne, R. (2008). Investigating faculty decisions to adopt Web 2.0 technologies: Theory and empirical tests. *The internet and higher education*, *11*(2), 71-80.

Altanopoulou, P., & Tselios, N. (2017). Assessing acceptance toward wiki technology in the context of higher education. *International Review of Research in Open and Distance Learning*, *18*(6), 127–149.

Alkhayat, L., Ernest, J., & LaChenaye, J. (2020). Exploring Kuwaiti preservice early childhood teachers' beliefs about using web 2.0 technologies. *Early Childhood Education Journal*, 1-11.

Arslan, K. (2019). Attitudes and perceptions of pedagogical formation physical education students about web 2.0 tools and factors for successful adaptation of these tools. *European Journal of Education Studies*, *6*(6), 101-114.

Baas, P. (2010). *Task-technology fit in the workplace (Affecting employee satisfaction and productivity).* Rotterdam- Netherlands: (MSc Business Administration), Erasmus University.

Bhattacherjee, A. (2001a). An empirical analysis of the antecedents of electronic commerce service continuance. *Decision Support Systems, 32*(2), 201–214.

Bhattacherjee, A. (2001b). Understanding information systems continuance. An expectation–confirmation model. *MIS Quarterly, 25*(3), 351–370.

Baydaş, Ö., & Yilmaz, R. M. (2017). A Model for pre-service teachers' intention to use interactive white boards in their future lessons. *Journal of Higher Education and Science*, *7*(1), 59-66.

Butler, J. (2012). Grappling with change: Web 2.0 and teacher education. In D. Polly, C. Mims, & K. A. Persichitte (Eds.), *Developing technology-rich teacher education programs: Key issues* (pp. 135–150). Hershey, PA: IGI Global.

Cakir, R., Yukselturk, E., & Top, E. (2015). Pre-service and in-service teachers' perceptions about using Web 2.0 in education. *Participatory Educational Research*, *2*(2), 70-83.

Castañeda, J. A., Muñoz-Leiva, F., & Luque, T. (2007). Web acceptance model (WAM): Moderating effects of user experience. *Information & management*, *44*(4), 384-396.

Chen, J., Wang, M., Kirschner, P. A., & Tsai, C. C. (2018). The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: A meta-analysis. *Review of Educational Research*, *88*(6), 799-843.

Cheon, J., Song, J., Jones, D. R., & Nam, K. (2010). Influencing preservice teachers' intention to adopt web 2.0 services. *Journal of Digital Learning in Teacher Education*, *27*(2), 53–64.

Cheung, R., & Vogel, D. (2013). Predicting user acceptance of collaborative technologies : An extension of the technology acceptance model for e-learning. *Computers & Education*, *63*, 160–175.

Cilliers, L. (2017). Wiki acceptance by university students to improve collaboration in higher education. *Innovations in Education and Teaching International*, *54*(5), 485–493.

Conole, G. (2010). Facilitating new forms of discourse for learning and teaching: harnessing the power of Web 2.0 practices. *Open Learning: The Journal of Open, Distance and e-Learning, 25*(2), 141-151.

Davis, F. D. (1986). *A technology acceptance model for empirically testing new end-user information systems: Theory and results* (Doctoral dissertation). Massachusetts Institute of Technology.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management Science, 35*(8), 982-1003.

Dillenbourg, P. (1999). Introduction: what do you mean by ''collaborative learning''? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches (Advances in learning and instruction series)* (pp. 1 – 19). Amsterdam: Pergamon.

Duffy, T. M., & Cunningham, D. J. (1996). Constructivism: Implications for the design and delivery of instruction. In David H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 693-719). New York: Macmillan.

Esteban-Millat, I., Martínez-López, F. J., Pujol-Jover, M., Gázquez-Abad, J. C., & Alegret, A. (2018). An extension of the technology acceptance model for online learning environments. *Interactive Learning Environments*, *26*(7), 895–910.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: an introduction to theory and research.* Reading, MA: Addison-Wesley.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th Edition). New York: McGraw-Hill.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50.

Granić, A., & Marangunić, N. (2019). Technology acceptance model in educational context: A systematic literature review. *British Journal of Educational Technology*, *50*(5), 2572-2593.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–152.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th Ed.). nj: Prentice-Hall.

Hew, K. F., & Cheung, W. S. (2013). Use of Web 2.0 technologies in K-12 and higher education: The search for evidence-based practice. *Educational research review*, *9*, 47-64.

Huang, W. H. D., Hood, D. W., & Yoo, S. J. (2013). Gender divide and acceptance of collaborative Web 2.0 applications for learning in higher education. *Internet and Higher Education*, *16*(1), 57–65.

Huang, Y. M. (2017). Exploring the intention to use cloud services in collaboration contexts among Taiwan's private vocational students. *Information Development*, *33*(1), 29–42.

Hsu, C. L., & Lin, J. C. C. (2008). Acceptance of blog usage: The roles of technology acceptance, social influence and knowledge sharing motivation. *Information & management*, *45*(1), 65-74.

Ifinedo, P. (2017). Examining students' intention to continue using blogs for learning: Perspectives from technology acceptance, motivational, and social-cognitive frameworks. *Computers in Human Behavior*, *72*, 189–199.

International Society for Technology in Education (2017). *ISTE standards for educators.* Retrieved 4th March 2020 from https://www.iste.org/standards/for-educators#startstandards

Jimoyiannis, A., Tsiotakis, P., Roussinos, D., & Siorenta, A. (2013). Preparing teachers to integrate web 2.0 in school practice: Toward a framework for pedagogy 2.0. *Australasian Journal of Educational Technology*, *29*(2), 248–267.

Johnson, D. W., & Johnson, R. T. (1996). Cooperation and the use of technology. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 1017-1044). New York: Macmillan.

King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & management*, *43*(6), 740-755.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in human behavior*, *19*(3), 335-353.

Kul, Ü. & Çelik, S. (2018). Investigating changes in mathematics teachers' intentions regarding web 2.0 technology integration. *Acta Didactica Napocensia, 11*(2), 89-104

Matthews, R. S., Cooper, J. L., Davidson, N., & Hawkes, P. (1995). Building bridges between cooperative and collaborative learning. *Change, 27,* 34–40.

Mcloughlin, C., & Lee, M. J. W. (2007). *Social software and participatory learning : Pedagogical choices with technology affordances in the Web 2.0 era Introduction : Social trends and challenges*. Paper presented at the Ascilite Singapore 2007. Retrieved from https://ascilite.org/conferences/singapore07/procs/mcloughlin.pdf

Moreno, V., Cavazotte, F., & Alves, I. (2017). Explaining university students' effective use of e-learning platforms. *British Journal of Educational Technology*, *48*(4), 995-1009.

OECD (2019). *OECD Skills Outlook 2019: Thriving in a Digital World*. OECD Publishing: Paris.

Park, S. Y. (2009). An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning. *Educational Technology & Society, 12*(3), 150–162.

Pence, H. E. (2007). Preparing for the real web generation. *Journal of Educational Technology Systems, 35*(3), 347−356.

Prensky, M. (2001). Digital natives, digital immigrants part 1. *On the Horizon, 9*(5), 1–6.

Redecker, C. (2009). *Review of learning 2.0 practices: study on the impact of web 2.0 innovations of education and training in Europe.* European Commission. Retrieved from https://publications.jrc.ec.europa.eu/repository/bitstream/JRC49108/jrc49108.pdf

Resta, P., & Laferrière, T. (2007). Technology in support of collaborative learning. *Educational Psychology Review*, *19*(1), 65-83.

Sadaf, A., Newby, T. J., & Ertmer, P. A. (2012). Exploring factors that predict preservice teachers' intentions to use web 2.0 technologies using decomposed theory of planned behavior. *Journal of Research on Technology in Education*, *45*(2), 171–196.

Sadaf, A., Newby, T. J., & Ertmer, P. A. (2016). An investigation of the factors that influence preservice teachers' intentions and integration of web 2.0 tools. *Educational Technology Research and Development*, *64*(1), 37–64.

Tarhini, A., Hone, K., Liu, X., & Tarhini, T. (2017). Examining the moderating effect of individual-level cultural values on users' acceptance of E-learning in developing countries: a structural equation modeling of an extended technology acceptance model. *Interactive Learning Environments*, *25*(3), 306-328.

Teo, T. (2009). Modelling technology acceptance in education: A study of pre-service teachers. *Computers & Education*, *52*(2), 302-312.

Teo, T., Lee, C. B., & Chai, C. S. (2008). Understanding pre-service teachers' computer attitudes: applying and extending the technology acceptance model. *Journal of Computer Assisted Learning, 24*(2), 128-143.

Teo, T., Ursavaş, Ö. F., & Bahçekapili, E. (2012). An assessment of pre-service teachers' technology acceptance in Turkey: A structural equation modeling approach. *Asia-Pacific Education Researcher*, *21*(1), 191-202.

Teo, T., Sang, G., Mei, B., & Hoi, C. K. W. (2019). Investigating pre-service teachers' acceptance of Web 2.0 technologies in their future teaching: a Chinese perspective. *Interactive Learning Environments*, *27*(4), 530–546.

UNESCO (2008). *ICT competency standards for teachers: policy framework.* Retrieved 4th March 2020 from https://unesdoc.unesco.org/ark:/48223/pf0000156210

Ursavaş, Ö. F., Şahin, S., & Mcilroy, D. (2014). The role of discipline in determining Turkish pre-service teachers' behavioral intentions to use ICT *Education and Science, 39*(175), 136 – 153.

Usoro, A., Echeng, R., & Majewski, G. (2014). A Model of Acceptance of Web 2.0 in learning in higher education: a case study of two cultures. *E-Learning and Digital Media*, *11*(6), 644-653.

Valtonen, T., Kukkonen, J., Kontkanen, S., Sormunen, K., Dillon, P., & Sointu, E. (2015). The impact of authentic learning experiences with ICT on pre-service teachers' intentions to use ICT for teaching and learning. *Computers and Education*, *81*, 49–58.

Venkatesh, V. (2000). Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. *Information Systems Research*, *11*(4), 342–365.

Venkatesh, V., & Davis, F. D. (2000). Theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science*, *46*(2), 186–204.

Wang, C. S., & Huang, Y. M. (2016). Acceptance of cloud services in face-to-face computer-supported collaborative learning: a comparison between single-user mode and multi-user mode. *Innovations in Education and Teaching International*, *53*(6), 637–648.

Wong, K. T., Russo, S., & McDowall, J. (2013). Understanding early childhood student teachers' acceptance and use of interactive whiteboard. *Campus-Wide Information Systems, 30*(1), 4-16.

Yadegaridehkordi, E., Shuib, L., Nilashi, M., & Asadi, S. (2019). Decision to adopt online collaborative learning tools in higher education: A case of top Malaysian universities. *Education and Information Technologies*, *24*(1), 79–102.

Yilmaz, R. M., & Baydas, O. (2016). Pre-service teachers' behavioral intention to make educational animated movies and their experiences. *Computers in Human Behavior*, *63*, 41–49.

Yueh, H. P., Huang, J. Y., & Chang, C. (2015). Exploring factors affecting students' continued Wiki use for individual and collaborative learning: An extended UTAUT perspective. *Australasian Journal of Educational Technology*, *31*(1), 16–31.

A Comparison of Classification Performances between the Methods of Logistics Regression and CHAID Analysis in accordance with Sample Size

Mehmet Şata[1], Fuat Elkonca[2]
[1]Agri Ibrahim Cecen University
[2]Mus Alparslan University

A Comparison of Classification Performances between the Methods of Logistics Regression and CHAID Analysis in accordance with Sample Size[*]

Mehmet Şata[1][†], Fuat Elkonca[2]
[1] Agri Ibrahim Cecen University
[2] Mus Alparslan University

Abstract

The aim of the study is to analyze how classification performances change in accordance with sample size in logistic regression and CHAID analyses. The dataset used in this study was obtained by means of "Attentional Control Scale." The scale was applied to 1824 students and the analyses were done by randomly choosing the samples from the dataset. Nine classification criteria were determined in order to evaluate classification performances of logistic regression and CHAID analyses, and the results were interpreted in consideration of these criteria. As a result of the analyses, it was found that classification performance in logistic regression showed no change as sample size increased, and performed a better classification in small sample size (N= between 25 and 900) than CHAID analysis. On the other hand, in the method of CHAID analysis it was seen that classification performance improved as sample size increased, and provided stronger findings in large sample size (N= 1000 and above). Moreover, in classification studies logistic regression analysis yielded more reliable results, and CHAID analysis provided stronger classifications. The results of this study are considered to suggest researchers to select the methods in classification studies based on sample size.

Key words: Logistic regression, CHAID analysis, Classification, Sample size.

Introduction

Classification is a method which is commonly used in scientific studies as well as daily life due to its benefit to problem solution (Köktürk, 2012). The decisions for placing students to a certain academic program, determining the individuals with psychopathology and the customers with a credit risk in a bank are examples for the studies indicating the importance of group membership. Many practitioners in various disciplines use different statistical methods to estimate the group membership belonging to any property (Finch & Schneider, 2007). In the classification studies, different results were obtained because of the reasons such as the existence of a great number of classification algorithm, each algorithm's having different parameters within itself, each algorithm's having more than one type, different purposes of different algorithms, the use of different data source, algorithms' support for different data types and the dependence of pre-treatments of data on the practitioner (Akpınar, 2000; Berry & Linoff, 1997).

In classification studies, the methods based on data mining are mostly used. The data mining techniques are influential in estimating determining crucial data classifications and data tendency by classifying especially large datasets. Among these methods, Logistic Regression, Decision Trees and Artificial Neural Networks are often used in classification studies (Kıran, 2010). Logistic regression and CHAID analyses of decision trees will be mentioned below as they were used within the scope of the study.

Logistic regression assumes that there is a logit relation between dependent and independent variables; therefore, logistic regression can provide non-linear models. The reasons for common preference of logistic regression in social sciences are that they impose no restriction on variable's being continuous or non-continuous; there is no condition for the likelihood function distribution of independent variables; there is no

---

obligation for a linear relation between independent and dependent variable; there are many statistical package programs (Tabachnick & Fidell, 2013).

In logistic regression analysis, there is no assumption regarding the distribution of independent variables. However, when logistic regression analysis is used, there are some assumptions. First, the sample rates of variables in the analysis is an important necessity. Second, logistic regression analysis uses goodness of fit tests to evaluate model-data fit. The goodness of fits tests includes the values expected for each cell in the dataset of the combination of discrete variables. If any expected frequency in cells is very low (usually ef<5), the strength of the analysis is very low. Third, there should not be a problem with multicollinearity. Last, the extreme values of independent variables should be carefully examined (Tabachnick & Fidell, 2013).

CHAID analysis is one of the methods in decision trees which are commonly used in data mining. CHAID analysis uses chi-square (χ2) if a dependent variable is discontinuous and F statistics if it is a continuous as it is concerned especially with the relation between independent variables and their interactions. As it is known, chi-square test statistics deals with the dependence between the variables. Due to this property, it is consistent that CHAID analysis establishes mathematical models on a chi-square ground. In addition, CHAID analysis helps to provide objective and robust results in the evaluation of the effects of sociodemographic variables in the sample and the sub factors of assessment instrument on dependent variables (Kayri & Boysan, 2007).

Classification studies are often seen in the disciplines of education, medicine and banking. In classification studies, there are a great number of models and different algorithms belonging to these models. The answers to the questions which one of these algorithms provide more accurate results, which algorithms are more successful in certain disciplines will increase the practices' success and enhance proficiency of the work. Therefore, an evaluation of algorithms by comparison is of great importance. However, classification studies are usually conducted with one sample, and also as a result of the difference in sample size classification studies provide different results. Therefore, in classification studies it is not determined which method is more effective in accordance with sample size. For example, classification performances of logistic regression, artificial neural networks and decision trees methods were compared based on small samples by Sabzevari, Soleymani and Noorbakhsh (2007), and logistic regression was indicated the most successful method. On the other hand, in the study conducted by Kıran (2010) decision trees were found to make a better classification than logistic regression. In the study by Neuilly, Zgoba, Tita, and Lee (2011) decision trees were found to have a lower classification rate than logistic regression. In another study by Heckert and Gondolf (2005), binary and multinomial logistic regression among logistic regression methods were compared with CHAID, Exhaustive CHAID, CART and QUEST analyses among decision trees methods. As a result of the analyses, logistic regression analysis has a better classification rate than decision trees. In addition, in other several classification studies, these methods were found to have different classification methods (Ekici, 2012; Karakış, 2009; King, Feng & Sutherland, 1995; Kurt & Türe, 2005; Zurada & Lonial, 2005). These differing results in the literature cause a confusion regarding which method is better and complicate the method preference of researchers in classification studies.

Therefore, in this study the classification performances of these methods from small sample size (n=25) to large sample size (n= 1800) were analyzed, and it is aimed to lead researchers which method to use in accordance with their sample size. In this sense, the study is expected to contribute to the literature. Considering that there is no empirical study comparing classification performances in accordance with sample size, and there are only studies with simulative data based on simulation methods (Dolgun, 2014), the current study is considered to have great importance for the relevant literature. In this study, it is hypothesized that logistic regression provides better results in small sample size and classifies within the acceptable error limit (α=0.05), CHAID analysis yields better results in a large sample size and classifies within the acceptable error limit (α=0.05).

Method

Research Type

In this study, how classification performances of logistic regression and CHAID analysis changed in accordance with sample size. As the relevant study determines the existent situation, it is a type of descriptive study.

Study Group

The study group of the study is composed of total 1824 students of 700 females and 1124 males in the high schools of Batman Provincial Directorate of National Education through convenience sampling which is one of

the non-random sampling methods. The samples were randomly created by the study group. While the samples were created, the condition for being five people minimum per each cell in the dataset for independent variable which is an assumption of logistic regression is taken into consideration. As there are five different variables in the study, the smallest sample size was determined as 25 and selected randomly from the total dataset. After this procedure, the other sample sizes were determined as multiples of 25 and it continued to reach 1800 participants. In this way, 72 samples in total were created and both analysis methods were applied to these samples.

Data Tools

In this study, the "Attentional Control Scale" developed by Fajkowska and Derryberry (2010) and translated into Turkish by Akın et al. (2013) was used as an instrument for data collection. Before the scale was applied, the necessary permission was obtained from the researchers translating into Turkish. Moreover, "Personal Information Form" was developed by the researcher to collect demographic information in accordance with the aim of the study. The items were scaled from negative to positive in a 4-point Likert scale. The points to be taken from the scale change between 20 and 80. The high score of total points from the scale indicate students' high level of attentional control while the low score of total points from the scale indicate students' low level of attentional control. The reliability and validity analyses were done and the reliability of the scale was found Mcdonald ω 0.894 (%CI 0.887 – 0.900). The confirmatory factor analysis was done for validity and it was found RMSEA = 0.077, SRMR = 0.057, IFI = 0.95, RFI = 0.94, NNFI = 0.94 and CFI = 0.95. In other words, the data collection instruments were found to be reliable and valid.

Data Analysis

In the study, logistic regression and CHAID analyses were conducted in order to determine classification performances in accordance with the sample size. For the analysis, SPSS (Version 25.0) and Mplus (Version 7) and Origin (Version 8) package programs were utilized. Polynomial fit was drawn to the graphics by means of origin package program in order to make these obtained graphics more understandable for interpretation. This polynomial fit was gained by creating a polynomial regression model between dependent and independent variables.

Two-step cluster analysis was used to categorize the scale scores as the total scores from the Attentional Control Scale is continuous. In this way, the total score of the scale was formed into categorical two clusters (high level and low level). The main reason for transferring total continuous score of the scale into categorical score is that CHAID analysis provided better results in categorical data (Pehlivan, 2006), and it is aimed to compare classification performances with logistic regression.

Before the analyses were conducted, multivariate normal distribution hypothesis, outlier, missing value and the multicollinearity were analyzed. As a result of the analyses, it is found that there are no outliers and multicollinearity, but there is a missing value and multivariate normal distribution could not be ensured.

For missing value, a new value was attained by using the EM (Expectation Maximization) algorithm. As there is no obligation to counter to normal distribution hypothesis of both CHAID and logistic regression analyses, analysis were done without any procedure for multivariate normality. In the current study, the same model was used to provide consistent results in the considered analysis methods.

Classification Criteria

In classification studies, the performances of analysis methods were determined by means of certain criteria. When the literature is examined, it is seen that there are various criteria. The most used criteria are as follows: positive likelihood rate (PLR), negative likelihood rate (NLR), Type I error rate, Type II error rate, confidence level, power of test, sensitivity, specificity and total accurate classification percentage. The classification table is used to calculate these parameters as shown in Table 1 (Koyuncu, 2015).

Table 1. Classification Table

Observed (Test Result) State		Expected (Actual) State				
		Positiive	N	Negative	N	Total
	Positive	Correct Positive (Confidence Level)	a	Incorrect Positive (Type I Error)	c	a+c
	Negative	Incorrect Negative (Type II Error)	b	Correct Negative (Power of Test)	d	b+d
	Total		a+b		c+d	a+b+c+d

Positive likelihood rate (PLR) is a classification performance measurement by combining sensitivity and specifying classification studies. It is obtained by dividing correct positive rate into incorrect positive rate. This value indicates how many incorrect positive results were provided for each correct positive result in classification study. It is usually demanded for PLR's being as high as possible (Deeks & Altman, 2004; Grimes & Schulz, 2005; Medcalc, 2018, s.222). PLR values are calculated as given below:

$$PLR = \text{Sensitivity}/(1\text{-Specificity}) \qquad (1)$$

Negative likelihood rate (NLR) is obtained by dividing incorrect negative rate into correct negative rate. This value indicates how many correct negative results were provided for each incorrect negative result in classification study. It is usually demanded for NLR's being as low as possible (Deeks & Altman, 2004). NLR values are calculated as given below:

$$NLR = (1\text{-Sensitivity})/\text{Specificity} \qquad (2)$$

Type I error rate (α) is an error rate when the test result is accepted as positive though it is negative in reality. Type I error is also known as α error (Cohen, 1988, p.4). It is accepted to have low values in practice. Considering that Type II error rate will be high as the Type I error rate is lower, it is necessary to use large samples in order to lower both error types (Tan, 2016, p.265). Type I error rate is calculated in 2x2 tables as follows:

$$\alpha = c/(a+b+c+d) \qquad (3)$$

Type II error rate (β) is an error rate when the test result is accepted as negative though it is positive in reality (Tan, 2016, p.265). Type II error is also known as β error (Cohen, 1988, p.5). It is accepted to have low values in practice similar to Type I error rate. In the studies, while Type II error rate is not stated while Type I error rate is calculated. Type II error rate is calculated in 2x2 tables as follows:

$$\beta = b/(a+b+c+d) \qquad (4)$$

The confidence level of the test is the likelihood of not having Type I error, and the power of the test is the likelihood of not having Type II error (Tan, 2016, p.265). The confidence level of the test is obtained by subtracting Type I error rate from 1, while the power of the test is obtained by subtracting Type II error rate from 1.

Sensitivity is described as the positive likelihood though it is positive in reality (Deeks & Altman, 2004; Medcalc, 2018, p.222). At the same time, it is known as an accurate positive rate. Sensitivity is calculated in 2x2 tables as follows:

$$\text{Sensitivity} = a/(a+b) \qquad (5)$$

Specificity is described as the negative likelihood as it is negative in reality (Deeks & Altman, 2004; Medcalc, 2018, p.222). At the same time, it is known as an accurate negative rate. Specificity is calculated in 2x2 tables as follows:

$$\text{Specificity} = d/(c+d) \qquad (6)$$

In the current study, classification performances of logistic regression and CHAID analyses were compared in accordance with the nine classification criteria mentioned above. The graphics were used for comparison and polynomial fit were drawn to the graphics to make it more understandable.

Findings

In the created regression model for both CHAID and logistic regression analyses, the grouped state (high control/low control) of total score obtained from the scale as dependent variable were used. As independent variable, the students' sex, age, state of living with their families, monthly income of their families and use of smart phone were included in the model. This created regression model was randomly selected and was tested in total 72 samples. Then, the results were provided by graphics for each classification criteria. First, the positive likelihood in accordance with sample size in CHAID and logistic regression analyses is shown in Figure 1.
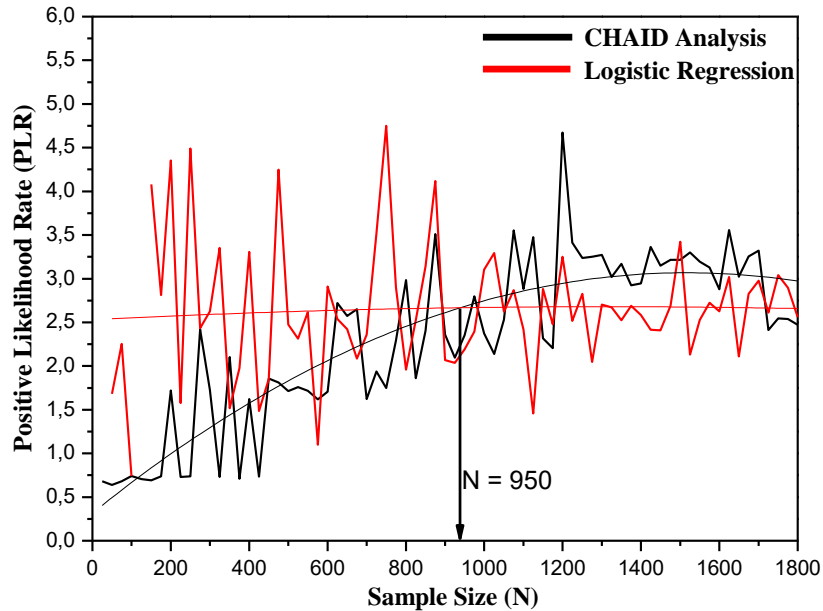
Figure 1. Positive likelihood rate for both analyses

When Figure 1 was examined, logistic regression analysis showed no change in accordance with sample size; however, in CHAID analysis PLR value is seen to increase as sample size increases. It is clear that logistic regression provides more desired results in classification studies in sample size up to 1000. Therefore, the sample size can be argued an important factor in selecting the method of analysis used in classification studies. The change in negative likelihood rate (NLR) in accordance with sample size, which is another classification criterion, is shown in Figure 2.
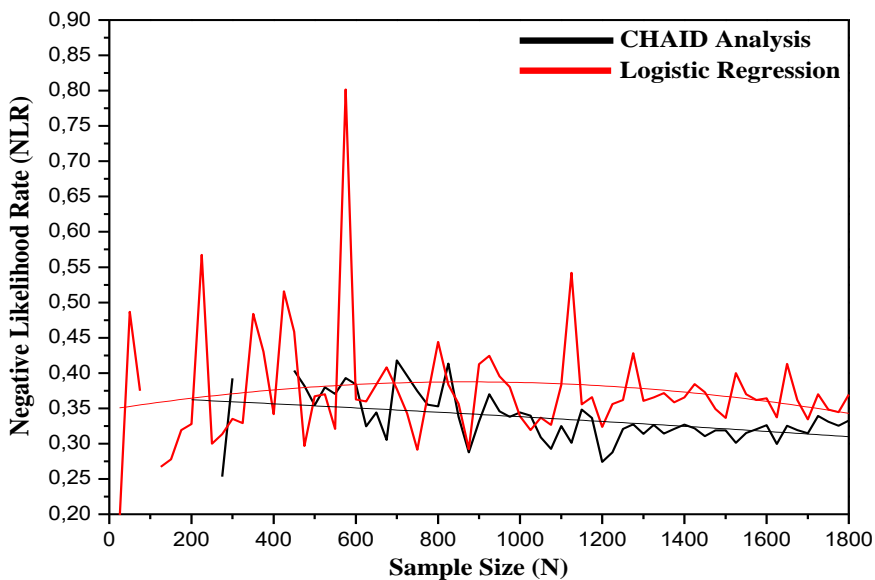
Figure 2. Negative likelihood rate for both analyses

When Figure 2 was examined, both analysis methods did not show much change as sample size increased. CHAID analysis is seen to have discrete values when sample size is approximately at 600. The main reason is derived from the fact that specificity percentage of in all the samples except two samples in CHAID analysis is zero. After the analyses of PLR and NLR values, the change of Type I error rate, which is a criterion often used in classification studies, in accordance with sample size was analyzed and the obtained values are shown in Figure 3.
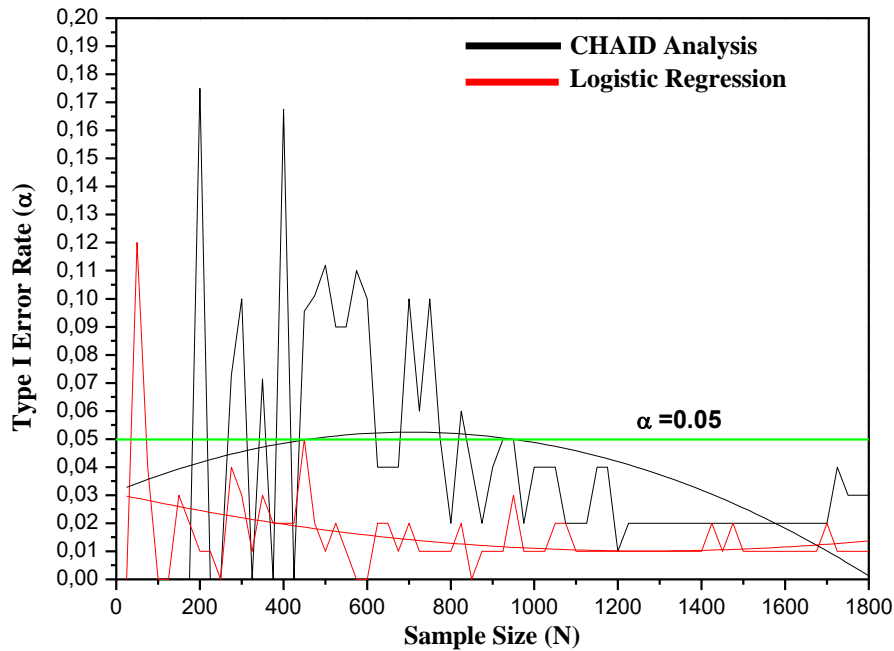
Figure 3. Type I error rate for both analyses

In Figure 3, it is seen Type I error rate decreases in logistic regression as the sample size increases while in CHAID analysis Type I error rate increases to a certain value and then decreases. Moreover, when the sample size is approximately between 440 and 920, it is seen that the error rate obtained from CHAID analysis exceed the level $\alpha = 0.05$ which is usually accepted as a range in social sciences. In the study, the change of Type II error rate in accordance with sample size was analyzed, and the graphic regarding the results is shown in Figure 4.
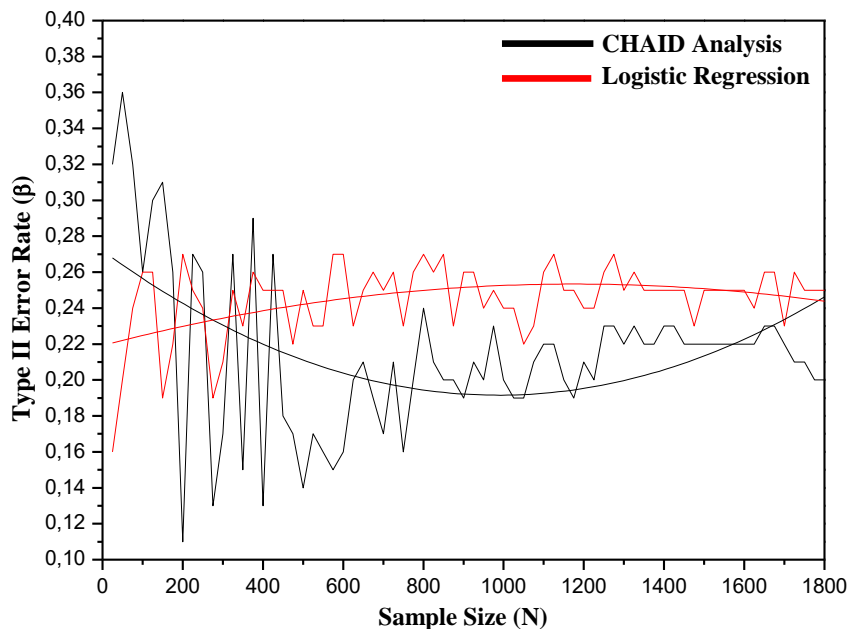
Figure 4. Type II error rate for both analyses

When Figure 4 was analyzed, it is seen Type II error rate did not show much change in logistic regression as the sample size increased while in CHAID analysis Type II error rate decreased and then increased once again. The reason is considered to derived from the considerably high level of specificity percentage while CHAID analysis classifies to a certain sample size (approximately 600 samples). After analyzing the change of Type I and Type II error rates in accordance with sample size, how the confidence level of the test changed was analyzed, and the results are shown in Figure 5.
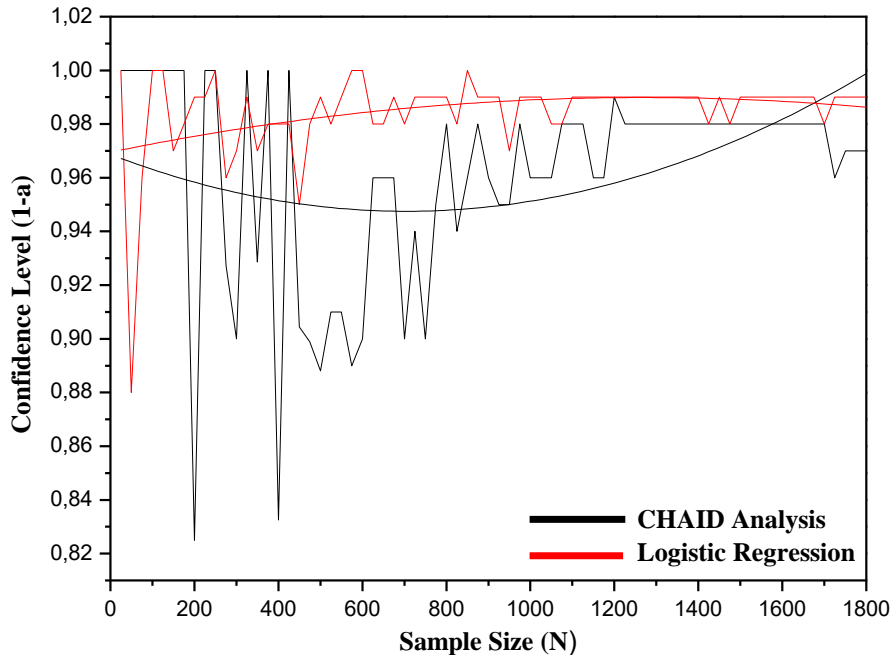
Figure 5. The confidence level of the test for both analyses

When Figure5 was analyzed, the confidence level of the classification through logistic regression analysis was found to be higher. However, as the sample size increased, the confidence level of the classifications through CHAID analysis was found to increase and provide more reliable classifications after n= 1600 samples than logistic regression. In the study, the change of power of classification tests in accordance with sample size was analyzed and the results are shown in Figure 6.
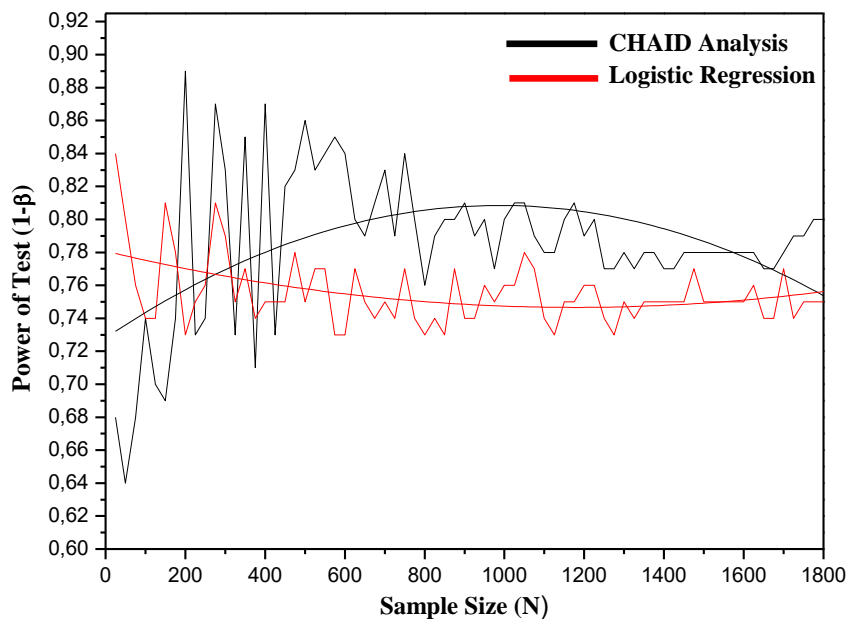
Figure 6. The power of test for both analyses

When Figure 6 was analyzed, the power of classification in logistic regression did not show much change as the sample size increased, while it first increased and then decreased in CHAID analysis. It is seen that CHAID

analysis has a weak power at approximately n= 400. In the study, the change of specificity percentages in accordance with sample size was analyzed, and the results are shown in Figure 7.
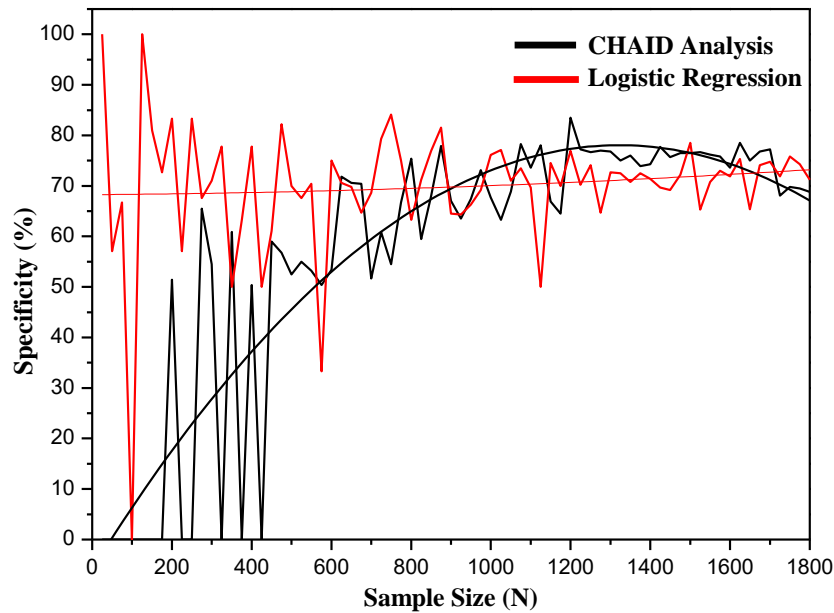
Figure 7. Specificity percentages for both analyses

When Figure 7 was analyzed, it is seen that specificity was very low in CHAID analysis with small samples and logistic regression was almost independent from sample size. In classification studies when CHAID analysis was used, it should be remembered that specificity would be low at size between 0 and 700 samples. If a specificity is an important criterion in a classification study and the sample size is small, the preference of logistic regression might be more reliable and valid. In the study, the change of classification sensitivity of both analysis methods was analyzed, and the results are shown in Figure 8.
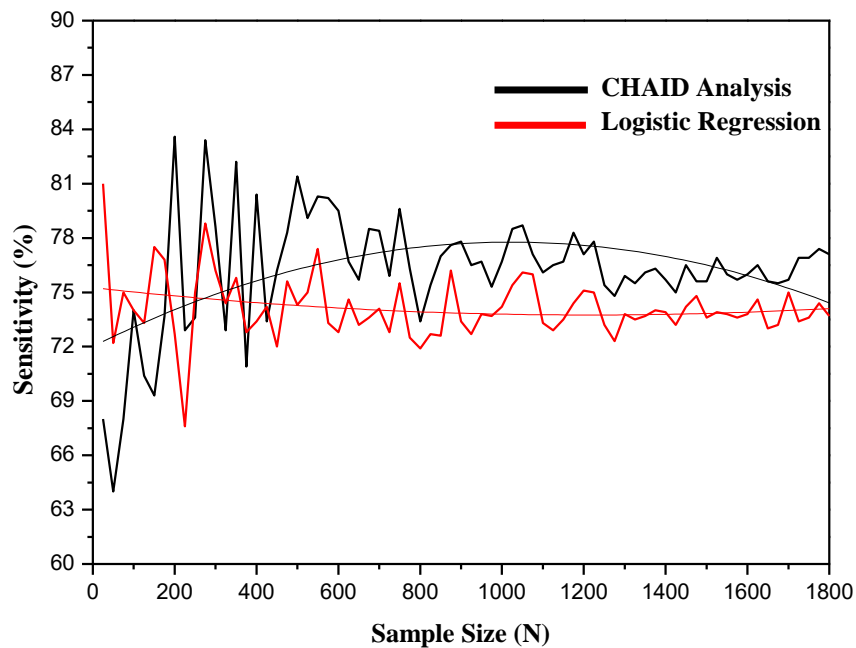
Figure 8. Sensitivity percentages for both analyses

As seen in Figure 8, as the sample size increased, the sensitivity percentage in CHAID analysis increased, logistic regression hardly changed and was independent from sample size. Last, the change of total classification percentages of both analysis methods was analyzed, and the results are shown in Figure 9.
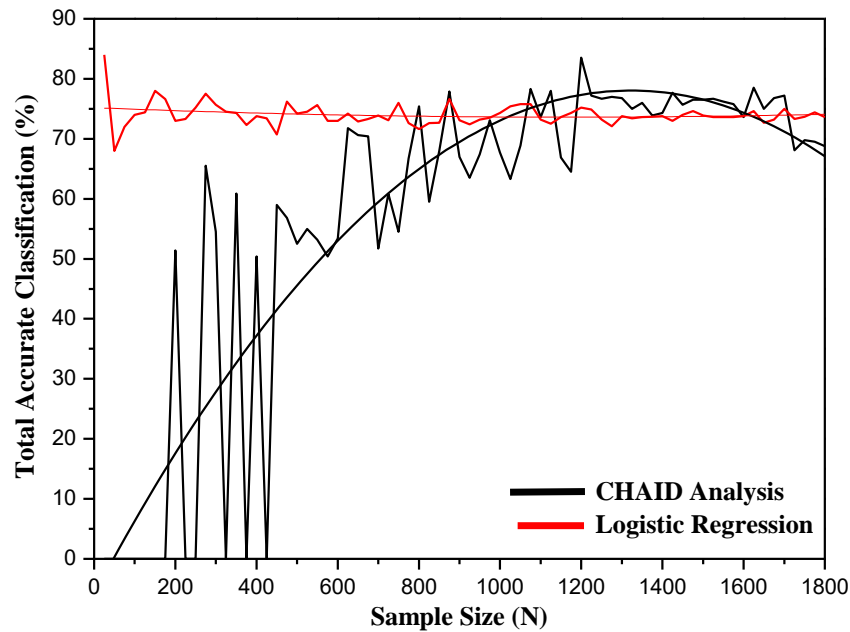
Figure 9. Total accurate classification percentages for both analyses

When Figure 9 was analyzed, it is seen that accurate classification percentage of CHAID analysis increased as the sample size increased and there was no considerable change in logistic regression. Moreover, it is found that logistic regression provided more accurate classification at small samples.

Discussion, Conclusion and Suggestions

The aim of this study is to compare classification performances of CHAID and logistic regression analysis in accordance with sample size. Therefore, nine classification criteria were determined to compare classification performances of both analyses. In addition, at the beginning of the study based on the literature the hypothesis was developed that logistic regression analysis made better classification with small samples and CHAID analysis made better classification with large samples.

In a scientific study, "good sample" rather than large sample" should be aimed (Balcı, 2015, p.105). This is related to be systematic and sensitive in the process of selection. The sample should be as small as possible, but should represent the population at a high level in order to accomplish its goal (Balcı, 2015, p.105). Therefore, the sample size especially in classification studies in scientific studies is a crucial factor. However, inappropriate sample size causes inaccurate or subjective results. This leads us to take incorrect decisions and use resources inefficiently (Stafford et al., 2006).

Based on the findings of the study, the sample size in accordance with the analysis methods used in classification studies can be considered as an important factor. It is observed that logistic regression analysis should be preferred in classification studies rather than decision trees methods particularly in small samples. In a similar way, in the study of Nemes et al., (2009) which the classification performance of the logistic regression in accordance with sample size was analyzed, logistic regression has more accurate classification with small samples. In the study conducted by Demidenko (2007), the likelihood rate in logistic regression was seen to decrease as sample size increased.

When the literature was analyzed, there are various studies which compare classification performances of logistic regression and decision trees. However, it was observed that classification performances of the methods used in these studies showed variance among studies (Ekici, 2012; Brewer, 2012; Neuilly et al., 2011; Kıran, 2010; Çakır, 2008; Heckert & Gondolf, 2005). Therefore, it has become difficult to choose the method to be used in the classification studies. In this sense, the current study provides evidence for which method researchers would choose in accordance with the sample size they can reach. According to the analyses, it is seen more appropriate to choose logistic regression analysis in small sample sizes and one of both analyses in large sample sizes.

In classification studies, as Type I error rate is an important criterion it is required to determine the threshold value of Type I error rate. In the current study, Type I error rate was analyzed and the threshold value was identified as α = 0.05. As a result of the analyses, logistic regression classified below the identified error rate, but CHAID analysis was seen to classify above the error rate at small samples. Considering this finding, the use of logistic regression is seen to be more appropriate with small samples and in the studies whose error rate is important. When the international literature was examined, in classification studies it is seen that Type I error rate was taken into consideration while the studies at national level do not consider Type I error rate in classification studies. In the current study, CHAID analysis exceed the α = 0.05 level for some sample groups. Therefore, the accuracy of the classification is suspected. It is understood that Type I and Type II error rates should be consulted in classification studies.

The power of classification of the used analysis is also an important criterion in classification studies. Particularly at large datasets, the speed of classification methods becomes an important factor. Therefore, the obtained findings suggested that CHAID analysis had stronger classification performance. Similarly, decision trees were concluded to classify more quickly in the study by Çakır (2008).

When the graphics, in which the nine criteria within the scope of the study, were analyzed, it was found that there was much fluctuation in both analysis methods in small sample sizes. This fluctuation causes doubt in the accuracy of the classification studies with small samples (between 0 and 400). Therefore, it is considered necessary to determine the sample size before conducting classification study.

There are some limitations of the study. First, although the dependent variable is continuous by nature, it was transformed into an artificial categorical form by means of two-step cluster analysis. Therefore, the analysis was conducted by the classification accuracy of two-step clustering analysis. The second limitation is that this study is limited to merely two of the classification methods. Finally, the distribution of the data used in this study is not normal. Therefore, it is useful for further studies to consider these limitations and conduct their research.

To sum up the changes of both analysis methods in accordance with sample size for the considered nine criteria within the scope of the study;

- It was observed that positive likelihood rate in logistic regression showed no change and positive likelihood rate in CHAID analysis increased as the sample size increased. In this case, at the beginning of the study, it can be accepted that the CHAID analysis classifies better in large samples.
- Considering the negative likelihood rate, there was no significant change in both analysis methods.
- Considering the change of Type I error rate, which is an important criterion in classification studies, in accordance with sample size, it was observed that Type I error rate decreased in both analysis methods despite increasing sample size. However, it should not be forgotten the reliability and validity of the classification done with the related samples were weak because CHAID analysis exceeded error limits in some sample sizes.
- Considering another criterion Type I error rate, it was seen that logistic regression was independent from sample size and CHAID analysis decreased as sample size increased. Bulut (2015) stated that Type I and Type II error rates decreased as sample size increased. It can be argued that the current study was supported with the literature.
- Considering the confidence level of the test, it was seen that it showed no change as sample size increased while it increased in CHAID analysis. Moreover, when the confidence level was examined, logistic regression was found to have a higher level.
- Considering the power of the test, logistic regression showed no change in accordance with sample size, while it increased in CHAID analysis. Moreover, it was found that CHAID analysis had stronger classification performance than logistic regression. Therefore, if the power of the test is important in classification studies, CHAID analysis might be useful; on the other hand, if the confidence level of the test is important, logistic regression might be useful.
- Last, considering specificity, sensitivity and total accurate classification, logistic regression showed no change in accordance with sample size while CHAID analysis increases. In addition, it was observed that logistic regression made a better classification with small samples.

To conclude, it was observed that logistic regression showed no change, but CHAID analysis showed change as sample size increased. The hypothesis which logistic regression made a better classification with small samples within acceptable error limits (α = 0.05) can be accepted. In a similar vein, the hypothesis which CHAID analysis made a better classification with large samples within acceptable error limits (α = 0.05) can be accepted.

Based on the results obtained from the study, these suggestions can be made:
- In this study, two analysis methods used in the classification studies were used. There are many methods used in classification studies ; therefore, similar studies can be conducted for different analysis methods.
- It is expected that logistic regression would be useful in a study which the confidence level is important (for example, clinical studies), and CHAID analysis would be useful in a study which the power of the test is important (for example, large datasets).
- Considering the limitations of the current study, further studies might contribute to the literature.

Acknowledgements or Notes

References

Akın, A., Kaya, Ç., Uysal, R., Çardak, M., Çitemel, N., Özdemir, E., & Gülşen, M. (2013). *Dikkat Kontrol Ölçeği Türkçe Formu: Geçerlik ve Güvenirlik Çalışması [The Turkish version of the attentional control scale:the validity and reliability study].* Paper presented at VI. National Graduate Education Symposium. Retrieved from http://www.academia.edu/download/43723223/Eitim_Modelinin_renci_zerindeki_Etkilili20160314-25744-1i99q7c.pdf#page=19

Akpınar, H. (2000). Veri tabanlarında bilgi keşfi ve veri madenciliği [Knowledge discovery and data mining in databases]. *Istanbul Business Research*, 29(1), 1-22. Retrieved from https://dergipark.org.tr/tr/pub/ibr/archive

Balcı, A. (2015). *Sosyal bilimlerde araştırma yöntem, teknik ve ilkeler[Research methods, techniques and principles in social sciences]*. Ankara: Pegem Akademi.

Berry M., & Linoff G., (1997). Data Mining Techniques for Marketing Sales and Customer Support. *John Wiley & Sons.*

Brewer S. L. (2012). *An empirical comparison of logistic regression to decision tree induction in the prediction of intimate partner violence reassault.* (Doctoral dissertation). Retrieved from https://www.proquest.com/

Bulut, N. (2015). *İzleme amaçlı klinik araştırmalarda öngörülen ölçütlere göre örneklem büyüklüğünün belirlenmesi [*Determination of sample size by criterias proposed on monitoring in clinical research]. (Master thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/

Çakır, Ö. (2008). *Veri madenciliğinde sınıflandırma yöntemlerinin karşılaştırılması "bankacılık müşteri veri tabanı üzerinde bir uygulama"[Comparison of classification methods in data mining "an application on banking customer database"].* (Doctoral dissertation). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* NJ: Erlbaum Hillsdale.

Deeks, J. J., & Altman, D. G. (2004). Diagnostic tests 4: likelihood ratios. *Bmj*, 329(7458), 168-169. https://doi.org/10.1136/bmj.329.7458.168

Demidenko, E. (2007). Sample size determination for logistic regression revisited. *Statist. Med.,* 26, 3385–3397. https://doi.org/10.1002/sim.2771

Ekici, E. (2012). *Farklı sınıflandırma yöntemlerinin karşılaştırılması ve bir uygulama[An application on the comparison of various classification methods].* (Master thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/

Fajkowska, M. & Derryberry, D. (2010) . Psychometric properties of Attentional Control Scale: The preliminary study on a Polish sample. *Polish Psychological Bulletin*, 41(1), 1-7. https://doi.org/10.2478/s10059-010-0001-7

Finch, H., & Schneider, M. K. (2007). Classification accuracy of neural networks vs. discriminant analysis, logistic regression, and classification and regression trees. *Methodology*, 3(2), 47-57. https://doi.org/10.1027/1614-2241.3.2.47

Grimes, D. A., & Schulz, K. F. (2005). Refining clinical diagnosis with likelihood ratios. *The Lancet*, 365(9469), 1500-1505. https://doi.org/10.1016/S0140-6736(05)66422-7

Heckert, D.A., & Gondolf, E.W. (2005). Do multiple outcomes and conditional factors improve prediction of batterer reassault? *Violence and Victims,* 20 (1), 3-24. https://doi.org/10.1891/vivi.2005.20.1.3

Karakış, R., (2009). *Yapay sinir ağları ve lojistik regresyon yöntemleri ile meme kanseri koltuk altı lenf nodu*

durumunun belirlenmesi[Prediction of the axillary lymph node status in breast cancer using artificial neural network and logistic regression analysis methods]. (Master thesis).   Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/

Kayri, M., & Boysan, M. (2007). Araştırmalarda CHAID analizinin kullanımı ve baş etme stratejileri ile ilgili bir uygulama[Using Chaid analysis in researches and an application pertaining to coping strategies]. *Ankara University Journal of Faculty of Educational Sciences.* 40(2), 133-149. https://doi.org/10.1501/Egifak_0000000146

King, R. D., Feng, C., & Sutherland, A. (1995). Statlog: comparison of classification algorithms on large real-world problems. *Applied Artificial Intelligence an International Journal*, 9(3), 289-333. https://doi.org/10.1080/08839519508945477

Kıran, Z. B. (2010). *Lojistik regresyon ve CART analizi teknikleriyle sosyal güvenlik kurumu ilaç provizyon sistemi verileri üzerinde bir uygulama[An application on pharmacy provision system data of social security institution by logistic regression and CART analysis technics].* (Master thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/

Köktürk, F. (2012). *K-en yakın komşuluk, yapay sinir ağları ve karar ağaçları yöntemlerinin sınıflandırma başarılarının karşılaştırılması[comparing classification success of k-nearest neighbor, artifical neural network and decision trees].* (Doctoral dissertation). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/

Koyuncu, M. S., (2015). *Psikolojik ölçeklerde ROC analizi yöntemiyle standart belirleme[Standard determination in psychological scales using ROC analysis].* (Master thesis).   Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/

Kurt, İ. & Türe, M.(2005). Tıp öğrencilerinde alkol kullanımını etkileyen faktörlerin belirlenmesinde yapay sinir ağları ile lojistik regresyon analizi'nin karşılaştırılması[Comparison of artificial neural networks and logistic regression analysis in determining factors affecting alcohol consumption among medicine students]. *The Balkan Medical Journal.* 22(3), 142-153. Retrieved from https://dergipark.org.tr/en/pub/bmj/issue/3749/49838

Medcalc. (2018). *Software manual*. Retrieved from https://www.medcalc.org/download/medcalcmanual.pdf

Nemes, S., Jonasson, J.M., Genell, A., & Steineck, G. (2009). Bias in odds ratios by logistic regression modelling and sample size. *BMC Medical Research Methodology*, 56(9), 1-5. https://doi.org/10.1186/1471-2288-9-56

Neuilly, M. A., Zgoba, K. M., Tita, G. E., & Lee, S. S. (2011). Predicting recidivism in homicide offenders using classification tree analysis. *Homicide Studies,* 15(2), 154-176. https://doi.org/10.1177/1088767911406867

Pehlivan, G. (2006). *CHAID analizi ve bir uygulama[CHAID analysis and an application].* (Master thesis). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/

Sabzevari, H., Soleymani, M., & Noorbakhsh, E. (2007). A comparison between statistical and data mining methods for credit scoring in case of limited available data. *In Proceedings of the 3rd CRC Credit Scoring Conference (pp. 1-5).*

Stafford, J.D., Kaminski, R.M. , Reinecke K.J., & Gerard, P.D., (2006). Multi-stage sampling for large scale natural resources surveys: a case study of rice and waterfowl. *Journal of Environtmental Management*, 78, 353-361. https://doi.org/10.1016/j.jenvman.2005.04.029

Tabachnick, B.G. & Fidell, L.S. (2013). *Multivariate statistics.* New Jersey: Pearson Education Inc.

Tan, Ş. (2016). *SPSS ve excel uygulamalı temel istatistik-1[Basic statistics-1 with SPSS and excel application].* Ankara: Pegem Akademi. https://doi.org/10.14527/9786053183877

Zurada, J., & Lonial, S. (2005). Comparison of the performance of several data mining methods for bad debt recovery in the healthcare industry. *Journal of Applied Business Research*, 21(2), 37-54. https://doi.org/10.19030/jabr.v21i2.1488

Şeref Tan[1], Sami Pektaş[2]
[1]Gazi University
[2]Niğde Ömer Halisdemir University

To cite this article:

Examining the Invariance of a Measurement Model by Using the Covariance Structure Approach

Şeref Tan[1*], Sami Pektaş[2]
[1]Gazi University
[2]Niğde Ömer Halisdemir University

Abstract

The primary aim of the present study is to examine the measurement invariance of the structural equating model constructed on the numerical and verbal abilities test for sixth grade students across gender, amount of weekly pocket money and students' perceptions of the sufficiency of their pocket money. The secondary aim is to illustrate the use of the IBM AMOS-24 software package step by step with examples to address invariance using the covariance structural analysis approach. The research data were collected from 2304 sixth grade students enrolled in public schools within the Keçiören and Pursaklar suburbs in Ankara. The covariance structures analysis approach was employed during the examination of the measurement model invariance. The study revealed that invariance was achieved in terms of configural, measurement (in both measurement weights and measurement residuals) and structural invariance with respect to all subgroups.

Key words: Configural invariance, Measurement invariance, Structural invariance, Structural equation modeling, Ability.

Introduction

In studies where measurements are used to compare different groups, it is important to ensure measurement invariance. If there is a situation of obtaining biased measurement results for a subgroup, determining this situation allows the research to be interpreted the findings more accurately. In this study, an example of how the measurement invariance study can be performed by using IBM AMOS-24 for an ability test scores is presented. For this purpose, first theoretical information about ability test and then some theoretical information about measurement invariance analysis are explained in this section.

The concept of ability has been defined in various ways. The concept of ability, which refers to mental computation activities, can be categorized based on the common properties of separate factors essential for individuals to carry out mental operations. The mental power required for each of these identified groups is called ability. It is defined as being inherited, a boundary enveloping learning, and the power to accommodate the effect originating from external factors (Turkish Language Association [TDK], 2016). As ability conveys differences among individuals and reflects the development of a process, the measurement of abilities is considered important. Thus, tests that could create the opportunity to examine individuals' abilities and identify differences among individuals were developed. Ability tests typically measure knowledge and skills acquired over long periods of time whereas so-called achievement tests are often "subject/topic" specific and may require more recent targeted study to perform well (Benson, 2008; Ghanizadeh, 2017; Kaufman, 2018). Another stream of literature shows that noncognitive skills are important determinants of performance in achievement tests (McGrew, 2005; Borghans, Golsteyn, Heckman & Humphries 2016). Researchers consider that the tests, which measure cognitive skills, abilities and intelligence, will show a positive correlation with achievement points. In this case, the scores obtained by achievement tests are used as the points that represent cognitive ability (Berkowitz & Stern, 2018; Kyllonen & Kell, 2018). In addition to ability tests that can measure more than one ability, there are also ability tests that measure specific characteristics (R. Atkinson, Atkinson, & Hilgard, 1995). Ability tests are categorized into two groups based on the distinction between individuals' abilities as general and specific. While ability tests are observed to have a homogeneous structure when intelligence tests are examined, they also appear to display a heterogeneous structure since they also measure such characteristics as language, number and reasoning as well as all the characteristics of intelligence (Özgüven, 2007).

---

[*] Corresponding Author: *Şeref Tan, sereftan4@yahoo.com*

Specific ability is defined as the power individuals embody to realize behaviors with specific characteristics. It is individuals' power to benefit from this condition since it gives them the opportunity to achieve certain things in different areas of ability (Yeşilyaprak, 2007). On the other hand, general ability can be defined as the general capacity that includes such abilities as solving problems, solving arithmetic computations, thinking abstractly, reasoning, establishing association between words, finding synonyms, and can more or less affect all the behaviours of an individual. This ability is equivalent to the general intelligence, also known as g factor (g), proposed by Charles Spearman and, hence, it can also be defined as general mental power (Sak, 2014, p. 42). Cattell-Horn-Carroll theory classifies cognitive abilities into three levels, which are narrow abilities, broad abilities and general ability. Cognitive ability refers to the mechanisms of mental capacity such as remembering, problem solving than actual knowledge (McGrew, 2005; Floyd, Keith, Taub & McGrew, 2007)

Measurement of the same characteristic across different groups during the construction stage of the tests developed is examined under the heading of measurement invariance. As Millsap & Tein (2004) pointed out that, the extension of the analysis to the multiple-population case is less well known especially for ordered-categorical data in the literature on factor analysis. When we want to compare different groups, it should be proved that the scores obtained from the scale are not biased. As Camilli (2006) pointed out that measurement invariance contributes to validity evidence in that scores from a tool are subject to issues of bias and lack of fairness if invariance does not hold. Because of this reason, as Chung and others, (2016) mentioned, further investigation is needed to answer the question of whether the scale items perform similarly across subgroups and one way to examine this question is through assessing the measurement invariance of a scale. There are studies conducted on measurement invariance to decide whether or not items on a test express the same meaning across different groups (Arana, Rice & Ashby, 2018; Başusta & Gelbal, 2015; Camerota, Willoughby, Kuhn & Blair, 2018; Chavez, Shrout, Garcia, Forno & Celedón, 2018; Gaddy, Casner & Rosinski, 2018).

Wicherts (2016) states that measurement invariance is a fundamental problem in identifying whether or not population norms are valid for sub-groups as well. In more than half of the studies he examined, he concluded that in intelligence tests, there was no measurement invariance across the groups by ethnic origin, gender, educational condition, and age. He underlined the fact that measurement invariance is very important for the validity of neurocognitive tests in clinical, educational and professional practices. Wicherts and Dolan (2010) stated that the fair use of measurement invariance across groups was very important in intelligence tests and other psychological tests. They mentioned that there is common belief that the factor loadings in confirmatory factor analysis are sufficient to ensure measurement invariance. At this point, they indicate that in constructing measurement invariance by means of confirmatory factor analysis, there is a need for a statistical test for the equation of the groups measured. Blankson and McArdle (2013) used cognitive performance tests measuring episodic memory and mental status and tested these by using multilevel modeling for two-factor structural invariance, confirmatory factor analyses and longitudinal data. In a study they conducted with 244 undergraduate students, Bailey, Neigel, Dhanani and Sims (2018) applied two spatial ability (spatial visualization and projection) measurements on a computer- and a paper-based format. By ensuring the measurement invariance in both paper- and computer-based tests to measure spatial ability, they aimed to ensure reliability. It was found that based on the type of test implemented, measurement invariance could not be ensured and that the way the test was implemented had an impact on different types of errors. Furthermore, when compared to the computer-based tests, the paper-based tests were found to be more reliable. Since the existing proof cannot reach the same structure in a reliable way, in such tests as ability tests, they mentioned the necessity of conducting measurement invariance.

During the process of testing structures that want to be measured, the condition where individuals from different subgroups have equal chance of achieving a certain score is referred to as measurement invariance (Watson, Thompson & Adam, 2007). For equal measurements, the connection between the observed and latent variable should be the same (Drasgow & Kanfer, 1985). For a measurement model to have the same structure across different groups, the factor loadings of the items in a scale, and the correlations and variances among the identified factors, should be the same. During the examination of between-groups measurement invariance of the measurement model constructed, the model constructed in each phase is built on the model constructed in the previous phase. Accordingly, the measurement invariance examined at a certain phase is examined based on the model in the previous phase, in with fewer restrictions are placed, by using the research data to test whether a significantly lower level of model fit is displayed. If it displays a good level of model fit with the data – as good as that of the previous model in which more restrictions were placed – then, it is believed that the more complex model can explain the data. The examination conducted shows that the measurement invariance in that phase was realized (Cheung & Rensvold, 2002). Measurement invariance proposed by Milfont and Fischer (2010) is addressed under seven titles, namely configural, metric, scalar, error variances, factor variances, factor

covariances and invariance of factor means. Vandenberg and Lance (2000) also address measurement invariance under seven terms: configural, metric, scalar, uniqueness, factor invariance, factor covariance and invariance of factor means. On the other hand, in some research, measurement invariance is addressed under three titles: configural, metric and scalar invariance (Campbell, Berry, Joe & Finney, 2008; Xu & Tracey, 2017). In some other studies, generally measurement invariance is addressed under four headings, namely configural, metric, scalar and strict invariance (Meredith, 1993; Wu, Li and Zumbo, 2007).

In this regard, within the scope of research in the literature defined as configural, metric, scalar and strict invariance, measurement invariance is addressed under four headings; namely unconstrained model (configural invariance), measurement weighted model (metric invariance), structural covariances model (scalar invariance) and measurement errors model (strict invariance).

> As stated by Byrne (2016: 227-228), *"In seeking evidence of multigroup equivalence, researchers are typically interested in finding the answer to one of five questions. First, do the items comprising a particular measuring instrument operate equivalently across different populations? In other words, is the measurement model group-invariant? Second, is the factorial structure of a single instrument or of a theoretical construct equivalent across populations? Third, are certain paths in a specified causal structure equivalent across populations? Fourth, are the latent means of particular constructs in a model different across populations? Finally, does the factorial structure of a measuring instrument replicate across independent samples drawn from the same population? This latter question addresses the issue of cross-validation."*

Configural invariance refers to whether or not the constructed model is the same across all groups. As Chung and others (2016) pointed out, configural invariance is the fact that factor structures between groups are equivalent. In other words, configural invariance tests that the same pattern of item-factor loadings exists across the groups being compared, which requires that the same items have nonzero loadings on the same factors (Chung and others, 2016). The model consisting of constant and free parameters is equal across the groups in the model at the step. Since it the most fundamental structure, it is also referred to as unconstrained model; it is also known as the initial model in measurement invariance analyses. To observe whether the other steps of invariance are ensured, comparisons are made based on the configural invariance values (Cheung & Rensvold, 2002; Vandenberg & Lance, 2000).

Metric invariance refers to the equalivance among regression coefficients, that is, factor loadings. As Chung and others (2016) mentioned, metric invariance, additionally requires that unstandardized factor loadings are the same across groups. Metric invariance identifies the invariance of the factor loadings across the groups; that is, it determines whether or not the responses given to the latent variables are equivalent. In addition to the factor loadings, it also refers to the equivalence across the factor loadings (Byrne, 2016; Meredith, 1993).

Scalar invariance is based on the equivalence of factor covariances across groups. It is the model where all factor loadings, factor variances and factor covariances are constrained. It is a kind of invariance where factor covariances are equalized across the groups after configural and metric invariances are ensured (Cheung & Rensvold, 2002; Meredith, 1993).

Strict invariance is also referred to as invariant uniqueness. It is based on the principle that the error terms are equivalent across the comparative groups. It is based on the error variance equivalence after configural, metric, and scalar invariances are ensured. It is a type of invariance where all factor loadings, factor variances, factor covariances and error variances are constrained (Cheung & Rensvold, 2002; Meredith, 1993; Vandenberg & Lance, 2000; Wu, Li & Zumbo, 2007).

As we mentioned before, if measurement invariance is not provided, the results would be biased. For this reason, measurement invariance is an important property that should be examined when developing or using a measurement tool. Measurement invariance can be used with tests and scales based on two fundamental theories such as the measurement invariance structural equating model and item response theory. In the present article, the stage of identifying measurement invariance was realized by utilizing the covariance structural analysis (COVS) approach. Only the variances in observed variables and the covariances among observed variables are made use of in covariance structural analysis.

Aim of the Study

The present study aimed to examine the measurement invariance of the structural equating model constructed on the numerical and verbal abilities tests for sixth grade students across gender, amount of weekly pocket money, and students' perceptions of the sufficiency of their pocket money by using the covariance structural analysis (COVS). This study also aims to illustrate the use of the IBM AMOS-24 software package step by step with examples to address invariance using the covariance structural analysis approach.

Method

The Research Model

The aim was to construct a structural model of the students' numerical and verbal ability tests and the sub-factors comprising these tests. Thus, in this respect, the research is based on a correlational study. It is descriptive in nature as it examines the measurement invariance aspect of the structural equating model based on the groups formed based on student variables.

The Study Group

The study group was comprised of 2304 sixth grade students enrolled in public schools within the Keçiören and Pursaklar suburbs in 2016-2017 academic year, Ankara. Of these students, 1146 (49.7%) were female students, while 1158 (50.3%) were male students. With respect to amount of pocket money, 608 (26.4%) students had a low (0-5 TL), 1118 (48.5%) students had a moderate (6-20 TL), and 578 (25.1%) students had a high amount of pocket money (21 TL and above). In terms of students' perceptions regarding the sufficiency of their pocket money, 1914 (83.1%) students stated that it was sufficient, whereas 390 (16.9%) students claimed that it was not sufficient.

Data Collection Instruments

To measure students' numerical and verbal abilities within the scope of the present study, the data were collected by means of the "Numerical Ability Test" and "Verbal Ability Test", developed by Pektaş (2018) within the scope of his Ph.D. dissertation study. This test is used to show how to perform measurement invariance analysis. In this test, the fact that it contains the actual data set is important for more realistic analysis compared to the simulation data. The numerical test consisted of 45 multiple choice test items and four sub-factors: patterns (*pattern*), finding the four operations based on the symbols (*operation's symbol*), finding the symbols used in the four operations (what is the *symbol*) and problem solving (*problem*). The verbal ability test consisted of two sub-factors, vocabulary and inter-word relationship, and of 45 multiple-choice items. When the test statistics of the pilot study for the numerical ability test were examined, it was observed that the difficulty level of the test was at moderate level with a difficulty index of 0.51, and that it was found to sufficiently discriminate between students with high and low levels of numerical ability with an average upper-lower group item discrimination index value of 0.43. The KR-20 reliability coefficient of the numerical ability test was 0.90, which shows that the measurements had a high degree of reliability. When the test statistics of the pilot study for the verbal ability test was examined, it was observed that the difficulty level of the test was again at a moderate level with a difficulty index of 0.58, and that it was found to sufficiently discriminate between students with high and low levels of verbal ability with an average upper-lower group item discrimination index value of 0.40. The KR-20 reliability coefficient of the verbal ability test was 0.85, which shows that the measurements had a high degree of reliability. During the confirmatory analysis phase of the four-observed variables of numerical ability test, the fit indices of the first level with these 4-observed variable, the fit indexes of this CFA model were found to be as follows: $\chi^2_{937} = 3018.356$; $\chi^2/df = 3.22$; CFI=.971, GFI=.932, AGFI=.924 and RMSEA=.033. In the confirmatory analysis phase of the two-observed variables of the verbal ability test, the fit indices of this CFA model were observed as the following: $\chi^2_{940} = 2838.217$; $\chi^2/df = 3.02$; CFI=.976, GFI=.939, AGFI=.933 and RMSEA=.030. When these are compared to the criteria defined by Byrne (2013) and Schermelleh-Engel, Moosbrugger & Müller (2003), they are considered acceptable and they display a perfect fit.

Data Analysis

The data collected from the 6th grade students were entered into the IBM SPSS-25 package program. In the examination of the measurement model invariance, the covariance structural analysis approach was utilized. The data were examined for missing data and outliers, and the mahalanobis distance values were examined. Tests for normality and multicollinearity were also conducted. The covariance matrix, the asymptotic covariance matrix and average vectors were computed for each subgroup. In the study, maximum likelihood approach was used as the parameter estimation method. During the construction of the measurement model and the invariance test phase, the IBM AMOS-24 package program was utilized. For the comparison of the measurement invariance in the measurement model, the changes in the criterion of CFI (ΔCFI) values were taken into consideration. In the first stage of the measurement model, the multivariate normality assumption was tested for each subgroup, and each kurtosis value obtained for each group was observed to be below 1.00. The difference between the CFI value in the configural model and the CFI value of the models tested in the later stages was less than .01. Based on the conditions for ensuring measurement invariance, this has been accepted as proof for the presence of measurement invariance (Cheung & Rensvold, 2002). Also a difference of less than .01 in the ΔCFI index supports the less parameterized model (Chung and others, 2016). The measurement invariance approach and the interpretations explained in the present article can be analyzed in models constructed with such measurement tools as ability tests, achievement tests, scales, psychological tests with the aim of comparing different groups. During the analyses, the operations done via the IBM AMOS-24 package program are explained as follows:

IBM AMOS-24 operations for configural invariance. In the first step, the groups are defined by selecting the *Manage Groups.* Function from the *Analyze* menu in the *AMOS program:* (***Analyze***, ***Manage Groups***, In Group Name Box **type** the name of the first group as ***Male***, click ***New*** icon and **type** the name of the second group as ***Female***, then click ***Close***). In the second step, the data files for male and female sub-groups are assigned by using the *Select Data File(s)* icon or by using function from the ***File, Data Files*** menu.
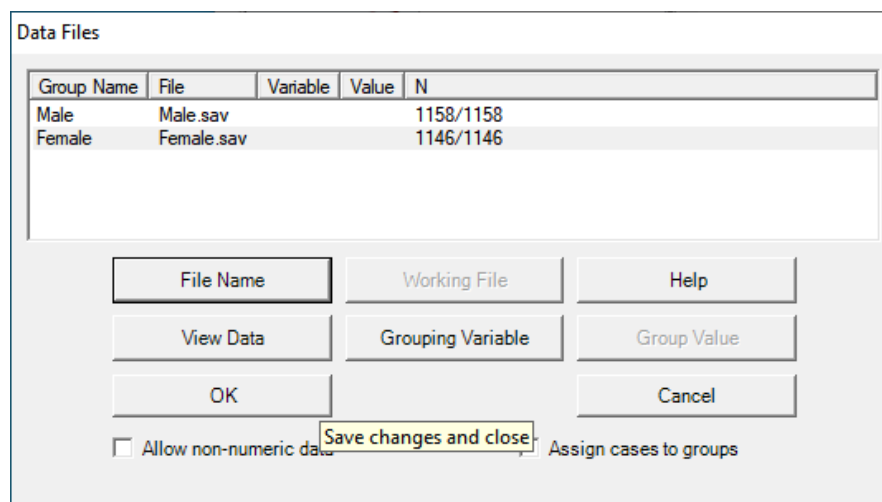
In third step, the ***Emulisrel6*** box is clicked by selecting ***Estimation*** from ***Analysis Properties*** in the ***View*** menu. In final step, the analysis is run by selecting ***Calculate Estimates*** from the ***Analyze*** menu.

IBM AMOS-24 operations for configural and structural invariance. Until the stage of making the predictions, as an addition to the operations mentioned above, the parameters to be predicted in the model are labeled manually or automatically. For automatic labeling, the ***Multiple Group Analysis*** function is selected from the ***Analyze*** menu. Then, the parameters to be constrained are selected in the *Multiple-Group Analysis* dialog box.

In final step, the analysis is run by selecting ***Calculate Estimates*** from the ***Analyze*** menu.

Findings and Interpretations

Measurement Model

Measurement invariance was tested both for the observed 6 sub-scale scores (For the numerical latent variable: *pattern, operation's symbol*, *what is the operation* and *problem* observed variables; for the verbal latent variable: *vocabulary* and *inter word relationship* observed variables) and for the latent structure related to these subscales (the relationship between numerical ability and verbal ability). The baseline model was displayed in Figure 1 below.
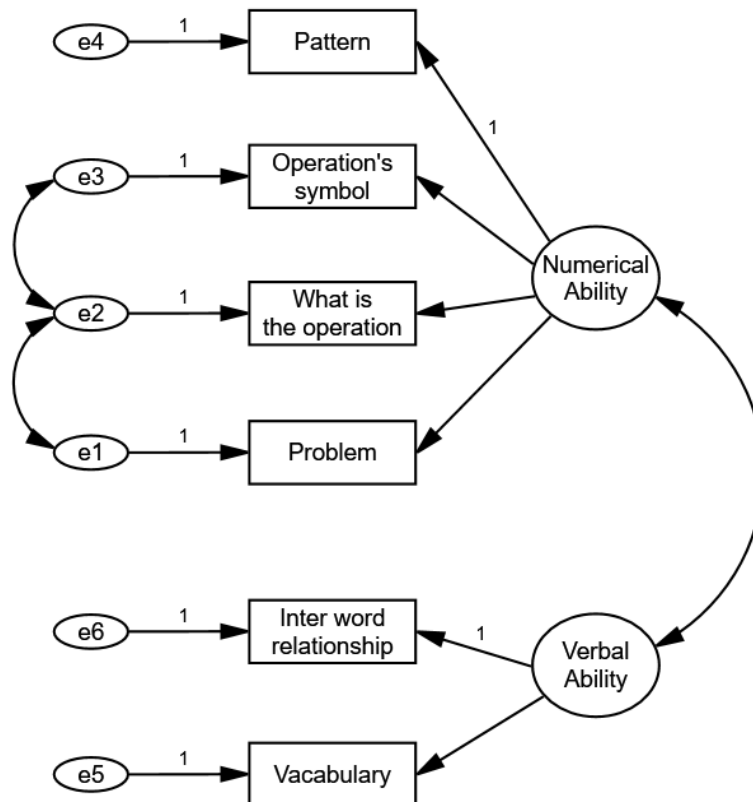
Figure 1. The baseline model for the multiple-group invariance of the numerical and verbal ability measurement model.

Measurement Invariance by Gender

The Baseline Model for the Measurement Invariance

The first stage, of the measurement invariance analysis to be conducted in various stages, is the identification of an appropriate baseline measurement model for each group. The baseline model is portrayed in Figure 1. In the baseline model of the measurement model based on 6th grade students' gender, the fit indices for female students were found to be as follows: χ^2_6 =4.351; χ^2/df =.725; GFI=.999; CFI=1.000 and RMSEA=.000. As for male students, they were found to be as follows: χ^2_6 =15,335; χ^2/df =2.556; GFI=.996; CFI= .997 and RMSEA=.037. In the baseline model, theoretically reasonable two covariance links (e1 with e2 and e2 with e3) between residuals of the same factor were used in order to have better-fit indexes for all subgroups in the study. In conclusion, the baseline model in Figure 1 displayed a high level of model fit for both female and male students.

Configural Invariance of the Measurement Model for Gender

As stated by Byrne (2016), to ensure configural invariance, factor loading patterns and the number of factors should be similar for each group. The measurement model based on students' gender has ensured configural invariance: χ^2_{12} =19.686; χ^2/df =1.641; GFI=.997; CFI=.999 and RMSEA=.017. That is, in this unconstrained measurement model, the factor structure for these two populations based on gender was found to be similar.

The unstandardized estimated parameters (regression weights, covariances, and variances) of Male and Female groups for configural invariance are given in Tables 1a, 1b and 1c, below.

Table 1a. Regression weight estimates of gender for configural model

Regression Weights			Male	Female
Problem	<---	Numerical Ability	.981**	1.093**
Pattern	<---	Numerical Ability	1.000	1.000
Vocabulary	<---	Verbal Ability	.651**	.786**
Inter word relationship	<---	Verbal Ability	1.000	1.000
What is the Operation	<---	Numerical Ability	.509**	.574**
Operation Symbol	<---	Numerical Ability	.690**	.698**

**:p<.01

Table 1b. Covariance estimates of gender for configural model

Covariance			Male	Female
Numerical Ability	<-->	Verbal Ability	9,308**	7,963**
e1	<-->	e2	1,013**	,720**
e2	<-->	e3	,501**	,445**

**:p<.01

Table 1c. Variance estimates of gender for configural model

Covariance	Male	Female
Numerical Ability	7,985**	7,367**
Verbal_Ability	21,469**	16,005**
e1	7,233**	7,240**
e2	2,385**	2,337**
e3	2,295**	2,885**
e4	4,969**	4,365**
e5	7,890**	5,803**
e6	5,642**	6,645**

**:p<.01

Configural, Measurement and Structural Invariance of the Measurement Model for Gender

As it was given by Byrne (2016), in the measurement and structural invariance test, the focus is more on which parameters in the measurement model and its structural constituents are equivalent in both groups. In this part of the analysis, progressively, first *measurement weights* are constrained, then *structural covariances* are constrained and finally *measurement errors* are constrained. The IBM AMOS-24 output path diagram of the unstandardized estimated parameters when all the model parameters are constrained equal is given in Figure 2. The results obtained for the measurement invariance by gender in terms of factor loadings (measurement weights), structural covariances and measurement errors are presented in Table 2, below.
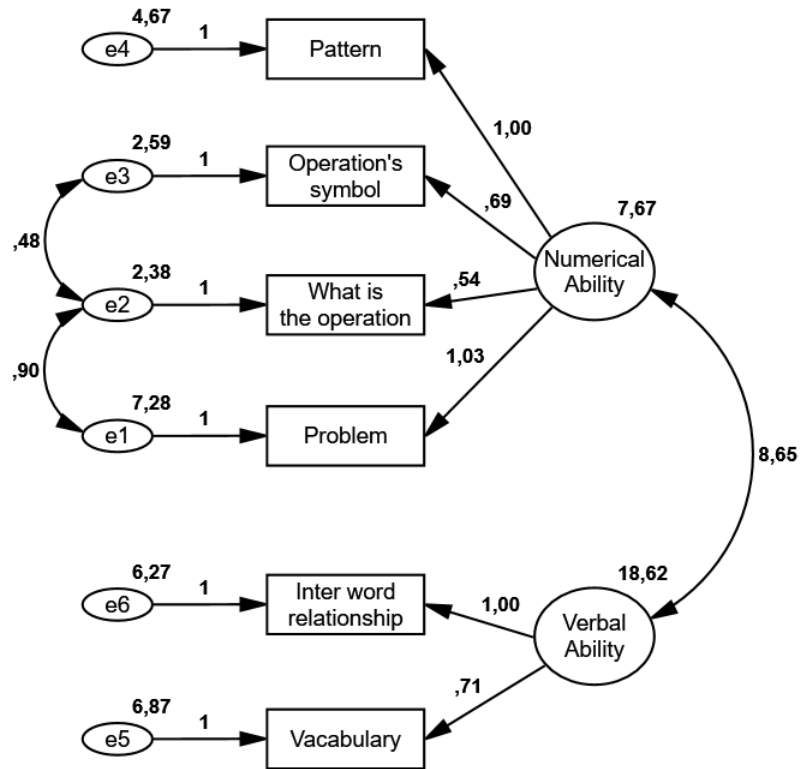
Chi square = 63,715, Df= 27, GFI= ,991, CFI= ,994, RMSEA= ,024

Figure 2. Output path diagram for configural, measurement weight, structural covariance, and measurement error invariance of the measurement model for gender.

Table 2. Measurement and structural invariance results by gender.

Model	Number of Parameters	χ^2	df	X^2/df	CFI	ΔCFI	RMSEA
1. Unconstrained (Configural)	30	19.686	12	1.641	.999		.017
2. Measurement Weights	26	33.334	16	2.083	.997	**0.002**	.022
3. Structural Covariances	23	37.196	19	1.958	.997	**0.002**	.020
4. Measurement Errors	15	63.715	27	2.36	.994	**0.005**	.024

Note:
Unconstrained Model: All the parameters are predicted freely.
Measurement Weights Model: All *Factor loadings* are constrained (equated).
Structural Covariances Model: All Factor loadings + *factor variances and covariances* are constrained (equated).
Measurement Errors Model: All Factor loadings + factor variances + factor covariances + *error variances* are constrained (equated).

According to the unconstrained model used for configural invariance, as initially only the four factor loadings (measurement weights) predicted for the measurement model are to be defined as being equivalent for the two groups, the number of parameters estimated by the measurement model in which the factor loadings are equated is reduced by 4 when compared to the configural model and, hence, the predicted number of parameter reduced to 26. In addition, owing to structural variances and covariances, numerical and verbal latent variables are to be defined as being equivalent for the two latent variances and 1 covariance; thus, the number of predicted parameters is reduced by 3, yielding 23 parameters. Finally, since 6 error variances and two error covariances are to be predicted once for each group, the parameters to be predicted are reduced by 8, yielding 15 parameters. As can be seen from Table 2, there is an increase in degrees of freedom as much as a decrease in the number of parameter predicted in the model. As can be observed in Table 2, according to the unconstrained (used in configural invariance) model, the changes in CFI in the models obtained by constraining, in sequence, measurement weights, structural covariances and error variances, were less than .01. Hence, the measurement model has ensured configural, measurement and structural invariance across gender.

Measurement Invariance by Amount of Weekly Pocket Money

The Baseline Model for Measurement Invariance

Based on students' amount of weekly pocket money, the constructed baseline model yielded the following fit indices for the group with low amount of pocket money: χ^2_6 =6.334; χ^2/df =1.056; CFI=1.000 and RMSEA=.010; for the group with a moderate amount of pocket money: χ^2_6 =10.924; χ^2/df =1.821; CFI= .998 and RMSEA=.027; and for the group with a high amount of pocket money: χ^2_6 =5.616; χ^2/df =0.939; CFI=1.000 and RMSEA=.000. Hence, the measurement model has ensured a high level of model fit for all three groups- a low, moderate, high amount of pocket money–in the baseline model.

Configural Invariance of the Measurement Model for Amount of Weekly Pocket Money

The measurement model has ensured configural invariance with respect to students' amount of weekly pocket money: χ^2_{18} =22.873; χ^2/df =1.271; CFI=.999 and RMSEA=.011.

Configural, Measurement and Structural Invariance of the Measurement Model for Amount of Weekly Pocket Money

The results obtained for the measurement invariance by students' amount of weekly pocket money in terms of factor loadings, structural covariances and measurement errors are presented in Table 3 below.

Table 3. Measurement and structural invariance results by amount of weekly pocket money.

Model	Number of Parameters	χ^2	df	χ^2/df	CFI	**ΔCFI**	RMSEA
1. Unconstrained (Configural)	45	22.873	18	1.271	.999		.011
2. Measurement Weights	37	30.314	26	1.166	.999	**0.000**	.008
3. Structural Covariances	31	36.276	32	1.134	.999	**0.000**	.008
4. Measurement Errors	15	57.596	48	1.2	.998	**0.001**	.009

As can be observed in Table 3, according to the unconstrained model, the changes in CFI in the models obtained by constraining, in sequence, measurement weights, structural covariances and error variances, were less than .01. Hence, the measurement model has ensured configural, measurement and structural invariance based on amount of weekly pocket money.

Measurement Invariance by Sufficiency of the Amount of Weekly Pocket Money

The Baseline Model for Measurement Invariance

The fit indices that the baseline model of the measurement model based on whether students' amount of weekly pocket money was sufficient yielded the following fit indices for students who claimed it was insufficient: χ_6^2=3.148; χ^2/df =.525; CFI=1.000 and RMSEA= .000; for students who claimed it was sufficient: χ_6^2=10.560; χ^2/df =1.76; CFI=.999 and RMSEA=.020. Hence, the baseline model in Figure 1 displayed a high level of model fit for both groups who found the amount of their weekly pocket money either sufficient or insufficient.

Configural Invariance of the Measurement Model for Sufficiency of the Amount of Weekly Pocket Money

The fit indices that the measurement model yielded for configural invariance with respect to the status of whether the amount of students' weekly pocket money was sufficient are as follows: χ_{12}^2=13.708; χ^2/df =1.142; CFI=1.000 and RMSEA=.008 Hence, configural invariance has been ensured.

Configural, Measurement and Structural Invariance of the Measurement Model for Sufficiency of the Amount of Weekly Pocket Money

The results obtained for the measurement invariance by the sufficiency of students' amount of weekly pocket money in terms of factor loadings, structural covariances and measurement errors are presented in Table 4 below.

Table 4. Measurement and structural invariance results by sufficiency of the amount of weekly pocket money.

Model	Number of Parameters	X^2	df	X^2/df	CFI	ΔCFI	RMSEA
1. Unconstrained (Configural)	30	13.708	12	1.142	1.000		.008
2. Measurement Weights	26	13.948	16	0.872	1.000	**0.000**	.000
3. Structural Covariances	23	16.282	19	0.857	1.000	**0.000**	.000
4. Measurement Errors	15	36.810	27	1.363	.998	**0.001**	.013

As can be observed in Table 4, according to the unconstrained model, the changes in CFI in the models obtained by constraining, in sequence, measurement weights, structural covariances and error variances, were less than .01. Hence, the measurement model has ensured configural, measurement and structural invariance based on sufficiency of the amount of weekly pocket money.

Conclusion and Discussion

This study investigates the measurement invariance of the numerical ability and verbal ability tests across some subgroups, explained above. For this purpose, the configural, measurement (both measurement weights and measurement residuals) and structural invariance across the sub-groups was tested for the multiple-group measurement invariance of the constructed measurement model by using the covariance structural analysis approach. In numerous decisions to be taken in the field of education, the variable of ability is used as an important measurement. Using measurement invariance to test whether models constructed with psychological tests like ability operate in the same way in different subgroups and to report the findings are of utmost importance in order to avoid biases in serious decisions to make. As seen in the explanatory examples, the IBM AMOS-24 program is a very practical and user-friendly for measurement invariance and provides very detailed outputs regarding the invariance stages.

Having examined the measurement model invariance with respect to configural, measurement, and structural invariance across the groups in terms of students' gender, the amount of pocket money they received from their family, and their perceptions of whether the amount of pocket money was sufficient, the present study reached at the conclusion that both configural and measurement and structural invariances were ensured. The measurement model constructed with the data obtained from the numerical ability and verbal ability student tests developed by Pektaş (2018), by using the covariance structural analysis and has perfectly met all the possible measurement invariances. Studies that examined different sub-groups regarding such ability tests as

psychological tests can be encountered in the related literature. Wicherts (2016) mentioned that neurocognitive test batteries, such as the up-to-date version of Wechsler's Adult Intelligence Scale (WAIS), were used in population-based norms. The main fundamental question focused on whether the implemented test batteries operated in the same way with different sub-groups as gender, age, educational background, socioeconomic status, ethnicity, mother tongue, and race. Based on the studies they reviewed, Wicherts and Dolan (2010) reported that overlooking group variations based on ethnicity in intelligence tests can lead to a high degree of bias in terms of minorities, and that in comparisons based on group differences in intelligence tests, it is essential to initially analyze measurement invariance. In their study titled 'Health and Retirement Study/Asset and Health Dynamics among the Oldest Old' (HRS/AHEAD), Blankson and McArdle (2013) aimed to test the invariance of cognitive variables across ethnic origin, gender and time. The analyses were done using a selected sub-sample of the HRS/AHEAD data set. Metric invariance was ensured across time. By means of measurement invariance, the invariance of cognitive talent measurements based on HRS/AHEAD was better understood. Since measurement invariances were provided in our examples, partial measurement invariance analyzes or additional analyzes for measurement invariance sources were not included.

As a conclusion, the results of this study provide evidence that the measurement invariance requirement for valid group comparisons has been satisfied. Proving the invariances of errors as well in the measurement model constructed to test numerical and verbal abilities also proves that the reliability of these numerical and verbal ability tests developed for sixth grade students does not vary across the sub-groups examined. The present study on students also proves that the characteristics examined as a group variable or the particular group one had fallen in did not create any bias in terms of sub-scale and scale scores in the measurement of numerical and verbal abilities. Finally, the measurement invariance approach and the interpretations explained in this article can be applied to measurement tools in all studies aiming to compare different groups. Hence, as it explains and exemplifies measurement invariance at the same time, the article also sheds light on an important issue for measurements that will be used as support for important decisions to be taken.

References

Arana, F. G., Rice, K. G., & Ashby, J. S. (2018). Perfectionism in Argentina and the United States: Measurement structure, invariance, and implications for depression. *Journal of Personality Assessment*, *100*(2), 219-230. https://doi.org/10.1080/00223891.2017.1296845

Atkinson, R., Atkinson, R., & Hilgard, E. (1995). *Psikolojiye giriş II [Introduction to Psychology II].* İstanbul: Sosyal.

Bailey, S. K., Neigel, A. R., Dhanani, L. Y., & Sims, V. K. (2018). Establishing measurement equivalence across computer-and paper-based tests of spatial cognition. *Human factors*, *60*(3), 340-350.https://doi.org/10.1177/0018720817747731

Başusta, N. B., & Gelbal, S. (2015). Gruplararası karşılaştırmalarda ölçme değişmezliğinin test edilmesi: PISA öğrenci anketi örneği [Testing measurement invariance in intergroup comparisons: PISA student survey sample]. *Hacettepe University Journal of Education*, *30*(4), 80-90.

Benson, N. (2008). Cattell-Horn-Carroll cognitive abilities and reading achievement. *Journal of Psychoeducational Assessment*, *26*(1), 27-41. https://doi.org/10.1177/0734282907301424

Berkowitz, M., & Stern, E. (2018). Which cognitive abilities make the difference? Predicting academic achievements in advanced STEM studies. *Journal of Intelligence*, *6*(4), 48. https://doi.org/10.3390/jintelligence6040048

Blankson, A. N., & McArdle, J. J. (2013). Measurement invariance of cognitive abilities across ethnicity, gender, and time among older Americans. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *70*(3), 386-397. https://doi.org/10.1093/geronb/gbt106

Borghans, L., Golsteyn B. H., Heckman J. J., & Humphries J. E. (2016). What grades and achievement tests measure. *Proceedings of the National Academy of Sciences, 113* (47), 13354–13359. https://doi.org/10.1073/pnas.1601135113

Byrne, B. M. (2013). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Psychology Press.

Byrne, Barbara M. (2016) *Structural equation modeling with AMOS*: *Basic concepts*, applications, and programming. Routledge.

Camerota, M., Willoughby, M. T., Kuhn, L. J., & Blair, C. B. (2018). The childhood executive functioning inventory (CHEXI): Factor structure, measurement invariance, and correlates in US preschoolers. *Child Neuropsychology*, *24*(3), 322-337. https://doi.org/10.1080/09297049.2016.1247795

Camilli, G. (2006). Test fairness. In R. L. Brennan's (Ed.), Educational measurement (pp. 221–256). Westport: Praeger.

Campbell, H., Barry, C. L., Joe, N. J., & Finney, J. S. (2008). Configural, metric and scalar invariance of the modified achievement goal questionnaire across African American and White university students. *Educational and Psychological Measurement, 68*(6), 988-1007. https://doi.org/10.1177/0013164408318766

Chavez, L. M., Shrout, P. E., García, P., Forno, E., & Celedón, J. C. (2018). Measurement invariance of the adolescent quality of life-mental health scale (AQOL-MHS) across gender, Age and Treatment Context. *Journal of Child and Family Studies*, *27*(10), 3176-3184. https://doi.org/10.1007/s10826-018-1158-5

Cheung, G.W. & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Chung, H., Kim, J., Park, R., Bamer, A. M., Bocell, F. D., & Amtmann, D. (2016). Testing the measurement invariance of the University of Washington Self-Efficacy Scale short form across four diagnostic subgroups. *Quality of life research*, *25*(10), 2559-2564. https://doi.org/10.1007/s11136-016-1300-z

Drasgow, F., & Kanfer, R. (1985). Equivalence of psychological measurement in heterogeneous populations. *Journal of Applied Psychology, 70*(4), 662-680. http://dx.doi.org/10.1037/0021-9010.70.4.662

Floyd, R. G., Keith, T. Z., Taub, G. E., & McGrew, K. S. (2007). Cattell-Horn-Carroll cognitive abilities and their effects on reading decoding skills: g has indirect effects, more specific abilities have direct effects. *School Psychology Quarterly*, *22*(2), 200-233. https://doi.org/10.1037/1045-3830.22.2.200

Gaddy, M. A., Casner, H. G., & Rosinski, J. (2018). Factor structure and measurement invariance of the Brief Addiction Monitor. *Journal of substance abuse treatment*, *90*(7), 29-37. https://doi.org/10.1016/j.jsat.2018.04.010

Ghanizadeh, A. (2017). The interplay between reflective thinking, critical thinking, self-monitoring, and academic achievement in higher education. *Higher Education*, *74*(1), 101-114. https://doi.org/10.1007/s10734-016-0031-y

Kaufman, A. S. (2018). *Contemporary intellectual assessment: Theories, tests, and issues*. Guilford Publications.

Kyllonen, P., & Kell, H. (2018). Ability Tests Measure Personality, Personality Tests Measure Ability: Disentangling Construct and Method in Evaluating the Relationship between Personality and Ability. *Journal of Intelligence*, *6*(3), 32. https://doi.org/10.3390/jintelligence6030032

McGrew, K. S. (2005). The Cattell-Horn-Carroll Theory of Cognitive Abilities: Past, Present, and Future. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (p. 136–181). The Guilford Press.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525-543. http://dx.doi.org/10.1007/BF02294825

Millfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research, 3*(1), 111-121.

Millsap, R., & Tein, J. Y. (2004). Model specification and identification in multiple-group factor analysis of ordered-categorical measures. *Multivariate Behavioral Research*, *39*(3), 479-515. DOI: 10.1207/S15327906MBR3903_4

Özgüven, İ. E. (2007). *Psikolojik testler [Psychological tests]*. Ankara: Nobel.

Pektaş, S. (2018). *The effects of differential item functioning determination methods on test parameters estimates, decision studies, g and phi coefficients*. (PhD Theses). Ankara: Gazi University Graduate School of Educational Sciences Department of Measurement and Evaluation in Education.

Sak, U. (2014). *Üstün zekâlılar: Özellikleri tanılanmaları eğitimleri* [The gifted: Traits identification training]. Ankara: Vize.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online*, *8*(2), 23-74.

Türk Dil Kurumu (2016). *Yetenek [Talent]*. It is obtained on 26.10.2019 at *http://www.tdk.gov.tr/index.php?option=com_gts&arama=gts&guid=TDK.GTS.5c27e8237eb1a0.18983 524* .

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

Watson, A. M., Thompson, L. F., & Meade, A. W. (2007, April). Measurement invariance of the job satisfaction survey across work contexts. paper presented at the 22nd annual meeting of the society for industrial and organizational psychology, New York. It is obtained at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.583.5341&rep=rep1&type=pdf .

Wicherts, J. M. (2016). The importance of measurement invariance in neurocognitive ability testing. *The Clinical Neuropsychologist*, *30*(7), 1006-1016. https://doi.org/10.1080/13854046.2016.1205136

Wicherts, J. M., & Dolan, C. V. (2010). Measurement invariance in confirmatory factor analysis: An illustration using IQ test performance of minorities. *Educational Measurement: Issues and Practice*, *29*(3), 39-47. https://doi.org/10.1111/j.1745-3992.2010.00182.x

Wu, D. A., Li, Z. & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research & Evaluation, 12*(3), 1-26.

Xu, H., & Tracey, T. J. G. (2017). Use of multi-group confirmatory factor analysis in examining measurement invariance in counseling psychology research. *The European Journal of Counselling Psychology*, *6*(1). 75-82. http://dx.doi.org/10.23668/psycharchives.2041

Yeşilyaprak, B. (2007). *Eğitimde rehberlik hizmetleri [Counseling services in education]*. Ankara: PegemA.

Mediator Role of the Need for Social Approval in the Relationship between Perfectionism and Codependency: A Structural Equation Modeling Study

Burcu Karaşar[1]
[1]Amasya University

Mediator Role of the Need for Social Approval in the Relationship between Perfectionism and Codependency: A Structural Equation Modeling Study[*]

Burcu Karaşar[1†]
[1]Amasya University

Abstract

The purpose of the study is to test the model developed in relation to the mediator role of the need for social approval in the relationship between perfectionism and codependency. The study was conducted on a total of 188 students (144 females and 44 males) attending the Education Faculty of Amasya University. In the study, Spann-Fischer Codependency Scale, Need for Social Approval Scale and Frost Multidimensional Perfectionism Scale were used. The theoretical model proposed in the study was tested with the structural equation model. In the analysis of the collected data, LISREL 8.51 was used. The goodness-of-fit indices calculated for the model were found to be good. The goodness-of-fit indices found for the structural model are as follows: RMSEA=0.07, Standardized RMR=0.08, NNFI=0.90, CFI=0.91 and IFI=0.91. It was found that the structural model constructed to reveal the mediator role of the need for social approval in the relationship between perfectionism and codependency has a good fit and the need for social approval was found to have a partial mediator role in the model.

Keywords: Perfectionism, codependency, need for social approval, mediator.

Introduction

Codependency is a subject which has been investigated only in recent years in Turkey. While it is a subject widely studied in the field of health in recent years (Ançel, 2012; Ançel, Yuva and Öztuna, 2012; Ölçüm and Büyükkayacı Duman, 2017; Öz, 1998; Özdemir and Buzlu, 2019), there is a limited amount of research carried out on it in the field of education (Reisoğlu, Yazıcı and Aydın, 2018; Tanhan and Mukba, 2014). Ançel (2017) stressed that the prevalence of codependency in Turkey is not known as the subject has not been researched widely enough. When the international literature is reviewed, it is seen that a large number of studies have focused on the subject (Hillborg, 1995; Martsolf, Sedlak and Doheny, 2000; Chang, 2010; Reyome and Ward, 2007; Parker, Faulk and Lobello, 2003; Chang, 2018; Bereza and Isaeva, 2018; Bacon et al., 2018; Zielinski et al., 2019).

Beattie (2012) defined the characteristics of codependency as follows: Codependent people strive for other people's contentment rather than their own contentment. While they try to give something to other people throughout their lives, they feel sad as they are not given anything. These people who tend to devote themselves to others have feelings of pity and guilt towards people with problems. They also tend to be self-sacrificing.

When we look at the history of the concept of codependency, it is seen the concept came into question after alcohol dependence was seen as a physical illness in the 1930s. Since the 1950s, it has become a notable concept with the recognition of the fact that some other members of the family were also affected from the person addicted to alcohol. After 1980s, studies on the concept started to increase. It has also been shown that there is also a codependency relationship between patients with chronic disease and their caregivers (Ançel, 2017). Therefore it can be said that codependency was initially used to illustrate dysfunctional relationship between the alcoholic person and his/her spouse, but was subsequently used to define dysfunctional relationships (Lindley, Giordano and Hammer, 1999).

---

[*] This research was orally presented in International Congress of Science Culture and Education (29 Oct-02 Nov 2019).

[†] Corresponding Author: *Burcu Karaşar, burcukarasar@gmail.com*

Ançel (2012) defined codependency as the pathological relationship in which the interdependence between the individual who is dependent on the care of the other and the caregiver is supported, increased and maintained. As a result of this relationship, the individual feels approved and strong. Tanhan and Mukba (2014) describe one person's developing a relationship dependent on another person due to his/her personality traits and behavioural patterns in interpersonal relationships as codependency. Interpersonal attitudes specific to codependency were defined as "diseased dependency" by Karen Horney (Crothers and Warren, 1996). Despite the difficulties in making a precise definition of codependency, codependent people are defined as those who care too much on what is happening around and lose their own feelings and thoughts while trying to control other people's lives (Lindley, Giordano and Hammer, 1999). DesRoches (2000) pointed out that codependency is not only a psychological concept but also a life-threatening, destructive and health-impairing disease. Ançel (2017) stated that having codependent behaviours is not a disease. He stated that it is a new form of a person traditionally considered to be good, self-sacrificing and considerate of others more than himself/herself.

Codependency can be explained by schemes of self-sacrifice and emotional deprivation from among the schema domains involved in the schema therapy. Young, Klosko and Weishaar (2009) stated that one's belief in the self-sacrifice scheme is as follows: The individual believes that if he / she meets all the needs of other people and does not care about his / her needs, others will accept him / her and will not feel loveless. The emotional deprivation scheme is also seen in people who have the self-sacrifice scheme. These people focus more on meeting others' needs rather than their own.

Hollabaugh (1995) stated that codependency can be explained through concepts from the social psychology literature. The concepts of self-presentation and impression management explain how people try to influence others' perceptions (Goffman, 1959; Vohs, Baumeister and Ciarocco, 2005; Murphy, 2007; Schlenker, 2012). Vohs, Baumeister, and Ciarocco (2005) stated that one of the most basic skills in social life is that the individual presents himself/herself pozitively to other people. On the basis of these explanations, it can be said that codependency seems to be related to the social approval. The need for social approval is defined as a concept in which the opinions of others are considered important. People having a strong need for social approval are those who try to make a positive impression on others (Karaşar, 2014). Horney (2019) defined the need for neurotic love and approval as the need to keep others pleasant and to be loved and approved by them. This need includes acting in accordance with the expectations of others, and seeing others' wishes and thoughts as superior to their own. It also includes fear of self-disclosure and fear of hostility of others. These explanations regarding the need for social approval are similar to the characteristics of codependency. Codependent people feel the need for approval (Yates and McDaniel, 1994; Beattie, 2012; Tanhan and Mukba, 2014; Friel and Friel, 2010).

Particularly in eastern societies, making sacrifices, altruism and devotion are characteristics strongly approved. Such characteristics are reinforced by the society (Ançel, 2017). In Turkey, characteristics such as making sacrifices and codependency are culturally promoted from early ages (Ançel, 2012). In addition to these characteristics, a great importance is attached to what other people think about what we do and say in our culture (Karaşar, 2014). The concern felt about what others think about what we do and say is generally accompanied by codependency in our culture. Tanhan and Mukba (2014) pointed out that characteristics such as suppressing their feelings, need for approval and lack of self-confidence are commonly seen in Turkey. Yates and McDaniel (1994) stated that symptoms such as suppressing their problems and feelings, perfectionism, low self-esteem, fear, anxiety and depression are seen in these individuals. At the same time, self-neglect of these individuals may result in various health problems.

Ançel (2017) stated that the need for empowerment in codependent people is met with the approval and appreciation of the other person. With the approval the individual receives, he/she increases his/her self-esteem. Codependent people value others' thoughts more than their own. They have behaviors to please others. They have expectations to be approved by people (DesRoches, 2000).

Codependency also seems to be related to perfectionism. Yates and McDaniel (1994) and Ançel (2017) argue that one of the characteristics of codependent people is perfectionism. Codependent people hope to do everything perfectly. They are trying to show that they are good enough for others. The codependent individual in the role of the saviour shows obsessive thoughts and behaviours. Together with perfectionism, being obsessive ensures the continuity of controlling behaviours. Codependents striving to control events avoid making mistakes (Ançel, 2017). Yates and McDaniel (1994) stated that one of the main characteristics of codependency on which experts agree is perfectionism and the need to control other people.

Perfectionism also appears to be related to the need for social approval. According to Antony and Swinson (2009), perfectionist individuals are those who experience extreme anxiety about other people's thoughts, and

these people have a high need for approval. There is a limited amount of research on this subject in the literature (Hewitt and Flett, 1991; Sherry, 2002; Karaşar and Öğülmüş, 2016a). In a study by Hewitt and Flett (1991), it was found that socially prescribed perfectionism was related to the social approval seeking sub-dimension of the Irrational Beliefs Test. In their research on young adults, Karaşar and Öğülmüş (2016a) found that self-directed perfectionism predicted the need for social approval. The findings of these studies show that the need for social approval and perfectionism are related.

Given the delineations above, it seems to be important to explain the relationships between codependency, perfectionism and the need for social approval. Although there is a great deal of research on the subject in the field of health, the number of such studies in the field of education is highly limited (Tanhan and Mukba, 2014; Reisoğlu et al, 2018). Especially the fact that the teaching profession is based on self-sacrifice and devotion can cause tendency towards codependency. Teachers can develop codependency by putting the needs of their students and parents ahead of their own needs. In this connection, Koehler (2010) stated that teachers' overload and fatigue are the greatest evidence for codependency. He also stated that teachers tend to overwork for the satisfaction of other people. It is also emphasized that the existence of external reward systems to evaluate teachers may also affect this. Considering the negative consequences of codependency, it can be said that the studies on pre-service teachers may be important in terms of preventing future problems in their professional lives. For this reason, investigation of the subject in relation to pre-service teachers will help fill an important void in the literature. In this regard, the aim of the this study is to investigate the mediator role of the need for social approval in the relationship between codependency and perfectionism.

METHOD

Research Model

In the current study, the quantitative and relational model was used to investigate the relationships between variables. In the testing of the model developed in the study, the structural equation model was employed.

Study Group

The study group of this research is comprised of 188 students (144 females and 44 females) attending the Education Faculty in Amasya University. Random sampling method was used in the research. In the research, pre-service teachers were informed about the study. After giving information, scales were applied in classrooms.

Data Collection Tools

In the study, the Spann-Fischer Codependency Scale, the Need for Social Approval Scale and the Frost Multidimensional Perfectionism Scale were used to collect data.

Spann-Fischer Codependency Scale: The Turkish adaptation studies of the scale developed by Fischer, Spann and Crawford (1991) were carried out by Tanhan and Mukba (2014). It is a six-point Likert type scale. High scores taken from the scale show that codependency is high. The internal consistency coefficient of the scale measuring the characteristics of codependency such as experiencing difficulty in expressing feelings and over-focusing on external world was found to be $\alpha=0.65$. The common variance of the scale was found to be 48.69%. In this study, cronbach alpha coefficient was found .73 for codependency. In this research, the goodness-of-fit indices found for the codependency are as follows: $x^2/sd=232.06/103=2.23$, RMSEA=0.08, Standardized RMR=0.08.

The Need for Social Approval Scale: It was developed by Karaşar and Öğülmüş (2016b). This scale consisted of 25 items. High scores taken from the scale show that the need for social approval is high. Sub-dimensions of the need for social approval are (i) sensitivity to others' judgements, (ii) leaving a positive impression and (iii) social withdrawal. The scale items are responded on a five-point Likert scale ranging from "Strongly Agree" to "Strongly Disagree". The internal consistency coefficients calculated for the sub-dimensions of the scale were found to be varying between .80 and .83. The internal consistency coefficient calculated for the whole scale is

.90. In order to test the convergent validity of the Need for Social Approval Scale, its correlation with the "The Fear of Being Negatively Evaluated Scale Short Form" was checked and the correlation was found.79. The Need for Social Approval Scale's pretest-posttest reliability was found to be .90. In this study, cronbach alpha coefficient was found .92 for social approval. In this research, the goodness-of-fit indices found for the need for social approval are as follows: $x^2/sd= 523.96/272= 1.92$, RMSEA=0.07, Standardized RMR=0.06, NNFI=0.96, NFI=0.93, CFI=0.96, IFI=0.96.

Frost Multidimensional Perfectionism Scale: Turkish adaptation studies of the scale developed by Frost, Marten, Lahart and Rosenblate, (1991) into Turkish were conducted by Özbay and Mısırlı Taşdemir (2003). The scale consists of 35 items. The response options to the scale items range from "Strongly Disagree" to "Strongly Agree". The scale was found to be consisted of 6 sub-dimensions explaining 47.8% of the variance. The reliability and validity studies of the scale were conducted on university students by Kağan (2011). As a result of the confirmatory factor analysis, it was concluded that the original six-factor structure of the scale is also valid for the Turkish sampling. Internal consistency coefficient was found to be .91. The internal consistency coefficients of the sub-scales were found to be ranging from .64 to .94. The fifteen-day classroom total correlation coefficient was found to be .82 for total scores. Kağan (2011) stated that order sub-dimension in the Turkish sampling should be evaluated independently and that the score taken from this sub-dimension should be evaluated separately from the general scores of the scale. Therefore, the order sub-dimension was discarded from the current study and the other five sub-dimensions were included in the analyses. In this study, cronbach alpha coefficient was found .89 for perfectionism.Cronbach alpha coefficient was found .70 for order sub-dimension. In this research, the goodness-of-fit indices found for the perfectionism are as follows: $x^2/sd=1175.43/545= 2.15$, RMSEA=0.07, NNFI= 0.92, Standardized RMR=0.08, CFI=0.93, IFI=0.93.

Data Analysis

In the study, the theoretical model developed to explain the mediator role of the need for social approval in the relationship between pre-service teachers perfectionism and codependency was tested with the structural equation modelling. In the analysis of the data, SPSS 20 and LİSREL 8.51 program packages were used.

Findings

In this section, firstly the correlations between the variables are presented. Then the structural equation modelling and analysis results related to the mediator role of the need for social approval in the relationship between perfectionism and codependency are presented. In Table 1, the correlation between variables and descriptive statistics are presented.

Table 1. Means, standard deviations and inter-correlations of the perfectionism, need for social approval and codependency.

	Mean	**SD**	**1**	**2**	**3**
1.Codependency	48.38	11.63	1		
2.Social approval	69.48	17.31	.41**	1	
3.Perfectionism	72.16	19.01	.34**	.39**	1

** $p < .01$

As can be seen in Table 1, there is a medium and positive correlation between the need for social approval and codependency. A positive and medium correlation was found between perfectionism and codependency. Moreover, there is a positive and medium correlation between perfectionism and the need for social approval.

Chi-Square=532.77, df=249, P-value=0.00000, RMSEA=0.078
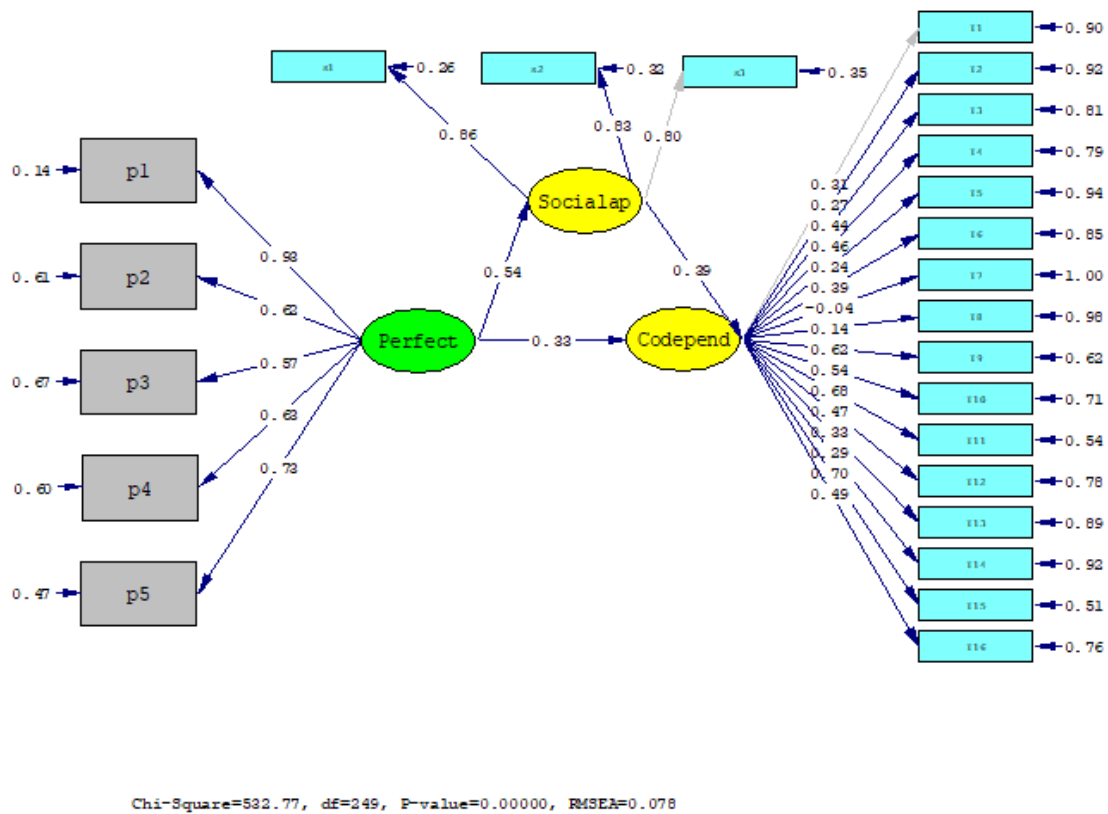
Figure 1: The mediator role of the need for social approval in the relationship between perfectionism and codependency

When the model presented in Figure 1 and the mediator role of the need for social approval in the relationship between perfectionism and codependency were tested, the relationships established between perfectionism and the need for social approval (t= 6.63), between the need for social approval and codependency (t= 2.95) and between perfectionism and codependency (t= 2.72) were found to be statistically significant. Chi-square value for the scale was found to be significant (χ^2=532.77, sd= 249, p = .00). When the ratio of the Chi-square value to degree of freedom (χ^2 /sd = 2.13) is below 3, it shows a perfect fit. When the goodness-of-fit indices of the model were calculated, they were found to be as follows: RMSEA=.07, Standardized RMR = 0.08, NNFI= .90, CFI=.91 and IFI=.91. When the modification suggestions for the model are examined, it can be said that there is no modification that can make a significant contribution to the χ^2 value. These values obtained as a result of the analysis show that the model has a good fit. When the RMSEA value is ≤.06, it indicates a good fit (Hu and Bentler, 1999) and when it is ≤.08, it indicates an acceptable fit (Sümer, 2000). Thus, the RMSEA value found in the current study can be argued to be an acceptable value. According to Karaca (2013), when RMR value is <.05, it indicates a perfect fit, when it is <.08, it indicates a good fit, when it is <.10, it indicates an acceptable fit. Thus, the RMR value found in the current study can be said to be an acceptable value. Sümer (2000) stated that when NNFI value is ≥.95, it indicates a perfect fit, when it is .90-.94, it indicates an acceptable fit. Thus, the NNFI value found in the current study can be said to be an acceptable fit. When IFI value is ≥.90, it indicates an adequate fit (Bryne,1998; Hoyle and Panter, 1995; Ullman, 2001; cited in Şimşek, 2007). Thus, the IFI value found in the current study can be said to be an acceptable value.

Table 2: The effect of the independent variables on the dependent variable

Dependent variable	Total effect
Codependency	0.39*Need for social approval + 0.33*Perfectionism
Need for Social Approval	0.54*Perfectionism

In Table 2, it is seen that the need for social approval and perfectionism variables positively affect codependency. The need for social approval is seen to positively affect perfectionism.

Discussion and Suggestions

In the current study, the need for social approval was taken as the mediator variable and its relationship with the pre-service teachers' perfectionism and codependency was investigated. When the model and the mediator role of the need for social approval in the relationship between perfectionism and codependency were tested, it was found that the correlations between perfectionism and the need for social approval, between the need for social approval and codependency and between perfectionism and codependency were found to be statistically significant. The goodness-of-fit indices found in the current study show that the structural model established has a good fit. The results of the current study show that with increasing perfectionism, the need for social approval also increases, and with increasing need for social approval, codependency is strengthened. At the same time, these results show that there are direct and indirect interactions between perfectionism and codependency.

The higher the level of individuals' perfectionism is, the greater their need for social approval is. As a result, they may exhibit codependency by ignoring their needs while trying to please other people. There are various studies focusing on the relationship between perfectionism and the need for social approval (Hewitt and Flett, 1991; Sherry, 2002, Karaşar and Öğülmüş, 2016a). Antony and Swinson (2009) stated that perfectionist individuals are individuals who have excessive anxiety about the thoughts of other people and that these people may need high levels of approval. In the current study, perfectionism was found to be directly and indirectly correlated with codependency. Yates and McDaniel (1994) and Ançel (2017) stated that one of the characteristics of codependent people is perfectionism. People with a high tendency towards perfectionism may strive to make a positive impression in the eyes of other people. This may cause them to make an excessive effort to become a better person. At the same time, they are more likely to exhibit obsessive behaviours as they are more interested in others' becoming perfect and focused on other people's shortcomings.

A direct correlation was found between the need for social approval and codependency. In the literature, it has been stated that codependent people have a higher need for social approval (Beattie, 2012; Tanhan and Mukba, 2014; Yates and McDaniel, 1994). The need for social approval and codependency are similar concepts. Through the influence of culture, the need for social approval and codependency may be increased. People with a high need for social approval and who want to be perceived as a good person in the eyes of others may be directed to the exhibition of codependent behaviours in order to create a positive impression and avoid negative evaluations. Codependency involves the role of the saviour who neglects himself/herself to help other people. The reasons behind this behaviour may be being remembered as a good person, pleasing others, and creating a positive impression. These factors related to the need for social approval indicate a strong need for social approval in the emergence of codependency. Culturally, the valuing others' thoughts, efforts to avoid criticism and efforts to make a good impression are common in Turkish society. Individuals who need a high level of social approval to be seen as a good person may exhibit codependency behaviours in order to be more accepted by the society without being aware of this. Therefore, they can forget their own desires and needs while living for others. It can be said that codependency can be seen at higher levels especially in those who grow up in a cultural structure where one's caring about oneself more than others is seen to be negative.

The findings of the current study should be evaluated on the basis of various limitations. The first limitation of the study is that the study was conducted on the students attending the Education Faculty of Amasya University. The model constructed in the current study can be tested with broader samplings from different universities. In light of the findings of the current study, psycho-education programs can be developed for pre-service teachers to better understand codependency. At the same time, future research can explore codependency in relation to some other concepts such as self-esteem, depression, anxiety and health problems among university students. Besides pre-service teachers, in-service teachers should also be included in such studies. The number of female

students in the research is more than the number of male students. This is a limitation of the research. In future studies, this subject can be studied in larger sample groups.Finally, more comprehensive studies can be conducted to explain the reasons for codependency considering various variables related to culture and family.

References

Ançel, G. (2012). Karşılıklı bağımlılık kavramı: hemşirelikle ilişkisi ve karşılıklı bağımlılığı belirleme araçları [The concept of codependency: the relationship with the nursing and assessment tools]. *Hemşirelikte Araştırma Geliştirme Dergisi, 1,* 70-78.

Ançel G., Yuva E., & Öztuna D.G., (2012). Eş-bağımlılık ve işyerinde mobing arasındaki ilişki [The relationship between co-dependency and mobbying/bullying]. *.Anatolian Journal Of Psychiatry, 13,*104-109.

Ançel, G. (2017). Kişilerarası ilişkilerde bağımlılık. Alter Yayıncılık.Ankara.

Anderson, S. C. (1994). A critical analysis of the concept of codependency. *Social Work, 39,* 6. 677– 685.

Antony, M. M., & Swinson, R. P. (2009). *When perfect isn't good enough: strategies for coping with perfectionism.* Oakland, CA: New Harbinger Publications

Bacon, I., McKay, E., Reynolds, F., & McIntyre, A. (2018). The lived experience of codependency: an interpretative phenomenological analysis. *International Journal of Mental Health and Addiction.* 1–18.

Beattie, M. (1992). *İlişki bağımlılığına son.* (Çev. Öztürk, F. Nagehan Öztürk). İstanbul: Ovvo.

Bereza, Z.V., & Isaeva, E.R. (2018). The effect of psychotherapeutic treatment in co-dependent relatives of drug-dependent patients on the efficiency of the treatment-rehabilitation process. *Bulletin of the South Ural State University series. Psychology.* 11:2, 54-62.

Chang, S. (2010). *Codependency among college students in the United States and Taiwan: a cross-cultural study.* Doctor of Philosophy. Ohio University, USA.

Chang, S. (2012). A cultural perspective on codependency and its treatment. *Asia Pacific Journal of Counselling and Psychotherapy, 3:1,* 50-60.

Chang, S. (2018). Testing a model of codependency for college students in Taiwan based on Bowen's concept of differentiation. *International Journal of Psychology,* 53:2, 107-116.

Crothers, M., & Warren, L.W. (1996). Parental antecedents of adult codependency. *Journal Of Clinical Psychology,* 52(2). 231-239.

Desroches, B. (2000). *Reclaiming your self.the coodependent's recovery plan.*Wellness Institute,Inc. Gretna.America.

Fischer, J. L., Spann, L., & Crawford, D. (1991). Measuring codependency. *Alcoholism Treatment Quarterly, 8,* 87-100.

Friel, J., & Friel, L. (2010). *Adult children the secrets of dysfunctional families.*Health Communications, Inc. Deerfield Beach, F.L. America.

Frost, R. O., Marten, P., Lahart, C., & Rosenblate, R. (1990). The dimensions of perfectionism. *Cognitive Theraphy And Research,* 14(5), 449-468.

Goffman, E. (1959). *The presentation of self in everyday life.* Garden City, Ny: Doubleday.

Hemfelt, R., Minirth, F., & Meier, P. (1989). *Love is a choice: recovery for codependent relation-ships.* Nashville, TN: Thomas Nelson Publishers.

Hewitt, P. L., & Flett, G. L. (1991). Perfectionism in the self and social contexts: conceptualization, assessment, and association with psychopathology. *Journal Of Personality And Social Psychology, 60,* 456-470.

Hillborg, V.L.B. (1995). *The relationship between the level of codependent behavior and the level of differentiation of self among nursing students.* Masters Theses. 224. Http://Scholarworks.gvsu.edu/theses/224.

Hollabaugh, L. C. (1995). *The social construction of the codependency construct: college students' evaluation of "codependent" characteristics in themselves and others.* Doctor of Philosophy. Department of Psychology in the Graduate School Southern Illinois University at Carbondale. Texas A&M University. Https://Search.Proquest.Com/Docview/304214185?Accountid=25074

Horney, K. (2019). *Kendi kendine psikanaliz. [Self-Analysis*]. Totem Yayınları.İstanbul.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55.

Kağan, M. (2011). Frost çok boyutlu mükemmelliyetçilik ölçeğinin Türkçe formunun psikometrik özellikleri [Psychometric properties of the Turkish version of the Frost Multidimensional Perfectionism Scale]. *Anatolian Journal of Psychiatry.* 12:192-197.

Karaca, E. (2013). Kadınların çalışma yaşamına katılımlarına yönelik tutum ölçeğinin (kçykytö) bir yapısal eşitlik modeli ile geçerlik ve güvenirliğinin incelenmesi [Investigation of the Reliability and Validity of Attitude Scale on Labor Force Participation of Women with a Structral Equation Model]. *International Conference on Eurasian Economies,* 380-388.

Karaşar, B. (2014). *Sosyal onay ihtiyacının çeşitli değişkenler açısından incelenmesi [Examining the need for social approval in terms of different variables among university students]*.Doctoral Thesis. Ankara University Graduate School of Educational Sciences. Ankara.

Karaşar,B., & Öğülmüş,S. (2016a). Üniversite öğrencilerinde sosyal onay ihtiyacının çeşitli değişkenler açısından incelenmesi [Examining the need for social approval in terms of different variables]. *Journal of Uludağ University Faculty of Education, 29 (2),* 2016, 469-495.

Karaşar, B., & Öğülmüş, S. (2016b). Sosyal onay ihtiyacı ölçeği: geçerlik ve güvenirlik analizi [Need For Social Approval Scale: The Validity and Reliability]. *Ege Journal of Education, 17 (1).* 84-104.

Koehler, M. (1991). The Teacher As Codependent. *The Clearing House. A Journal Of Educational Strategies, Issues And Ideas, 65:1*, 9-10.

Lindley, N.R., Giordano, P.J., & Hammer, E.D. (1999). Codependency: Predictors and psychometric issues. *Journal Of Clinical Psychology, 55 (1),* 59-64.

Marks, A.D.G., Blore, R.L. Hine,D.W., & Dear, G.E. (2012). Development and validation of a revised measure of codependency. *Australian Journal Of Psychology, 64.* 119–127.

Martsolf, D. S., Sedlak, C. A., & Doheny, M. O. (2000). Codependency and related health variables. *Archives Of Psychiatric Nursing, 14*(3), 150–158.

Murphy, N. (2007). Appearing smart: the impression management of intelligence, person perception accuracy, and behavior in social interaction. *Personality And Social Psychology Bulletin, 33 (3).* 325-339.

Ölçüm, H.İ., & Büyükkayacı Duman, N.(2017). Hemşirelerde kök aileyle ilişkiler ve karşılıklı bağımlılık [Family of origin relations and codependency in nurses]. *Journal of Academic Research in Nursing,* 3(2):60-65.

Öz, F. (1998). Bağlaşıklık ve hemşirelik. [Codependency and nursing]. *Kriz Dergisi,* 6(1). 67-73.

Özbay, Y. & Mısırlı-Tasdemir, Ö. (2003). Çok boyutlu mükemmeliyetçilik ölçeği: geçerlik ve güvenirlik çalışması. VII. National Congress Of Psychological Counseling And Guidance, Malatya.

Özdemir, N., & Buzlu, S.(2019). Codependency in nurses and related factors. *Annals of Medical Research, 26(7):*1145-51.

Parker, F.M.., Faulk, D., & Lobello, S.G.(2003). Assessing codependency and family pathology in nursing students. *Journal Of Addictions Nursing,* 14:2, 85-90.

Reisoğlu, S., Yazıcı, H., & Aydın, H. (2018). Rehber öğretmen adaylarının psikolojik sağlamlık düzeylerini yordamada ilişki bağımlılıklarının, kişilik özelliklerinin ve bazı demografik değişkenlerin rolü. 4th International Contemporary Educational Research Congress. Congress Proceedings (Full Texts). Muğla.

Reyome , N.D., & Ward, K. S. (2007). Self-reported history of childhood maltreatment and codependency in undergraduate nursing students. *Journal Of Emotional Abuse*, 7:1,37-50.

Schlenger, B.R. (2012). *Self presantation.* In M.R.Leary & J.P. Tangney (Eds), Handbook of self and identity (542-570).(Second Edition). New York: The Guilford Press.

Sherry, B. A. (2002). Perfectionism dimensions, perfectionistic dysfunctional attitudes, need for approval, and depression symptomps in adult psychiatric patients and young adults.(Masters thesis, The University Of British Columbia). Retrieved from https://pdfs.semanticscholar.org/af05/b73b0fa0422fce2c8137a34bd6f51d11f19b.pdf

Özbay,Y., & Mısırlı Taşdemir, Ö. (2003). Çok boyutlu mükemmeliyetçilik ölçeği:geçerlik ve güvenirlik çalışması.VII.Ulusal Psikolojik Danışma ve Rehberlik Kongresi.Malatya.

Sümer, N. (2000). Yapısal eşitlik modelleri: temel kavramlar ve örnek uygulamalar [Structural Equation Modeling:Basic Concepts and Applications]. *Turkish Psychological Review, 3(6),* 49-74.

Şimşek, Ö. F. (2007). *Yapısal eşitlik modellemesine giriş. Temel ilke ve lisrel uygulamaları.* [Introduction to structural equation modeling. Basic principles and applications of lisrel]. Ankara: Ekinoks Yayınları.

Tanhan, F., & Mukba, G. (2014). Spann-Fischer ilişki bağımlılığı ölçeği'nin Türkçe'ye uyarlama çalışmasına ilişkin psikometrik bir analiz [A Psychometric analysis concerning the Turkish version of the Spann-Fischer codependency scale]. *Pamukkale University Journal of Education, 36(2),* 179-189.

Vohs, K., Baumeister, R., & Ciarocco, N. J. (2005). Self-regulation and self-presentation: regulatory resource depletion impairs impression management and effortful self-presentation depletes regulatory resources. *Journal Of Personality And Social Psychology,* 88(4):632-57.

Yates, J.G. & Mcdaniel, J.L. (1994). Are you losing yourself in codependency. *The American Journal Of Nursing, 94;* 4. 32-36.

Young, J. E., Klosko, J. S., & Weishaar, M. E. (2009). *Şema terapi.*[Schema Therapy]. (Çev. T.V.Soylu). İstanbul: Litera Yayıncılık.

Zielinski, M., Bradshaw, S., Mullet, N., Hawkins, L., Shumway, S., & Chavez, M.S. (2019). Codependency and prefrontal cortex functioning: preliminary examination of substance use disorder impacted family members. *The American Journal on Addictions*, 28:5. 367-375.

Effects of Dynamic Geometry Software on Students' Geometric Thinking Regarding Probability of Giftedness in Mathematics

Bilal Özçakır[1], Duygu Özdemir[2], Yasemin Kıymaz[3]
[1]Alanya Alaaddin Keykubat University
[2]İstanbul Aydın University
[3]Kırşehir Ahi Evran University

To cite this article:

Effects of Dynamic Geometry Software on Students' Geometric Thinking Regarding Probability of Giftedness in Mathematics

Bilal Özçakır[1*], Duygu Özdemir[2], Yasemin Kıymaz[3]
[1]Alanya Alaaddin Keykubat University
[2]İstanbul Aydın University
[3]Kırşehir Ahi Evran University

Abstract

Gifted students have potential to improve countries and this potential can be revealed and developed in schools where they spend most of their times with other regular students. However, these classrooms have some limitations for them; hence, they need some differentiated activities. Usage of dynamic geometry in mathematics lessons could be an opportunity to provide differentiated activities. Therefore, the aim of this study is to explore effects of mathematics lessons integrated with dynamic geometry activities on students' van Hiele geometric thinking levels controlling their probability for mathematical giftedness. Participants of the study were fifty-three fifth graders from a private school in Marmaris, Turkey. These students were grouped in terms of their probability of mathematical giftedness. Seven dynamic geometry activities about properties of line segments, triangles and quadrilaterals were developed and implemented in classrooms with tablets. Van Hiele geometric thinking level test was administered to the participants as pre-test and post-test. Results showed that dynamic geometry activities help students to move from geometric thinking level about recognizing shapes with visual clues to higher level about geometrical properties of shapes, namely relationship among shapes and their properties. Moreover, interaction between their probability for mathematical giftedness and improvements in geometric thinking levels were found. This study may contribute both to the gifted education and mathematics education fields by exploring the improvements in geometric thinking level and differentiated opportunities for gifted students. As a suggestion, a more comprehensive experimental study with larger samples so as to obtain generalize the findings could be conducted as a further study.

Key words: Mathematically gifted, Mathematics education, Dynamic geometry activities, Geometric thinking levels.

Introduction

Recently, gifted students are seen as having the potential to improve countries with the help of their ability to solve problems in creative ways as well as their property of leadership (Hannah, James, Montelle & Nokes, 2011; Maryland, 1973). According to National Association for Gifted Children (NAGC, 2005) gifted students "shows or has the potential for showing, an exceptional level of performance in one or more areas of expression" (p. 4) and this property makes them important for the future of the countries. Therefore, developmental and educational needs of gifted students were seen as the vital points that should be given necessary importance in educational environments. Likewise, most countries accepted the issue about enhancing gifted students' potential as their social requirement (Trna, 2014). Giftedness of children was thought to be related not only with intelligence or IQ scores of relevant tests but also with creativity and other various factors. Three Ring Conception of Giftedness Model of Renzulli (1979) addressed students' ability levels, their motivations on learning phase and their creativity as some factors related with giftedness of children. In line with this, Sternberg (1997) mentioned about high levels in analytical, creative and practical facets as factors affected giftedness in his Triarchic Theory. Some researchers also referred to some characteristics to identify giftedness of children. For instance, Davis and Rimm (2004) referred to high level of success and superiority in language usage, quick learning with enjoyment, memory retention, problem solving ability, reasonable curiosity and high level of thinking, attention levels in tasks and interests to contents as indicators for giftedness of children. Moreover, Baykoç (2014) stated that giftedness could exist as a capacity related to genetics as well as

---

[*] Corresponding Author: *Bilal ÖZÇAKIR*, bilalozcakir@gmail.com

being related with environmental factors. Therefore, giftedness could be developed and trained within appropriate environment.

Mathematical Gifted Students and Gifted Education

In some research, giftedness of a child is referred as giftedness on specific domain or on any domain (Mayer, 2005). Therefore, giftedness on mathematics could be considered as one of the specific domains of giftedness. Among gifted students, mathematical gifted students are the ones who could see the world from the eyes of mathematics (Krutetski, 1976) and differ from other students in the abilities like spontaneous formation of problems, flexibility in handling data and originality in interpretation (Greenes, 1981). Although there is not still a common definition for mathematical giftedness (Pitta-Pantazi, Christou, Kontoyianni & Kattou, 2011), some studies mentioned that mathematical gifted children could organize tasks, use new statements in patterns and study advanced concepts better than regular children (Davaslıgil, 2004; Miller, 1990). Moreover, Johnson (2000) stated that examining quality of thinking about mathematical reasoning is the other way of determining mathematical giftedness. From another perspective, it is a misconception that conducting calculations does not directly and solely indicate mathematical giftedness; in fact, comprehension of mathematical ideas is the indicator of mathematical giftedness (Karaduman, 2010). Similarly, Sheffield (1994) defined features of mathematical giftedness as speed of understanding, higher level of ability in questioning, and seeing cause-effect relations in mathematical constructs. Mathematical gifted children differ from regular children for features of learning speed, understanding levels and interest levels to mathematics (Dağlıoğlu, 2004). Furthermore, mathematical creativity is also referred as another feature for differentiating mathematical gifted children by many other authors (Leikin, 2009; Sriraman, Haavold & Lee, 2013).

Apart from how they are defined, it is a reality that gifted students spend their times in the school with other regular students (Baykoç, 2011). Although these regular classroom environments are required for their social and emotional needs, these environments have some limitations in terms of their cognitive needs (Baykoç, 2014). When the lessons were not appropriate to their pace of learning and differentiated abilities and interests, the students face with boredom, lack of enjoyment or negative disposition towards mathematics (Maxwell, 2001; Park & Park, 2006). Thus, mathematical gifted students need differentiated activities that meet their differentiated needs, support their interest and ability domains, and motivate them toward mathematics tasks in classrooms (Baykoç, 2010; Pierce et al., 2011). As in the case in other countries, gifted students in Turkey need special activities or services that are outside the needs of regular students (Aydemir & Çakıroğlu, 2013). There are some institutions that support gifted students in Turkey. For example, Science and Art Centers (BİLSEM) under the auspices of MoNE and children universities as well as some private institution/centers for gifted children try to provide opportunities for gifted students in the remaining time from their school hours. There is a great need for learning and meeting their needs both in schools and in these institutions where they have only opportunity to be nurtured. Thus, opportunities to be offered to these students in any environments should be increased and enriched by means of research based studies. Based on this need, some studies related with mathematical gifted students concentrated on the modifications, differentiation or inclusion on the math curriculum/instruction to analyze their effects on students' achievement and to better fit the requirement of gifted students (Gavin, Casa, Adelson, Carroll, Sheffield & Spinelli, 2007; Thomas, 2019; Tieso 2003; Ysseldyke, Tardrew, Betts, Thill, & Hannigan, 2004). Additionally, some studies (Deringöl & Davaslıgil, 2020; Erdogan & Yemenli, 2019; Kamarudin, Kamarulzaman, & Ishak, 2018) examined mathematically gifted students' views and attitudes. For example, Hammer (2002) explored precocious mathematics students' attitudes when they were not challenged appropriately. Likely, Martin and Pickett (2013) mentioned their implementation of differentiated instruction and its effects on the improvements of the mathematical gifted students' motivation and engagement. However, as Ysseldyke et al. (2004) indicates, studies in literature reveal the problem that most gifted students deprive of learning environments which enables them to construct their own learning. Since gifted students have differentiated needs in line with their differentiated properties, learning tasks provided to mathematical gifted students should support differentiated characteristics like challenging, entertaining and so on (Özdemir, 2018). At that point, technology integration is one of these characteristics that both helps to construct and lead their own learning and a dimension suggested for the differentiated needs of mathematically gifted students (Özdemir, 2016).

Geometry Education and Technology

Geometry is one of the fields of school mathematics. Since our environment, with which we are surrounded consists of many geometric shapes and objects, geometry education has its own importance. Geometry can be

considered as an important skill of doing mathematics (Suydam, 1985) and geometry education provides opportunities to enhance logical thinking abilities, spatial insight about physical environment and knowledge for understanding higher level mathematics (Suydam, 1985). Literature review revealed that the van Hiele geometric thinking hierarchy has commonly been considered to describe learners' knowledge and thinking about two-dimensional geometry (Battista, 2002; Olkun, Sinoplu & Deryakulu, 2005; Özçakır & Çakıroğlu, 2019). Van Hiele explained factors related to enhancing logical thinking, spatial insight and understanding higher level of geometry within van Hiele geometric thinking hierarchy (Usiskin, 1982). In this hierarchy, geometric concepts were arranged into levels in accordance with prerequisite concepts suitable to students' geometric thinking. This theory of geometric thinking is consisted of five levels as visualization, analysis, informal deduction, formal deduction and rigor (Crowley, 1987). The geometric thinking levels within this hierarchy are organized as;

- Level 0 – Visualization: visual clues about geometric figures, recognizing figures depends on visual information about figure.
- Level 1 – Analysis: geometric properties have their own value, recognizing includes simple definition and properties of figure.
- Level 2 – Informal Deduction: properties can form a family of figures, interrelationships between figures based on their similar or different properties.
- Level 3 – Deduction: going beyond identifying properties of figures and relationships among them, proofs can be constructed, using postulates or axioms and definitions.
- Level 4 – Rigor: learner could go beyond Euclidean geometry and can work in different geometric and axiomatic systems.

National Council of Teachers of Mathematics (NCTM, 2010) suggests that students should proceed first level of van Hiele geometric thinking hierarchy at kindergarten to second grade, second level at third grade to fifth grade and third level before graduated from middle school. Hence, it was suggested that students should have achieved first three levels of this geometric thinking hierarchy at middle school in order to understand high school mathematical concepts (Cansız-Aktaş & Aktaş, 2012). However, progresses between these geometric thinking levels are related with educational experiences of learners rather than their ages or maturations. Therefore, teaching geometry with experiment-based activities could foster students' understanding in geometry (Fidan & Türnüklü, 2010; Özçakır, 2013; Tan-Şişman & Aksu, 2012). Moreover, students can learn geometric concepts sufficiently when learning environments are prepared in line with their knowledge on geometry and their geometric thinking levels (Choi-Koh, 1999). Therefore, understanding learners' knowledge and linking with van Hiele hierarchy is important for developing suitable learning activities, materials and so instructions, to provide them learning environments in which they could advance through the levels of van Hiele with these learning opportunities (Malloy, 2002).

Computer technology can provide these learning opportunities for learners since educational technology can be helpful to provide tasks and tools included with multiple representations of concepts dynamically linked together. It also offers learning environment with different opportunities to concretize an abstract concept of mathematics with digital dynamic contents and virtual objects (Özçakır, 2013). Therefore, these tools provide students a digital environment for exploring and identifying mathematical concepts and relationships within or among mathematical objects (Thomas & Holton, 2003). Dynamic geometry software is one of the technological tools used in mathematics education. Dynamic geometry software permits students to interact with mathematical constructs so they can examine different examples of the constructs with dynamic features. In other words, students have an opportunity to investigate mathematical objects like in a laboratory for mathematics (Tabach, 2011). In this environment, students can manipulate dynamic geometric objects and observe changes in multiple representation of the objects provided by hot links among these representation and real-time measures (Laborde, Kynigos & Strasser, 2006). Therefore, they can make tests and observe changed and unchanged parts of the objects among various manipulation, can record and conjecture constructs and theorems with dynamic geometry software.

Although, in various research, dynamic geometry activities are found to engage students' understanding of geometry and to provide a learning environment to foster their geometric thinking (Gawlick, 2005; Karakuş & Peker, 2015; Özçakır, 2013), there still needs to understand effects of dynamic geometry activities on both gifted and regular students' geometric thinking levels. Since technology enables the students to continue with their own learning pace (Özçakır, 2013), it may be seen as a useful way for mathematical gifted students to meet their needs in classrooms (Johnson, 2000; Periathiruvadi & Rinn, 2012). Based on this idea, in instructional phase, using educational technology for learning tasks could be an opportunity to develop abilities of mathematical gifted students. In order to increase the motivation of mathematical gifted students in lessons,

usage of dynamic geometry can be seen as a valuable opportunity. It also enables students to examine the geometric concepts with details as one of the needs of gifted students in mathematics classrooms. Furthermore, this may help them to use their giftedness potential and proceed with their own pace. Therefore, this study aimed to explore effects of mathematics lessons integrated with dynamic geometry activities on students' van Hiele geometric thinking levels controlling their probability for mathematical giftedness.

When studies (Levenberg & Shaham, 2014; Shillor, 1997, Taylor, 2008) about gifted and geometry education analysed, the lack, importance and the need in this area can be seen. For example, Tanahan's (2006) study with high school geometry teachers across California highlights that although teachers in high school geometry understand the need of differentiated education for gifted students, they do not significantly differentiate their instruction. Besides, Casa et al. (2017) indicate the need for more challenge and special activities for the students in kindergarten. A similar study to this study conducted by El-Demerdash (2010), the effectiveness of an enrichment program using dynamic geometry software on developing mathematically gifted high school students' geometric creativity can also be seen as an indication for such kind of studies in also middle school level. Moreover, some researchers (Johnson, 2000; Siegle, 2004) mention about the role of technology integration in gifted education due to the ability and motivation of gifted students in technology usage. Even, Kontostavlou and Drigas (2019) mentioned about proved effectiveness of technology usage in special education and provided a report about studies related with using technology for gifted education. That is, because technology allows to proceed at their own pace (Kaput, 1992; Özçakır, 2013), technology integration can be seen as one of the valuable characteristics for the task of mathematically gifted students (Özdemir, 2016).

Method

Research Design

This study was designed as pre-test – post-test experimental research methodology without control group in accordance with the aim of this study and regularities of the school of the students. This method enabled the researchers to compare students' pre- and post-scores as well as their giftedness score. Since this school, which we had conveniently, did not allow us to administrate different instructional methods for student groups, there was no control group and all student groups learnt mathematics using dynamic geometry software within our experimental study. This study focused on exploring effects of mathematics lessons integrated with dynamic geometry activities on regular and mathematical gifted students' geometric thinking. Seven dynamic geometry activities about 5th grade mathematics objectives were developed, and students were engaged with dynamic geometry activities in lessons.

Participants
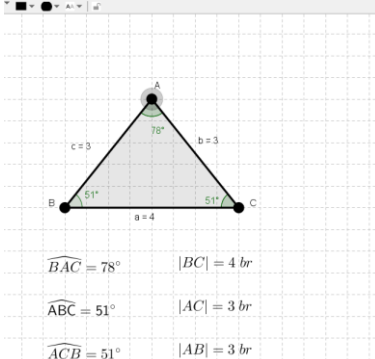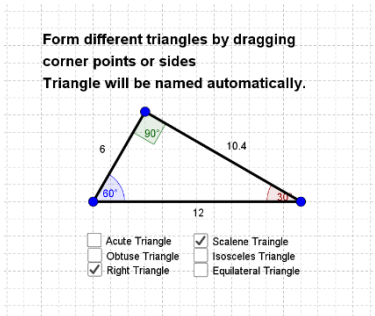
In this study, one of the researchers was mathematics teacher of the three classes of students and so convenient sampling was used, which formed our working group. That is, twenty-one female and thirty-three male 5th grade students from three classrooms of a school in Marmaris, Turkey involved in this study. However, since this school only allowed similar treatment for all classes in the same grade level, all students participated in the treatment. Therefore, forming a control group was not possible in the context of this study. These students had generally experienced using tablets for educational purposes and so they had some background for technology supported learning. The participants were grouped for data analysis in terms of their probability of mathematical giftedness by using Test of Mathematical Abilities for Gifted Students (TOMAGS) (Ryser & Johnsen, 1998). Students' scores in the TOMAGS were used to form three different mathematically giftedness levels as low, average and high probability of mathematical giftedness levels. According to these scores, 25 students were grouped in low probability group, 11 students were grouped in average probability group and 17 students were grouped in high probability of giftedness group.

Research Procedure

In this study, seven dynamic geometry activities were developed in accordance with Middle School Mathematics Curriculum for 5th grade (Ministry of National Education [MoNE], 2018). The activities were about concepts related with properties of line and line segments, and properties and types of triangle and quadrilaterals such as rectangle, parallelogram, rhombus and trapezoid. While designing and implementing these activities, GeoGebra dynamic geometry software was used as a tool for learning since this dynamic geometry

software supports multilanguage included Turkish Language and runs on multiplatform included mobile devices like tablet, PCs and mobile phones. The activities were designed as in line with learning objectives in the curriculum (MoNE, 2018) and as easy as possible to use GeoGebra as a learning tool via tablets. Since all needed figures provided to students, they did not have to construct any geometric figures in these activities. They only moved or relocated the points and line segments via touching and dragging in these activities. Moreover, in order to design suitable activities for students in different levels of van Hiele hierarchy, activities were designed to allow students to manipulate geometric figures dynamically by preserving their basic properties so that they have an opportunity to observe changed and unchanged constructs within figures. These activities were evaluated for their appropriateness by four researchers with doctoral degree in the field of Elementary Mathematics Education and two elementary mathematics teachers. According to their feedbacks, some changes in design were made and the activities were made ready for administration stage. Before administrating the activities to the classroom, a pilot study was conducted with two, 6th and 5th grade, students from a different school and the points and instructions that need to be clarified were reorganized to take the last form of the activities, some of which were briefly described in Table 1. That is, through the study, these dynamic geometry activities were used in mathematics lessons, which were designed by the researchers via GeoGebra dynamic geometry software.

Table 1: Some of the dynamic geometry activities designed in this study

Concept	Description	Sample images
Line segments	There were several line segments in four different positions and lengths. For each line segment one can move or drag points of an orange line segment to make parallel and in same length to other black one. If one successfully completed task dynamic geometry activity confirms results with a green tick.	
Triangle - Properties	Three dragging points to explore angle – side relationships for a triangle. One can stretch sides by dragging points and observe changes in angle measures and side lengths.	
Triangle - Types	There are three dragging points to explore properties of types of triangles. One can drag corners or sides to change appearance of the triangle and dynamic geometry activity automatically named triangle type.	

Rectangle	There are three dragging points to explore a rectangle: A – rotate rectangle B – stretch rectangle vertically C – stretch rectangle horizontally One could do some measurements to examine the rectangle and extract its properties.	
Parallelogram	There are three dragging points to explore a parallelogram: A – change lengths of a and c sides B – stretch parallelogram vertically or horizontally C – change length of b and d sides One could do some measurements to examine the parallelogram and extract its properties.	
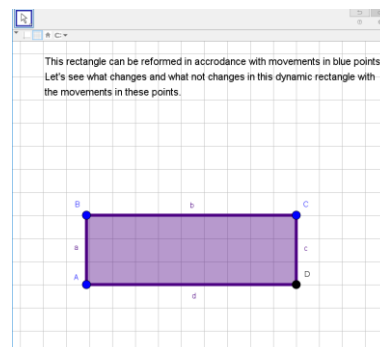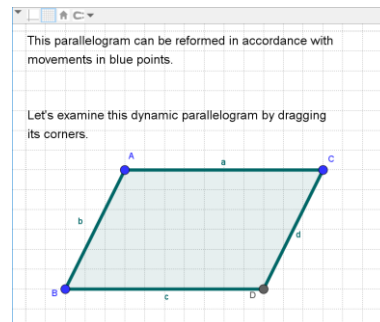
During the study, at first, an hour preparation course for using GeoGebra on tablet PCs was introduced to these students. Moreover, the first activity, which is about line segments, was used to make students familiar with the usage of GeoGebra in learning environments. Then, the dynamic geometry activities were implemented during six-hour lessons with tablet PCs in classrooms. Since one of the researchers was teacher of these students, this researcher had full control of teaching processes during the study. In these activities, students did not have any technical difficulties, so they easily used GeoGebra as a learning tool. In the learning process, they were observed as active participants in learning environment, and they allowed to share ideas about their explorations freely. In general, students dealt with learning activities and GeoGebra as well as sharing ideas and engaging discussions throughout the research. In the meanwhile, the teacher who was also one of the researchers, guided students if they struggle in both using tablet and engaging learning activities, while monitoring students' works and giving feedback on their progress.

Instruments

In this study, a test to determine students' probability of mathematical giftedness, which was TOMAGS, was administered to all students at the beginning of the study. Items in TOMAGS aim to identify children who are gifted in mathematics (Ryser & Johnsen, 1998). Therefore, the TOMAGS could be used as an identification instrument for mathematical giftedness due to its strong validity and reliability scores, as greater than .80 which can be seen in Table 2 (Ryser & Johnsen, 1998).

Table 2. TOMAGS validity and reliability scores (Ryser & Johnsen, 1998, p. 28).

	Content Sampling		Time sampling	Scorer	Average
	Normal	Gifted			
TOMAGS Intermediate	0,88	0,86	0,94	0,99	0,93

The TOMAGS includes 47 open-ended questions with a degree of difficulty to test the limits of students in a problem-solving environment. These questions cover learning domains of middle school mathematics curriculum such as numbers, geometry, measurement, and statistics and probability (Ryser & Johnsen, 1998). Moreover, three curriculum standards of NCTM were represented in construction of TOMAGS such as mathematical problem solving, mathematical communication and mathematical reasoning. In the context for this research, these standards were also applicable since they are aligned with basic skills for mathematics defined in mathematics curriculum of Turkey (MoNE, 2013). Results of the TOMAGS demonstrated high reliability in all three types of error because reliability coefficients approximating or exceeding .80 are found reliable and .90 or above found as most desirable (Ryser & Johnsen, 1998; Salvia & Ysseldyke, 1995). Similarly, in order to

evaluate the reliability and validity of the Turkish version of TOMAGS, Özdemir (2016) evaluated the test and concluded to the appropriateness of the adapted version of the test to determine mathematically giftedness of students in Turkey.

In this study, while identifying probability of mathematical giftedness, students' quotient scores were calculated based on their raw scores obtained in the test as well as their age in terms of month and year. Then, based on these quotient scores, probability of mathematical giftedness was determined in line with the guidelines of Ryser and Johnsen (Ryser & Johnsen, 1998, p.17) as low, average and high probability of mathematical giftedness within the context of this study. Thus, the participants of the study who are 21 girls and 32 boys, were grouped in terms of their scores in the TOMAGS. That is, among those fifty-three students, seventeen students were identified as high probability of mathematical giftedness, eleven students were identified as average probability of mathematical giftedness and remaining twenty-five students were identified as low probability of mathematical giftedness as seen in Table 3.

Table 3. Probability of giftedness for participants

| Gender | Probability of Mathematical Giftedness | | | Total |
	Low	Average	High	
Female	11	3	7	21
Male	14	8	10	32
Total	25	11	17	53

The participants were also pre-tested at the beginning of the study and post-tested at the end of the study by Van Hiele geometric thinking level test, which was developed by Usiskin (1982) to define students' van Hiele geometric thinking levels. This test was translated into Turkish by Duatepe (2004). The Van Hiele geometric thinking level test is consisted of 25 multiple-choice items. In this test, each 5-item were related with van Hiele Geometric Thinking Levels, respectively. Therefore, in this study, first fifteen items were used since these three levels are related with the intended grade level in accordance with the suggestion of NCTM (Fidan & Türnüklü, 2010).

Data Analysis

Scoring of the Van Hiele geometric thinking level test is criterion-based which suggested by Usiskin (1982). Usiskin clarified that if a student correctly answered three of five questions for each level, this student is considered as achieved the related level. Thus, for this test, students were assigned a weighted sum score in the following manner in Table 4. According to this scoring method, for general assessment for van Hiele geometric thinking levels of the group mean value, the intervals for placement were considered such that 1,00 - 2,50 points for Level 0, 2,51 - 5,00 points for Level 1 and 5,01 - 7,00 points for Level 2.

Table 4. Scoring van Hiele Level Test

Score	Criteria
0 Point	If at most two of first five question are correct
1 Point	If three of first five questions are correct
2 Points	If three of second five questions are correct
3 Points	If three of third five questions are correct

Data collected through TOMAGS was used to determine probability of mathematical giftedness of students and so, according to this data from administration of TOMAGS three group for students were formed. These groups constituted independent variable for data analysis. On the other hand, data collected through pre-test and post-test administrations of the Van Hiele geometric thinking level test provided scores of students regarding their geometric thinking levels. This data formed dependent variables for data analysis. Since this study included three groups of students and their scores for the Van Hiele geometric thinking level test over time, 3x2 mixed model repeated measure analysis of variance (ANOVA) was considered as appropriate to analyse this data. Besides, the mixed model repeated measure ANOVA is powerful to reveal the main effect of intervention by disregarding groups and interaction effect between intervention and groups of students. Based on this, preliminary analyses were conducted, and it was seen that the data satisfied assumptions of the mixed model ANOVA. Therefore, the mean scores from the pre-test and post-test administration of the Van Hiele geometric thinking level test in terms of groups formed by probability of mathematical giftedness were analysed through mixed model repeated measure ANOVA.

Results

The data were handled, and findings were reported in terms of van Hiele geometric thinking levels of students and their probability of mathematical giftedness. First of all, fifth grade students' scores regarding van Hiele geometric thinking level have been presented in accordance with their pre-test and post-test results from Van Hiele geometric thinking level test in terms of probability for mathematical giftedness, as a starting point for data analysis in Table 5.

Table 5. Descriptive statistics regarding scores of students on van Hiele geometric thinking level test

Test	Giftedness Probability	N	M	SD
Pre-test	Low	25	1,48	0,87
(M=1,72; SD=0,97)	Average	11	2,27	1,01
	High	17	1,71	0,99
Post-test	Low	25	2,44	1,96
(M=3,91; SD=2,47)	Average	11	4,27	2,24
	High	17	5,82	1,88

Descriptive data on Table 5 showed that van Hiele geometric thinking scores of students changed positively after intervention, since before the intervention students generally were placed in Level 0 (M=1,72; SD=0,97), they were placed in Level 1 of the van Hiele hierarchy after study (M=3,90; SD=2,47). Students' placements for the van Hiele hierarchy in terms of their probability for mathematical giftedness changed also positively between pre-test and post-test administrations. In detail, at the beginning of the study, all students who have low probability of mathematical giftedness (M=1,48; SD=0,87), average probability of mathematical giftedness (M=2,27; SD=1,01) and high probability of mathematical giftedness (M=1,71; SD=0,99) were defined at Level 0 of this hierarchy. After students' completion of the dynamic geometry activities, their status of van Hiele hierarchy changed as in the following manner. That is, students with low probability were still categorized as at Level 0 (M=2,44; SD=1,96), while students with average probability were categorized as at Level 1 (M=4,27; SD=2,24). Additionally, students with high probability of mathematical giftedness were stated at Level 2 (M=5,82; SD=1,88) of this hierarchy.

In order to investigate effects of intervention with dynamic geometry activities on van Hiele geometric thinking levels of students in terms of their probability of mathematical giftedness, their pre-test and post-test scores on the Van Hiele geometric thinking level test were analysed through 3x2 mixed model repeated measure ANOVA (Table 6).

Table 6. Mixed model ANOVA results for main effect and interaction

Source of Variance	SS	df	F	p
Tests	132,010	1	105,841	0,000
Tests * Giftedness	50,694	2	20,322	0,000
Error	62,362	50		

The output of this mixed model ANOVA analysis included two main results about the data. First of all, according to mixed model repeated measure ANOVA results, there was a significant main effect of intervention with dynamic geometry activities on students' gains about geometric thinking levels between pre-test and post-test results of Van Hiele geometric thinking level test (F(1, 50)=105.84, p<0.05). That is, in a general manner, intervention with dynamic geometry activities about geometry concepts had a significant effect on students' van Hiele geometric thinking levels even if we ignore their probability of mathematical giftedness (Table 7).

Table 7. Distribution of van Hiele Levels for main effect

Test	M	95% CI	Frequencies for levels		
			Level 0	Level 1	Level 2
Pretest	1,72	1,45 - 1,98	34 (64%)	19 (36%)	0 (0%)
Posttest	3,91	3,23 - 4,59	14 (26%)	20 (38%)	19 (36%)

In the Table 7, it was revealed that, at the beginning of the study, most of the students were at Level 0 of van Hiele geometric thinking hierarchy regarding mean of their pre-test scores (M=1.72, SD=0.97) and they were able to reach Level 1 of van Hiele hierarchy at the end of the study according to mean of their scores in post-test administration of the Van Hiele geometric thinking level test (M=3.91, SD=2.47). Therefore, this main effect

implies that the intervention with dynamic geometry activities could help students to move on from visualization level of van Hiele geometric thinking hierarchy to analysis level and also, even for some students, to informal deduction level as seen on the Table 7.

Furthermore, according to mixed model repeated measure ANOVA results on Table 6, there was also a significant interaction between probability of giftedness in mathematics of students and differences in their scores from pre-test and post-test administration of the Van Hiele geometric thinking level test ($F=2$, $50)=20.322$, $p<0,05$). This interaction effect signifies that students' probability of mathematical giftedness had influences on their gains from the intervention with dynamic geometry activities about geometry concepts in terms of van Hiele geometric thinking levels (Figure 1).
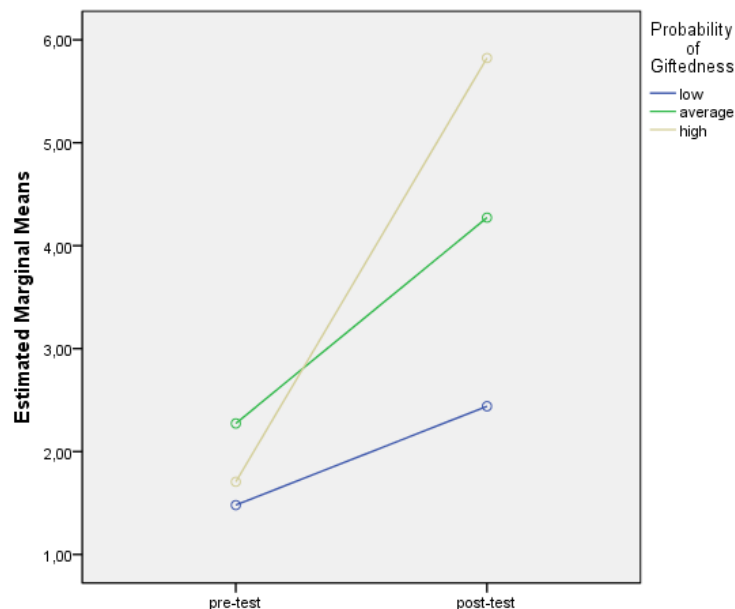
Figure 1. Estimated marginal means for van Hiele geometric thinking levels test

Students' placements for van Hiele geometric thinking hierarchy changed incrementally regarding their probability for giftedness in mathematics between pre-test and post-test administration of the Van Hiele geometric thinking level test. In detail, mean scores for Van Hiele geometric thinking level test administrations of low probability of mathematical giftedness students changed from 1,48 (SD=0.87) to 2,44 (SD=1.96), scores of students with average probability changed from 2,27 (SD=1.01) to 4,27 (SD=2.24) while scores of students with high probability changed from 1,71 (SD=0.99) to 5,82 (SD=1.88) due to intervention for dynamic geometry activities about geometry concepts. Hence, the estimated marginal means expressed nature of the interaction between students' probability of giftedness and their scores in Van Hiele geometric thinking level test administrations. The graph in Figure 1 clearly shows that all three groups of students benefited from the intervention positively. However, students with low probability of mathematical giftedness benefited from intervention at minimum level while students with high probability benefited from intervention at maximum level in terms of their scores in Van Hiele geometric thinking level test. Additionally, students with higher probability of mathematical giftedness performed more incremental progress from other groups of students (Table 8).

Table 8. Distribution of van Hiele Levels for main effect

Giftedness Probability	Test	M	Frequencies for levels		
			Level 0	Level 1	Level 2
Low	Pretest	1,48	19 (76%)	6 (24%)	0 (0%)
	Posttest	2,44	13 (52%)	9 (36%)	3 (12%)
Average	Pretest	2,27	4 (36%)	7 (64%)	0 (0%)
	Posttest	4,27	1 (9%)	6 (55%)	4 (36%)
High	Pretest	1,71	11 (65%)	6 (35%)	0 (0%)
	Posttest	5,82	0 (0%)	5 (29%)	12 (71%)

According to the interaction effect, intervention with dynamic geometry activities could help students to move between levels of van Hiele hierarchy but differently in terms of their probability of giftedness in mathematics

as described on Table 6. In all groups, some students accomplished to reach level 2 of van Hiele geometric thinking hierarchy but in different rates. 12% of students in low probability of giftedness group places at level 2, while 36% of students in average probability of giftedness group and 71% of students in high probability of giftedness groups places at this level.

Discussion

In this study, examining the effects of mathematics lessons integrated with dynamic geometry activities on students' van Hiele geometric thinking levels controlling their probability for mathematical giftedness was aimed. The results of the study enlightened the effects of mathematics lessons supported by dynamic geometry activities on geometric thinking as well as interactions among students' probability of mathematical giftedness and their gains for geometric thinking.

First of all, dynamic geometry activities in this study helped to move forward in levels of van Hiele hierarchy for fifth grade students. Many studies stated that dynamic geometry supported learning activities have great influences on geometric thinking (Karakuş & Peker, 2015; Özçakır, 2013). According to results, mathematics lessons supported by dynamic geometry activities helped students to make transitions from visualization level to analysis level of geometric thinking hierarchy regarding difference on mean scores from pre-test and post-test administration of Van Hiele geometric thinking level test. The main effect of intervention implied that all students benefited from education supported with dynamic geometry activities if we disregard their probability of mathematical giftedness and consider them as one group of students. Since these dynamic geometry activities provide students a dynamic learning environment to explore mathematical constructs like doing experiments for mathematics (Gawlick, 2005), their geometric thinking levels improved. This dynamic feature of the activities allows students to make experiments for mathematics by dragging, making manipulations, exploring quadrilaterals in different positions, and realizing unchanged and changed properties in these movements (Fidan & Türnüklü, 2010; Özçakır, 2013; Selçik & Bilgici, 2011). Therefore, dynamic feature of dynamic geometry activities gave students an opportunity of working with dynamic figures and so inspecting the same concept with numerous different drawings because a dynamic figure always preserves its basic properties and make these basic properties solid. In other words, with dynamic figures, a student can change the figure easily while maintaining its basic features. Therefore, students in this study, had access to all of these possible variations of triangle, rectangle, parallelogram, rhombus and trapezoid.

The other crucial point highlighted in the findings of the study is that students having high probability of mathematical giftedness were moved from visualization level to higher level of geometric thinking. Most of these students were in informal deduction level of van Hiele hierarchy at the end of the study that when compared to the low or average probability of students. In other words, most of the students who could reach to the informal deduction level at the end of the study were students having high probability of giftedness in mathematics. These findings coincided with the idea that mathematical gifted students require differentiated materials that are not restricted only with the curriculum requirements and the tasks that they can discover the ideas (Johnson, 2000). Moreover, as stated by Özdemir's (2016) study, technology integrated tasks, may enable them to follow with their own speed and provides an individualized pathway to them. As seen in the findings of the study, other students could only move at most one level, mathematical gifted students could go further. This can be interpreted in this way that such technology integrated tasks might be a good way to meet the differentiated needs of mathematical gifted students (Baykoç, 2010; Pierce et al., 2011). Furthermore, when it was seen in the Vygotsky's (1980) point of view, with regular classroom activities, mathematical gifted students have to follow other students' zone of proximal development by proceeding at the same pace with others. However, it is also essential that they are scaffolded in their own zone of proximal development to move further in mathematical concepts. Additionally, these findings also coincide with the Edwards's (2006) idea that, geogebra may be seen an effective tool for gifted students' materials due to its easiness in planning of enrichment activities. Similarly, El-Demerdash (2010) concluded that enrichment programs supported by dynamic geometry activities has a positive effect on mathematically gifted students' geometric creativity. Moreover, the other crucial point revealed from the analysis was such that the mean score of students having high probability of giftedness was not so high in the beginning of the study while they had a heavy increase at the end of the study. Even, the students who could reach to the level 2 was mostly high probability students. However, low level of these students' means scores when compared to other groups of students was remarkable in the pre-test scores. At that point, what needs to be discussed is the issue that with regular instruction and tasks, mathematical gifted students may get bored and they may not use their full potential (Johnson, 2000). However, by means of the study, they could find an interesting opportunity to go beyond in geometry.

Conclusion

Based on these, findings of the study could provide some valuable hints both to the area of theory and practice. Initially, providing students a learning environment, which includes dynamic geometry learning materials, gives them to access different variations of the same geometric figures thereby, explore mathematical constructs and make discoveries via dynamic interactions. This learning environment could help students to focus on properties of geometric constructs among different orientations of the same figures. Results of this study reflected that dynamic geometry activities helped students to move from shape properties to geometrical properties, namely relationship among shapes, and even from recognizing figures based on visualization to realizing relationship among shapes. Furthermore, the role of these activities on gifted students' performance is another issue that needs to be highlighted. Due to the fact that, both research and application area requires various examples that can be provided to gifted students, this study could be seen as an opportunity for this.

In conclusion, dynamic geometry activities are powerful medium for improving students' geometric thinking, especially for gifted students. Teachers can use activities consisted of dynamic figures while teaching quadrilaterals concepts or other many geometric concepts in order to engage their students in situations where they make mathematical experiments. Moreover, this study is limited to the participants of the present study that a more comprehensive similar study could be conducted with different samples so as to generalize the findings of the study. Moreover, a longitudinal and experimental study that can reveal the effects of such intervention to students' geometric thinking can be conducted for the further research study.

Acknowledgements

References

Battista, M. T. (2007). The development of Geometric and Spatial Thinking. In F. K. Lester, Jr., (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp.843-908). Charlotte, NC: Information Age Publishing.

Baykoç, N. (2010). Üstün ve özel yetenekli çocuklar ve eğitimleri [Education of gifted and specialized children]. In N. Baykoç (Ed.), *Öğretmenlik programları için özel eğitim [Special education for teaching programs]* (pp. 301-337). Ankara: Gündüz Eğitim ve Yayıncılık.

Baykoç, N. (2011). Üstün ve özel yetenekli çocuklar ve eğitimleri (Education of Gifted and Specialized Children). In N. Baykoç (Ed.), *Özel gereksinimli çocuklar ve özel eğitim [Special Needs Students and Special Education]* (pp.359-386). Ankara: Eğiten Kitap.

Baykoç, N. (2014). *Üstün yetenek gelişimleri ve eğitimleri [Education and Developments of Gifted]*. Ankara: Vize Yayıncılık.

Cansız - Aktaş, M., & Aktaş, D. Y. (2012). Investigating Explanation of Relations Between Quadrilaterals in the 7th Grade Elementary School Mathematics Curriculum, Textbooks And Students' Workbooks. *e-Journal of New World Sciences Academy NWSA-Education Sciences, 1C0548, 7*(2), 848-858.

Casa, T. M., Firmender, J. M., Gavin, M. K., & Carroll, S. R. (2017). Kindergartners' achievement on geometry and measurement units that incorporate a gifted education approach. *Gifted Child Quarterly, 61*(1), 52-72.

Crowley, M. L. (1987). The van Hiele Model of the Development of Geometric Thought. In M. M. Lindquist (Ed.), *Learning and Teaching Geometry, K-12, Yearbook of the National Council of Teachers of Mathematics* (pp.1-16). Reston, Va.: National Council of Teachers of Mathematics.

Dağlıoğlu, H. E. (2004). Okul öncesi eğitim kurumuna devam eden beş altı yaş grubunda ve matematik alanında üstün yetenekli olan çocukların sosyodemogrofik özellikler bakımından incelenmesi [Investigation of the children in the age group of five and six and who are gifted in mathematics in terms of sociodemogrophic characteristics]. In M.R. Şirin, A. Kulaksızoğlu & A.E. Bilgili (Eds), *1. Türkiye Üstün Yetenekli Çocuklar Kongresi Seçilmiş Bildiriler Kitabı* (pp. 247-261). İstanbul: Çocuk Vakfı Yayınları.

Davaslıgil, Ü. (2004). Yüksek matematik yeteneğinin erken kestirimi [Early estimation of high math ability]. In M. R. Şirin, A. Kulaksızoğlu, A.E. Bilgili (Eds.), *1. Türkiye Üstün Yetenekli Çocuklar Kongresi Seçilmiş Bildiriler Kitabı* (pp. 263-283). İstanbul: Çocuk Vakfı Yayınları.

Duatepe, A. (2004). *The Effects of Drama Based Instruction on Seventh Grade Students' Geometry Achievement, Van Hiele Geometric Thinking Levels, Attitude Toward Mathematics and Geometry* [Unpublished doctoral dissertation]. Middle East Technical University, Ankara, Turkey.

Deringöl, Y., & Davaslıgil, Ü. (2020). The effect of differentiated mathematics programs on the mathematics attitude of gifted children. Mojes. *Malaysian Online Journal of Educational Sciences, 8*(1), 27-37.

Davis, G. & Rimm, S. (2004). *Education of the gifted and talented (5th ed.).* Needham Heights, MA: Allyn & Bacon.

Edwards, J. A., & Jones, K. (2006). Linking geometry and algebra with GeoGebra. *Mathematics Teaching, 194,* 28-30.

El-Demerdash, M. (2010). *The effectiveness of an enrichment program using dynamic geometry software in developing mathematically gifted students' geometric creativity in high schools* [Unpublished doctoral dissertation]. University of Education, Karlsruhe.

Erdogan, A., & Yemenli, E. (2019). Gifted students' attitudes towards mathematics: a qualitative multidimensional analysis. *Asia Pacific Education Review, 20*(1), 37-52.

Fidan, Y. & Türnüklü, E. (2010). İlköğretim 5. Sınıf Öğrencilerinin Geometrik Düşünme Düzeylerinin Bazı Değişkenler Açısından İncelenmesi. [Investigation of Geometric Thinking Levels of Primary School 5th Grade Students in Terms of Some Variables]. *Pamukkale University Journal of Education, 27,* 185-197.

Gavin, M. K., Casa, T. M., Adelson, J. L., Carroll, S. R., Sheffield, L. J., & Spinelli, A. M. (2007). Project M3: Mentoring mathematical minds - A research-based curriculum for talented elementary students. *Journal of Advanced Academics*, *18*(4), 566-585.

Gawlick, T. (2005). Connecting arguments to actions - Dynamic geometry as means for attainment of higher van Hiele levels. *Zentralblatt für Didaktik der Mathematik, 37*(5), 361-370.

Greenes C. (1981). Identifying the gifted student in mathematics. In J., A. Hatch (Ed.) *Doing qualitative research in education settings*. Albany: State University of New York Press.

Hannah J., James A., Montelle C., & Nokes J. (2011). Meeting the needs of our best and brightest: curriculum acceleration in tertiary mathematics. *International Journal of Mathematical Education in Science and Technology, 42*(3), 299-312.

Johnson, D. T. (2000). *Teaching mathematics to gifted students in a mixed-ability classroom.* Reston, VA: Eric Clearinghouse.

Karaduman, G. B. (2010). Üstün yetenekli öğrenciler için uygulanan farklılaştırılmış matematik eğitim programları [Differentiated mathematics education programs for gifted students]. *Hasan Ali Yücel Eğitim Fakültesi Dergisi*, *13*(1), 1-12.

Karakuş, F. & Peker, M. (2015). The effects of dynamic geometry software and physical manipulatives on pre-service primary teachers' van Hiele levels and spatial abilities. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 6*(3), 338-365.

Kamarudin, M. F. B., Kamarulzaman, M. H. B., & Ishak, N. M. (2018). The Relationship between Gender, Age, and Attitude toward Mathematics among Malaysian Gifted Students. *The Educational Review, USA, 2*(8), 410-416.

Kontostavlou, E. Z., & Drigas, A. S. (2019). The Use of Information and Communications Technology (ICT) in Gifted Students. International Journal of Recent Contributions from Engineering, *Science & IT (iJES), 7*(2), 60-67.

Krutetski, V. A. (1976). *The psychology of mathematical abilities in school children*. Chicago: The University of Chicago Press.

Laborde, C., Kynigos, K. H. & Strasser, R. (2006). Teaching and learning geometry with technology. In A. Gutierrez, P. Boero (Eds), *Handbook of research of the psychology of mathematics education: Past, present and future* (pp. 275-304), Rotterdam, The Netherlands: Sense Publishers.

Leikin, R. (2009). Exploring mathematical creativity using multiple solution tasks. In R. Leikin, A. Berman and B. Koichu (Eds.), *Creativity in mathematics and the education of gifted students* (pp. 129-145). Rotterdam, the Netherlands: Sense Publishers.

Levenberg, İ & Shaham, C . (2014). Formulation of Word Problems in Geometry by Gifted Pupils. *Journal for the Education of Gifted Young Scientists, 2*(2), 28-40.

Malloy, C. (2002). *The van Hiele Framework in Navigating Through Geometry in Grades 6-8*. Reston, VA: NCTM.

Maryland, S. P. (1972). *Education of gifted and talented*. Washington DC: US Office of Education.

Maxwell, K. (2001). Positive learning dispositions in mathematics. *ACE Papers,* 11, 30-39.

Mayer, R. E. (2005). The scientific study of giftedness. *Conceptions of Giftedness,* 2, 437-447.

Miller, R. C. (1990). *Discovering mathematical talent* (ERIC Digest No. E482). Reston, VA: ERIC Clearinghouse on Disabilities and Gifted Education.

Ministry of National Education [MoNE]. (2018), *Matematik Dersi Öğretim Programı (Mathematics Curricula Program)*. Ankara: MEB.

National Association for Gifted Children (2005). *What is gifted?*. Retrieved from http://www.nagc.org/index.aspx?id=574&an

National Council of Teachers of Mathematics. (2010). *Curriculum and Evaluation Standards for School Mathematics*. Reston, VA: Author.

Olkun, S., Sinoplu, N. B., & Deryakulu, D. (2005). Geometric Explorations with Dynamic Geometry Applications based on van Hiele Levels. *International Journal for Mathematics Teaching and Learning,* 6, 1-12.

Özçakır, B. (2013). *The Effects of Mathematics Instruction Supported by Dynamic Geometry Activities on Seventh Grade Students' Achievement in Area of Quadrilaterals* [Unpublished master thesis], Middle East Technical University, Ankara, Turkey.

Özçakır, B. & Çakıroğlu, E. (2019). Effects of dynamic geometry activities on seventh graders' achievement in area of quadrilaterals. *International Journal for Mathematics Teaching and Learning 20*(2), 257-271.

Özdemir D. (2016). *Design and development of differentiated tasks for 5th and 6th grade mathematical gifted students* [Unpublished doctoral dissertation]. Ankara: Middle East Technical University.

Özdemir, D. (2018). Matematikte üstün yetenekli ilköğretim öğrencilerinin okullarındaki matematik derslerine ilişkin algıları [Perceptions of Gifted Students in Mathematics Related to Mathematics Courses in Their Schools]. *Kastamonu Eğitim Dergisi, 26*(1), 153-160.

Park, H. S., & Park, K. H. (2006). Analysis on the mathematical disposition of the mathematical gifted students in the middle school of Korea. *Journal of the Korea Society of Mathematical Education Series, 10*(2), 125-134.

Periathiruvadi, S. & Rinn, A. N. (2012). Technology in gifted education: *A Review of best practices and empirical research, Journal of Research on Technology in Education, 45*(2), 153-169.

Pierce R. L., Cassady J.C., Adams C. M., Neumeister K. L. S., Dixon F. A., & Cross T. L. (2011). The effects of clustering and curriculum on the development of gifted learners' math achievement. *Journal for the Education of the Gifted*, *34*(4), 569–594.

Pitta-Pantazi D. P., Christou C., Kontoyianni K., & Kattou M. (2011). A model of mathematical giftedness: integrating natural, creative, and mathematical abilities. *Canadian Journal of Science, Mathematics and Technology Education, 11*(1), 39-54.

Renzulli, J. S. (1979). *What makes giftedness? Reexamining a definition*. USA: Chronicle Guidance Publications.

Ryser, G.R. & Johnsen, S.K. (1998). *Test of mathematical abilities for gifted students.* Austin, TX: Pro-ed.

Salvia, J., & Ysseldyke, J. E. (1995). Using student portfolios in assessment. NJ: Houghton Mifflin Company.

Selçik, N. & Bilgici, G. (2011). The effects of the Geogebra software on students' academic achievement. *Kastamonu Eğitim Dergisi, 19*(3), 913-924.

Sheffield, L. J. (1994). *The development of gifted and talented mathematics students and the National Council of Teachers of Mathematics Standards* (Research Monograph No. 9404). Storrs: National Research Center on the Gifted and Talented, University of Connecticut.

Shillor, I. (1997). Gifted Mathematicians Constructing Their Own Geometries-Changes in Knowledge and Attitudes. *Gifted Education International, 12(2)*, 102–105.

Sriraman, B., Haavold P., & Lee K. (2013). Mathematical creativity and giftedness: a commentary on and review of theory, new operational views, and ways forward. *Mathematics Education, 45*(1), 215–225.

Sternberg, R. J. (1997). A Triarchic View of Giftedness: Theory and Practice. In N. Coleangelo & G. A. Davis (Eds.), *Handbook of Gifted Education* (pp. 43–53). Boston, MA: Allyn and Bacon.

Suydam, M. N. (1985). The Shape of Instruction in Geometry. Some Highlights from Research, *Mathematics Teacher*, 7, 481-486.

Tabach, M. (2011). A Mathematics Teacher's Practice in a Technological Environment; A Case Study Analysis Using Two Complementary Theories. *Technologgy, Knowledge and Learning, 16*(3), 247-265.

Tall, D. (1998). Information Technology and Mathematics Education: Enthusiasms, Possibilities & Realities. In C. Alsina, J. M. Alvarez, M. Niss, A. Perez, L. Rico, A. Sfard (Eds), *Proceedings of the 8th International Congress on Mathematical Education* (pp. 65-82), Seville: SAEM Thales.

Tanahan, E. (2006). *Actual and ideal instructional practices in California high school gifted geometry education* [Unpublished doctoral dissertation], University of Southern California.

Taylor, C. (2008). *Promoting Mathematical Understanding through Open-Ended Tasks; Experiences of an Eighth-Grade Gifted Geometry Class* [Unpublished Doctoral Dissertation]. Georgia State University, Georgia.

Tan-Şişman, G., & Aksu, M. (2012). Sixth Grade Students' Performance on Length, Area, and Volume Measurement. *Education and Science*, 37, 141-145.

Thomas, M. O. & Holton, D. (2003). Technology as a tool for teaching undergraduate mathematics. In A. Bishop, M. A. K. Clements, J. Keitel-Kreidt, & F. K. S. Leung (Eds.) *Second international handbook of mathematics education* (pp. 351-394), Netherlands: Springer.

Thomas, R. J. (2019). *Mathematics Achievement among Fourth Grade Gifted Students in Inclusion and Pull-Out Programs* [Unpublished doctoral dissertation]. Grand Canyon University.

Tieso, C. (2003). Ability grouping is not just tracking anymore. *Roeper Review,* 26, 29-38.

Trna, J. (2014). IBSE and gifted students. *Science Education International, 25*(1), 19-28.

Usiskin, Z. (1982). *van Hiele levels and achievement in secondary school geometry*. CDASSG Project. Chicago, US: The University of Chicago.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes,* MA: Harvard University Press.

Ysseldyke, J., Tardrew, S., Betts, J., Thill, T., & Hannigan, E. (2004). Use of an Instructional Management System to Enhance Math Instruction of Gifted and Talented Students. *Journal for the Education of the Gifted, 27*(4), 293-310.

The Effect of Student-Centered Approaches on Students' Creative Thinking Skills: A Meta-Analysis Study

Tufan Aytaç[1], Sultan Selen Kula[1]

[1]Kırşehir Ahi Evran University

The Effect of Student-Centered Approaches on Students' Creative Thinking Skills: A Meta-Analysis Study

Tufan Aytaç[1*], Sultan Selen Kula[1]
[1]Kırşehir Ahi Evran University

Abstract

The aim of this study is to determine the effects of student-centered approaches (SCA) on students' creative thinking skills (CTS). Meta-analysis method is used in the study. The scope of the study consists of Master and PhD theses and research articles addressing this issue in Turkey. As a result of the screening, it is seen that there are 104 studies between 1990 and 2020 in line with inclusion criteria. The total number of samples within the scope of these studies is 6434 students. 3215 of these students are in the experimental group and 3219 are in the control group. According to the results of the study and the random effects model, a moderate (d=0.73; [0.63; 0.82]) statistically significant, positive effect size is determined according to SCA in favor of the students' CTS. In other words, it is seen that when student-centered approaches and methods are applied in the lessons, creative thinking skills of students are positively affected. As a result of the moderator analysis, it is determined that the effect sizes of the studies differed according to education stage (p=0.00) and lessons (p=0.00). In the study, it is concluded that SCA is a meaningful variable that positively affects students' CTS.

Key words: Student centered, learner centered, creative thinking, meta-analysis.

Introduction

In the twenty-first century, it is expected to raise individuals with advanced thinking skills who can question, produce and use knowledge in the solution of problems. Raising this globally valued human profile has taken its place among the educational objectives of all countries. The question "*How can we raise individuals with advanced thinking skills?*" has led to many new approaches in education. One of the best ways to use thinking skills effectively is to make individuals active in their learning processes. In this context, student-centered approaches (SCA) such as active learning, learning by living and learning to learn shape today's education.

Instead of using a single teaching method or certain techniques, SCA transform the role of the teacher from the person who conveys information to person who create learning environments that facilitate learning. It is not possible for teachers who have adopted SCA to use a single teaching method in lessons (Darsih, 2018). Methods that facilitate learning and activate students is within the scope of SCA (Blumberg, 2008). Student-centered teaching is an approach that focuses on student learning rather than what the teacher does, respects student learning needs, learning styles and strategies, and puts the student at the center of learning (Blumberg, 2008; Brown, 2003). In student-centered classes, students work individually, in pairs or with small groups on different tasks and projects. Successful control of such a classroom environment is possible with comprehensive planning and effective classroom management (Brown, 2003). SCA allows the student to make sense of the content in student-centered classroom settings and allows the teacher to help students in this regard. In addition, SCA, which develops higher order thinking skills that are suitable for learning objectives and students' characteristics, is frequently used in effective classroom settings (Weimer, 2002). Trilling and Fadel (2009) and Yalçın (2018) state that using SCA in classroom improves students' thinking skills. In this context, it can be said that there is a significant relationship between SCA and thinking skills.

---

[*] Corresponding Author: *Tufan Aytaç, tufana60@gmail.com*

Other than basic skills, twenty-first century skills consist of three main skill areas: a) learning and innovation skills, b) knowledge, media and technology skills, and c) life and career skills. Learning and innovation skills consist of four sub-skills: critical thinking and problem solving, communication, cooperation and creativity. These skills are seen as the key to lifelong learning and creative thinking. Creativity means producing, evaluating, explaining and implementing new ideas. Creative thinking is a higher order thinking skill that aims to solve problems, present alternative ideas and solutions, synthesize or make sense through re-meaning, and use imagination (Kylonen, 2012; Özçelik, 2019; Trilling & Fadel, 2009; Yalçın, 2018). In the document Turkey's Education Vision 2023, emphasis is placed on an educational approach based on these four soft skills of 21st Century which are to be acquired by students (MEB, 2018). On a global scale, it is stated in many countries from Finland to Singapore that 9 of every 10 skills that will shape the future in 2020 and beyond are creativity or other creativity related skills. Creative thinking skills are at the top of the skills students need to learn (Trilling & Fadel, 2009; World Economic Forum, 2016).

SCA is expected to develop higher order thinking skills such as creative, critical and reflective thinking (American Psychological Association [APA], 1997; Perry & Karpova, 2017). It is generally known that creativity skills of students are not given importance in educational environments. Accordingly, it is seen that teachers keep the level of creativity and general intelligence equal (Sternberg, 2003). However, there are studies showing that the student's general intelligence or academic achievement and creativity are not directly related (Gajda, Karwowski & Beghetto, 2017; Kim, 2005). Torrance (1966) defines creativity as "*being sensitive to problems, malfunctions, lack of information, lost items, incompatibility, seeking solutions to difficulties and predicting*". So creative thinking can be defined as mental processes that lead to an invention, solution or synthesis in any field (Vanderbos, 2006). In short, individuals with higher CTS can produce different or original ideas in a wide variety of fields and make predictions based on existing situation. While creativity evokes both mental and performance-based activities, creative thinking evokes mostly cognitive activities (Doğan, 2005). Creative thinking individual produces new, original, unexpected, appropriate, useful and adaptable concrete objects and ideas (Sternberg & Kaufman, 2010). The importance of knowledge and motivation is emphasized in creativity (Runco & Chand, 1995). In order to talk about creativity in any subject, we encounter two important factors; the level of knowledge on the subject and the desire to discover. The motivation mentioned in creativity is mostly intrinsic motivation. When the student is free to choose the tasks in the teaching process, his motivation increases because this makes the task meaningful to the student. Thus, in order to increase the creativity of students, it is necessary to take advantage of SCA, which will make them active in the teaching process.

The aim of pedagogy, which develops innovation and creative thinking skills, is to provide students with the necessary equipment to be creative, be prepared for problems to be able to analyze and manage information and work with it. Teachers who want to see the effects of using SCA to develop students' CTS, provide feedback and guide their learning to help students think creatively about content in terms of achievement and content. Students are encouraged to become independent active learners. Student centered approaches and methods are used in this process (Drapeau, 2014; Özçelik, 2009).

Educational approaches that center the student include many different teaching strategies, methods and techniques. It is seen that meta-analysis studies on the subject are concentrated on the effects of these student-centered strategies, methods and techniques on students' academic success. There are many meta-analysis studies examining the effect of inquiry-based learning (Aktamış, Hiğde & Özden, 2016; Furtak, Seidel, Iverson & Briggs, 2012; Wang, Huang, Tsay, Lee, Lin & Kao, 2011), constructivist approach (Arık & Yılmaz, 2020; Erdamar, Aytaç, Demir & Demir, 2015; Tutal, Kaçire & Atabey, 2016), project based learning (Ayaz, 2014; Kaşarcı, 2013; Tutal, Kaçire & Atabey, 2016), multiple intelligence theory (Ayaz, 2014; Baş, 2016; Biçer, 2017; Çelik, 2013; Kaplan, Duran & Baş, 2015; Tutal, Kaçire & Atabey, 2016), active learning (Batdı, 2014), technology-based methods (Batdı, 2014b; Çelik, 2013; Çırak-Kurt, Yıldırım & Cücük, 2018; Demir & Başol, 2014; Kaya & Öçal, 2018; Li & Ma, 2010; Merchant, Goetz, Cifuents, Keeney-Kennicutt,& Davis, 2014), STEM (Becker & Park, 2011; Belland, Walker, Kim & Lefler, 2017), creative drama (Biçer, 2017; Cantürk Günhan, 2016; Çelik, 2013; Erden, Aytaç & Erden, 2016; Lee, Patall, Cawthon & Steingut, 2015; Tutal, Kaçire & Atabey, 2016; Özbey & Sarıkaya, 2019) collaborative learning (Biçer, 2017; Camnalbur & Mutlu-Bayraktar, 2018; Capar & Tarim, 2015; Çelik, 2013; Johnson, Johnson & Stanne, 2000), game based learning (Biçer, 2017), problem based learning (Çelik, 2013; Dağyar, 2014; Dochy, Segers, Bossche, & Gijbels, 2003; Tutal, Kaçire & Atabey, 2016), brain based learning (Gözüyeşil & Dikici, 2014), argumentation technique (Karakuş & Yalçın, 2016; Tutal, Kaçire & Atabey, 2016), concept maps (Okursoy-Günhan, 2019) and student centered

approaches (Tutal, Kaçire & Atabey, 2016; Yeşilpınar Uyar & Doğanay, 2018) on students' academic achievement.

It is seen that meta-analysis studies in the literature focus on different variables as well as the effects of SCA on students' academic achievement. It is noteworthy that there are also meta-analysis studies investigating the effect of research and inquiry-based teaching on students' scientific process skills (Akkaya, 2019); project-based learning on science lesson attitude (Ayaz, 2014; Kaşarcı, 2013); 5E learning model (Ayaz, 2015) and constructivist approach on lesson attitudes (Ayaz & Şekerci, 2016); cooperative learning on mathematics attitude (Capar & Tarim, 2015); technology-based methods on learning mathematics (Özdemir, Aslaner & Açıkgül, 2020) and foreign languages (Chiu, Kao & Reynolds, 2012; Peterson, 2010); web-based problem solving (Kuo, Chen and Hwang, 2014); argumentation technique on scientific process skills (Karakuş & Yalçın, 2016); creative drama on motivation and social skills (Özbey & Sarıkaya, 2019) and skill development (Cruz, Lian & Morreau, 1998; Ulubey, 2018). However, when the studies conducted in Turkey and other countries, no meta-analysis study is found to examine the effects of SCA on students' CTS.

The question *"is there a relationship between SCA applied in the class and the students' CTS?"* has been asked for a long time, but there is no clear answer yet. At the conceptual level, scientists claimed that creativity and learning represent interrelated phenomena (Beghetto, 2016; Sawyer, 2012; Vygotsky, 2004). It is seen that there are many studies in Turkey in recent years dealing with the effects of SCA on thinking skills of students. The increasing number of studies dealing with the effect of various student-centered approaches, methods and techniques on students' creative thinking skills (CTS) has revealed the need to compile these results and synthesize them by taking into account the sample numbers. However, it is seen that there are no meta-analysis studies that deals with the effects of SCA on students' CTS. In addition, it is seen that there are no holistic and comparative studies that examine the effects of SCA on students' CTS. In this context, determining whether SCA has an impact on the students' CTS is the problem of this study. The aim of this study is to determine the effects of SCA on students' CTS.

Method

Research Design

Research model of the study is meta-analysis method, which is a method of systematically analyzing and synthesizing the data of quantitative studies carried out independently on the same subject. In the analysis of the data, one of the group comparison meta-analysis methods, Group Difference Method is used (Card, 2012; Cumming, 2012, 205). In a meta-analysis study; single and independent quantitative studies related to the same research question and subject are selected according to the inclusion criteria and the data obtained from these studies are synthesized with advanced statistical methods and their effect sizes are determined and interpreted (Dinçer, 2014; Ellis, 2012:5). The purpose of a meta-analysis is to reach a synthesis with the least error by comparing the quantitative data obtained from empirical studies conducted in different places and times on the same subject, combining them with appropriate methods, increasing the number of samples and thereby reducing the confidence interval of the overall result (Cumming, 2012; Hartung, 2008). Meta-analysis process steps are shown in Figure 1 (Dinçer, 2014, 11).
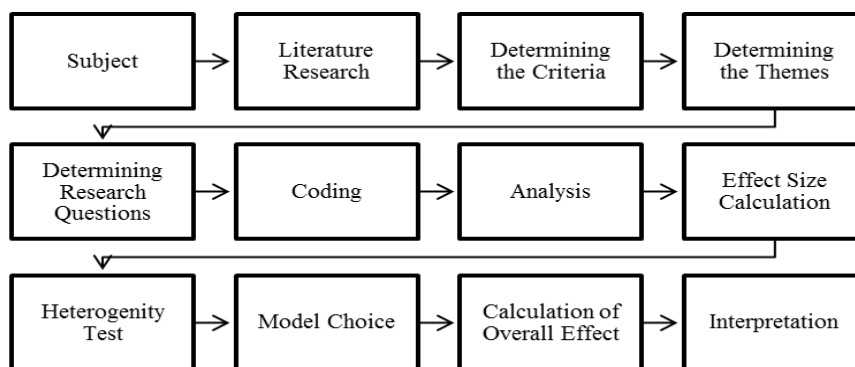
Figure1. Meta-analysis process steps

Data Collection

Master's and PhD theses, and research articles (conference paper, book chapter etc.) which discuss the subject of this study in Turkey constitute the main data sources and scope of this study. In order to have access to the relevant studies, the keywords "creativity", "creative thinking", "learner/student centered", "student centered strategies", "teaching strategies and methods" are searched both in Turkish and English on Web of Science, ERIC, ULAKBİM, EBSCOhost, Google Scholar and YOK National Thesis Center databases. After the search, it is determined that there are 104 studies in accordance with the inclusion criteria from 187 studies conducted on the subject of the study. In this study, studies focusing on experimental and quasi-experimental studies are selected which especially have pretest-posttest, and comparisons between groups. The inclusion criteria used in the selection of the studies included in the research are given in Table 1.

Table 1. Inclusion Criteria of the Studies

Inclusion Criteria
1. *Published or unpublished study sources*: Master's and PhD theses and research articles published in the literature are covered.
2. *Appropriateness of the dependent and independent variables in the studies to the meta-analysis study*: In order to reach the effect size in the meta-analysis studies it is taken into account that the included studies are empirical studies and that SCA is used as an independent variable and CTS is used as a dependent variable.
3. *Sample group*: Studies involving students who are studying in the formal education institutions from preschool education up to postgraduate education in Turkey are included.
4. *Containing the quantitative data required for meta-analysis*: It is taken into account that studies contain quantitative data (mean, standard deviation, number of samples, p value, etc.) in order to calculate the effect sizes required for the meta-analysis study.
5. Studies carried out in Turkey between 1990 and 2020 are taken into account.
6. Studies in both Turkish and English languages with a sampling from Turkey are included in scope.

The process of determining the studies included in the meta-analysis according to the inclusion criteria mentioned above is given in Table 2.

Table 2. The Process of Determining the Studies Included in the Scope of Meta-analysis

1. Number of studies reached within the context of determined keywords	2. Number of studies not included in the scope according to exclusion criteria	3. Number of studies in accordance with inclusion criteria
187 studies (34 PhD thesis, 98 Master thesis, 55 Articles)	83 studies (11 PhD thesis, 51 Master thesis, 21 Articles)	104 studies (23 PhD thesis, 47 Master thesis, 34 Articles)

Reliability of the Research. In a meta-analysis study, reliability between coders is important in the coding phase of the studies related to the reliability of the results. For this purpose, a coding protocol and form containing the identity, content and data of the study are created. In order to ensure the reliability of coding, the studies that will be included in the meta-analysis must be coded by at least two coders (Cooper, 2009; Wilson, 2009). The reliability of the codings was calculated using the formula "Reliability=Consensus/(Consensus + Disagreement) x 100" (Miles and Huberman, 1994) and was 96%. Values of 70% and above obtained from this formula are sufficient for reliability (Yıldırım & Şimşek, 2011). It can be said that the codings performed in this framework are reliable.

Validity of the research. All studies that meet the inclusion criteria for the meta-analysis are searched on all available databases and their inclusion is an indication of the validity of the research (Petticrew, & Roberts, 2006). In this context, each of the 104 studies included in the meta-analysis are examined in detail, and it is confirmed that the validity and reliability of the data collection tools used in studies are provided. Therefore, it can be said that this meta-analysis study is also valid.

Reporting. "Preferred Reporting Items for Systematic Reviews and Meta Analysis (PRISMA)" writing guide, which is a protocol used for systematic review and meta analysis, is used in the reporting of this study (Moher, Liberati, Tetzlaff & Altman, 2009). Turkish translation of the PRISMA checklist is made by Aşık & Özen (2019); Turkish version published on the official website of PRISMA is used (Aşık & Özen, 2019; Prisma Checklist, 2019: Figure 2).
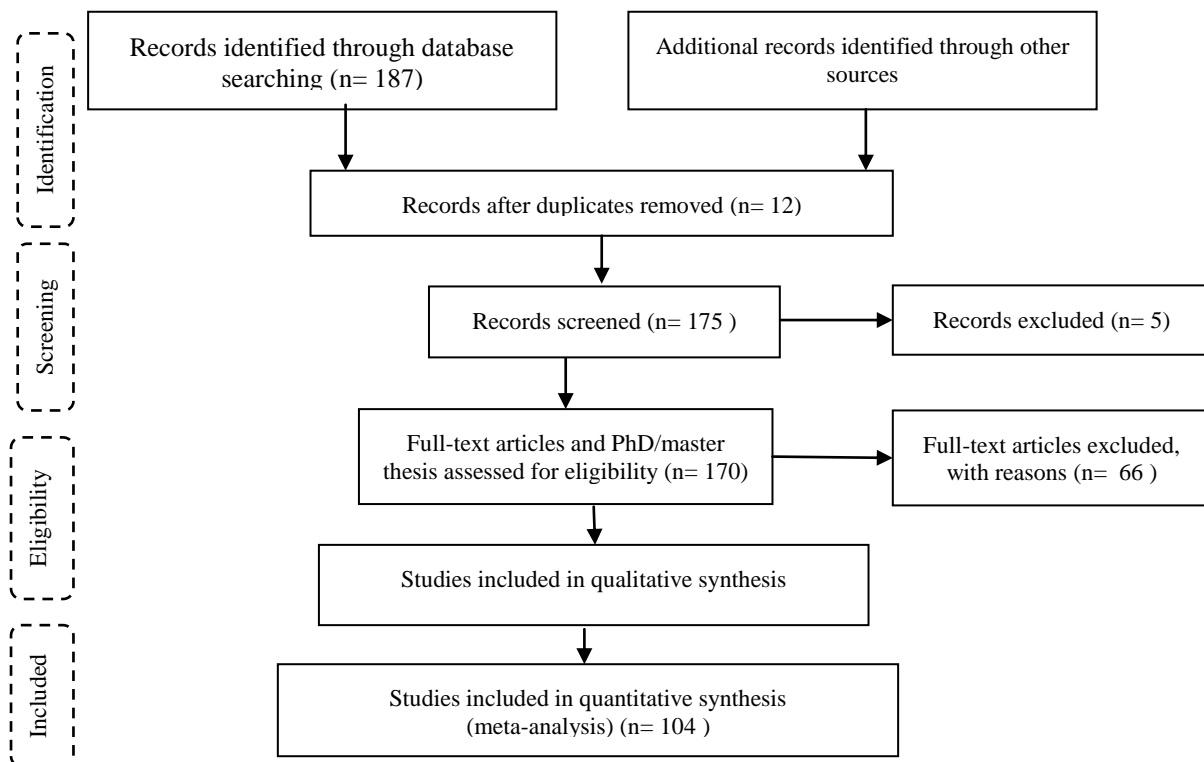
Figure 2. Prizma Flow Diagram for Meta-analysis

Data Analysis

For statistical calculations of this study, CMA [Comprehensive Meta Analysis] Ver. 2 software is used. In this meta-analysis study, random effects model is used to calculate the overall effect size. In the random effects model, universe effect sizes are assumed to vary from study to study, and studies included in the meta-analysis are considered to be part of a distribution. When studies are obtained from published literature, it is appropriate to use random effects model. When the samples of the studies within the scope of meta-analysis show a heterogeneous structure, random effects model can be used. Furthermore, if the researcher wants to generalize the universe and has a representative sample of the universe, it is appropriate to select the random effects model (Borenstein, Hedges, Higgins, & Rothstein, 2013; Başol, 2016; Başol, Doğuyurt & Demir, 2016; Ellis, 2012). Thus, it is predicted that it would be more appropriate to use the random effects model since the studies vary both in patterns and variables, that is, they are heterogeneous. In the study, students in the classes where SCA are applied are taken as experimental group and students in the classes where SCA are not applied are taken as control group. Therefore, the positive effect size is interpreted in favor of the students in the classes where SCA are applied, and the negative effect size is interpreted in favor of the students in the classes where SCA are not applied.

Publication Bias

In this study, publication bias is calculated by using Funnel plot, Orwin's Fail-Safe N., Duval and Tweedie's Trim and Fill method, Egger's test, and Kendall's Tau coefficient (Borenstein, Hedges, Higgins, & Rothstein, 2009; Cooper, 2009). As seen in Figure 2, the majority of 104 studies included in the study are located towards the top of the figure and highly close to the combined effect size. In this sense, the funnel plot shows that there is no publication bias for the studies included in the study (Borenstein et. al., 2009). It can be said that there is no publication bias in this study according to the evidence that pointers of effect sizes and sample sizes take the form of a funnel, the peak point is at the same level as the actual effect size, and the effect sizes are gathered symmetrically at the top (Bakioğlu & Özcan, 2016; Borenstein et al., 2013). When the funnel plot is examined, it is seen that the studies are mostly within the boundaries of the graph and the number of studies showing asymmetric distribution is low. In other words, the majority of the studies in the distribution are not

concentrated on only one side. When the funnel plot is examined, it is seen that the studies are mostly within the boundaries of the graph and the number of studies showing asymmetric distribution is low. In other words, the majority of the studies in the distribution are not concentrated on only one side.
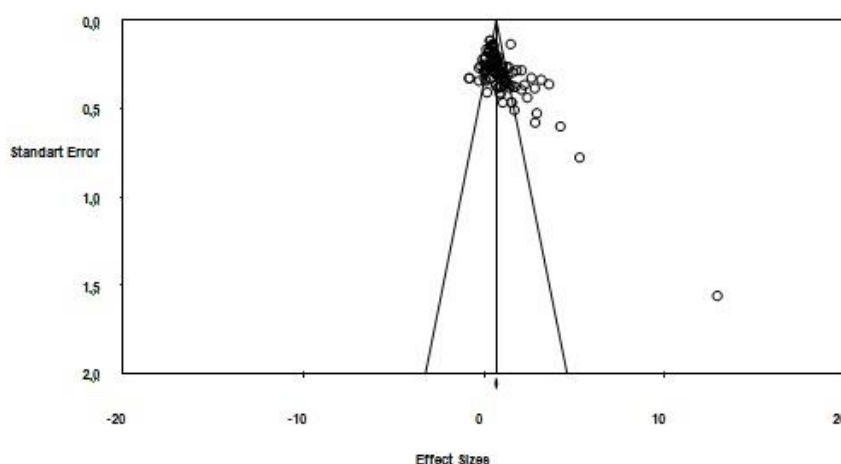
Figure 3. Funnel Plot

Publication bias test results of the studies included in the meta-analysis are given in Table 3. Orwin's Fail-Safe N calculation is also made to test publication bias. Orwin's Fail-Safe N calculates the number of studies that may be missing in a meta-analysis (Borenstein et al., 2009). As a result of this analysis, Orwin's Fail-Safe N is calculated as 6991. The number of studies required for the average effect size of 0.88 found in meta-analysis to reach 0.01 level (trivial), that is almost zero effect level, is 6991. 104 studies that are determined according to inclusion criteria are all of the studies conducted for this research question in Turkey. Apart from these, since 6991 studies are unlikely to be reached, this result is accepted as another indication that there is no publication bias in this meta-analysis.

The fact that the result of the Egger's test ($p=0.08$) was not significant is accepted as another indication that there is no publication bias in this meta-analysis. Another method, Kendall's Tau Coefficient is found to be 6.28 and $p=0.41$, so it is statistically revealed that there is no publication bias since the expectation that the p-value does not make a significant difference, i.e., greater than 0.05 is met. According to the results of Duval and Tweedie's Trim and Fill Method, when 29 precedent studies are included, the 0.88 average effect size found in the meta-analysis has changed to 0.49. This change can't be ignored, but it can be accepted that the reported effect size is reliable According to other tests, it may say that there is no publication bias, but a change from .88 to .49 should also be taken into consideration.

Results

Findings acquired from researches within the scope of meta-analysis study (random effects model, and moderator analysis) are given in this chapter.

Findings of Effect Sizes Combined by Fixed and Random Effects Model and Homogeneity Test Results
The mean of effect sizes regarding the effects of SCA on students' CTS which is combined according to the fixed and random effects model, standard error, and lower and upper limits in accordance with 95% confidence interval are given in Table 4.

Table 3. Combined Findings of the Effect Size Meta-Analysis of Studies by Fixed and
Random Effects Model and Homogeneity Test

Model/Sub-dimension				Homogeneity				
Random effects	k	Effect Size	Standart Error	Lower Limit	Upper Limit	*P*-Value	Q-Value	I^2
RBL	1	0.09	0.21	-0.32	0.50	0.66		
AL	4	1.53	0.43	0.69	2.36	0.00		
AR	2	0.37	0.47	-0.56	1.30	0.44		

STH	1	1.37	0.36	0.66	2.07	0.00		
BBL	2	1.28	0.27	0.74	1.81	0.00		
TE	20	0.85	0.16	0.54	1.16	0.00		
DI	8	1.18	0.32	0.55	1.82	0.00		
CL	3	0.42	0.18	0.06	0.77	0.02		
GBL	3	1.25	0.65	-0.03	2.53	0.06		
CS	2	0.81	0.22	0.37	1.25	0.00		
ST	9	1.62	0.40	0.82	2.41	0.00		
PBL	13	0.53	0.13	0.28	0.78	0.00		
PJTL	8	0.82	0.21	0.41	1.23	0.00		
STEM	7	0.67	0.15	0.37	0.97	0.00		
TBL	7	0.70	0.10	0.51	0.88	0.00		
CD	9	1.19	0.37	0.46	1.91	0.00		
CWR	5	1.03	0.21	0.61	1.45	0.00		
DIMENSIONS	104	0.73	0.04	0.63	0.82	0.01	37.88	86.17

SCA: Thinking education (TE), Problem based learning (PBL), Creative drama (CD), Story telling (ST), Differentiated instruction (DI), Project Based Learning (PJTL), Science, technology, engineering, and mathematics (STEM), Technology-based learning (TBL), Creative writing-reading (CWR), Cooperative learning (CL), Game based learning (GBL), Argumentation (AR), Research-inquiry based learning (RBL), Active learning (AL), Brain-based learning (BBL), Case study (CS), Six thinking hats (STH).

As seen in Table 3, the average effect size value of the effect size values of the studies included in the study according to the random effects model according to the SCA variable is calculated as ES=0.73. According to the random effects model, the students in the SCA applied classes show that CTS is higher than the classes in which traditional teacher-centered approaches are used. In the interpretation of the effects size, Cohen (1988, 1992) proposed the standards for interpreting OR, SMD, and r shown in Table 4 (Bakioğlu & Özcan, 2016: Littell, Corcoran & Pillai, 2008). The effect size value of this study, ES = 0.73, can be evaluated as a medium effect size according to Cohen's classification.

Table 4. Cohen's (1988) Standards for Interpreting Effect Size (ES)

ES Metric	Small Effect	Medium Effect	Large Effect
OR	1.5	2.5	4.3
SMD	0.2 0	0.50	0,80
R	0.1	0.25	0.4

OR: odds ratio; SMD: standardized mean difference; R: correlation coefficient.

It is seen that, among SCA used in the teaching-learning process, approaches and methods that have the greatest impact on students' CTS are; story telling (d=1.62), active learning (d=1.53), six thinking hats (d=1.37), brain-based learning (d=1.28), game-based learning (d=1.25) and creative drama (d=1.19). On the other hand, it is seen that research-inquiry based learning (d=0.09), argumentation (d=0.37) and cooperative learning (d=0.42) have a lower level of effect on students' CTS.

Homogeneity Test and Q and I^2 Statistics
For homogeneity test, Q is calculated as 37.88 (Q=37.88). From the chi-square table, 16 degrees of freedom is found to be 7.96 at 95% significance level. Since the Q-statistic value (Q = 37.88) exceeds the critical value of the chi-square distribution with 16 degrees of freedom ($\chi 2$ 0.95=7.96), the absence hypothesis of the distribution of effect sizes was rejected in the fixed effects model.
Recently, the I^2 index has been proposed to quantify the degree of heterogeneity in a meta-analysis (Huedo, Sanchez, Julio, Fulgencio and Botella (2006). I^2 shows the the ratio of the total variance of the effect size. I^2 reveals the percent variability of effect size estimates due to heterogeneity (Bakioğlu ve Özcan, 2016: 197). In addition, the 86.17 % value obtained as a result of calculating the I^2 value indicates high level of heterogeneity. The I^2 statistic describes the percentage of total variation across studies that is attributable to heterogeneity rather than chance (Higgins et al., 2003). A value greater than 25% is considered to reflect low heterogeneity,

50% moderate, and 75% high heterogeneity. In the interpretation of I^2, 25% indicates low level of heterogeneity; 50% indicates medium level of heterogeneity; and 75% indicates high level of heterogeneity (Cooper et. al., 2009, 263; Higgins et al., 2003). As a result of homogeneity tests (Q and I^2) for SCA variable, since there is a level of heterogeneity which is close to high level between studies, moderator analyzes are carried out to determine the possible causes of this heterogeneity.

Results of the Moderator Analysis According to SCA Variable
Results of the moderator analysis which was carried out in order to reveal the reasons of heterogeneity occurring as a result of SCA variable are given in Table 5.

Table 5. Categorical Moderator Results Related to the Effects of SCA on students' CTS

Moderator	k	d	SE	%95 CI	Q	p
Publication Type					3.77	0.15
MA	47	0.88	0.10	[0.67; 1.09]		
PhD	23	1.13	0.17	[0.78; 1.48]		
Article	34	0.74	0.09	[0.55; 0.93]		
Education Stage					22.10	0.00
Preschool	19	1.59	0.23	[1.14; 2.04]		
Primary	55	0.88	0.09	[0.70; 1.05]		
Secondary	9	0.31	0.20	[-0.08; 0.70]		
Higher	21	0.60	0.10	[0.40; 0.80]		
Lessons					130.5	0.00
Information Technologies	11	0.50	0.10	[0.31;0.70]		
Biology	1	0.54	0.15	[0.25;0.84]		
Geography	2	-0.79	0.23	[-1.25;-0.33]		
Thinking Education	2	1.22	0.25	[0.73;1.71]		
Philosophy	2	1.03	0.34	[0.35;1.70]		
Science	18	0.52	0.11	[0.32;0.73]		
Physics	1	0.62	0.20	[0.23;1.01]		
Visual Arts	3	0.92	0.20	[0.52;1.30]		
Life Science	2	0.49	0.42	[-0.33;1.30]		
English	1	0.19	0.41	[-0.61;0.99]		
Mathematics	10	0.95	0.21	[0.54;1.36]		
Preschool	19	1.60	0.23	[1.15;2.04]		
Assessment and Evaluation	1	2.64	0.33	[2.00;3.28]		
Free Activity	3	0.55	0.13	[0.29;0.80]		
Social Studies	9	1.69	0.46	[0.80;2.58]		
All Lessons	1	0.33	0.16	[0.01;0.64]		
Turkish	12	0.99	0.18	[0.65;1.34]		
Creative Drama	5	0.48	0.19	[0.10;0.86]		
Intelligence Games	1	1.41	0.27	[0.88;1.92]		
Region of the Study					5.22	
Mediterranean	9	0.88	0.13	[0.61; 1.14]		
Eastern Anatolia	12	0.91	0.20	[0.52; 1.30]		
Aegean	7	0.91	0.48	[-0.04; 1.87]		
Southeast	2	0.30	0.57	[-0.82; 1.44]		
Central Anatolia	31	1.12	0.14	[0.84; 1.39]		
Black Sea	13	0.74	0.14	[0.45; 1.03]		
Marmara	30	0.80	0.10	[0.59; 1.01]		
School Type					2.52	
State	86	0.88	0.07	[0.73; 1.02]		
Private	17	0.87	0.16	[0.55; 1.19]		
Science and Arts Center	1	1.69	0.50	[0.69; 2.69]		

P.S.: k=number of studies, d= Cohen's d

(SOF), SE=Standard Error CI=Confidence
Interval, Q= Heterogeneity between studies,
Comparison analyzes are made for studies with
sub-dimension numbers of 2 or more. *p<.05

As a result of the moderator analysis, it is determined that the effect sizes of the studies differ according to the education stage (*p*=0.00) and lessons (*p*=0.00). The results of the studies dealing with the preschools in terms of education stage show that the students' CTS are higher (*d*=1.59) in favor of SCA variable. In terms of lessons, especially in preschool education stage (*d*=1,60), it is seen that when SCA are applied in Social Studies (*d*=1,69), Assessment and Evaluation (*d*=2,64), Turkish (*d*=0,99) and Mathematics (*d*=0, 95) lessons, students' CTS increased more compared to other lessons. As an interesting result, it is observed that the implementation of SCA in geography lessons does not have any effect on students' CTS (*d*=-0.79).

It is determined that the effect sizes of studies according to the publication type (*p*=0.15), school type (*p*=0.28), region of the study (*p*=0.51) does not differ significantly.

Conclusion and Discussion

In this study, 104 effect sizes are calculated from 104 studies constituting a sample of 6434. According to the results of the study, a statistically significant moderate effect size is determined according to the random effects model (*d*=0.73; [0.63; 0.82]) in favor of the students in the experimental group according to the SCA variable. This result is a medium level according to the classifications of Cohen (1988). The results of this meta-analysis study show that SCA practices are a significant variable that positively affects students' CTS.

According to the random effects model in the context of SCA variable, the result of a moderate positive difference in favor of the students' CTS is in parallel with the results of the researches conducted by Akar & Şengil Akar (2013), Akkılıç (2018), Akran & Aşıroğlu (2018), Aksoy (2018), Aktamış, Hiğde & Özden (2016), Altındağ et. al., (2012), Arkan Sezgin & Baysal (2019), Atalay (2014), Avcu (2014), Aydın (2011), Bacak (2008), Batdal Karaduman (2012), Birişçi & Karal (2011), Bulut (2015), Chung & Ro (2004), Çetinkaya (2014), Çolakoğlu (2018), Dere & Ömeroğlu (2018), Engin (2019), Kaya (2018), Terzi (2019), Tezci (2002), Tut (2018), Ülger & İmer (2013), Üret (2019), Yıldırım (2018), Yıldırım & Akman (2020), Yıldız (2012), Yiğit & Erdoğan (2008). The view that the use of SCA in lessons is effective in the development of higher order thinking skills of students, which also include CTS (Brown & Freeman, 2000; Galton, et. al., 2009; Pascarella, et al., 2013) supports this meta-analysis study. In the national and international literature, a meta-analysis study examining the effect of SCA on students' CTS is not found. In this context, it is not possible to compare the results of this study with the findings of other meta-analysis studies. No meta-analysis study dealing with the effect of various SCA on the CTS or a study on the holistic effect is not found in the literature.

It is seen in both individual studies and in this meta-analysis study that students' CTS increase in a classroom where SCA are used. However, in the moderator analysis conducted to reveal which approaches and methods are more effective, it is seen that approaches and methods that have the most effect among SCA on students' CTS are respectively; storyline, active learning, six thinking hats, brain-based learning, game-based learning and creative drama. The common point of these approaches are that they provide individual and group work environment, develop multiple learning skills, and bring in skills such as collaboration, discussion and empathy. In the meta-analysis study conducted by Hsen-Hsing Ma (2009), it is seen that problem solving and communication skills stand out among the important variables that affect creativity. In this context, especially active learning, brain-based learning, creative drama and game/storyline methods can be considered meaningful in the context of having more of these two skills. On the other hand, the effects of research-inquiry based learning, argumentation, and cooperative learning approaches on students' CTS are slightly lower than other methods. In the meta-analysis study conducted by Ören and Sarı (2019), the effect of research-inquiry based learning on students' higher order thinking skills is moderate and positive (ES=0.66), while in this meta-analysis study, the effect on CTS which is among higher order thinking skills is insignificant and low. This may be due to the fact that there is the only one study included in the scope of meta-analysis (Şensoy & Yıldırım, 2017). Generally, it is seen that research, interaction and activity-based learning positively affect students' higher order thinking skills, especially CTS (Aktamış, Hiğde & Özden, 2016). The findings of this meta-analysis study also support this view. In the context of this meta-analysis study, it can be said that learning environments where approaches and methods such as constructivist approach, storyline, educational games, active learning and drama are applied, the opportunities offered to students have a positive effect students' CTS. In these learning

environments, students actively work (Aydın, 2011) make inquiries, scientific method process is applied, and in this way, many higher order thinking skills, especially scientific process skills (Karakuş & Yalçın, 2016), are developed (Perry & Karpova, 2017).

Another result obtained in the study is that in terms of education stage, students' CTS is higher in favor of SCA in preschool education stage. Theorists (Hertzog, 2008; Özçelik, 2019; Vygotsky, 2004) that deal with the developmental processes of CTS state that the creativity of younger children is more developable and they use these skills more to make sense of the world around them. In the study conducted by Yaşar and Aral (2010), the finding that preschool education stage has a high impact on students' CTS also supports the results of this meta-analysis study. In the study conducted by Krauksta, Rozenvalde and Ciekurs (2016); It is revealed that active learning approaches involving activities carried out in extra-curricular environments and storyline and game activities contributed to the development of CTS in the preschool education stage. In the study conducted by Atay (2009), it is stated that students in preschool education stage have higher level of CTS than students at other stages, and that other than the use of SCA, imagination, flexible thinking and learning speed/agility are also effective as a feature of their cognitive development period. These meta-analysis results reveal that students contribute more to the development of CTS compared to other education stages, especially in this stage where the activity-intensive learning-teaching process is carried out. As we move from preschool education to higher education, it is seen that students' CTS decrease. This situation can be explained by the increase in the effect of a teacher-centered and exam-oriented education approach in primary and secondary education stages and the less use of SCA. However, the developmental process does not follow a linear rise in the context of creativity, it is stated that there is a collapse in CTS after the 4th grade (Hong & Milgram, 2010; Raina, 1980; Torrance, 1968). In addition, it is known that the students in the upper classes get higher scores in creative thinking than the students in the lower grades among classes 1 to 5 (Hong & Milgram, 2010; Sak & Maker, 2006). Due to this developmental difference in creativity skills, it is thought that in this meta-analysis stud, results are in favor of preschool students. In addition, the finding that preschool teachers have high creative thinking tendencies in the study conducted by Eskidemir and Tezel (2019) supports the results of this meta-analysis study. In a sense, this can be evaluated as that the attitudes of the teachers are important in the development of the students.

In terms of scale type, Whetton and Cameron (2002) Creative Thinking Skill Scale "How Creative Are You?" which is used in 19 studies and The Torrance (1966) Creative Thinking Test scale "Test of Creative Thinking" used in 65 studies produced similar results. In the literature, it is seen that these two scales are often used to measure CTS in the world and in Turkey and that the validity/reliability is higher (Çetin, Üstündağ, Kerimoğlu and Beyazıt, 2015, Kim, 2011). It is seen that the effect size of the studies using these two scales (ES=0.70-0.79) is very close to the effect size found in this meta-analysis study (ES=0.73). However, there are significant differences between other scales in terms of effect sizes. In addition to the valid/reliable scales that measure students' CTS in a way that matches the results of the study, it is also seen that different measurement tools give results that are not close to each other (Carson, Peterson & Higgins, 2005; Dollinger, Urban & James, 2004; Haasse et al., 2018).

When the gender of the researcher is female, it is seen that the students' CTS are higher in favor of SCA. This can be explained by the fact that only 26 of the 104 studies included in the meta-analysis study are done by male researchers and that female researchers have more interest and motivation for learning and studying on this subject. In addition, the fact that studies in the preschool education stage and studies where the students have higher CTS are carried out by female researchers can be effective in this.

In terms of the moderator variable, it is observed that the students are more likely to increase their CTS when SCA are applied in lessons in preschool education stage and Social Studies, Philosophy, Assesment and Evaluation, Turkish, Mathematics and Visual Arts lessons in different education stages. Since teacher centered approaches and methods are generally used in these lessons, it is clear that when SCA are applied, it will contribute more to the development of higher order thinking skills of students (Perry & Karpova, 2017; Weimer, 2002). As an interesting result, it is observed that the application of SCA does not have any effect on students' CTS in geography lessons. Geography lessons in Turkey are carried out with many concepts taught with a complex structure (Akbulut, 2004). Although student centered approaches are tried to be used, it is difficult to increase creativity in schools without geography laboratories (Gardner et al., 1997; as cited in Akbulut, 2004). This interesting result of the study is thought to have originated from the conditions in which Geography lessons are carried out in Turkey.

It is determined that the effect size of the studies does not differ according to publication type, student group (normal-gifted), school type (state, foundation, science and arts center) and region of the study. Since no meta-

analysis studies that reveal the effect of SCA on students' CTS carried out in Turkey and abroad are found, it is impossible to compare results.

Furthermore, it is seen that the positive effect of SCA used in lessons on students' CTS continues within years of which meta-analysis studies are carried out (1990-2020). The adoption of the constructivist approach in the education programs in Turkey since 2005 has increased interest in the higher order thinking skills such as CTS. Therefore, the research carried out in Turkey dealing with CTS is known to increase in numbers after 2005 (Saracaloğlu et. al., 2014).

Recommendations
Recommendations for Researchers

- In the context of the results of this meta-analysis study, meta-analysis studies can be conducted by using variables that have effect on students' CTS other than SCA applications such as teacher quality, school culture, out-of-school activities, family and social environment.
- In the context of the results of this meta-analysis study, it can be suggested to make qualitative and quantitative studies on the factors that may have effect on students' CTS. Studies can be conducted to determine the effects of SCA on higher order thinking skills other than CTS (problem solving, meta-cognitive thinking etc.).
- It can be suggested to make meta-synthesis studies on the factors that may have effect on students' CTS.
- It can be suggested to make research conducted outside the Turkey in order to compare and to generalize the findings of the meta-analysis.

Recommendations for Implementers

- As a result of the meta-analysis study, it is observed that the learning-teaching approaches such as storyline, active learning, six thinking hats, brain-based learning, game-based learning and creative drama are more effective on students' CTS than other approaches. In this context, it can be suggested that teachers use these approaches more in the lessons for the development of students' CTS. In addition, it can be suggested to use more of SCA and activities that develop students' CTS in primary and secondary education.
- It can be suggested to use teaching methods and techniques based on social interaction in cognitive sense in order to develop students' creative thinking skills.

References
(The symbol of * refers to the studies included in the meta-analysis).
*Akar, İ., & Şengil Akar, Ş. (2013). The effectiveness of the creative reversal act (creact) on students' creative thinking: further evidence from Turkey. *The Turkish Online Journal of Educational Technology,* 12(4), 183-191.
Akbulut, G. (2004). Geography teaching and creative thinking. *C.Ü. Sosyal Bilimler Dergisi*, 28(2), 215-223.
Akkaya, M. (2019). *The effects of research inquiry-based teaching on scientific process skills: A national meta analysis study.* Ağrı İbrahim Çeçen University. Master Dissertation, Institute of Science, Ağrı.
*Akkılıç, G. (2018). *Investigation of the effect of activities based on six thinking hats technique in a philosophy class on creativity, critical thinking and academic self-efficacy.* Atatürk University. Master Dissertation, Institute of Educational Sciences, Erzurum.
*Akran, K. S., & Aşıroğlu, S. (2018). The effect of writing circle technique on prospective teachers' creative thinking skills. *Inonu University Journal of the Graduate School of Education*, 5(10). *Aksoy, M. (2018). Investigation of the effects of block play on children's creativity. Gazi University, Master Dissertation, Institute of Educational Sciences, Ankara.
Aktamış, H., Hiğde, E., & Özden, B. (2016). Effects of the ınquiry-based learning method on students' achievement, science process skills and attitudes towards science: a meta-analysis science. *Journal of Turkish Science Education*, 13(4), 248-261.
*Akyürek Tay, B. (2019). *The effect of patterned case method with band caricatures on success, permanence and creative thinkıng skills in social studies course.* Gazi University, PhD Dissertation, Institute of Educational Sciences, Ankara.
*Altındağ, C., Göksel, Ç., Koray, Ö., & Koray, A. (2012). *The effect of science and technology laboratory applications based on critical and creative thinking on problem solving and creativity.* X.Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi. Niğde: Türkiye (27-30 Haziran 2012)
*Aras, R. (2018). *Secondary school 6. grade social studies in problem based learning approach the effect of higher-order thinkıng skills and creativity level.* Erzincan Binali Yıldırım University, Master Dissertation, Institute of Social Sciences, Erzincan.

*Arkan Sezgin, K. & Baysal, Z. N. (2019). Examining the effects of thinking skills program developed for the primary school 3rd-grade students on their creativity levels. *HAYEF: Journal of Education,* 16(1); 19-47.APA (1997). *Learner-centered psychological principles: A framework for school redesign and reform.* Retrieved from https://www.apa.org/ed/ governance/bea/learner-centered.pdf.

Arık, S. & Yılmaz, M. (2020). The effect of constructivist learning approach and active learning on environmental education: A meta-analysis study. *International Electronic. Journal of Environmental Education,* 10(2), 44-84.

Aşık, Z., & Özen, M. (2019). *Meta-Analysis Steps and Reporting. Turkish Journal of Family Medicine and Primary Care*, 13(2), 232-240.

Atay, Z. (2009). *Okul öncesi eğitim kurumlarına devam eden 5-6 yaş öğrencilerinin yaratıcılık düzeylerinin yaş, cinsiyet ve ebeveyn eğitim durumlarına göre incelenmesi [Investigation of the creativity levels of 5-6 year old students attending preschool education ınstitutions according to age, gender and parental education status],* Master Dissertation, Selçuk University, Konya

*Atalay, Z. Ö. (2014). *The effect of differentiated social studies instruction on gifted students' academic achievement, attitudes, critical thinking and creativity.* Istanbul University, PhD Dissertation, Institute of Educational Sciences, Istanbul.

*Avcu, Y. E. (2014). *The effect of creative thinking activities on students' creative thinking and academic achievement "Geography Lesson Example".* Çanakkale Onsekiz Mart University. Master Dissertation, Institute of Educational Sciences, Çanakkale.

Ayaz, M. F. (2014). *The effect of the project-based learning approach on the academic achievements of the students in science lesson and their attitudes towards science lesson: A meta-analysis study.* Dicle University. PhD Dissertation, Institute of Educational Sciences, Diyarbakır.

Ayaz, M. F. (2015). The effect of 5E learning model on the attitudes towards lessons of the students: A meta-analysis study. *Electronic Journal of Education Sciences.* 4(7), 29-50.

Ayaz, M. F. & Şekerci, H. (2016). The effect of constructivist learning approach to attitudes of the students towards lessons: A meta-analysis study. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi,* 47, 46-63. Retrieved from https://dergipark.org.tr/tr/pub/dpusbe/issue/26801/286166

*Aydın, Z. (2011). *The effects of active learning activities on the 6th graders' attitude towards mathematics classes and their creative thinking abilities.* Gaziantep University, Master Dissertation, Institute of Social Sciences, Gaziantep.

*Bacak, S. (2008). *The effect of story-based learning approach on academic success and creativity of 5th grade social sciences course.* Celal Bayar University, Master Dissertation, Institute of Social Sciences, Manisa.

Bakioğlu, A. & Özcan, Ş. (2016). *Meta Analysis.* Ankara: Nobel Academy Publishing.

Baş, G. (2016). The effect of multiple intelligences theory-based education on academic achievement: A meta-analytic review. *Educational Sciences: Theory & Practice*, 16, 1833–1864.

*Başkök, B. (2012). *Investigation of the effect of creative writing studies on the seven grade students on the creativities and attitude of Turkish courses.* Akdeniz University, Master Dissertation, Institute of Social Sciences, Antalya.

Başol, G., Doğuyurt, M. F., & Demir, S. (2016). A content analysis and methodological evaluation of meta-analyses on Turkish samples. *International Journal of Human Sciences*, 13(1), 714-745.

*Batdal Karaduman, G. (2012). *The effect of differentiation geometry teaching on creative thinking, spatial ability level and access for primary school 5th grade gifted students.* Istanbul University, PhD Dissertation, Institute of Social Sciences, Istanbul.

Batdı, V. (2014). The effect of activity-based learning approach on academic achievement (A meta-analytic and thematic study). *e-International Journal of Educational Research*, 5(3), 39-55.

Batdı, V. (2014b). The effect of blended learning environments on students' academic achievement: a meta-analysis study. *Çankırı Karatekin Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 5(1), 287-302.

*Baylav Korkmaz, H. (2002). *The effect of project based learning on creative thinking, problem solving and academic risk taking levels in science education.* Hacettepe University, PhD Dissertation, Institute of Social Sciences, Ankara.

*Baz, B. (2019). *The effect of the program based on the children's literature texts on the creative thinking skills of primary school students.* Yıldız Teknik University, Master Dissertation. Institute of Social Sciences, İstanbul.

Becker, K & Park, K. (2011). Effects of integrative approaches among science, technology, engineering, and mathematics (STEM) subjects on students' learning: A preliminary meta-analysis. *Journal of STEM Education*, 12(5/6), 23-37.

Beghetto, R. A. (2016). Creative learning: A fresh look. *Journal of Cognitive Education and Psychology*, 15, 6–23.

Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research*, *87*(2), 309-344.

*Biber, M. (2006). *The effect of the discovering learning method on the creativity of elementary school students of 2nd stage mathematics course.* Dokuz Eylül University, Master Dissertation, Institute of Educational Sciences, İzmir.

Biçer, N. (2017). The influence of student-centered methods in Turkish language instruction on academic success: A meta-analysis study. *Universal Journal of Educational Research,* 5(4), 687-69.

*Birinci, E. (2008). *The effect of using project-based learning in adaptation and development of materials on teacher candidates' critical thinking, creative thinking and scientific-process skills.* Zonguldak Karaelmas University, Master Dissertation. Institute of Social Sciences, Zonguldak.

*Birişçi, S. & Karal, H. (2011). Effect of collaborative studies on prospective teachers' creative thinking skills while designing computer based material. *Ahi Evran Üniversitesi Eğitim Fakültesi Dergisi*, 12(2/Haziran), 203-219.

Blumberg, P. (2008). *A guide to learner-centered teaching*. San Francisco: Jossey-Bass.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex-UK: John Wiley & Sons Ltd.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2013). *Meta-Analize Giriş*. Çev. Serkan Dinçer, Ankara: Anı Yayıncılık.

Brown, K. L. (2003). From teacher-centered to learner-centered curriculum: Improving learning in diverse classrooms. *Education,* 124(1), 49-54.

Brown, M. N. & Freeman, K. (2000). Distinguishing features of critical thinking classrooms. *Teach. Higher Educ*. 5(3), 301-309.

*Bulut, D. (2015). *The effect of educational game design process on student's creativity.* Bahcesehir University, Master Dissertation, Institute of Educational Sciences, İstanbul.

*Camcı Erdoğan, S. (2014). *The effect of differentiated science and technology instruction based on scientific creativity on gifted and talented students' achievement, attitude and creativity.* Istanbul University, PhD Dissertation, Institute of Educational Sciences, Istanbul.

Camnalbur, M. & Mutlu Bayraktar, D. (2018). The effect of collaborative learning on academic achievement: A meta-analysis study. *Journal of the Human and Social Sciences Researches,* 7(2), 1149-1172.

Cantürk Günhan, B. (2016). The effect of drama based learning applied in turkey on success of mathematics: a meta-analysis study. *International Online Journal of Educational Sciences,* 8(2), 145-162.

Capar, G. & Tarim, K. (2015). Efficacy of the cooperative learning method on mathematics achievement and attitude: A meta-analysis research. *Educational Sciences: Theory & Practice,* 15(2), 553-559.

Card, N. A. (2012). *Applied meta-analysis for social science research*. New York: The Guilford Press, 7, 9.

Carson, S. H., Peterson, J. B., & Higgins, D. M. (2005). Reliability, validity, and factor structure of the creative achievement questionnaire. *Creativity Research Journal*, 17, 37-50. DOI:10.1207/s15326934crj1701_4

Chung, N. & Ro, G. (2004). The Effect of problem-solving instruction on children's creativity and self-efficacy in the teaching of the practical arts subject. *The Journal of Technology Studies.* 30(2), 116-122.

Cohen, J. (1988). *Statistical power analysis for behavioral sciences* (2nd ed.).Hillsdale, NY: Lawrence Erlbaum Associates.

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, 1 (3), 98-101.

Cooper, H. M. (2009). *Research synthesis and* meta-analysis*: A step-by-step approach* (4th ed.). Thousand Oaks, CA: Sage Publications.

Cruz, R. E., Lian, M. J. & Morreau, L. E. (1998) The effects of creative drama on social and oral language skills of children with learning disabilities. *Youth Theatre Journal*, 12(1), 89-95. DOI: 10.1080/08929092.1998.10012498

Cumming, G. (2012). *Understanding the new statistics*. New York: Routledge, Taylor and Francis Group.

Çelik, S. (2013). *The effect of alternative teaching methods used in elementary mathematics classes on academic success: A meta analysis study.* Eskişehir Osmangazi University. Master Dissertation, Institute of Educational Sciences, Eskişehir.

Çetin, Z., Üstündağ, A., Kerimoğlu, G. & Beyazıt, U. (2015). Examining the assessment tools used for creativity of children in Turkey and worldwıde. *H.Ü. Sağlık Bilimleri Fakültesi Dergisi*, 2(2), 30-49.

*Çetinkaya, Ç. (2014). The effect of gifted students' creative problem solving program on creative thinking. *Procedia - Social and Behavioral Sciences*, 116(2014), 3722–3726.

Chiu, Y. H., Kao, C. W., & Reynolds, B. L. (2012). The relative effectiveness of digital game-based learning types in English as a foreign language setting: A meta-analysis. *British journal of educational technology*, *43*(4), E104-E107.

Çırak-Kurt, S., Yıldırım, İ. & Cücük, E. (2018). The effects of blended learning on student achievement: a meta-analysis study. *H. U. Journal of Education*, 33(3), 776-802.

*Çoban, B. (2014). *The effects of problem based learning to success, creativity and transfer skills of the students.* Gazi University, Master Dissertation, Institute of Educational Sciences, Ankara.

*Çolakoğlu, S. (2018). *The effect of the teaching of the circle topic on the creative thinking skills of 7th grade students with the geogebra software*. Bayburt University, Master Dissertation, Institute of Social Sciences, Bayburt.

Dağyar, M. (2014). *The effects of problem based learning on academic achievement: A meta-analysis study.* Hacettepe University. PhD Dissertation, Institute of Educational Sciences, Ankara.

Darsih, E. (2018). Learner-centered learning: What makes it effective. *Indonesian EFL Journal,* 4(1), 33-42. DOI: 10.25134/ieflj.v4i1.796.

Demir, S. & Başol, G. (2014). Effectiveness of computer-assisted mathematics education (came) over academic achievement: A meta-analysis study. *Educational Sciences: Theory & Practice,* 14(5), 2026-2035.

*Dere, Z. & Ömeroğlu, E. (2018). A study on the effects of creativity training program on creative behaviors of children. *Cumhuriyet International Journal of Education,* 7(1), 1-15. Doi:10.30703/cije.333901.

Dinçer, S. (2014). *Eğitim bilimlerinde uygulamalı meta-analiz.* Ankara: PEGEM Akademi Yayıncılık.

Dochy, F., Segers, M., Bossche, P. & Gijbels, D. (2003). Effects of problem-based learning: a meta-analysis. *Learning and Instruction*, 13(5), 533-568.

Doğan, N. (2005). Creative thinking. Demirel Ö. (Ed.*), New Trends in Education* (2. Edition). Ankara: Pegem A.

Dollinger, S. J., Urban, K. K., & James, T. A. (2004). Creativity and openness: Further validation of two creative product measures. *Creativity Research Journal*, 16, 35-47.

Drapeau, P. (2014). *Sparking Student Creativity: Practical Ways to Promote Innovative Thinking and Problem Solving*. ASCD Learn Teaching.

*Düzen, Ü. (2019). *The effect of mathematical centered stem activities on the creative thinking skills of students and ınvestigation of student opinions.* Bayburt University, Master Dissertation, Institute of Social Sciences, Bayburt.

*Ekinci Işık, D. (2007). The effect of project based learning on academic success, creative thinking, permanance attitude levels towards life study course in knowledge of life teaching. Dokuz Eylül University, Master Dissertation, Institute of Educational Sciences, İzmir.

Ellis, P. D. (2012). *The essential guide to effect sizes* (5th edition). Cambridge-UK: Cambridge University Press.

*Engin, G. A. (2019). *The effect of integrated creative drama-art activities applied to 4-5 years old pre-school children on creative thinking skills.* Ankara University, Master Dissertation, Institute of Educational Sciences, Ankara.

Erdamar, G., Ayataç, T., Demir, E. & Demir, F. (2015). The Effects of Constructivist Learning Approach on Academic Success of Students: A Meta-Analysis. Education in the 21st Century: Theory and Practice, St. Kliment Ohridski University Pres.

Erden, H., Aytaç, T. & Erden, A. (2016). *Effect of Creative Drama on Students' Academic Achievement: A Meta-Analysis Study*. 2nd International Conference on Lifelong Education and Leadership for All, ICLEL 2016, July 21-23, 2016, Sakarya University&Liepaja University, Liepaja-LATVIA.

*Ersoy, E. & Başer, N. (2014). The effects of problem-based learning method in higher education on creative thinking. *Procedia-Social and Behavioral Sciences,* 116(2014) 3494-3498.

Eskidemir, M. S. & Tezel, Ş. F. (2019). Creative thinking dispositions of the preschool teachers. *OPUS Uluslararası Toplum Araştırmaları Dergisi*, 13 (19) , 311-331.

*Eşsizoğlu, G. (2013). *The effect of project based learning in social sciences on gifted students' achievement, critical thinking and creativity.* Istanbul University, Master Dissertation, Institute of Educational Sciences, İstanbul.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of educational research*, *82*(3), 300-329.

Gajda, A., Karwowski, M., & Beghetto, R. A. (2017). Creativity and academic achievement: A meta-analysis. *Journal of Educational Psychology, 109*(2), 269–299.

Galton, M., Hargreaves, L. & Pell, T. (2009). Group work and whole-class teaching with 11 to14 year olds compared. *Cambridge J. Educ.* 39(1), 119-140.

*Gizir Ergen, Z. (2013). *A study on the effects of the project approach on the creativity of children attending preschool.* Ankara University, PhD Dissertation, Institute of Science, Ankara.

*Gölcük, A. (2017). *The effect of science education that is supported by scıentıfıc storıes on students' creatıvıty and affectıve properties.* Hacettepe University, Master Dissertation, Institute of Educational Sciences, Ankara.

Gözüyeşil, E. & Dikici, A. (2014). The effect of brain based learning on academic achievement: A meta-analytical study. *Educational Sciences: Theory & Practice*, 14(2), 642-648.

*Güldemir, S. (2019). *The effect of stem activities on creativity in pre-school education.* Recep Tayyip Erdoğan University. Master Dissertation, Institute of Science, Rize.

*Günbatar, M. S. (2009). *The effect of web-assisted problem-based learning on students 'creative thinking skills and attitudes.* Yüzüncü Yıl University, Master Dissertation, Institute of Social Sciences, Van.

* Gündoğan, A., Arı, M. & Gönen, M. (2013). The effect of drama on the creative ımagination of children in different age groups. *Hacettepe University Journal of Education*, 28(2), 206-220.

*Güner Yüksel, S. (2019). *The effect of problem based learning method on students 'creative thinking and problem solving skills and academic achievements in social studies.* Gaziantep University. Master Dissertation, Institute of Educational Sciences, Gaziantep.

Haase, J., Hoff, E. V., Hanel, P. H. P. & Innes-Ker, A. (2018) A meta- analysis of the relation between creative self-efficacy and different creativity measurements. *Creativity Research Journal,* 30(1), 1-16.

*Hacıoğlu, Y. (2017). *The effect of science, techology, engineering and mathematics (stem) education based activities on prospective science teachers' critical and creative thinking skills.* Gazi University, PhD Dissertation, Institute of Educational Sciences, Ankara.

Hartung, J., Knapp, G., & Sinha, B. K. (2008). *Statistical meta-analysis with applications*. New Jersey: Wiley Publishing Inc.

Hertzog, N. B. (2008). *Ready for preschool*. Prufrock Press, Waco: TX.

*Hızır, B. (2014). *The effect of creative reading to creative thinking skill in primary and secondary school.* Muğla Sıtkı Koçman University. Master Dissertation, Institute of Educational Sciences, Muğla.

Higgins, J. P., Thompson, S. G., Deeks, J. J., Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *Bmj.* 327(7414), 557-60.

Hong, E. & Milgram, R. M. (2010) Creative thinking ability: Domain generality and specificity, *Creativity Research Journal*, 22(3), 272-287. DOI: 10.1080/10400419.2010.503535

Hsen-Hsing Ma (2009). The Effect Size of Variables Associated with Creativity: A Meta-Analysis. *Creativity Research Journal*, 21:1, 30-42.

Huedo-Medina, Tania; Sanchez-Meca, Julio; Marin-Martinez, Fulgencio; and Botella, Juan (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index?". *CHIP Documents*. 19.

Internet: Prisma Checklist Turkish Version. WEB: Retrieved from http://www.prismastatement.org/documents/PRISMA%20Turkish%20checklist.pdf

*Işık, A. D. & Saygılı, G. (2015). The effect of learning creative development techniques on the skills of creative thinking. *Bartin University Journal of Faculty of Education*, Special Issue on XIV. International Participation Symposium of Primary School Teacher Education (21-23 May, 2015), 133 – 139.

*İkikat, U. (2019). *The effect of enriched science course on the creative thoughts of primary school students.* İstanbul Aydın University, Master Dissertation, Institute of Social Sciences, İstanbul.

*İşeri, Ş. (2010). *Effect of class activities based on newspapers used in science and technology lessons on the attitude and creativity of students.* Kocaeli University. Master Dissertation, Institute of Science, Kocaeli.

Johnson, D. W., Johnson, R. T. & Stanne, M. B. (2000). *Cooperative learning methods: A meta-analysis.* University of Minnesota, Minneapolis, Minnesota.

*Kanlı, E. & Emir, S. (2013). The Effect of problem based learning on gifted and normal students' achievement and creativity levels. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education,* 7(2, December), 18-45.

Kaplan, A., Duran, M. & Baş, G. (2015). The effects of ınstruction based on multiple ıntelligence theory in mathematics course on academic achievement: A meta-analysis study. *Mersin University Journal of the Faculty of Education,* 11(3): 814-831.

*Kaplan Sayı, A. (2013). *The effect of differentiated foreign language instructıon on gifted students' achievement, critical thinking and creativity.* Istanbul University, PhD Dissertation, Institute of Social Sciences, Istanbul.

*Kar, T. (2010). *The effects of problem-based learning in linear algebra on the academic achievement, problem-solving skills and creativity of students.* Atatürk University, Master Dissertation, Institute of Science, Erzurum.

*Karaca, T. & Koray, Ö. (2017). The Effect of using the creative reversal act in science education on middle school students' creativity levels. *Euraisan Journal of Educational Research.* 67, 199-214. http://dx.doi.org/10.14689/ejer.2017.67.12

*Karadayı, Ş. İ. (2018). *The effect of creative education on preschool period on creative performance, administrative mental functions and emotional regulatory skills.* Ege University. Master Dissertation, Institute of Social Sciences, İzmir.

*Karakuş, M. (2000). *The effects of creative problem solving curriculum on the creativity levels of creativity in the second grade students who are from low socio-economic status.* Cukurova University, Master Dissertation, Institute of Social Sciences, Adana.

Karakuş, M. & Yalçın, O. (2016). The effect of the argümantation-based learning in science education to the academic achievement and scientific process skills: A meta analysis study. *Anadolu University Journal of Social Sciences*, 16(4), 1-20.

*Karataş, S. & Özcan, S. (2010). The effects of creative thinking activities on learners' creative thinking and project development skills. *Ahi Evran Üniversitesi Eğitim Fakültesi Dergisi,* 11(1), 225-243.

*Karataş Öztürk, S. (2007). *The effect of the learning approach based on creative thinking upon the creative thinking and problem solving skills of students.* Eskişehir Osmangazi University, Master Dissertation, Institute of Science, Eskişehir.

Kaşarcı, İ. (2013). *The effectiveness of project based learning on students' academic achievements and attitude: A meta-analysis.* Eskişehir Osmangazi University. Master Dissertation, Institute of Educational Sciences, Eskişehir.

Kaya, A. & Öçal, M. F. (2018). A meta-analysis for the effect of geogebra on students' academic achievements in mathematics. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education*, 12(2), 31-59.

*Kaya, H. İ. (2010). *Effects of the practices based on constructivist learning in teacher education on teacher candidates' tendencies of problem solving, critical thinking and creative thinking.* Atatürk University, PhD Dissertation, Institute of Social Sciences, Erzurum.

*Kaya, M. E. (2018). *Effects of stem applications on scientific arrangement and creativity of science teachers of science.* Erzincan Binali Yıldırım University. Master Dissertation, Institute of Science, Erzincan.

*Kıncal, R. Y., Avcu, Y. E. & Kartal, O. Y. (2016). The effects of creative thinking activities on learners' creative thinking skills and academic achievement. *Journal of Theoretical Educational Science,* 9(1), 15-37.

Kim, K. H. (2005). Can only intelligent people be creative? A meta-analysis. Journal of Secondary Gifted Education. 16(2-3), 57-66.

Kim, K. H. (2011). (2011). The creativity crisis: The decrease in creative thinking scores on the torrance tests of creative thinking. *Creativity Research Journal*, 23, 4, 285-295.

*Koray, Ö. (2004). The influence of science education based on creative thinking on creativity of preservice teachers. *Educational Administration in Theory & Practice,* Fall (40), 580-599.

* Koray, Ö. & Köksal, M. S. (2009). The effect of creative and critical thinking based laboratory applications on creative and logical thinking abilities of prospective teachers. *Asia-Pacific Forum on Science Learning and Teaching,* 10(1), 3-13.

*Kök, B. (2012). *The effect of differentiated geometry teaching on gifted and talented students in view of creativity, spatial ability and success.* Istanbul University, PhD Dissertation, Institute of Social Sciences, Istanbul.

*Köse Biber, S. (2014). *The effect of web-based learning supported by creativity improving activities on students' achievement and creativity.* Istanbul University, PhD Dissertation, Institute of Science, Istanbul.

Krauksta D., Rozenvalde, P. & Ciekurs, K. (2016). Outdoor Activities to Promote Creative Thinking in Preschool Children. *Society, Integration, Education* Proceedings of the International Scientific Conference. Volume II, May 27th - 28th, 2016. 396-403.

Kuo, F.R., Chen, N. S. and Hwang, G. J. (2014). A creative thinking approach to enhancing the web-based problem solving performance of university students. *Computers & Education.* Vol. 72, 220-230.

*Kuytan, Ş. (2019). *The effect of individual or group educational game development on pre-service teachers' creative thinking.* Mersin University, Master Dissertation, Institute of Educational Sciences, Mersin.

Kylonen, P. C. (2012). Measurement of 21st century skills within the common core state standards. Paper presented at the Invitational Research Symposium on Technology Enhanced Assessments, May 7-8.

Lee, B. K., Patall, E. A., Cawthon, S. W., & Steingut, R. R. (2015). The effect of drama-based pedagogy on preK–16 outcomes: A meta-analysis of research from 1985 to 2012. *Review of Educational Research*, *85*(1), 3-49.

Li, Q. & Ma, X. (2010). A meta analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review*, 22, 215-243.

Littell, J. H., Corcoran, J. and Pillai, V. (2008). *Systematic reviews and meta-analysis.* UK: Oxford University Press.

MEB. (2018). *2023 Eğitim Vizyonu Belgesi*. Retrieved from http://2023vizyonu.meb.gov.tr/ 11.01.2019.

Merchant, Z., Goetz, E. T., Cifuents, L., Keeney-Kennicutt, W. & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education,* 70, 29-40.

*Metin, Ö. (2019). *The effect of music story and tales on the creativity levels of children in pre-school education.* Marmara University, Master Dissertation, Institute of Educational Sciences, İstanbul.

Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. G., The PRISMA Group (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA Statement.* PLoS Medicine, 6 (7): e1000097.

Okursoy-Günhan, F. (2009). The effect of concept mapping on learner success: A meta-analytic study. Marmara University. Master Dissertation, Institute of Educational Sciences, İstanbul.

*Orçan, A. (2013). *The effect of science-fiction stories developed by comics technique, on creative thinking skills and physics attitudes of students.* Gazi University, Master Dissertation, Institute of Educational Sciences, Ankara.

*Öç, U. (2019). *Scientific process skills of science lab applications based on argumentation, their effects on the attitude for laboratory and the creativity.* Sivas Cumhuriyet University, Master Dissertation, Institute of Educational Sciences, Sivas.

*Ömeroğlu, E. (1990). *The effect of creative drama education on the development of verbal creativity of 5-6 year old children who attend kindergarten.* Hacettepe University, PhD Dissertation, Health Sciences Institute, Ankara.

Özbey, Ö. F. & Sarıkaya, R. (2019). Investigation of effectiveness on drama method studies in Turkey: A meta-analysis study. *Gazi Journal of Education Sciences (GJES),* 5, Özel Sayı, 231-253.

*Özcan, E. (2019). *Investigation of the effects of argumentation based science learning approach on the conceptual understanding of the students, their creative thought and the epistemological beliefs.* Marmara University, Master Dissertation, Institute of Educational Sciences, İstanbul.

Özçelik, A. D. Ö. (2019). *Eğitimde ve Endüstride 21. Yüzyıl Becerileri.* Editörler: Prof. Dr. Ayşe Dilek Öğretir Özçelik ve Dr. Öğr. Üyesi Mehmet Nur Tuğluk, 2. Baskı, Ankara: Pegem Akademi.

Özdemir, F., Aslaner, R. & Açıkgül, K. (2020). The effect of the computer-aided mathematics teaching on the students' attitudes towards mathematics: A meta - analysis study. *Inonu University Journal of the Graduate School of Education,* 7(13), DOI: 10.29129/inujgse.543534.

*Özdemir, S. M. & Çakmak, A. (2008). The effect of drama education on prospective teachers' creativity. *International Journal of Instruction*, 1(1), 13-30.

*Özgenel, M., Canpolat, Ö., Yağan, E. & Canlı, Ö. (2019). The effects of enriched workshop training given to pre-school students on creative thinking skills of students. *Problems of Education in the 21st Century,* 77(5), 616-635.

*Özyaprak, M. (2012). *The effect of different mathematics teaching on his, attitude and creativity for the gifted and talented students.* Istanbul University, PhD Dissertation, Institute of Social Sciences, Istanbul.

*Palamut, İ. (2008). *Effects of reading story on primary students creativity and academical success.* Dokuz Eylül University, Master Dissertation, Institute of Educational Sciences, İzmir.

Pascarella, E. T., Wang, J. S., Trolian, T.L. & Blaich, C. (2013). How the instructional and learning environments of liberal arts colleges enhance cognitive development. *Higher Educ.* 66, 569-583.

Perry, A. & Karpova, E. (2017). Efficacy of teaching creative thinking skills: A comparison of multiple creativity assessments. *Thinking Skills and Creativity*. 24(2017), 118–126.

Peterson, M. (2010). Computerized games and simulations in computer-assisted language learning: A meta-analysis of research. Simulating & Gaming, 41(1), 72-93.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences.* MA- USA: Blackwell Publishers Ltd.

Runco, M. A. & Chand, I. (1995). Cognition and creativity. *Educational Psychology Review*, 7(1), 243-26.

*Sak, U & Öz, Ö. (2010). The effectiveness of the creative reversal act (CREACT) on students' creative thinking. *Thinking Skills and Creativity,* 5(2010), 33–39.

Saracaloğlu, A. S., Gündoğgu, K., Altın, M., Aksu, N., Kozağaç, Z. B. & Koç, B. (2014). Analysis of articles published in the field of creative thinking skills after the year of 2000. *Adnan Menderes Üniversitesi Eğitim Fakültesi Eğitim Bilimleri Dergisi,* 5(2), 62-74.

*Sarmaşık Kaya, G. (2018). *Research regarding the effects of "effective thinking training" program on the ninth grade students" problem solving and creative thinking abilities.* İstanbul University, Master Dissertation, Institute of Educational Sciences, İstanbul.

Sawyer, R. K. (2012). *Explaining creativity: The science of human innovation* (2nd ed.). New York, NY: Oxford University Press.

*Sevgen, M. (2016). *The effect of social skills education given to children through creative drama on the social development and creativity of kindergarten children.* İstanbul Sabahattin Zaim University, Master Dissertation, Institute of Social Sciences, İstanbul.

*Siper Kabadayı, G. (2019). *The effects of robotic activities on pre-school children's creative thinking skills.* Hacettepe University, Master Dissertation, Institute of Educational Sciences, Ankara.

*Sönmez, B. (2016). *The effect of thinking training lesson on 6th grade elemantary school students' critical and creative thinking skills.* Anadolu Universty, PhD Dissertation, Institute of Educational Sciences, Eskişehir.

*Sözer Çapan, A. (2014). *Investigation of the effect of a 5-year children's creative skills of a training program prepared by the brain based learning approach*. Marmara Universty, PhD Dissertation, Institute of Educational Sciences, İstanbul.

Sternberg, R. J. (2003). Creative thinking in the classroom. *Scandinavian Journal of Educational Research,* 47(3), 325-338. DOI: 10.1080/00313830308595

*Şener, N., Türk, C. & Taş, E. (2015). Improving science attitude and creative thinking through science education project: a design, implementation and assessment. *Journal of Education and Training Studies,* 3(4), 57-67.

*Şendağ, S. & Erol, O. (2015). Use of 3D drawing software in promoting creativity in preschool. *Journal of Theoretical Educational Science*, 8(3), 316-336.

*Şensoy, Ö. & Yıldırım, H. İ. (2017). The Effect of research investigation based learning approach on creative thinking and scientific process skills. *Cumhuriyet International Journal of Education-CIJE*, 6(1), 34 – 46.

*Tedik, G. (2013). *The effect of creative drama activities in the fourth grade of prımary school on creatıvıty abilities of students*. Atatürk University, Master Dissertation, Institute of Educational Sciences, Erzurum.

*Terzi, H. (2019). *The effects of intelligence games on the creative thinking skills of 6th grade students*. Bayburt University. Master Dissertation, Institute of Social Sciences, Bayburt.

*Tezci, E. (2002). *The effects of cunstructivist instructional design on the success and creativity of fifth-year students in primary schools.* Fırat University, PhD Dissertation, Institute of Social Sciences, Elazığ.

Torrance, E. P. (1966). *Torrance test of creative thinking*. Bensenville, IL: Scholastic Testing Service

Trilling, B. & Fadel, C. (2009). *21st century skills: Learning for life in our times*. Francisco: Jossey-Bass.

*Tut, E. (2018). *Effect of game-based learning applications in 4th grade science class on students' academic achievement and creative thinking skills.* Ondokuz Mayıs University. Master Dissertation, Institute of Educational Sciences, Samsun.

Tutal, Ö., Kaçire, İ. & Atabey, E. (2016). *Öğretmen eğitiminde kullanılan alternatif öğretim yöntemlerinin akademik başarıya etkisi: Bir meta-analiz çalışması.* 4. Uluslararası Eğitim Programları ve Öğretim Kongresi, Antalya.

*Ulaş, A. H., Tedik, G. & Sevim, O. (2014). The effect of creative drama activities in the fourth grade of prımary school in Turkey on creativity levels of students. *A. Ü. Türkiyat Araştırmaları Enstitüsü Dergisi*, 52, 331-350.

Ulubey, Ö. (2018). The effect of creative drama as a method on skills: A meta-analysis study. *Journal of Education and Training Studies*, 6(4), 63-78.

*Umar, Ç. N. (2014). *The effects of differentiated curriculum with blended learning method on gifted students' academic achiecevement, critical thinking abilities and creative.* Istanbul University, PhD Dissertation, Institute of Educational Sciences, Istanbul.

*Uygun, D. (2017). *The effectiveness of creative reversal act on 7th grade students' creative thinking.* Eskişehir Anadolu University. Master Dissertation, Institute of Educational Sciences, Eskişehir.

*Uysal, M. E. (2009). *The effects of cooperative learning on access, critical thought and creativity skills in primary education Turkish course*. Dokuz Eylul University, PhD Dissertation, Institute of Educational Sciences, İzmir.

*Ülger, K. (2018). The Effect of Problem-Based Learning on the Creative Thinking and Critical Thinking Disposition of Students in Visual Arts Education. *Interdisciplinary Journal of Problem-Based Learning*, 12(1).

*Ülger, K. & İmer, Z. (2013). The effect of problem based learning (pbl) approach on students' creative thinking ability. *Hacettepe University Journal of Education*, 28(1), 382-392.

*Üret, A. (2019). *Effects of stem education on creativity levels of 5-year-old preschool children.* Yıldız Teknik University, Master Dissertation, Institute of Social Sciences, İstanbul.

Vanderbos, G. R. (2006). *APA Dictionary of Psychology*. Washington, DC: American Psychological Association.

Vygotsky, L. S. (2004). Imagination and creativity in childhood. *Journal of Russian and East European Psychology*, 42, 7–97.

Wang, J. R., Huang, B. Y., Tsay, R. F., Lee, K. P., Lin, S. W., & Kao, H. L. (2011). A Meta-analysis of inquiry-based ınstruction on student learning outcomes in Taiwan. *Asia-Pacific Education Researcher (De La Salle University Manila)*, 20(3).

Wilson, D. B. (2009). Systematic coding. H. Cooper, L. V. Hedges ve J. C. Valentine (Ed.), *The handbook of research synthesis and meta-analysis* içinde (2. bs., s. 159-176). New York: Russell Sage Foundation.

World Economic Forum (2016). The 10 skills you need to thrive in the Fourth Industrial Revolution. Retrieved from https://www.weforum.org/agenda/2016/01/the-10-skills-you-need-to-thrive-in-the-fourth-industrial-revolution/ 09.04.2020

Yalçın, S. (2018). 21. Yüzyıl becerileri ve bu becerilerin ölçülmesinde kullanılan araçlar ve yaklaşımlar. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 51(1), 183-201.

*Yaman, S. & Yalçın, N. (2005). Effectiveness on creative thinking skills of problem based learning approach in science teaching. *Elemantary-Online*, 4(1), 42-52.

*Yaman, Y. (2014). *Effects of brain based science teaching on gifted students' achievement, critical thinking, creativity and attitudes.* Istanbul University, PhD Dissertation, Institute of Educational Sciences, Istanbul.

*Yaprakdal, A. B. (2013). *The effect of teaching objects design on the critical, creative thinking and scientific process skills of the teacher candidates.* Marmara University, PhD Dissertation, Institute of Educational Sciences, Istanbul.

Yaşar, M. C. & Aral, N. (2010). The effects of early childhood education on creative thinking skills. *Kuramsal Eğitimbilim*, 3 (2), 201-209.

*Yaşar, M. C. & Aral, N. (2012). Drama education on the creative thinking skills of 61-72 months old pre-school children. *US-China Education Review*, 6(2012), 568-577.

*Yazıcı, E., Belen, H, B., Baydar, I. Y., Aksu, G. & Okutan, N. Ş. (2019). The effect of story based activities on children's creativity skills. *Adıyaman University Journal of Educational Sciences*, 9(2), 226-249.

Yeşilpınar Uyar, M. & Doğanay, A. (2018). The effect of student-centered strategies, methods and techniques on academic achievement: A meta-analysis study. *Mersin University Journal of the Faculty of Education.* 14(1), 186-209.

*Yıldırım, A. & Akman, B. (2020). The effects of creative problem solving activities on creativity of preschoolers. *Anadolu Journal of Educational Sciences International*, 10(1), 399-431.

Yıldırım, A. ve Şimşek, H. (2011). *Sosyal bilimlerde nitel araştırma yöntemleri [Qualitative research methods in the social sciences].* Ankara: Seçkin Yayıncılık.

*Yıldırım, M. (2018). *The effect of context-based storyline method in science education on acadeıic achievement, creativity and attitudes towards science.* Abant İzzet Baysal University, PhD Dissertation, Institute of Educational Sciences, Bolu.

*Yıldız, Z. (2012). The effect of project based learning approach to high school students' level of creative thinking, problem solving, taking academic risk. Gazi University, Master Dissertation, Institute of Educational Sciences, Ankara.

*Yılmaz, O. (2006). *Effects of project-based learning on students' academic achievement, creativity and attitudes in 7th grade social science lesson.* Zonguldak Karaelmas University, Master Dissertation, Institute of Social Sciences, Zonguldak.

*Yılmaz, R. M. (2014). *Effects of three dimensional storytelling developed with augmented reality technology on narrative skill and creativity.* Atatürk University, PhD Dissertation, Institute of Educational Sciences, Erzurum.

*Yiğit, E. Ö. & Erdoğan, T. (2008). The effect of studying method applied in the social science course on the creative thinking levels of primary school students. *Ç.Ü. Sosyal Bilimler Enstitüsü Dergisi,* 17(3), 399-416.

*Yurdakal, İ. H. (2019). The effects of creative reading practices on the skills of creative thinking. *PAU Journal of Education*, 47, 130-144.

*Yüksel, Y. B. (2018). *The impact of creativity education program for preschoolers on kids' creativity and audio reasoning.* Gazi University, Master Dissertation, Institute of Educational Sciences, Ankara.

Weimer, M. (2002). *Learner-Centered Teaching*. San Francisco: Jossey-Bass.

Review of Two-tier Tests in the Studies: Creating a New Pathway for Development of Two-tier Tests

Okan Sıbıç[1], Behiye Akçay[1], Merve Arık[2]
[1]İstanbul University-Cerrahpasa
[2]Sair Fevzi Kutlu Kalkanci Secondary School, Turkey

Review of Two-tier Tests in the Studies: Creating a New Pathway for Development of Two-tier Tests

Okan Sıbıç[1*], Behiye Akçay[1], Merve Arık[2]
[1]İstanbul University-Cerrahpasa
[2]Sair Fevzi Kutlu Kalkanci Secondary School, Turkey

Abstract

One of the diagnostic tests which is very valuable and used in education frequently is two-tier test. Two-tier tests are used for different purposes such as determining misconceptions, determining the comprehension level of students and etc.. Due to the wide and effective uses of two-tier tests, there are different studies in which two-tier tests are developed in science education. The aim of the present study is to review the studies in which two-tier tests are developed and used in different study and sample settings. Additionally, development steps of two-tier tests have been examined in depth. Through this review study, samples and subjects' areas to whom and which two-tier tests are applied; the style of the two-tier tests and development process of the tests have been examined heedfully. As a result of in-depth examinations of the studies, a new and effective way for developing two-tier tests is proposed at the final phase of the study.

Key words: Meta-analysis, Misconception, Science education, Two-tier tests

Introduction

As the constructivist approach has been integrated into today's educational programs, meaningful learning has been discussed among educators more and more. The aims of meaningful learning are not only what students learn but also how students learn (Ausubel, 1963). Inquisition of the cognitive schemes of learners and revealing misconceptions that students possess move the measurement and evaluation trends towards paying attention to the assessment process not the outcomes. Therefore, the use of different assessment tools has become significant in that they reveal students' learning thresholds, levels, and conceptual perceptions (Reeves & Okey, 1996). Such different measurement tools as portfolios, holistic and analytical rubrics, decoding tables, diagnostic branched tree, self-assessment, peer review, and structured grid gain importance and are used by educators; open-ended and multiple-choice (MC) tests are still continued to be used in assessment.

In MC tests, with item stems, students are presented at least three alternatives with one right choice which they are expected to choose (Tan, 2009; Tan, Kayabaşı & Erdoğan, 2002). In the literature, it is emphasized that MC tests are usually used by teachers and researchers since they have some advantages. Those advantages are pointed out as follows: they are easy to implement and score, do not require expertise in assessing, provide wide-scale usage areas and objective assessment and higher reliability is provided through them (Haladyna, 2004; Karataş, Köse & Coştu, 2003; Tan, 2009). Despite their advantages, MC tests also have some disadvantages. Since the reason of students' answers is not assessed in depth and because of the luck factor (Liu, 2010), MC tests are considered as weak tests. In addition, students' ideas are shaped by limited number of choices. Yıldırım (1999) emphasized that through MC tests, students' factual knowledge about new learning can be measured, but how they organize and synthesize the knowledge cannot be measured with them. Similarly, as disadvantages of MC tests, Tan (2009) pointed out that MC tests can be used to measure knowledge and application skills easily while they are not appropriate to measure synthesis skill, creativeness and producing ideas.

The fact that MC tests have disadvantages as well as advantages pushed researchers to create a new instrument which would eliminate the disadvantages of MC tests while maintaining their advantages. To this end, Treagust developed two-tier multiple-choice diagnostic tests in 1988 to minimize the weaknesses in assessment through

---

[*] Corresponding Author: *Okan Sıbıç, okansibic@istanbul.edu.tr*

MC tests (Treagust, 1995). Two-tier tests consist of two interrelated phases, the first of which consists of an item stem and a number of answer options. This phase (the first tier of the test) is similar to MC tests. The aim of this tier is to identify how an individual interprets scientific knowledge. The difference that separates two-tier tests from other tests is the second phase of two-tier tests and in this phase, students are expected to present the reason for the answer they give in the first phase (Treagust, 1988). Because they justify their answers, the second phase provides a sensitive and an effective way for the students to learn meaningfully, and also serves as an effective diagnostic tool for conceptual understanding and misconception (Tamir, 1989). In this respect, it can be said that it minimizes the criticism about hypothetical answers that is frequently directed at MC tests in the literature meanwhile it maintains the advantages of MC tests such as wide usage, easy scoring. Thus, two-tier tests are considered a practical and valuable way of assessment since they justify students' answers, decrease hypothetical answers, offer large-scale use, enable the process to be managed easily, enable easy scoring and present ideas on the way of students' thinking (Othman, Treagust & Chandrasegaran, 2008).

In addition, two-tier tests provide flexibility in terms of creating different assessment options in two phases as per personal preference. The following table (Table 1) provides two-tier test types that are commonly used in the literature:

Table 1. Two-tier test styles

First Tier	Second Tier
Multiple choice	Multiple choice
Multiple choice	Open-ended
True/False	Multiple choice
True/False	Open-ended

As seen in Table 1, in the first phase of two-tier tests, multiple choice tests or true-false questionnaires are usually used. The second tier of the test is from literature review or interviews the student may have multiple choice forms including student misconceptions. In addition, this second tier can also be arranged in an open-ended structure in order to measure students' reasoning abilities better and to determine if there are alternative concepts other than the previously identified misconceptions (Treagust & Mann, 1998). When the second tier of two-tier tests is structured in an open-ended way, there is no need to develop a distractor and the students are asked to explain the reason of the answer they chose in the first tier. On the other hand, if the second tier is designed as a multiple choice test, the misconceptions should be well defined, and the distractors should be structured accordingly. In the second tier, the aim is to investigate the justification of the answer given by the student in the first tier; it is necessary to follow a rigorous process while preparing the second-tier items in a valid and reliable manner (Karataş, Köse & Coştu, 2003).

Treagust (1988), who introduced two-tier tests into educational research, proposed a method of developing these tests, consisting of a total of ten steps under three main steps, namely, *defining the content*, *obtaining information about students' misconception* and *developing the diagnostic test*. Based on the recommendation of Treagust (1988), the steps of developing two-tier tests are explained in the following:

'Defining the content' is about drawing the boundaries of the subject or concepts; and it depends on four steps:

Step 1. Identifying propositional knowledge statement: In this step, you should discover a lot of propositional knowledge which is written in the textbooks and literature depending on the information available in the curriculum. These propositions should include all aspects of the relevant topic or concept.
Step 2. Developing a concept map: It is important to develop a map of concepts which relate to the topic under investigation was developed based on the procedure described by Novak (1980). As is the case with the development of propositional knowledge, this activity enables the researcher to carefully consider the nature of the content which has been selected for instruction (Treagust, 1988).
Step 3. Relating propositional knowledge: Concept map and propositional knowledge statements are directly related. Therefore, the overlap between these two structures serves as a kind of control mechanism for the internal consistency of the test to be prepared (Karatas, Köse & Costu, 2003).
Step 4. Validating the content: The propositional knowledge statements and the concept map are analysed, and the content is validated by science educators, secondary science teachers and science specialists with thorough knowledge of the subject matter.

'Obtaining information about students' misconception' involves developing diagnostic tests to evaluate students' misconceptions involves a thorough examination of the relevant literature dealing with cognitive structure and

conceptual change, interviews with students about their understanding of the science content and obtaining responses from open ended pencil and paper questions. And it consists of three steps as follows:

Step 5. Examining related literature: In this step, the literature based on the subject determination of misconceptions is reviewed. The data obtained from the review is used to develop the semi-structured or unstructured interview questions both for the development of the test and for the next step.

Step 6. Conducting unstructured student interview: In order to gain a broad perspective of students' understanding, unstructured interviews are held with students. These interviews help any areas of misunderstanding and misconceptions be identified and lead to the development of ideas for further probes by multiple choice questions with a free response.

Step 7. Developing multiple choice content items with free response: Multiple choice questions and distractors are developed based on the general misconceptions determined thanks to the literature review and analysis of unstructured interviews. Common misconceptions about these propositional statements are placed in distractors. After each MC question, a statement such as "because" or "explain your reason" is added to the phrase with a space allocated for them to give reasons. Then, this form (the first tier is MC and the second tier is open-ended) is delivered to students (Karatas, Köse & Coştu, 2003).

'Developing a diagnostic test' involves the development of two-tier test items, the first of which requires a content response and the second requires a reason for the response. And it consists of three steps as follows:

Step 8. Developing the two-tier diagnostic tests: The second tier of the test is arranged as MC based on the students' open-ended answers, which have been determined in step 7. Each justification option in the second tier should include the common misconceptions that students have in addition to the correct answer.

Step 9. Designing a specification grid: All knowledge of each of the questions and the concepts in the concept map should be associated with a specification grid for the developed two-tier test.

Step 10. Continuing refinements: At this stage, the implementation of the pilot study starts. Performing a substance analysis of the test with the pilot study and calculating its reliability are aimed. Necessary arrangements are made on the test by taking advantage of these results (Treagust, 1988).

Treagust (1985), Haslam and Treagust (1987), and Odom and Barrow (1995) stated that the development of two-tier alternative multiple-choice tools for defining students' concepts has huge potential to make substantial contributions to the field of alternative assessment. Students' conceptual status can be improved thanks to instructors' awareness, and items of two-tier tests can be also used during group discussions; it can provide useful information for curriculum revision; the development of instruments incorporates research findings that can be readily utilized in the classroom (Lin, 2004). Treagust (1988; 1995) stated that these tools are particularly useful guides to identify alternative concepts of students towards different concepts. It is observed that two-tier tests are used in different research cases in the field of science education. The three-step process described by Treagust (1988) in the development of these tests is illustrative, and alternative sub-steps and processes can be found in the literature.

The Aim of the Study

The aim of the study is to analyze the studies in which two-tier test development process was observed and implemented in the science education field; to examine the similarities and differences among the processes followed in different studies and to propose a new two-tier test development process model that can guide prospective studies within the scope of analysis and results.

In accordance with these purposes, the research questions determined are as follows:

1) In the related literature, what was the general aim of using two-tier tests and for which science field were two-tier tests developed by the researchers?
2) In the related literature, which steps are usually used in the process of developing two-tier tests for science lessons?
3) What can be an alternative two-tier test development process?

Method

In the present study, the researchers used meta-synthesis method which is also called as thematic content analysis (Walsh & Downe, 2005). Meta-synthesis study is a methodology in which qualitative and quantitative studies are used together, and related with identifying and understanding the themes and key points related with a topic which are found in the related literature (Bair, 1999 p.4). In the present study, in parallel to the nature of the meta-synthesis method, it is aimed to examine the studies which are conducted in the field of science education and related with two-tier tests qualitatively.

Sample

According to Walsh and Downe (2005), a meta-synthesis study has some steps. Those are:

(1) Making search for research articles
(2) Determining some criteria for the process of selection of the articles based on the purpose of the study
(3) Analysing and evaluating the studies
(4) Conceptualizing and comparing the selected studies, and
(5) Synthesizing the findings

In the present study, the researchers follow a same pathway to search, find and include the articles; and also, to present the research findings. In the following, the researchers represent the selection process of the articles by following the method of Walsh and Downe (2005) respectively.

To select the research articles which would be examined as per the aim of the present study, the researchers used the five steps that Walsh and Downe (2005) introduced for meta-synthesis study. Those were fulfilled as follows:

(1) For the first step, the researchers determined how to reach the research articles in line with the purpose of the study. For settling the sample of the study, the researchers used different electronic online databases that İstanbul University Library gives access to. İstanbul University Library provides access to 110 electronic online databases ranging from the field of education to that of health. In other words, the researchers have huge feasibility to reach nearly all articles that can be used as per the aim of the study.
(2) Second, in order to select convenient articles in line with the aim of the study, the researchers decided on using two different keywords: *two-tier tests and iki aşamalı testler (equivalence of two-tier tests in Turkish).* By using the databases of which İstanbul University Library is an electronic subscriber and typing those keywords, the researchers accessed 271 research articles. In the present study, it was only included the studies which were published until the last quarter of 2018. Those articles are indexed in *Social Sciences Citiation Index (SSCI), Education Resources Information Center (ERIC), EBSCO, Ulusal Akademik Ağ ve Bilgi Merkezi (ULAKBIM), DergiPark, Teacher References Center and JSTOR.*
(3) In step three and four, two researchers of the study independently examine each research article according to the purpose and the criteria for including the related research articles in the study. The criteria for research articles to be included in this study that the researchers of the present study determined are:
   a. Being related to the science education field,
   b. Developing two-tier tests within the study, and
   c. Showing the process of developing two-tier tests explicitly within the test.
   After the detailed examinations, the exact sample of the study consisted of 42 research articles. In Appendix-1, the table which consists of some information about those research articles were presented by the researchers.
(4) In the last step, the researchers examined 42 research articles according to their test focus, content, target participants, test style of each tier, and the process of developing the two-tier test in each related study.

Data Collection

In the study, as data collection tool, the researchers created two different forms to present different specifications related to the research questions. In the first form (Table 2), the researchers aimed to unroll and

give a general perspective about the aim, content, focus of the studies in which two-tier tests were used by the authors of the related studies. In addition, the researchers also examined the question types of the two-tier tests for both of their tiers.

Table 2. General Information about the Studies in Which Two-tier Tests Were Used

Article No	Author(s) Year	Content	Test Focus	Target Participants	Test Style	
					First Tier	Second Tier
A.1	aaa&bbb, 2017	ccc	ddd	eee	MC	MC
A.2						
A.3						
.						
.						
.						

By the help of the other form, researchers checked how the authors of the related studies developed their two-tier tests, in other words, it is presented that which steps were followed by the authors to develop the two-tier tests. In Table 3, the researchers presented the method they used to analyse each study.

Table 3. Two-tier development steps followed in the articles

Article No	Literature Review	Curriculum Analysis	Creating Propositional Knowledge Statements (PKS)	Creating Concept Map (CM)	Founding Relationships btw. CM and PKS	Content Validity (Expert Opnion)	Literature Review for Misconception (Misconp.)	Developing (Dev.) Interview Skills of Participants	Making Interviews for Determination of Misconcp.	Dev. Test with Open-ended second and Application	Dev. Two-tier Test Items	Creat. Table of Specifications	Taking Expert Opinion	Pilot Test	Reviewing the Test	Statistical Analysis
A.1	-	-	x	x	x	-	-	-	-	-	x	-	-	-	-	-
A.2																
A.3																
.																
.																
.																

*The steps which are shown with "**BOLD TYPE**" were proposed by Treagust (1988).

Data Analysis

In the present study, the research articles in which two-tier tests were used were analyzed by the researchers as follows:

(1) First of all, the researchers presented the distribution of the articles by years. By this way, it is presented in which years two-tier tests became popular and in which years they were not.
(2) Next, the general information about each article was examined and processed in Table 1. In this part, information, authors, science content (e.g. physics, biology etc.), test focus (e.g. determining misconception, defining concept, etc.), target participants and two-tier test styles were processed and presented in the table for each article.
(3) In this step, the researchers examined the two-tier test development process in each research article. The steps that were used by the corresponding authors are examined and presented in a table. Table 2 provides all the findings.

(4) Next, the researchers presented a general figure which contains all the steps of developing two-tier tests which were used in different studies (Figure 2).

(5) Lastly, the researchers provided interpretive explanations about the findings and a new method which can be used to develop two-tier tests was proposed (Figure 3) and explained implicitly.

Validity and Reliability of the Study

In order to be able to conduct a valid and reliable meta-synthesis study, it is required that the researchers should objectively present the methods used throughout the study to the readers. Additionally, the researchers explicitly show the readers the inclusion criteria of the articles in an appropriate way and should include at least 10 studies which are selected through purposeful sampling method to the study. Besides, they should examine each of the studies without disrupting the integrity. In this way, the credibility of the study was strengthened (Sandelowski, Docherty & Emden, 1997). For this study, to conduct a valid and reliable meta-synthesis study, the researchers followed different ways. First, the researchers of the present study independently examined each 271 research article in order to achieve credibility and conformability (Guba, 1981; Lincoln & Guba, 1985). After the examinations, they built 95% consensus on including 42 research articles to the present study. In order to build transferability, the researchers presented the whole process that they followed while conducting a meta-synthesis study in detail and explained what they did in each step explicitly. In addition, the researchers explicitly present all the details of the selected research articles.

Findings

As a result of extensive literature review and analyses, a total of 42 academic papers in the field of science education were identified and included in the study by taking into account the analysis criteria which were determined based on the scope of the study. Among the papers identified, the oldest one dates back to 2000 (Voska & Heikkinen, 2000) and the distribution of the total number of the papers included in the study by years is presented in Graph 1:

Graph 1. Distribution of studies by years
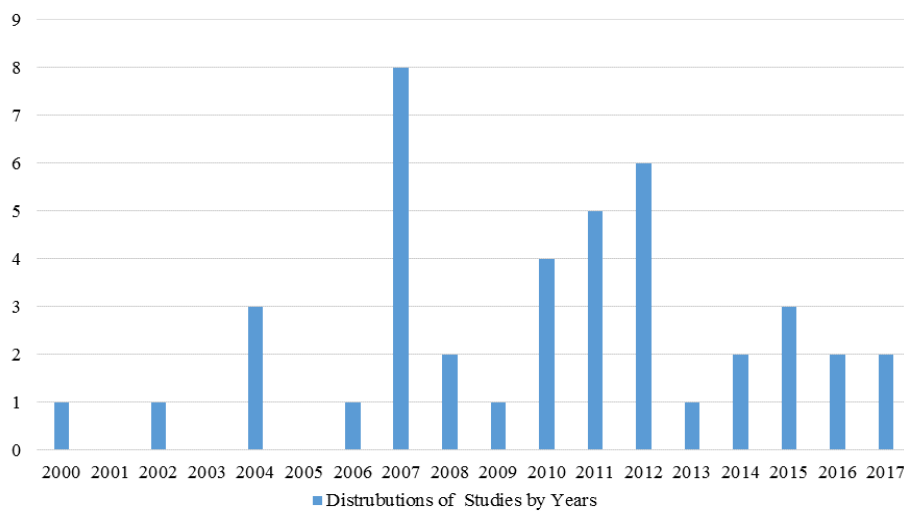
As understood from Graph 1, the two-tier test development studies in the field of science education became highly frequent between 2007 and 2011. Table 4 provides a detailed analysis of the studies examined:

Tablo 4. Detailed information about the studies

Article No	Author(s) (Year)	Content	Test Focus	Target Participants	Test Style	
					First Tier	**Second Tier**
A.1	Adadan and Savaşçı (2012)	Solubility Dissolution	Det. of Misconcept.	High School Students	MC	MC
A.2	Akkuş, Kadayıfçı and Atasoy (2011)	Chemical Equilibrium	Det. of Misconcept. – Concept. Iden.	High School Students	MC	MC
A.3	Alharbi, Treagust, Chandrasegaran and Won (2015)	Diffusion, Osmosis	Concept Iden.	University Students	MC	MC
A.4	Apaydın and Sürmeli (2006)	Natural Selection, Adaptation, Mutation	Concept Iden.	University Students	MC	MC
A.5	Canpolat and Pınarbaşı (2011)	Vaporization	Det. of Misconcept.	Middle School Students	MC	MC
A.6	Chandrasegaran, Treagust and Mocerino (2011)	Chemical Reactions	Det. of Comp. Level	High School Students	MC	MC
A.7	Chang et al. (2007)	Mechanics, Optics Electircity, and etc.	Det. of Misconcept.	Middle School Students High School Students	MC	MC
A.8	Chang, Lee and Yen (2010)	Combustion	Det. of Misconcept.	Primary School Students Middle School Students	YN	MC
A.9	Cheong, Johari, Said and Treagust (2015)	Alternative Energy	Concept Iden.	High School Students	MC	MC
A.10	Cheong, Treagust, Kyvele and Oh (2010)	Malaria	Det. of Comp. Level Det. of Misconcept.	Non-formal Education Students	MC	MC
A.11	Chiu (2007)	Chemistry Concepts	Det. of Misconcept.	Middle School Students High School Students	MC	MC
A.12	Cinici and Demir (2013)	Diffusion, Osmosis	Concept Iden.	High School Students	MC	MC
A.13	Çetin-Dindar and Geban (2011)	Acids and Bases	Det. of Comp. Level	High School Students	MC	MC
A.14	Dahsah and Coll (2008)	Stoichiometry	Det. of Comp. Level	High School Students	MC	MC
A.15	Demirci and Akdemir	Solid and Liquid's	Det. of Comp. Level	Middle School Students	MC	MC

	(2009)	Pressure	Det. of Misconcept.			
A.16	Eymur and Geban (2017)	Chemical Bonds	*No information provided..*	High School Students	MC	MC
A.17	Kao (2007)	Respiration	Det. of Misconcept.	Middle School Students High School Students	MC	MC
A.18	Lee (2007)	Cells and Batteries	Det. of Comp. Level	Primary School Students Middle School Students	TF	MC
A.19	Lin (2004)	Plant Cycle, Seed Germination	Det. of Comp. Level	High School Students	MC	MC
A.20	Lee, Liu and Linn (2011)	Physics, Chemistry, Biology Concepts	Det. of Comp. Level	Middle School Students	MC	OE
A.21	Loh, Subramaniam and Tan (2014)	Electrochemical Cells	Concept Iden.	High School Students	MC	MC
A.22	Monteiro, Nobrega, Abrantes and Gomes (2012)	Minerals	Det. of Misconcept.	High School Students	MC	MC
A.23	Moutinho, Moura and Vasconcelos (2016)	Seismic Activity	Det. of Comp. Level	University Students	TF	MC
A.24	Mutlu and Şeşen (2016)	Thermochemistry, Acids and Bases	Det. of Comp. Level	University Students	MC	MC
A.25	Nantawanit, Panijpan and Ruenwongsa (2012)	Plants	Det. of Comp. Level	High School Students	TF	MC
A.26	Ormancı and Özcan (2012)	Human Body Systems	Det. of Comp. Level	Primary School Students	MC	MC
A.27	Othman, Treagust and Chandrasegaran (2008)	Nature of Matter, Chemical Bonds	Det. of Comp. Level	Middle School Students High School Students	MC	MC
A.28	Özbayrak and Kartal (2012)	Compounds	Det. of Comp. Level	High School Students	MC	MC
A.29	Sia, Treagust and Chandrasegeran (2012)	Electrolyse	Det. of Comp. Level	High School Students	MC	MC
A.30	Siswaningsih, Firman and Khoirunnisa (2017)	Mole	Det. of Misconcept.	High School Students	MC	MC
A.31	Taber and Tan (2007)	İonization Energy	Det. of Comp. Level	High School Students	MC	MC
A.32	Tan, Goh, Chia and Treagust (2002)	İnorganic Chemistry	Det. of Comp. Level	High School Students	MC	MC
A.33	Treagust et al. (2010)	Matter	Det. of Comp. Level	High School Students	MC	MC
A.34	Tsai, Chen, Chou and Lain	Electric Circuits	Concept Iden.	High School Students	MC	MC

	(2007)					
A.35	Tsui and Treagust (2010)	Genetics	Det. of Comp. Level	High School Students	MC	MC
A.36	Uyulgan, Akkuzu and Alpat (2014)	Molecular Geometry	Det. of Comp. Level	University Students	MC	OE
A.37	Vitharana (2015)	Matter Transmission in Plants	Det. of Misconcept.	High School Students	MC	MC
A.38	Voska and Heikkinen (2000)	Principal of Le Chatelier	Det. of Misconcept.	University Students	MC	OE
A.39	Wang (2004)	Matter Transmission in Plants	Det. of Comp. Level	Primary, Middle, High School Students	MC	MC
A.40	Wang (2007)	Matter Transmission in Plants	Det. of Misconcept. Concept Iden.	Primary School Students Middle School Students	MC	MC
A.41	Yen, Yao and Chiu (2004)	Classification	Det. of Misconcept.	Primary, Middle, High School Students	MC	MC
A.42	Yen, Yao and Mintzes (2007)	Biodiversity	Concept Iden.	Primary, Middle, High School and University Students	MC	MC

Abbreviations

Concept Identification (Concept Iden.)
Determination of Comprehension Level (Det. of Comp. Level)
Determination of Misconceptions (Det. of Misconcept.)
Multiple Choice (MC)
Open Ended (OE)
True-False (TF)
Yes-No (YN)

Authors Abbreviations

A.7 - Chang et al. (2007) (Chang, Chen, Guo, Chen, Chang, Lin, Su, Lain, Hsu, Lin, Chen, Cheng, Wang and Tseng 2007)
A.33 – Treagust et al. (2010) (Treagust, Chandrasegaran, Crowley, Yung, Cheong and Othman 2010)

As Table 4 indicates, it was found that two-tier tests were developed mostly in the field of chemistry (46,66%) among the branches of science. In addition, in the fields of biology (33,33%), physics (13,33%), and earth science (6,66%), the studies in which two-tier tests were developed were observed (Graph 2). Among the findings, two-tier tests were mostly developed and implemented to the sample consisting of high school students (50%). In Graph 3, the distribution of sample groups in the 42 studies is presented in detail.

Graph 2. Subject Area Distribution of Studies in Which Two-tier Test Were Implemented

Subject Area

Physiscs · Chemistry · Biology · Earth Sciences
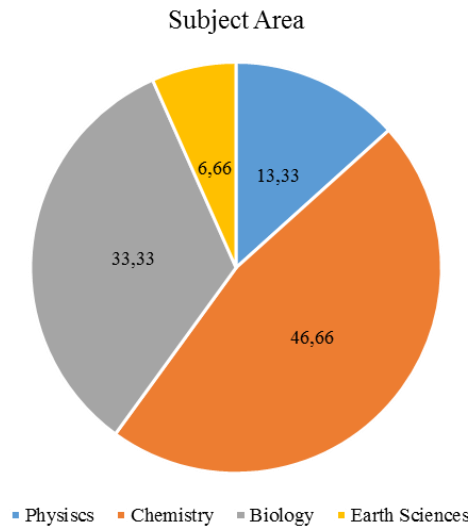
Graph 3. Sample Distribution

Sample Distrubition

University Students · High School Students · Middle School Students
Primary School Students · Non-Formal Education Students
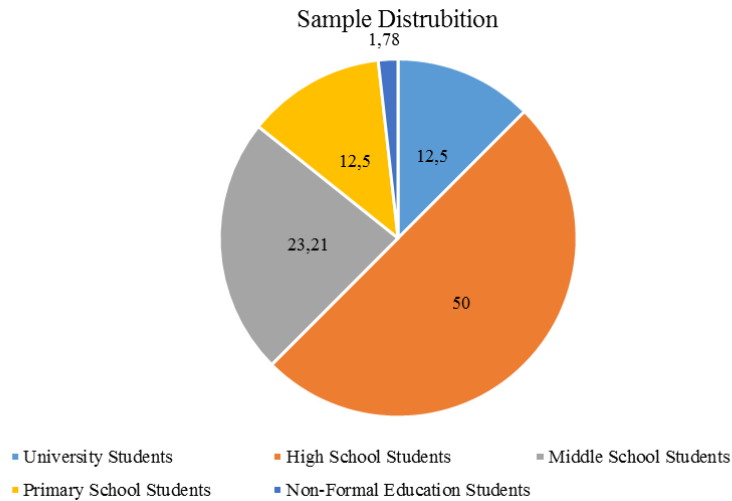
In the following, Figure 1 reflects for what aims the researchers used two-tier tests.
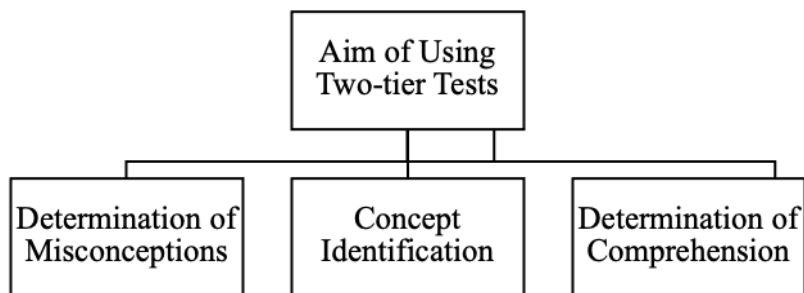
Figure 1. Aims of Using Two-tier Tests

Graph 4 shows the frequency status of the aims of using two-tier tests in the relevant academic papers.

Graph 4. Frequency Status of the Aims of Using Two-tier Tests

Aim of Implication the Two-tier Test

■ Determination of Misconception   ■ Concept Identification   ■ Determination of Comprehension
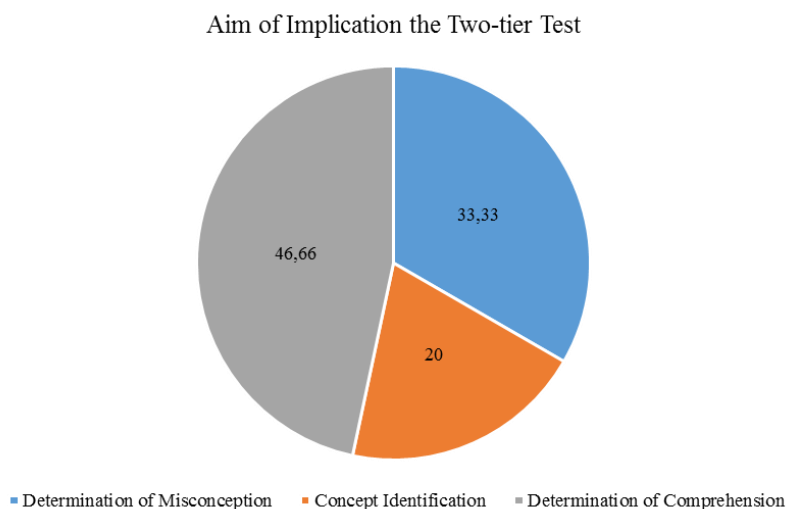
When Figure 1 and Graph 4 are examined, the aims of using two-tier tests are grouped under three headings which are 'concept identification (determination of the mean of concept for students) (20%), determination of comprehension (33,33%), and determination of misconception (46,66%)'.

In accordance with the second research question of the study, the two-tier test development process and the steps of creating the tests in the studies included, which is reflected in Table 5, were examined from the perspectives of the similarities and the differences in test development processes. In order to provide ease of reading, each article was coded as A.1, A.2, A.3, …., A.42 in Table 4, and in Table 5 those codes were directly used for the corresponding articles. In Table 5, all the findings about test development processes were reflected by the researchers.

Table 5. Two-tier Test Development Steps Used in the Selected Articles

Article No	Literature Review	Curriculum Analysis	Creating Propositional Knowledge Statements (PKS)	Creating Concept Map (CM)	Founding Relationships btw. CM and PKS	Content Validity (Expert Opnion)	Literature Review for Misconception (Misconp.)	Developing (Dev.) Interview Skills of Participants	Making Interviews for Determination of Misconcp.	Dev. Test with Open-ended second tier and Application	Dev. Two-tier Test Items	Creat. Table of Specifications	Taking Expert Opinion	Pilot Test	Reviewing the Test	Statistical Analysis
A.1	√	-	√	√	√	-	√	-	√	√	√	√	-	-	√	√
A.2	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.3	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.4	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.5	√	-	-	-	-	-	√	-	-	-	√	-	-	√	-	√
A.6	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.7	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.8	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.9	-	-	√	√	√	√	√	-	-	√	√	-	-	√	√	√
A.10	-	-	√	√	√	-	√	-	√	√	√	√	-	-	√	-
A.11	-	-	√	√	-	-	-	-	√	-	√	-	√	√	√	√
A.12	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.13	-	-	√	-	-	√	√	-	√	√	√	-	-	√	√	√
A.14	√	-	√	-	-	-	-	-	√	√	√	-	√	√	√	√
A.15	√	-	-	-	-	√	-	-	-	√	√	-	√	√	√	√
A.16	-	-	√	√	√	-	√	-	√	√	√	√	-	√	√	√
A.17	√	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.18	√	√	√	√	√	-	-	√	√	-	√	-	-	√	√	-

A.19	√	-	√	√	√	√	-	-	√	√	√	-	-	√	√	√
A.20	√	-	-	-	-	-	√	-	√	√	-	-	-	-	-	-
A.21	√	-	√	√	√	-	√	-	√	√	√	√	-	√	√	-
A.22	√	-	√	√	√	-	-	-	-	√	√	-	-	√	√	-
A.23	√	-	-	-	-	-	√	-	-	√	√	√	-	-	√	-
A.24	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.25	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.26	-	-	√	√	√	√	√	-	√	√	√	-	-	√	√	√
A.27	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.28	√	-	√	-	-	√	√	-	-	-	√	√	√	√	√	√
A.29	√	-	√	√	√	-	√	-	√	√	√	-	-	√	√	-
A.30	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.31	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.32	-	-	√	√	√	√	-	-	√	√	√	-	√	-	√	√
A.33	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.34	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.35	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.36	√	-	-	√	-	√	-	-	√	√	-	-	√	√	√	√
A.37	-	-	√	√	√	-	√	-	√	√	√	√	-	-	√	-
A.38	√	-	√	-	-	√	√	-	-	√	-	√	√	-	-	-
A.39	√	-	√	√	√	√	√	√	√	√	√	√	√	√	√	√
A.40	√	-	√	√	√	√	√	√	√	√	√	√	√	√	√	√
A.41	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-
A.42	-	-	√	√	√	√	√	-	√	√	√	√	-	-	√	-

*The steps which are shown with "BOLD TYPE" were proposed by Treagust (1988).

When the papers were examined, it was seen that the method that Treagust (1988) proposed was the most preferred (%40,5) method used in the development process of two-tier tests (articles: A.2, A.3, A.4, A.6, A.7, A.8, A.12, A.24, A.25, A.27, A.30, A.31, A.33, A.34, A.35, A.41, A.42). Furthermore, some researchers followed different steps in the process of developing two-tier tests, in other words, some researchers added different steps to the method of Treagust (1988). In addition to the two-tier test development process proposed by Treagust (1988), it was observed that the "literature review" step which is conducted at the beginning of the development process (A.1, A.5, A.14, A.15, A.17, A.18, A.19, A.20, A.21, A.22, A.23, A.28, A.29, A.36, A.38, A.39, A.40) was usually included in the development process by the researchers. This step is used for the purposes of defining the concepts, determining the connections, making a detailed literature review that includes the examination of books, teaching guides and curricula about the related branch of science etc. before establishing proposition statements and concept maps (Wang, 2004). All the steps of developing two-tier tests used in the related articles were presented in Figure 2.
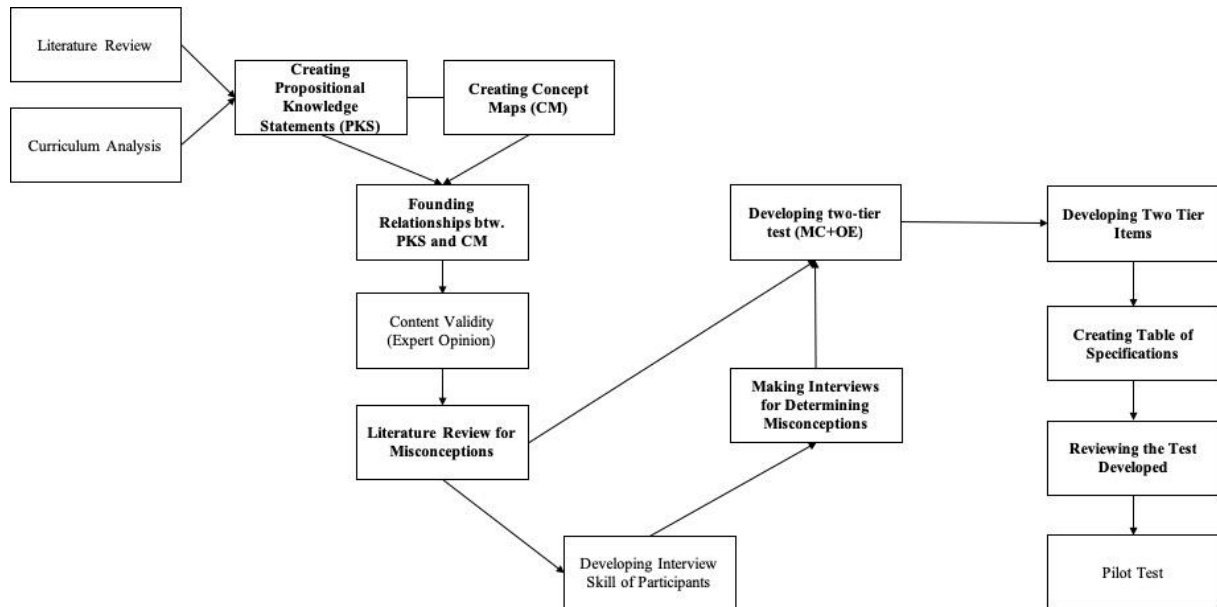
Figure 2. Summary of Two-tier Test Development Method

In a study, Lee (2007) emphasized that after the literature review step, 'curriculum analyses' should be carried out, too. In some studies (A.18, A.39, A.40), while developing the second tier of the two-tier test, it was aimed to increase the effectiveness of the interviews by conducting studies to improve the interviewing skills of the students before the interview in order to carry out successful interviews. Other steps proposed in addition to Treagust (1988)'s two-tier test steps are 'receiving expert opinion' and 'pilot test' steps which are to be used after forming the second tier of the two-tier test. In some of the studies examined (A.11, A.14, A.15, A.28, A.32, A.36, A.38, A.39, A.40), the researchers emphasized the necessity of getting expert opinions in order to examine the test in terms of some features (structure, format and etc.) after the development of the first and second parts of the two-tier test. On the other hand, it was observed in a number of studies (A.5, A.9, A.11, A.13, A.14, A.15, A.16, A.18, A.19, A.21, A.22, A.26, A.27, A.28, A.29, A.30, A.31, A.32, A.33, A.34, A.35, A.36, A.37, A39, A42) that after taking the expert opinion, the researchers emphasized the necessity of a "pilot test" in order to examine the test on a small sample group in terms of the purposes of readability, intelligibility, etc.

Discussion

In the present study, it was aimed to examine the studies in which researchers developed two-tier tests and explicitly explain how they developed those tests. Based on the detailed examination of the test development processes, the researchers of the present study proposed a new pathway which can be used for developing two-tier tests. The pathway is presented in Figure 3 in general. The arrows show the flow starting from the beginning. In the next part, the pathway is explained in detail.
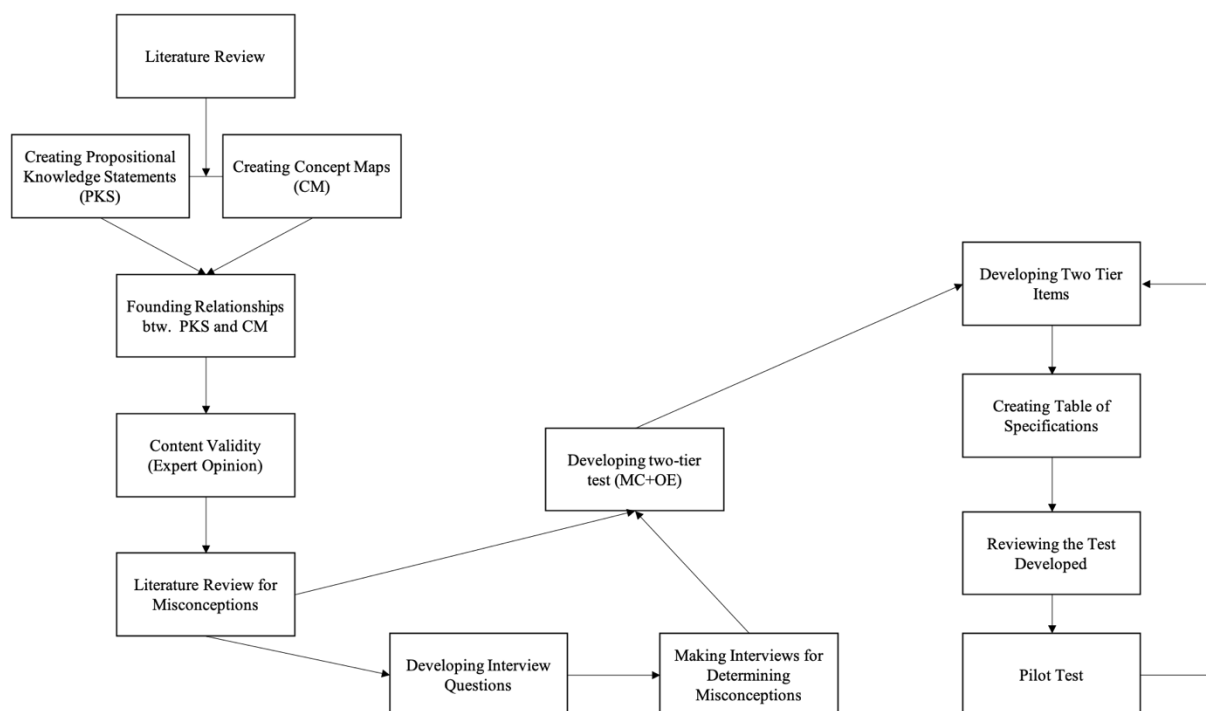
Figure 3. The New Model Suggestion for Two-tier Test Development

As a result of the detailed analysis, at the beginning of the test development process, it is proposed to carry out literature review, in which science books, teaching and teacher guides and programs, articles that can be found in the study area of the two-tier tests, and science programs are examined, before creating propositional knowledge statements and concept maps (Adadan & Savaşçı, 2012; Dahsah & Coll, 2008; Kao, 2007; Lee, 2007; Lin, 2004; Monteiro, Nobrega, Abrantes & Gomes, 2012; Moutinho, Moura & Vasconcelos, 2016; Sia, Treagust & Chandrasegaran, 2011; Voska et al., 2000; Wang, 2004; Wang, 2007). After conducting extensive literature review step, researchers must create propositional knowledge statements and concept maps and next, between propositional knowledge statements and concept maps, the relationships between these two should be found out as Treagust proposed in his study (1988). In the following step, it is proposed to provide content validity by taking expert opinions. Therefore, the relationships found out between concept maps and propositional knowledge statements can be strengthened.

After following the defined steps above, a researcher should follow some other steps explained in the next steps to create the second tier (justification) of the two-tier tests, which makes two-tier tests different from other tests. In order to determine the misconceptions which are used as alternatives or causes of the second tier, it was found that in some studies, only literature review (Cheang et al., 2015; Moutinho et al., 2016); in some studies, interviews with students (Chiu, 2007; Dahsah & Coll, 2008; Lee, 2007; Lin, 2004; Monteiro et al., 2012; Tan et al., 2002); and in some studies both interviews with students and literature review on the misconceptions about corresponding topics of the study (Adadan & Savaşçı, 2012; Chang et al., 2007; Chang et al., 2010; Eymur & Geban, 2017; Kao, 2007; Nantawanit et al., 2012; Othman et al., 2008; Taber & Tan, 2007; Treagust, 1988; Treagust et al., 2010; Tsai et al., 2007; Tsiu & Treagust, 2010; Sia et al., 2011; Wang, 2004; Yen et al., 2004; Yen et al., 2007) are followed by the researchers. As a result of detailed examinations, the researchers of the present study propose that to create alternatives to the second tier of two-tier tests, researchers should conduct both literature review and interviews with students to determine all the misconceptions about corresponding topics as much as possible. In the next step, researchers should create different two-tier items having multiple choice items in the first tier and open ended items in the second tier. Researchers should, then, examine the answers given by students to the open ended items qualitatively and develop reasons and alternatives for the second tier of the two-tier tests aimed to be developed. In order to reveal whether the developed test satisfies the content validity or not, all the items are placed in table of specifications. In addition, the two-tier test developed should be examined by field experts (Reviewing the Test Developed) from the perspectives of validity and reliability. Additionally, a pilot study should be conducted on a small size sample to examine usefulness, reliability and validity of the test and test items. In case of any mistaken items or items which needs

readjustments, researchers must turn back to the 'developing two-tier test items' step as shown in Figure 3 and by going through the corresponding item or items, the process might be continued.

The present study aims to create a new pathway for the development of two-tier tests which are widely used as diagnostics tests for misconceptions and for other purposes because of its power. It is thought that researchers can benefit from the method created as a result of synthesis of different studies described above while creating a new two-tier test. It is also thought that by following the steps mentioned above, the strength of the tests can be enhanced in terms of validity and reliability.

Acknowledgment

References

Adadan, E., & Savaşçı, F. (2012). An analysis of 16–17-year-old students' understanding of solution chemistry concepts using a two-tier diagnostic instrument. *International Journal of Science Education*, *34*(4), 513-544.

Akkus, H., Kadayıfçı, H., & Atasoy, B. (2011). Development and application of a two-tier diagnostic test to assess secondary students' understanding of chemical equilibrium concepts. *Journal of Baltic Science Education*, *10*(3).

AlHarbi, N. N., Treagust, D. F., Chandrasegaran, A. L., & Won, M. (2015). Influence of particle theory conceptions on pre-service science teachers' understanding of osmosis and diffusion. *Journal of Biological Education*, *49*(3), 232-245.

Apaydin, Z., & Sürmeli, H. (2006). Üniversite öğrencilerinin doğal seçilim, adaptasyon ve mutasyon ile ilgili görüşleri [University students' conceptions about natural selectıon, adaptation and mutation]. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi*, *22*, 31-46.

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. New York: Grune & Stratton.

Bair, C. R. (1999, November). Meta-synthesis. In *annual meeting of the Association for the Study of Higher Education, San Antonio, TX*.

Canpolat, N., & Pınarbaşı, T. (2011). Bazı kimya kavramlarına yönelik iki kademeli çoktan seçmeli bir testin geliştirilmesi ve uygulanması [Development and application of two-tier multiple choice instrument on some concepts of chemistry]. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, *13*(1), 55-80.

Chandrasegaran, A. L., Treagust, D. F., & Mocerino, M. (2011). Facilitating high school students' use of multiple representations to describe and explain simple chemical reactions. *Teaching Science: The Journal of the Australian Science Teachers Association*, *57*(4), 13-20.

Chang, H. P., Chen, J. Y., Guo, C. J., Chen, C. C., Chang, C. Y., Lin, S. H., ... & Chen, C. C. (2007). Investigating primary and secondary students' learning of physics concepts in Taiwan. *International Journal of Science Education*, *29*(4), 465-482.

Chang, J. M., Lee, H., & Yen, C. F. (2010). Alternative conceptions about burning held by Atayal indigene students in Taiwan. *International journal of science and mathematics education*, *8*(5), 911-935.

Cheong, I. P. A., Treagust, D., Kyeleve, I. J., & Oh, P. Y. (2010). Evaluation of students' conceptual understanding of malaria. *International Journal of Science Education*, *32*(18), 2497-2519.

Cheong, I.P.A., Johari, M., Said, H., & Treagust, D., F. (2015). What do you know about alternative energy? Development and use of a diagnostic insrument for upper secondary school science. *International Journal of Science Education, 37*(2), 1-27.

Chiu, M. H. (2007). A national survey of students' conceptions of chemistry in Taiwan. *International Journal of Science Education*, *29*(4), 421-452.

Cinici, A., & Demir, Y. (2013). Teaching through cooperative POE tasks: A path to conceptual change. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *86*(1), 1-10.

Çetin-Dindar, A., & Geban, O. (2011). Development of a three-tier test to assess high school students' understanding of acids and bases. *Procedia-Social and Behavioral Sciences*, *15*, 600-604.

Dahsah, C., & Coll, R. K. (2008). Thai grade 10 and 11 students' understanding of stoichiometry and related concepts. *International Journal of Science and Mathematics Education*, *6*(3), 573-600.

Demirci, N. & Akdemir, E. (2009). Katıların ve sıvıların basıncı konularında öğrencilerin kavrama düzeyleri ve kavram yanılgıları [Students' conceptual levels and misconceptions about solid and liquid pressure]. *Milli Eğitim, 182, 314-330.*

Eymur, G., & Geban, Ö. (2017). The collaboration of cooperative learning and conceptual change: enhancing the students' understanding of chemical bonding concepts. *International journal of science and mathematics education*, *15*(5), 853-871.

Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. New Jersey, Mahwah: Routledge.

Haslam, F., & Treagust, D. F. (1987). Diagnosing secondary students' misconceptions of photosynthesis and respiration in plants using a two-tier multiple choice instrument. *Journal of biological education*, *21*(3), 203-211.

Kao, H. L. (2007). A study of aboriginal and urban junior high school students' alternative conceptions on the definition of respiration. *International Journal of Science Education*, *29*(4), 517-533.

Karataş, Ö. F., Köse, S., & Coştu, B. (2003). Two-stage tests used in determining students' errors and understanding levels. *Pamukkale University Journal of Education, 13 (1)*, 54-69.

Lee, S. J. (2007). Exploring students' understanding concerning batteries—Theories and practices. *International Journal of Science Education*, *29*(4), 497-516.

Lee, H. S., Liu, O. L., & Linn, M. C. (2011). Validating measurement of knowledge integration in science using multiple-choice and explanation items. *Applied Measurement in Education*, *24*(2), 115-136.

Lin, S. W. (2004). Development and application of a two-tier diagnostic test for high school students' understanding of flowering plant growth and development. *International Journal of Science and Mathematics Education*, *2*(2), 175-199.

Liu, X. (2010). *Essentials of science classroom assessment*. Sage Publications.

Loh, A. S. L., Subramaniam, R., & Tan, K. C. D. (2014). Exploring students' understanding of electrochemical cells using an enhanced two-tier diagnostic instrument. *Research in Science & Technological Education*, *32*(3), 229-250.

Monteiro, A., Nóbrega, C., Abrantes, I., & Gomes, C. (2012). Diagnosing portuguese students' misconceptions about the mineral concept. *International Journal of Science Education*, *34*(17), 2705-2726.

Moutinho, S., Moura, R., & Vasconcelos, C. (2016). Mental models about seismic effects: students' profile based comparative analysis. *International Journal of Science and Mathematics Education*, *14*(3), 391-415.

Mutlu, A., & Şeşen, B. A. (2016). Evaluating of preservice science teachers' understanding of general chemistry concepts by using two tier diagnostic test. *Journal of Baltic Science Education*, *15*(1).

Nantawanit, N., Panijpan, B., & Ruenwongsa, P. (2012). Promoting students' conceptual understanding of plant defense responses using the fighting plant learning unit (fplu). *International Journal of Science and Mathematics Education*, *10*(4), 827-864.

Novak, J. D. (1980). Learning theory applied to the biology classroom. The American biology teacher, 42(5), 280-285.

Odom, A. L., & Barrow, L. H. (1995). Development and application of a two-tier diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction. *Journal of research in Science Teaching*, *32*(1), 45-61.

Ormancı, Ü., & Özcan, S. (2012). The effectiveness of drama method in unit "the systems in our bodies" in science and technology course: using two tier diagnostic test. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi (EFMED) (Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education)*, *6*(2), 153-182.

Othman, J., Treagust, D. F., & Chandrasegaran, A. L. (2008). An investigation into the relationship between students' conceptions of the particulate nature of matter and their understanding of chemical bonding. *International Journal of Science Education*, *30*(11), 1531-1550.

Özbayrak, Ö., & Kartal, M. (2012). Ortaöğretim 9. sınıf kimya dersi" bileşikler" ünitesi ile ilgili kavram yanılgılarının iki aşamalı kavramsal anlama testi ile tayini [Determination of misconceptions regarding "compounds'' chapter in secondary education 9th grade by two-tier conceptual understanding test]. *Buca Eğitim Fakültesi Dergisi, 32,* 144-156.

Reeves, T. C., & Okey, J. R. (1996). Alternative assessment for constructivist learning environments. *Constructivist learning environments: Case studies in instructional design*, 191- 202.

Sandelowski, M., Docherty, S., & Emden, C. (1997). Qualitative metasynthesis: issues and techniques. *Research in Nursing & Health, 20*(4), 365-371.

Sia, D. T., Treagust, D. F., & Chandrasegaran, A. L. (2012). High school students' proficiency and confidence levels in displaying their understanding of basic electrolysisconcepts. *International Journal of Science and Mathematics Education*, *10*(6), 1325-1345.

Siswaningsih, W., Firman, H., & Khoirunnisa, A. (2017, February). Development of Two-Tier Diagnostic Test Pictorial-Based for Identifying High School Students Misconceptions on the Mole Concept. In *Journal of Physics: Conference Series* (Vol. 812, No. 1, p. 012117). IOP Publishing.

Taber, K. S., & Tan, K. C. D. (2007). Exploring learners' conceptual resources: Singapore A level students' explanations in the topic of ionisation energy. *International Journal of Science and Mathematics Education*, 5(3), 375.

Tamir, P. (1989). Some issues related to the use of justifications to multiple-choice answers. *Journal of Biological Education*, 23(4), 285-292.

Tan, K. C. D., Goh, N. K., Chia, L. S., & Treagust, D. F. (2002). Development and application of a two-tier multiple choice diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 39(4), 283-301.

Tan, Ş. (2009). *Öğretimde ölçme ve değerlendirme KPSS el kitabı (3. baskı) [Measurment and evaluation in teaching: KPSS handbook (3ʳᵈ ed)]*. Ankara: Pegem Akademi Yayıncılık.

Tan, Ş., Kayabaşı, Y., & Erdoğan, A. (2002). *Planning and evaluation of teaching*. Ankara: Anı Yayıncılık.

Treagust, D. F. (1985). Diagnostıc tests to evaluate student misconceptions in science. *Science Education Association of Western Australia: Western Australian College of Advanced Education, Perth, October 1984*, 17.

Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International journal of science education, 10*(2), 159-169.

Treagust, D. F. (1995). Diagnostic assessment of students' science knowledge. *Learning science in the schools: Research reforming practice*, 1, 327-436.

Treagust, D. F., Chandrasegaran, A. L., Crowley, J., Yung, B. H., Cheong, I. P. A., & Othman, J. (2010). Evaluating students' understanding of kinetic particle theory concepts relating to the states of matter, changes of state and diffusion: a cross- national study. *International Journal of Science and Mathematics Education*, 8(1), 141-164.

Treagust, D. F., & Mann, M. (1998). A pencil and paper instrument to diagnose students' conceptions of breathing, gas exchange and respiration. *Australian Science Teachers Journal*, 44(2), 55-59

Tsai, C. H., Chen, H. Y., Chou, C. Y., & Lain, K. D. (2007). Current as the key concept of Taiwanese students' understandings of electric circuits. *International Journal of Science Education*, 29(4), 483-496.

Tsui, C. Y., & Treagust, D. (2010). Evaluating secondary students' scientific reasoning in genetics using a two-tier diagnostic instrument. *International Journal of Science Education*, 32(8), 1073-1098.

Uyulgan, M. A., Akkuzu, N., & Alpat, Ş. (2014). Assessing the students' understanding related to molecular geometry using a two-tier diagnostic test. *Journal of Baltic Science Education*, 13(6).

Vitharana, P. R. K. A. (2015). Student Misconceptions about Plant Transport--A Sri Lankan Example. *European Journal of Science and Mathematics Education*, 3(3), 275-288.

Voska, K. W., & Heikkinen, H. W. (2000). Identification and analysis of student conceptions used to solve chemical equilibrium problems. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 37(2), 160-176.

Walsh, D., & Downe, S. (2005). Meta-synthesis method for qualitative research: a literature review. *Journal of advanced nursing*, 50(2), 204-211.

Wang, J. R. (2004). Development and validation of a two-tier instrument to examine understanding of internal transport in plants and the human circulatory system. *International Journal of Science and Mathematics Education*, 2(2), 131-157.

Wang, J. R. (2007). Students' thinking and alternative conceptions of transport systems in plants: a follow-up study. *International Journal of Science and Mathematics Education*, 5(2), 307-328.

Yen, C. F., Yao, T. W., & Chiu, Y. C. (2004). Alternative conceptions in animal classification focusing on amphibians and reptiles: A cross-age study. *International Journal of Science and Mathematics Education*, 2(2), 159-174.

Yen, C. F., Yao, T. W., & Mintzes, J. J. (2007). Taiwanese students' alternative conceptions of animal biodiversity. *International Journal of Science Education*, 29(4), 535-553.

Yıldırım, C. (1999). *Eğitimde ölçme ve değerlendirme [Measurment and evaluation in education]*. Ankara: ÖSYM Eğitim Yayınları.

Comparison of Artificial Neural Networks and Logistic Regression Analysis in PISA Science Literacy Success Prediction

Ali Bozak[1], Eren Can Aybek[2]
[1]Republic of Turkey Ministry of National Education
[2]Pamukkale University

To cite this article:

Comparison of Artificial Neural Networks and Logistic Regression Analysis in PISA Science Literacy Success Prediction

Ali Bozak[1*], Eren Can Aybek[2]
[1]Republic of Turkey Ministry of National Education
[2]Pamukkale University

Abstract

The present study aims to determine which analysis technique-Artificial Neural Networks (ANNs) or Logistic Regression (LR) Analysis-is better at predicting the science literacy success of the 15-year Turkish students who participated in PISA research carried out in 2015 by using learning time spent on science, test anxiety, environmental awareness, environmental optimism, epistemological beliefs, inquiry-based science teaching and learning practices, instrumental motivation, and disciplinary climate in science classes as the predictor variables. For this purpose, the data from 5895 students who participated in the PISA 2015 test were analyzed. Models were developed using LR and ANNs, and the results were compared. As a result, although the classification performance of artificial neural network is significantly better compared to LR, it is understood that practical significance is low due to the intersection of AUC confidence intervals.

Keywords: Artificial Neural Networks, Logistic Regression, PISA, Science Literacy

Introduction

In today's world, education and training activities are considered not only as transferring information to students but also the acquisition of high-level skills such as applying the learned knowledge to real life problems, conducting team studies, and learning to learn. As the perspectives on education have recently changed, it is also valuable for students to gain these skills and to measure how much these skills are gained.

The research such as Trends in International Mathematics and Science Study (TIMMS), Progress in International Reading Literacy Study (PIRLS) are conducted worldwide to observe the results of investments made in the field of education and make comparisons with other countries. The Programme for International Student Assessment (PISA), which is implemented by the Organisation for Economic Co-operation and Development (OECD) countries, is one of the researches carried out for this purpose. PISA started to be implemented in 2000 and is held every three years with the participation of countries that are OECD members and non-member countries, as well. In PISA, in which the students who complete the compulsory education of the countries attend, the aim is to assess the level of students' basic knowledge and skills required to be successful in real life. 15-year-old students from seventh and above grade levels participate in PISA research (Republic of Turkey Ministry of National Education [MEB], 2016). In PISA, survey studies are also conducted on variables that may affect student performance. Through the items related to socio-cultural, economic, and educational fields, information is gathered about the students, parents, and schools.

The competencies of students in mathematics, science and reading are measured within the scope of PISA research. At this stage, the concept of *literacy* becomes prominent. Literacy is defined as the ability to use the knowledge and learning outcomes that they gained from the problems they face, to be able to make inferences about the subject by analyzing and to effectively communicate with others. In 2015 PISA research, mainly science literacy was measured. Science literacy is defined as the ability to deal with science-related issues (MEB, 2016).

It is crucial for policy makers to identify the characteristics that affect student success to make better decisions. In this context, PISA results are a very useful data source to discover student performance and variables that describe the performance. In this context, besides regression or structural equation models for predicting student performance, analyses such as ANN, support vector machines, decision trees, and data envelopment analysis

---

* Corresponding Author: *Ali Bozak, alibozak@gmail.com*

can also be used. LR and ANN analysis, which are used in this research, are discussed in detail in the next sections.

Logistic Regression

In LR analysis, prediction of the dependent variable is provided by using independent / predictor variables. The independent/predictor variables in this analysis can be continuous or categorical. However, the dependent variable must be categorical. Predicting of group membership is an example of a LR problem (Tabachnick & Fidell, 2013). In cases which the dependent variable is categorical, LR analysis is preferred rather than multiple regression analysis (Kline, 2011). LR is more flexible compared to other techniques. Contrary to most analyzes, the predictor variables do not have to be normally distributed. These variables can also be continuous, categorical, or mixed. However, while the negative predictor value can be produced in multiple regression analysis, LR does not produce negative values (Tabachnick & Fidell, 2013). There are two methods used in LR analysis. These are called standard and stepwise methods. Stepwise methods are two parts, forward and backward (Çokluk, 2010).

In cases where the dependent variable has two categories, binary LR is used. In LR, independent variables are determined through logarithmic operations. The LR equation with one independent variable is as follows (Rençber, 2018):

$$p(x) = \ln\left[\frac{E(y/x)}{1 + E(y/x)}\right] = \ln\left[\frac{\exp(\beta_0 + \beta_1 x)}{1 + exp(\beta_0 + \beta_1 x)}\right] \tag{1}$$

The value to be obtained from the formula varies between 0 and 1. β_0 indicates the constant in regression. β_1 regression coefficient specifies the effect of the independent variable on the dependent variable.

$p/(1- p)$ ratio refers to the probability of belonging to a group or not (odds). The logarithm of this value is called the logit value of logodds or p. This transformation can be formulated as follows (Retherford & Choe, 1993):

$$logit\ (p) = \log\left[\frac{p}{1 - p}\right] \tag{2}$$

When the equation (2) is considered as a correlation function and xs are placed to indicate independent variables, the following logit model emerges (Oğuzlar, 2005):

$$\log\left[\frac{P_i}{1 - P_i}\right] = z_i = \sum_{k=0}^{P} \beta_k x_{ik} \tag{3}$$

When the data is processed in this equation, the dependent variable increases to 1 or approaches to 0 by decreasing. If the resulting value is greater than 0.5, the solution will result in 1, and 0 if it is smaller (Rençber, 2018).

Artificial Neural Networks

ANNs are algorithms that mimic how neurons and the human brain work and are based on learning information such as the biological brain and creating new information automatically from what they learn. ANNs are a great interest for researchers as a branch of the science of artificial intelligence and it allows computers to learn by massively parallel computing systems (Jain, Mao, & Mohiuddin, 1996). Kim (2017) stated that the ANNs are actually a machine learning model. They are the models that try to imitate human intelligence, and these models were created by analyzing the structure of biological nerve cells (neurons). Neurons receive, process, and transmit information via biochemical reactions (Abraham, 2005). The general structure of a biological neuron is shown in Figure 1.
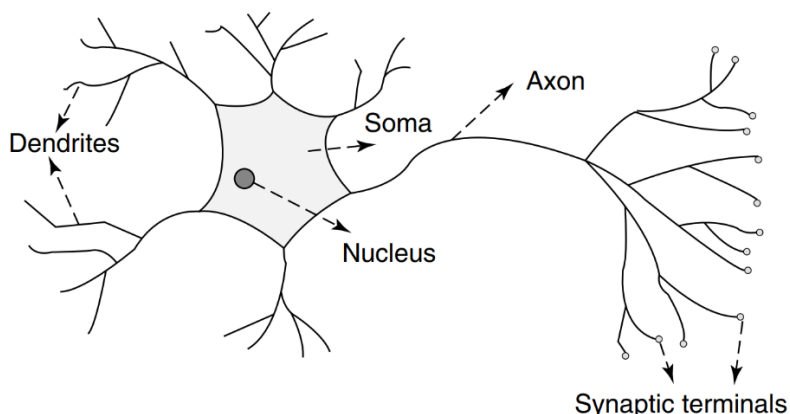
Figure 1. General Structure of the Biological Neuron (Abraham, 2005)

Dendrites in biological neurons receive input signals and transmit them to the soma. This information is processed in the soma. Axons convert these signals into output signals. Connections provide electrochemical contact between neurons. The basic principles of ANNs were first formulated by McCulloch and Pitts in 1943 based on assumptions listed as follows (Graupe, 2013):

- The activity of a neuron is all or none principle.
- The only significant delay in the nervous system is synaptic delay.
- The activity of any inhibitory synapse strictly prevents the simultaneous stimulation of the neuron.
- The structure of the connection network between neurons does not change in time.

ANNs are the structures made up of a large number of processing units connected together, similar to biological neurons. ANNs have interdependent processing elements (nodes). Each node is called an artificial nerve. An artificial neuron is shown in Figure 2.
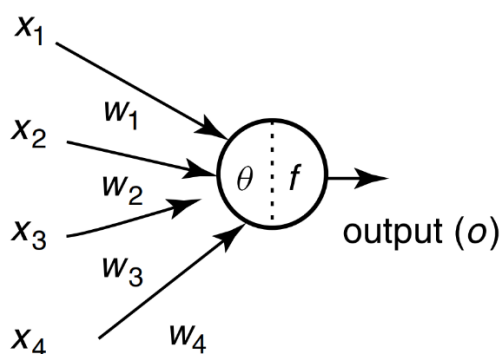
Figure 2. Artificial Neuron (Abraham, 2005)

Inputs: The data transmitted from the outside world or another neuron ($x_1, x_2 \ldots x_n$).

Weights: The data is processed by multiplying the weights. Weights are distributed randomly at the initial training phase (Henseler, 1995). Weights are called ($w_1, w_2 \ldots w_n$).

Summation function: It is the unit where weighted totals are calculated. Values from the inputs are summed up at this stage, multiplied by weights. The result is compared with the threshold value (Shanmuganathan, 2016). w as the weight value, x as the data entering the cell, n as the total value to indicate the number of inputs is calculated by the following formula (Graupe, 2013):

$$z_i = \sum_{i=1}^{n} w_{ij} x_i \qquad (4)$$

Activation Function: The result of the collection phase is sent to the output being passed through the activation function. "Activation function converts the neuron input to its activation … This allows the variation of input

conditions to affect the output…" (Radi & Hindawi, 2013, p. 185). The main reason for using an activation function is to provide the non-linear feature to the neural network. Without using the activation functions, a neural network can perform a limited number of operations to learn nonlinear functions. The output we expect neural networks to learn is rarely linear (Heaton, 2012). Table 1 shows the commonly used activation functions and mathematical formulas (Rençber, 2018, p. 55).

Table 1. *Activation Functions*

Activation Function	Mathematical Formulas
Linear Function	$F(x) = x$
Step Function	$F(x) = \left\{ \begin{array}{ll} 1, & x > 0 \\ -1, & x \leq 0 \end{array} \right\}$
Threshold Function	$F(x) = \left\{ \begin{array}{ll} 1, & x > 0 \\ 0, & x \leq 0 \end{array} \right\}$
Sigmoid Function	$F(x) = \dfrac{1}{1 + \exp(-x)}$
Hyperbolic Tangent Function	$F(x) = \tanh(x)$
Gaussian Function	$F(x) = \exp\left(-\dfrac{x^2}{2\sigma^2} \right)$

Output Function: The activation function generates an output value. This output value can be sent to another cell or it can be given out as a result (Veelenturf, 1995).

ANNs are made up of combined artificial nerve cells. Cells come together in the same direction in layers called input, hidden and output layers (Patterson & Gibson, 2017). These layers is shown in Figure 3.
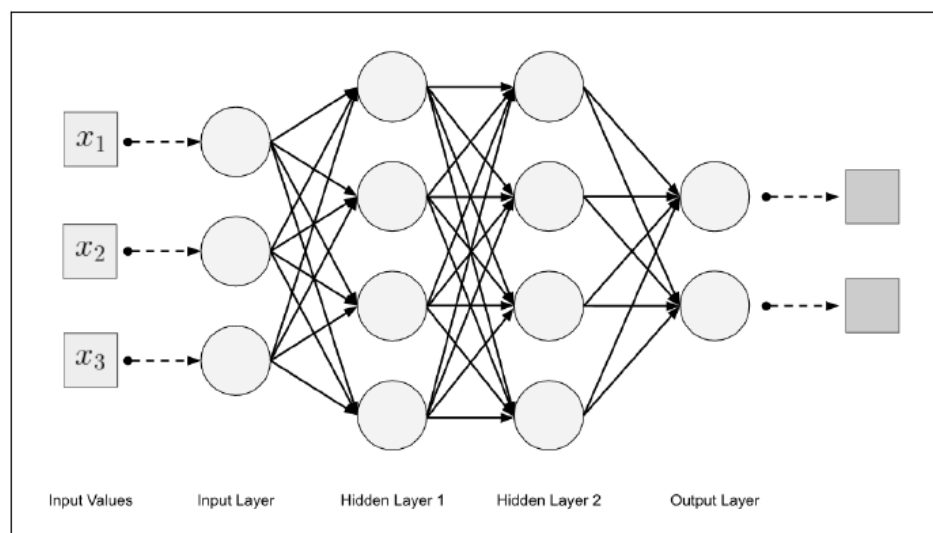
Figure 3. Artificial Neural Network Layers (Patterson & Gibson, 2017).

In the *Input Layer,* information coming from outside to the network is sent to *the Hidden Layer* which classifies the input vector and *Output Layer* retrieves the representations coming from the hidden layer (Hristev, 1998). In ANNs, the inputs are multiplied by weights and included in the function. Weight values must be determined for the network to learn correctly. This process is called *network training*. These values are randomly distributed when data first arrives in the neural network, they are recalculated as examples are given. In this way, suitable weight values are found. When the correct weight value is achieved, the artificial neural network reaches the level to generalize the problem represented by the samples. This is called network learning. The learning process of the network takes place through certain rules which are called learning rules. There are different learning rules such as Hebb rule, Hopfield rule, Delta rule, curved decline rule and Kohonen learning rule. After the completion of network training, it is tested whether it has completely learned or not. This process is called network testing. The network is asked to produce output by providing information that it does not see during training. The network produces results using the weights learned. The higher the results are, the better the network has learned (Elmas, 2018; Öztemel, 2006).

ANNs can be analyzed in three categories as supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, examples of input vectors are given, and the network learns through these examples (Shanmunagathan, 2016). Multilayer perceptron, least mean square, back propagation, and delta rule are some of the examples of supervised learning models (Akı & Karasulu, 2014).

ANNs are classified as single-layer or multi-layer according to layer architecture. Single layer networks have a single input layer and an output node. This simplest neural network is called a perceptron. Multilayer neural networks contain more than one computation layer. A perceptron contains an input and output layer. The input layer transmits data to the output layer and all calculations, and it is fully visible to the user. In multi-layered neural networks, there are additional middle layers between the input and output layers and these middle layers are called hidden layers because the calculations made are not visible to the user. The specific architecture of multi-layered neural networks is also called feedforward networks because the successive layers are supported by each other forward from the input (Aggarwal, 2018).

In the literature, there are some research which compare ANNs and LR to estimation the dependent / predictor variable in different fields (Alkan, Köklükaya & Subaşı, 2005; Biganzoli, Boracchi, Marianı & Marubini, 1998; Chiang, Zhang, & Zhou, 2006; Çakın & Özdemir, 2019; Dreiseitla & Ohno-Machadob, 2002; Eftekhar, Mohammad, Ardebili, Ghodsi & Ketabchi, 2005; Jefferson, Pendleton, Lucas & Horan, 1997; Kayri & Çokluk, 2010; Kurban, Kantar & Hocaoğlu, 2006; Kurt, Ture & Kurum, 2008; Manel, Dias, & Ormerod, 1999; Nefeslioğlu, Gökçeoğlu & Sönmez, 2008; Sarıkaya, 2019; Subaşı & Erçelebi, 2004; Tu, 1996; Vasconcelos, Silva, Tome, Alvim & Perelra, 2001; Yazıcı, 2018).

LR and ANNs are used in predicting student success as a predictor variable and there are studies on which method is more successful (Assi, Nahiduzzaman, Ratrout & Aldosary, 2018; Aydoğan & Zırhlıoğlu, 2018; Bahadır, 2016; Benzer & Benzer, 2017; Bezek Güre, Kayri & Erdoğan, 2019; Çırak & Çokluk, 2013; Delen, 2012; Gorr, Nagin, Szczypula, 1994; Hardgrave, Wilson & Walstrom, 1993; Ibrahim & Rusli, 2007; İnal & Turabik, 2017; Paliwal & Kumar, 2009; Pavlekovic, Bensic & Zekic-Susac, 2010; Teshnizi & Ayatollahi, 2015; Tosunoğlu & Apaydın, 2004). In the majority of the studies, it has been observed that ANNs create a better prediction model whereas some other studies have found no difference. This research, in which PISA 2015 data is used, is considered important in terms of its large data set to compare methods, and determining which variables are more important in predicting student success.

This study aims to compare ANNs and LR Analysis techniques at predicting the science literacy success of the 15-year Turkish students who participated in PISA research carried out in 2015. For this purpose, the following questions were considered:
1. According to the variables of students' learning time spent on science, test anxiety, environmental awareness, environmental optimism, epistemological beliefs, inquiry-based science teaching and learning practices, instrumental motivation, and disciplinary climate in science classes:
    a. What accuracy rate does the model created by LR produce in predicting the success of students' science literacy?
    b. What accuracy rate does the prediction model created by artificial neural network analysis produce in predicting students' success in science literacy?
2. Which model is better in predicting student achievement when ANNs and LR are compared?
3. What is the importance level of the variables included in the model in predicting science success?

Method

The variables used as predictors in the analysis are learning time spent on science (SMINS), test anxiety (ANXTEST), environmental awareness (ENVAWARE), environmental optimism (ENVOPT), epistemological beliefs (EPIST), inquiry-based science teaching and learning practices (IBTEACH), instrumental motivation (INSTSCIE), and disciplinary climate in science classes (DISCLISCI). LR and ANNs were used to determine the levels of predicting the success of science literacy of these variables and the estimation ratios were compared. Therefore, correlational model (Fraenkel, Wallen, & Hyun, 2012) has been adopted in the research. Predictor variables have been chosen according to their correlation with the science subtest score. The variables which have correlation above .10 are included in the analyzes. The only exception is test-anxiety which has correlation .08, yet it is considered as a potential important variable for the prediction of achievement.

Sample

The population of the students (aged 15) with which the PISA research conducted in 2015, is 1,324,089 and the population of the students that can take PISA test is 925,366 in Turkey. The sample of this study is 5895 students who participated in PISA 2015 test. In PISA, the school sample is determined using stratified random sampling. According to Statistical Regional Unit Classification in Turkey (SRE) Level 1; 5,895 students from 187 schools representing the 12 regions in 61 cities. When the number of students is compared, the result is that the most participation is from the Istanbul region and the least is from the Eastern Black Sea region (MEB, 2016).

Table 2. *Distribution of Students by Regions in PISA 2015*

Region Code	Region Name	Number of Participants	Percentage of Participant Students
TR1	İstanbul	1,070	18.15
TR2	Western Marmara	245	4.16
TR3	Aegean	707	11.99
TR4	Eastern Marmara	510	8.65
TR5	Western Anatolia	553	9.38
TR6	Mediterranean	817	13.86
TR7	Central Anatolia	334	5.67
TR8	Western Black Sea	303	5.14
TR9	Eastern Black Sea	194	3.29
TRA	Northeastern Anatolia	199	3.38
TRB	Central Eastern Anatolia	276	4.68
TRC	Southeastern Anatolia	687	11.65
TOTAL		*5,895*	
		100	

Data Collection Tool

In PISA research, literacy levels of students are measured in different areas such as reading, mathematics, collaborative problem solving, and science. In addition, the affective features associated with academic achievement score is assessed. Items applied to students includes the topics such as how motivated students were to learn about science, instrumental motivation towards science subjects, science self-efficacy (MEB, 2016). Computer-based tests are used in the PISA research. Each student has two hours to complete the test. Test consists of essay type and multiple-choice items. There is no fixed test form for students, i.e. different students take different item combinations. Moreover, students are asked to apply questionnaires about themselves, their families, and school environment (OECD, 2018). In this way, the students' willingness to learn science, whether they find science learning useful, how good they are at solving science problems and difficulties are investigated.

Data Analysis

First, students were ranked according to their science literacy levels, and 27% of the most successful group and 27% of the most unsuccessful group was determined. Accordingly, 1341 students were included in each group. Then, LR, via Jamovi 1.2.0.0 (The Jamovi Project, 2019) software, was used to determine whether students are in the lower or upper 27% according to their PISA 2015 science literacy scores (PV1SCIE) to be able to estimate with the variables SMINS, ENVAWARE, ENVOPT, IBTEACH, EPIST, ANXTEST, INSTSCIE and DISLISCI. To predict the same variable, an artificial neural network was created through Multi-Layer Perceptron (MLP) with the same predictor variables and which group the students belong to was estimated. *The standardized* method was used to the rescaling of continuous variables, and 90% of the data set was used for training and 10% of the data set was used for the test. This proportions are used because of two reasons. First, the data already split to 27% groups which has already reduced the amount of data, and second, generally full data set is used for LR in educational research. These proportions are used for LR and ANNs comparability purpose and both techniques used the same training data. Initial seed value was determined as 2081980 for reproducibility of the results. The lowest number of hidden layers is set to 1, and the maximum number of hidden layers to 50. In training, *batch* method was used. The artificial neural network has been tested through IBM SPSS Statistics v20. PROC v.1.16.1 (Robin et al., 2011) was used on R 3.6.0 (R Core Team, 2019) to compare the classification performance of both LR and ANNs.

Findings

According to the LR conducted to estimate whether the students are in the lower or upper 27% groups according to PISA 2015 science literacy scores, it was found that the *full-model* with all predictor variables was significant against the *constant-only-model* (χ^2_8=1169; p<.05). This shows that the predictor variables are able to distinguish between students with low and high science literacy. In addition, when the *pseudo R^2* values of Negelkerke, McFadden, Cox and Snell, which are calculated for how much of the group membership of the students are estimated with the tested regression model, the following values are obtained respectively: .314, .353 and .471. R^2 values calculated in LR tend to be smaller than the values calculated in multiple linear regression models; therefore, it is considered sufficient to have *McFadden R^2* value between .20 and .40 (Alpar, 2011; Tabachnick & Fidell, 2013). When the classification accuracy of the students was analyzed, it was found that a satisfactory classification was made even though no achievement variable was used. 76.5% of the students in the unsuccessful group; 79.8% of students in the successful group and 78.2% of all students in general was classified correctly. The Receiver Operating Characteristic (ROC) curve obtained for classification is given in Figure 4. The area under this curve [Area Under the Curve] (AUC) value was calculated as .856 [.842 - .871]. This value indicates that students are classified very well with predictor variables. Cantor and Kattan (2000) indicated the cut-off values for AUC as .80, good; .65, fair, and .50, poor. Regression coefficients and odds ratio values for predictor variables are given in Table 3.
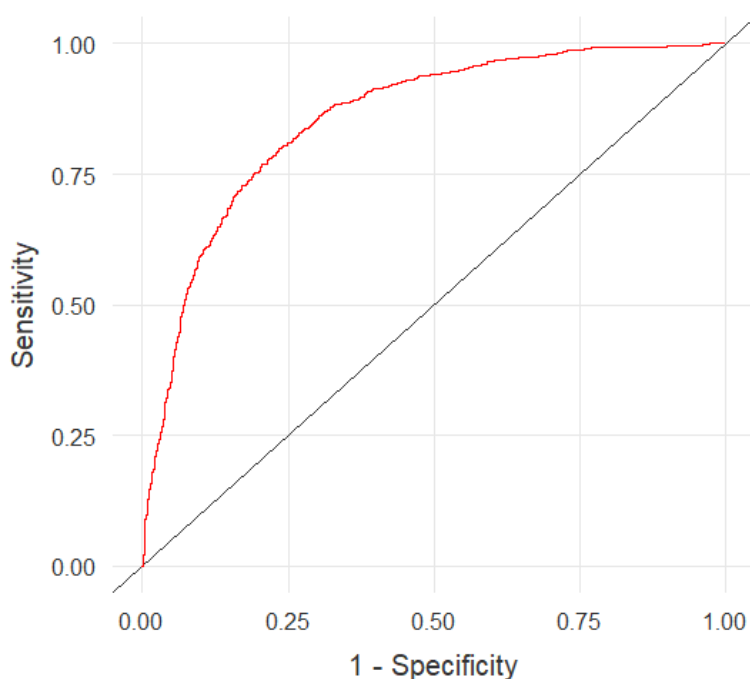
Figure 4. ROC Curve for Logistic Regression Model

Table 3. *Logistic Regression Analysis to Estimate Student Success*

Variables	B	SE	Z	Odds Ratio	95% Confidence Interval for Odds Ratio	
					Lower	Upper
Intercept	-1.896	.130	-14.54**	.150	.116	.194
SMINS	.007	5.27e-4	14.09**	1.007	1.006	1.008
ENVAWARE	.340	.037	9.22**	1.405	1.307	1.511
ENVOPT	-.538	.038	-14.18**	.584	.542	.629
IBTEACH	-.300	.044	-6.78**	.741	.679	.808
EPIST	.404	.045	9.05**	1.497	1.372	1.634
INSTSCIE	.185	.054	3.41**	1.203	1.082	1.338
ANXTEST	-.298	.048	-6.16**	.742	.675	.816
DISLISCI	.133	.054	2.47*	1.142	1.028	1.269

** p < .01; * p < .05

In Table 3, there are lower and upper limits of the 95% confidence interval for the regression coefficients, Z and Odds Ratio values for the eight predictor variables. Accordingly, all variables have a significant role in predicting students' position in the lower or upper group according to PISA 2015 science literacy success. In a one-unit increase in each predictor variable, 1 was subtracted from the e^B coefficients and multiplied by 100 to determine what change is expected in the group belonging odds of the students (Hair, Black, Babin, & Anderson, 2014). In this way, with a one-unit increase in the ENVOPT variable, a decrease of approximately 41.6% can be expected in the odds of students to be in the unsuccessful group ($[e^{-0.538} - 1] .100 = -41.61$). One-unit increase in the EPIST variable, an increase of approximately 49.8% can be expected in the odds of students to be in the unsuccessful group ($[e^{0.404} - 1] .100 = 49.78$).

According to the MLP results, it was observed that a satisfactory accuracy classification was made in both training and test sets, even any success related variable was not used when estimating the students' group. In the training set, it was found that 77% of students in the unsuccessful group, 80.7% of students in the successful group and 78.9% of all students in general can be classified correctly. In the test set, 80.4% of the students in the unsuccessful group, 88.1% of the students in the successful group and 84.3% of all students in general were classified correctly. The ROC curve in Figure 5 was obtained for the training and test sets. The area under this curve (AUC) was found to be .868 [.854 - .881] for both training and test sets. The fact that the AUC value obtained is greater than .80 indicates that the students are also classified very well with MLP.
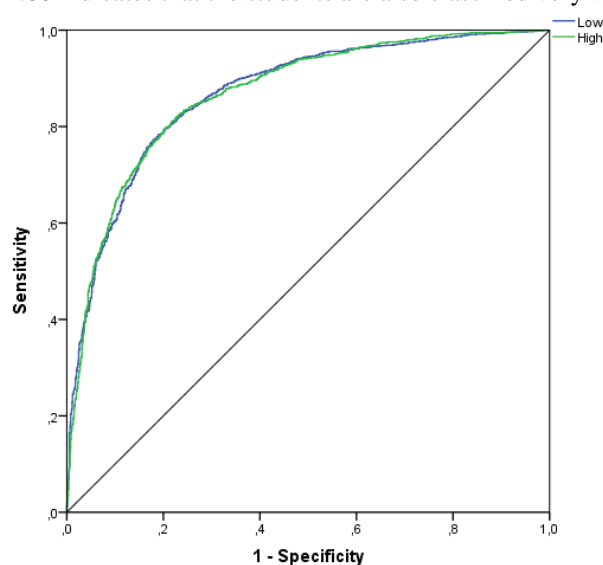
Figure 5. ROC Curve for Multi-Layer Perceptron

Correct and incorrect classification rates for estimating the lower or upper groups according to LR and ANNs according to PISA 2015 science literacy are summarized in Table 4.

Table 4. Correct Estimation Percentages for *Logistic Regression and Artificial Neural Network*

	Observed	Estimated		Correct Estimation Percentage
		Unsuccessful	Successful	
Logistic Regression	Unsuccessful	1026	315	76.5
	Successful	271	1070	79.8
	Total			78.2
ANN (Training)	Unsuccessful	923	275	77.0
	Successful	231	967	80.7
	Total			78.9
ANN (Test)	Unsuccessful	115	28	80.4
	Successful	17	126	88.1
	Total			84.3

As can be seen in Table 4, multilayer perceptron has been able to correctly classify the group that students belong to above 80%, especially in the test data. In addition, in comparing LR and classification performance of ANNs AUC values were compared as suggested by DeLong, DeLong, and Clarke-Pearson (1988) and a significant difference was found between the two AUC values (Z = 3.636; p <.05). However, when 95% probability confidence intervals of AUC values are analyzed, it is seen that confidence intervals are [.842 - .871] for LR and [.854 - .881] for artificial neural network, and therefore two confidence intervals intersect. In this context, although the classification performance of the artificial neural network is found to be significantly better than LR, the significance level is practically low due to the intersection of AUC confidence intervals. In the prediction of which group to the students belong to according to their science success with ANNs, the order of importance of the predictor variables is given in Figure 6.
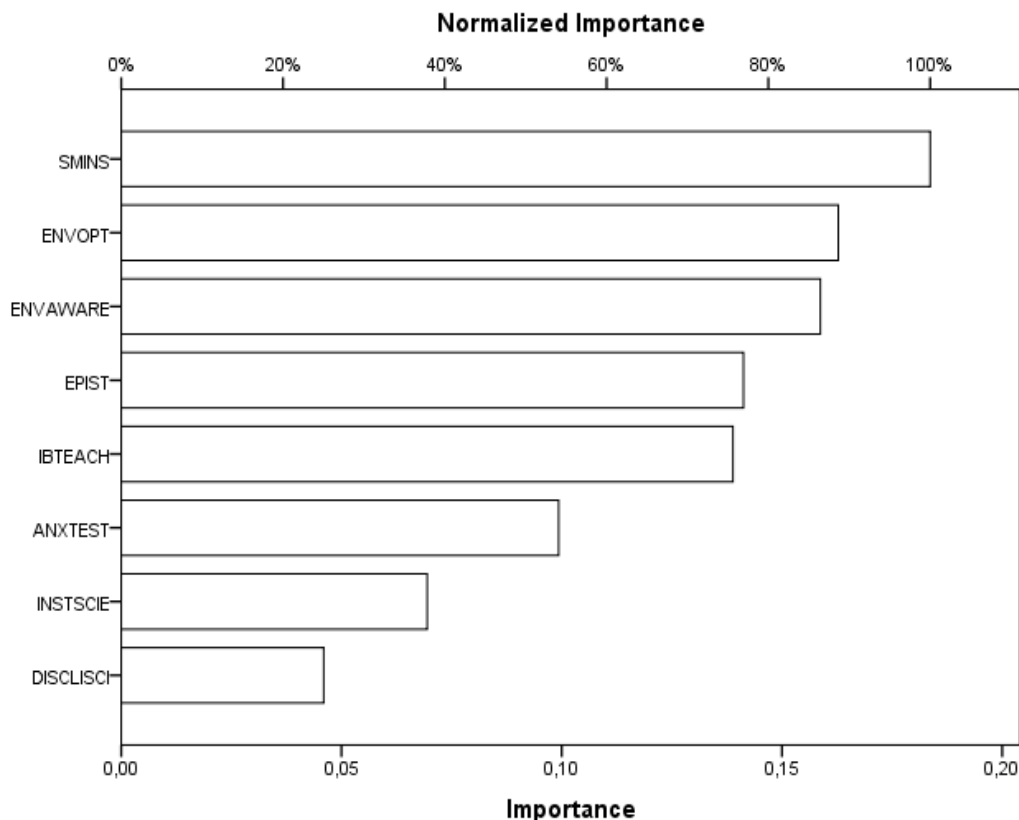
Figure 6. Order of Importance of Predictor Variables

According to the science literacy scores of the students, it is seen that SMINS is the most important variable in predicting if they are in the lower or upper groups with ANNs, and the least important variable is DISLISCI. Among the variables included in this study, it was concluded that the most important factor on student success is the time allocated to learning. In addition, students' interest in environmental problems and inquiry-based science education were found to be important variables that explain success. It is concluded that the variable that has the least important variable that explain student achievement is the disciplinary climate in the classroom.

Conclusion

LR and ANNs are two methods that can be used to predict a variable. Although it has been reported in the literature that ANNs perform better in the vast majority of studies comparing the two methods (Acar Boyacıoğlu & Kara, 2007; Alkan et.al, 2005; Assi et.al, 2018; Bahadır, 2016; Benzer & Benzer, 2017; Bezek Güre et.al, 2019; Biganzoli et.al, 1998; Budak & Erpolat, 2012; Chiang et.al, 2006; Çakın & Özdemir, 2019; Çırak & Çokluk, 2013; Delen, 2012; Flitman, 1997; Ibrahim & Rusli, 2007; Jefferson et.al, 1997; Kayri & Çokluk, 2010; Keskin Benli, 2005; Kurt et.al, 2008; Pradhan & Lee, 2010; Rençber, 2018; Sarıkaya, 2019; Subasi & Erçelebi, 2005; Teshnizi & Ayatollahi, 2015; Tolon & Güneri Tosunoğlu, 2008; Vasconcelos et.al, 2001; Walczak & Sincich, 1999; Yeşilnacar & Topal, 2005), there are also studies reporting that LR predicts at a higher level (Kurban et.al, 2007) or reaches the finding that they are equally successful (Tosunoğlu & Apaydin, 2004).

In addition, there are some studies in the literature that conclude that when two analyses are compared, there is no clear advantage, and the success of the method depends on the situation worked on (Dreiseitla & Ohno-Machadob, 2002; Eftekhar et.al, 2005; Gorr et.al, 1994; Hardgrave et.al, 1994; Manel et.al, 1999; Nefeslioğlu et.al, 2008; Pavlekovic et.al, 2010; Tu, 1996).

In the literature, it is seen that when comparing LR and ANNs' success conditions, the results show that the comparisons are carried out based on percentage values. In this study, after the percentage values were determined, the significance of the difference between the AUC values obtained for classification performance was also analyzed. Accordingly, which groups the students belong to can be estimated correctly with logistics regression at 78.2%, with ANNs at 84.3%. In addition, when comparing the classification performances, ANNs were shown to perform better. However, when the confidence intervals were considered, it was found that this difference was low in practice.

The results show that the most important variable in predicting students' science literacy was the time allocated to science education, and the least important variable was the disciplinary climate in the classroom. Other variables were environmental optimism, environmental awareness, epistemological beliefs, inquiry-based science teaching, test anxiety, and instrumental motivation in terms of the order of importance. When the variables are analyzed, the sensitivity of the student to environmental issues (ENVAWARE and ENVOPT) comes to the fore in predicting science success. In addition, it is expected that preference of inquiry-based science learning (IBTEACH) will increase the expected science literacy score in PISA.

Recommendations

This research was conducted on science literacy score with 2015 PISA results. ANNs and LR models can be used for different subtests of PISA. In addition, different variables can be included in the study to analyze their role in prediction. However, in the present study, only MLP were used as artificial neural network model. Different artificial neural network models or studies to test the predictive power of activation functions can be done. Variables such as computational cost and analysis time required were excluded from the purpose of this research. Therefore, further studies that compare LR and ANNs in terms of analysis time and required computational power are important in determining which method is more efficient to solve the same problem. It will also be useful to consider this finding in the studies to be carried out to increase student success. While developing policies on school and education system, it is especially important to pay attention to student responsibilities and effective teaching techniques, teacher competencies and classroom management.

Acknowledgements or Notes

References

Abraham, A. (2005). Artificial Neural Networks. In P.H. Sydenham and R. Thorn (Eds.), *Handbook of Measuring System Design* (pp. 901-908). John Wiley & Sons. doi:10.1002/0471497398.mm421

Acar Boyacıoğlu, M. & Kara, Y. (2007). A performance comparison between artificial neural networks and multivariate statistical methods in forecasting financial strength rating in Turkish banking sector. *Dokuz Eylül University Faculty of Economics and Administrative Sciences Journal*, *22*(2), 197-217

Aggarwal, C. C. (2018). *Neural networks and deep learning*. Switzerland: Springer International Publishing AG

Akı, K. & Karasulu, B. (2014). A review on the approaches and systems based on neuro fuzzy synergism in the soft computing (review). *Cumhuriyet University Faculty of Science Journal (CSJ)*, *35*(2), 54-86

Alkan, A., Koklukaya, E. & Subaşı, A. (2005). Automatic seizure detection in EEG using logistic regression and artificial neural network. *Journal of Neuroscience Methods*, *148*, 167–176.

Alpar, R. (2011). *Uygulamalı Çok Değişkenli İstatistiksel Teknikler [Application of Multivariate Statistics]*. Ankara: Detay Yayıncılık.

Assi, K. J., Nahiduzzaman, K. M., Ratrout, N. T. & Aldosary, A. S. (2018). Mode choice behavior of high school goers: Evaluating logistic regression and MLP neural networks. *Case Studies on Transport Policy*. doi: 10.1016/j.cstp.2018.04.006

Aydoğan, İ. & Zırhlıoğlu, G. (2018). Estimation of Student Successes by Artificial Neural Networks. *Yüzüncü Yıl University Journal of Education Faculty*, *15*(1), 577-610.

Bahadır, E. (2016). Using neural network and logistic regression analysis to predict prospective mathematics teachers' academic success upon entering graduate education. *Educational Sciences: Theory & Practice*, *16*(3), 943-964

Benzer, S. & Benzer, R. (2017). Examination of international PISA test results with artificial neural networks and regression methods. *The Journal of Defense Sciences, 16*(2).

Bezek Güre, Ö., Kayri, M. & Erdoğan, F. (2019). Predicting factors affecting PISA 2015 mathematics literacy via radial basis function artificial neural network. *Journal of Engineering and Technology, 3*(1), 1-11

Biganzoli, E., Boracchi, P., Mariani, L. & Marubini, E. (1998). Feed forward neural networks for the analysis of censored survival data: A partial logistic regression approach. *Statistics in Medicine, 17*, 1169 – 1186

Budak, H. & Erpolat, S. (2012). Comparison of artificial neural networks and logistic regression analysis in credit risk estimation. *AJIT-e: Online Academic Journal of Information Technology*, *3*(9), 24–30. doi: 10.5824/1309-1581.2012.4.002.x

Cantor, S.B. & Kattan, M.W. (2000). Determining the area under the ROC curve for a binary diagnostic test. *Medical Decision Making, 20,* 468-470. doi: 10.1177/0272989X0002000410

Chiang, W.K., Zhang, D. & Zhou L. (2006). Predicting and explaining patronage behavior toward web and traditional stores using neural networks: a comparative analysis with logistic regression. *Science Direct Decision Support Systems*, *41*, 514 – 531. doi:10.1016/j.dss.2004.08.016

Çakın, E. & Özdemir, A. (2019). Estimation of the efficiency of technology development regions by artificial neural networks and logistics regression analysis on the basis of data envelopment analysis. *Hacettepe University Journal of Economics and Administrative Sciences*, *37*(2), 271-293

Çırak, G. & Çokluk, Ö. (2013). The usage of artificial neural network and logistic regression methods in the classification of student achievement in higher education. *Mediterranean Journal of Humanities*, *3*(2), 71-79

Çokluk, Ö. (2010). Lojistik regresyon analizi: Kavram ve uygulama. [Logistic regression analysis: Theory and practice]. *Educational Sciences: Theory & Practice*, *10*(3), 1357-1407

Delen, D. (2012). Predicting student attrition with data mining methods, *J. College Student Retention*, *13*(1), 17-35

DeLong, E., DeLong, D., & Clarke-Pearson, D. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics, 44*(3), 837-845. doi:10.2307/2531595

Dreiseitla, S. & Ohno-Machadob L. (2002). Logistic regression and artificial neural network classification models: A methodology review. *Journal of Biomedical Informatics*, *35*, 352–359

Eftekhar, B., Mohammad, K., Ardebili, H.E., Ghodsi, M. & Ketabchi E. (2005). Comparison of artificial neural network and logistic regression models for prediction of mortality in head trauma based on initial clinical data. *BMC Medical Informatics and Decision Making*, *5*(3). doi:10.1186/1472-6947-5-3

Elmas, Ç. (2018). *Yapay zekâ uygulamaları [Applications of artificial intelligence]*. Ankara: Seçkin Yayıncılık

Flitman, A. M. (1997). Towards analyzing student failures: Neural networks compared with regression analysis and multiple discriminant analysis. *Computers Ops Res*, *24*(4), 367-377

Fraenkel, J.R., Wallen, N.E., & Hyun, H.E. (2012). *How to Design and Evaluate Research in Education.* New York: McGraw Hill.

Gorr, W.L., Nagin, D. & Szczypula, J. (1994). Comparative study of artificial neural network and statistical models for predicting student grade point averages. *International Journal of Forecasting*, *10*, 17-34

Graupe, D. (2013). *Principles of artificial neural networks*. Singapore: World Scientific Publishing Co. Pte. Ltd.

Hair,J.F., Black, W.C., Babin, B.J. & Anderson, R.E. (2014). *Multivariate Data Analysis.* Essex: Pearson.

Hardgrave, B. C., Wilson, R. L. & Walstrom, K. A. (1994). Predicting graduate student success: a comparison of neural networks and traditional techniques. *Computers Ops Res*, *21*(3), 249-263

Heaton, J. (2012). *Introduction to the math of neural networks*. Heaton Research. Inc, Chesterfield

Henseler, J. (2011). Back propagation. In P.J. Braspenning, F. Thuijsman, & A.J.M.M. Weijters (Eds.), *Artificial Neural Networks* (pp. 37-66). Springer-Verlag Berlin Heidelberg

Hristev, R.M. (1998). *The ANN Book.* Retrieved from http://www.pdg.cnb.uam.es/cursos/Complutense/ Complutense2004/pages/12_NeuralNetworks/Hritsev_The_ANN_Book.pdf

Ibrahim, Z. & Rusli, D. (2007). Predicting students' academic performance: comparing artificial neural network, decision tree and linear regression. *21st Annual SAS Malaysia Forum,* September 2007, Shangri-La Hotel, Kuala Lumpur

İnal, H. & Turabik, T. (2017). Determination of predictive power of some factors affecting mathematics achievement via artificial neural networks. *Uşak University Journal of Educational Research*, *3*(1), 23-50

Jain, A. K., Jianchang Mao, & Mohiuddin, K. M. (1996). Artificial neural networks: a tutorial. *Computer, 29*(3), 31–44. doi:10.1109/2.485891

Jefferson, M. F., Pendleton, N., Lucas, S.B. & Horan, M. A. (1997). Comparison of a genetic algorithm neural network with logistic regression for predicting outcome after surgery for patients with non-small cell lung carcinoma. *Cancer, 79*, 1338–1342.

Kayri, M. & Çokluk, Ö. (2010). Using multinomial logistic regression analysis in artificial neural network: an application. *Ozean Journal of Applied Science, 3*(2), 259 – 268

Keskin Benli, Y. (2005). Bankalarda mali başarısızlığın öngörülmesi lojistik regresyon ve yapay sinir ağı karşılaştırması. [Forecasting financial failure in banks: A comparison of logistic regression and artificial neural networks] *The Journal of Industrial Arts Education Faculty of Gazi University, 16*, 31-46

Kim, P. (2017). *MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence.* Apress.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York: The Guildford Press

Kurban, M., Kantar, Y., & Hocaoğlu, F. O. (2007). Analysis of the power generation states of wind-solar power plant model by using logistic regression and perceptron models. *Süleyman Demirel University Journal of Natural and Applied Sciences*, *11*(1), 91-95

Kurt, I., Ture, M., & Kurum, A. T. (2008). Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. *Expert Systems with Applications*, *34*, 366 – 374

Manel, S., Dias, J. M. & Ormerod, S. J. (1999). Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: A case study with a Himalayan river bird. *Ecological Modelling*, *120*, 337 – 347

Nefeslioglu, H. A., Gokceoglu, C., & Sonmez, H. (2008). An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps. *Engineering Geology*, *97*, 171–191

Oğuzlar, A. (2005). Lojistik regresyon analizi yardımıyla suçlu profilinin belirlenmesi [Determination of criminal profile with the help of logistic regression analysis]. *International Journal of Economics and Administrative Sciences, 19*(1), 21-34

Organisation for Economic Co-operation and Development [OECD]. (2018). *PISA 2015 PISA Results in Focus*. doi: https://doi.org/10.1787/aa9237e6-en

Öztemel, E. (2006). *Yapay Sinir Ağları [Artificial Neural Networks]*. İstanbul: Papatya Yayıncılık

Paliwal, M. & Kumar, U. A. (2009). A study of academic performance of business school graduates using neural network and statistical techniques. *Expert Systems with Applications*, *36*, 7865–7872

Patterson, J. & Gibson, A. (2017). *Deep Learning a Practitioner's Approach*. Sebastopol: O'Reilly Media

Pavlekovic, M., Bensic M. & Zekic-Susac M. (2010). Modeling children's mathematical gift by neural networks and logistic regression. *Expert Systems with Applications*, *37*, 7167–7173

Pradhan, B. & Lee, S. (2010). Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. *Environmental Modelling & Software*, *25*, 747–759

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Radi, A. & Hindawi, S.K. (2013). Applying artificial neural network hadron – Hadron collisions at LHC. In K. Suzuki (Ed.) *Artificial Neural Networks: Architectures and Applications* (pp. 183-201). Rijeka: InTech.

Rençber, Ö. F. (2018). *Sınıflandırma Problemlerinde Çoklu Lojistik Regresyon, Yapay Sinir Ağları ve ANFIS Yöntemlerinin Karşılaştırılması: İnsani gelişmişlik endeksi üzerine uygulama [Comparison of Multinomial Logistic Regression, Artificial Neural Network, and ANFIS Methods in Classification Problems: An application on human development index]*. Ankara: Gazi Kitapevi

Republic of Turkey Ministry of National Education [MEB], (2016), *PISA 2015 Ulusal Raporu [PISA 2015 National Report]*, Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü, Ankara.

Retherford, R. D. & Choe M. K. (1993). *Statistical Models for Causal Analysis*. New York: A Wiley-Interscience Publication

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*, 77. doi: 10.1186/1471-2105-12-77

Sarıkaya G. (2019). Forecasting Bist national-100 index by using artificial neural network and regression models. *Education and Society in the 21st Century*, *8*(23), 325-340

Shanmuganathan, S. (2016). Artificial neural network modeling: An introduction. In S. Shanmuganathan & S. Samarasinghe (Eds.) *Artificial Neural Network Modeling* (pp. 1-14). Switzerland: Springer.

Subasi, A. & Erçelebi, E. (2005). Classification of EEG signals using neural network and logistic regression. *Computer Methods and Programs in Biomedicine*, *78*, 87-99

Tabachnick, B.G. & Fidell, L.S. (2013). *Using Multivariate Statistics.* Pearson: Boston.

Teshnizi, S. H. & Ayatollahi, S. M. T. (2015). A comparison of logistic regression model and artificial neural networks in predicting of student's academic failure. Acta Inform Med, 23:5, 296-300

The Jamovi Project (2019). *Jamovi (Version 1.0)* [Computer Software].

Tolon, M. & Güneri Tosunoğlu, N. (2008). Analyzing consumer satisfaction data: A comparison of artificial neural networks and regression analysis. *Gazi University Journal of the Faculty of Economics and Administrative Sciences*, *10*(2), 247-259

Tosunoğlu, N. & Apaydin, A. (2004). Öğrenci başarılarının sınıflandırılmasında lojistik regresyon analizi ve sinir ağları yaklaşımı [Logistic regression and artificial neural networks on the classification of student success]. *The Journal of Gazi University Faculty of Tourism. 1*, 170-188.

Tu, V. J. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemid*, *49*(11), 1225 -1231

Vasconcelos, M. J. P., Sllva, S., Tome, M., Alvim, M. & Perelra J. M. C. (2001). Spatial prediction of fire ignition probabilities: Comparing logistic regression and neural networks. *Photogrammetric Engineering & Remote Sensing*, *67*(1), 73-81.

Veelenturf, L.P.J. (1995). *Analysis and Applications of Artificial Neural Networks.* London: Prentice Hall.

Walczak, S. & Sincich, T. (1999). A comparative analysis of regression and neural networks for university admissions. *Information Sciences*, *119*, 1-20

Yazıcı, M. (2018). Comparison of discriminant analysis, logistic regression, and artificial neural networks in credit risk analysis. *Journal of Finance Letters*, *109*, 91-106

Yesilnacar, E. & Topal, T. (2005). Landslide susceptibility mapping: A comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). *Engineering Geology*, *79*, 251–266

Mentoring Inservice Teachers to Support their Inclusive Science Teaching Practices for Students with Visual Impairment

Gulistan Yalcin[1], Tugba Kamali Arslantas[1]
[1]Aksaray University

To cite this article:

Mentoring Inservice Teachers to Support their Inclusive Science Teaching Practices for Students with Visual Impairment

Gulistan Yalcin[1*], Tugba Kamali Arslantas[1]
[1]Aksaray University

Abstract

The Accessible Science for Students with Visual Impairment (ASVI) mentoring program aimed at developing innovative teaching methods for teaching science effectively to third-grade and fourth-grade students with visual impairment (VI). In order to achieve this, the program aimed to guide classroom or science teachers to develop or adapt instructional materials based on the objectives of the science class and also the needs of students with VI. The study was conducted during 2019-2020 at Aksaray University, in Turkey, with the participation of 10 faculty members (as the mentors) and 23 inservice teachers (as the mentees). The main purpose of the study was to investigate the inservice teachers' perceptions (as mentees) about the faculty members' mentoring roles, and to understand the mentors' self-perceptions. This research was designed as a qualitative case study. The study's findings revealed that mentors as the focus of the current study, they had the opportunity to implement procedures based on effective mentoring, and were thereby able to help the inservice teachers develop professionally in their preparations. Mentoring was highlighted as a means of overcoming some of the problems that the inservice teachers' faced in their educating of students with special needs.

Keywords: Inclusive education, Visually impaired, Mentoring, Inservice teacher education

Introduction

Whilst the basis for inclusion practices was enacted in the United States of America in 1975 under Public Law (PL) 94-142, inclusion practices in Turkey were initiated much later, in 1997, with Delegated Legislation No. 573 (Batu & Kırcaali-İftar, 2005). Fairchild and Henson (1993) defined inclusion practice as keeping students with special needs together with their typically developing peers as much as possible, and providing an educational environment which best meets their needs with minimal limitations. This practice has been defined in the Turkish Special Education Services Regulation (Milli Eğitim Bakanlığı [Ministry of National Education], 2018) as maintaining interaction between students with special needs and their typically developing peers at the same level, and providing the necessary support in order to ensure that they receive education alongside their peers.

According to the 2018-2019 statistical data from the Turkish Ministry of National Education (Milli Eğitim Bakanlığı [Turkish Ministry of National Education], 2020), there are 1,260 inclusion students in formal preschool education, 115,556 inclusion students in primary education, and 130,624 inclusion students in secondary education throughout Turkey. Although the exact number of students with visual impairment (VI) in inclusive classrooms is unknown, the number of students included in inclusion practices in Turkey has increased. The biggest reason for this increase is that most schools dedicated to VI have ceased to offer the accommodation services previously offered in the past. Although the educational rights of students with VI is protected by law in most countries, Turkey included, those rights are unlikely to be upheld fully due to several reasons. It has been similarly reported in the literature that many obstacles exist for students with VI attending classrooms based on inclusion practices (Bardin & Lewis, 2008; Gray, 2009; Metatla, 2017; Morelle, 2016; Morelle & Tabane, 2019; Okonkwo, Fajonyomi, Omotosho Esere, & Olawuyi, 2017; Ramrathan & Mzimela, 2016). Some of these obstacles are a lack of educational services support aimed at classroom teachers, their access to educational materials with Braille or large-font text, and limited access to specialized equipment and assistive technological tools. Also, teachers working with students with VI are not equipped with

---

adequate knowledge or skills in the education of students with VI (Porter & Lacey, 2008; Smith, Kelley, Maushak, Griffin-Shirley, & Lan, 2009).

Background

Students with VI face numerous academic challenges in accessing the required information in their educational courses due to high levels of visual content, mostly in science and mathematics courses when compared to their sighted peers. There are three main reasons for the academic challenges that students with VI experience. First, visual signs do not convey meaning to students with VI, mostly because they do not have preformed mental images with which to refer back to like their sighted peers (Jones, Minogue, Oppewal, Cook, & Broadwell, 2006). In addition, students with VI are less ready for science-based courses (Darrah, 2013; Kolitsky, 2014), and therefore experience certain difficulties in these classes. Second, two-dimensional images of three-dimensional concepts may prove difficult for students with VI to interpret (Bogner, Wentworth, Ristvey, Yanow, & Wiens, 2006). Third, information provided within a single diagram may require various tactile graphics and verbal descriptions in order to be understood by students with VI (Bogner, Hurd, Wentworth, Ristvey, & Arens, 2011). Therefore, the use of assistive technologies, actual physical objects and real-time talking about real experiences is of vital importance for the students with VI in science-based courses (Aykut & Özmen-Güzel, 2010; Hasper et al., 2015; Rule, Stefanich, Boody, & Peiffer, 2011).

In order to deal with the challenges that students with VI face in science education, there are certain critical issues that should be provided for such as individualized instruction, accessibility, material adoption, and the use of assistive technologies. Students with VI may require individualized learning strategies (Kamali Arslantas, Yildirim, & Altunay Arslantekin, 2019; Koenig & Holbrook, 2000; Lohmeier, 2009) that enable them to discover through the use of their other senses as well as their remaining vision, if any, and proper educational materials developed according to their needs (Altunay Arslantekin, 2012; Yalcin & Altunay Arslantekin, 2019). It is therefore necessary to meet the needs of students with VI in their learning activities in order to ensure that they face similar conditions to their sighted peers. To this end, Expanded Core Curriculum, as developed by Hatlen in 1996 and adopted as a popular approach in the education of students with VI, meets the needs of students with VI in both their academic and daily life. The curriculum focuses on teaching compensatory skills to students with VI so as to enable them to become involved in each phase of their education (Lohmeier, 2009; Sapp & Hatlen, 2010), and thereby enjoy equal educational opportunities with their sighted peers.

Science courses should be accessible to all students, whether disabled or able-bodied (Atika, Ediyanto, & Kawai, 2018; Ediyanto, Atika, Hayashida, & Kawai, 2017). Additionally, science courses are expected to prepare students to enter further or higher education, business life, and also for social life (Mundilarto, 2002). However, experts have stated that students with VI do not succeed in science courses to the same level as their sighted peers (Yalcin, 2020). In order to be successful in science courses, students with VI require certain arrangements to their educational environment, as well as particular materials made available within inclusion classrooms and other school environments according to their individual needs (Altunay Arslantekin, 2012; Rule, 2011; Toenders, de Putter-Smits, Sanders, & den Brok, 2017). Şahin and Yorek (2009) conducted interviews with science teachers in order to learn about their experiences with students with VI, with some teachers having stated that students with VI need more time throughout the learning process than their sighted peers, and that they need adaptation of certain materials such as tactile learning materials. In his study Rule (2011), prepared tactile materials in order to teach earth and space concepts and conducted instructions with those materials. He figured out that these tactile materials are effective in learning of students with VI. Similarly, Toenders et al., (2017) stated that adapting physics materials based on the needs of students with VI, it is possible to ensure students with VI can access and understand physics concepts.

Students with VI may require materials that have been prepared in different formats for science classes (Jones & Broadwell, 2008), such as a Braille barometer, Braille thermometer, Braille-labelled laboratory equipment, a human anatomical model, three-dimensional (3D) materials, 3D models of cells and DNA, and a tactile anatomy atlas in their science classes (Aslan, 2016). Although many of these materials are supplied from the Directorate General of the State Supply Office in Turkey, it is known that teachers have limited access to them. For this reason, many teachers are expected to adapt their normal teaching materials that they use with their sighted students according to the needs of students with VI (Yalcin, 2020).

With the improvements seen in technology, it has become more possible to create the required methodological and pedagogical conditions for students with VI. Technology usage has been introduced with the Expanded Core Curriculum, and the use of assistive technology has been shown to be highly effective with VI students in

accessing information and in the pursuit of an independent daily life (Altunay Arslantekin, 2012; Yalcin & Altunay Arslantekin, 2019). Studies in the literature support the effectiveness of assistive technology application in teaching science and mathematics since it ensures the active participation of students with VI in classes (da Mota Silveira & Martini, 2017; Koehler, Wild, & Tikkun, 2018; Ludi, Canter, Ellis, & Shrestha, 2012; Lunney, 1995; Nees & Berry, 2013; Negrete, Lisboa, Peña, Dib, & Vargas, 2020; Supalo, Humphrey, Mallouk, Wohlers, & Carlsen, 2016). Supalo et al. (2016) reported that students with VI can effectively participate in technology-supported science laboratories. Similarly, Koehler et al. (2018) used 3D printed models in their science class for students with VI, and indicated that adaptations in the curriculum can be effective means for teaching students with VI. However, the use of assistive technology in the education of students with VI requires comprehensive teacher training. As Abner and Lahm (2002) pointed out, classroom teachers can struggle in using assistive technologies effectively since many are unprepared and are in need of training. However, in a study conducted by Moreland (2015), the findings showed that teachers who were specifically trained in assistive technology usage have more positive attitudes toward technology usage in the classroom.

Full attendance of students with VI in science-based classes depends on adaptations having been applied to the curriculum and teaching strategies, assistive technology usage in class, materials having been developed according to needs of the students with VI, and cooperation between experts in the field and science or classroom teachers (Yalcin, 2020). The literature has shown that students with VI can learn science subjects to the same level as their sighted peers if certain changes are applied to the teaching and classroom materials (Betts & Cross, 2010; Fraser & Maguvhe, 2008; Kızılaslan, Zorluoğlu, & Sözbilir, 2020; Şahin & Yorek, 2009; Urquhart, 2012).

One of the most important elements in inclusion practices are the teachers themselves. It is expected that teachers should possess adequate knowledge about the learning characteristics of the inclusion student in the classroom, and to be able to organize their teaching methods in line with their students' needs. However, it is also known that classroom and science teachers have limited information and skills regarding the education of students with VI. Therefore, the literature has emphasized that, for an effective inclusion practice, supportive educational services should be provided by experts in the form of training these teachers (Fraser & Maguvhe, 2008; Koehler & Wild, 2019; Rosenblum, Ristvey, & Hospitál, 2019). Koehler and Wild (2019) defined the content of support offered by experts to classroom teachers in six steps in order to better support students with VI in their science-based classes. These steps are to guide classroom teachers on how to ensure student participation in the classroom, to adapt materials and provide the necessary support according to the students' needs, to ensure that students can readily access the course, and to provide the necessary prerequisite skills to students prior to them having to learn the science, to guide the teacher on how best to evaluate students with VI, and being a model for the classroom teacher by teaching in small groups. In the study by Rule et al. (2011), the results of the education program given to secondary school teachers on attitudes towards students with VI in science and mathematics classes and material adaptations were examined. The findings revealed that when materials (tactile, auditory) were prepared based on students' needs, students with VI were as successful in lessons as their sighted peers; whilst teachers who acquired the necessary knowledge and skills through working with students with VI found themselves ready to work with these students.

Mentoring as part of inservice teacher education is one of the strategies that can be applied in order to support teachers' skills in learning to teach (Wang, 2002). Notably, the literature states that mentoring contributes to teachers' professional and personal growth (Ambrosetti & Dekkers, 2010; Hudson, 2004; Mena, Hennissen, & Loughran 2017). In the current study, mentoring was implemented as an intervention in order to contribute to the teachers' inclusive science teaching practices.

Mentoring

In the literature, mentoring has been similarly defined by several researchers. Van Dijk (2008) defined mentoring as the process of transferring essential job-related skills, attitudes, and behaviors from one person to another. Godshalk and Sosik (2003) and Luecke (2004) defined mentorship as a relationship in which a more skilled and experienced person helps to inform and guide a less knowledgeable and less skilled person. Accordingly, the purpose of mentoring is that a mentee interprets the knowledge and experiences transferred from a mentor to contribute to their own self-development (Klasen & Clutterbuck, 2002). In the current form, mentoring is a non-hierarchical process in which both mentor and mentee can benefit (Ambrosetti & Dekkers, 2010) since both parties can learn from the process and experience personal gain (Forde & O'Brien, 2011; Hudson, 2013).

There are two important elements within the mentoring process, that of the "mentor" and the "mentee." A "mentor" is someone with a greater level of experience, whose role it is to guide, support, and nurture a less experienced person, who is their "mentee" (Truter, 2008). Certain qualities that a mentor should possess have been previously defined by researchers in the literature (Janse van Rensburg & Roodt, 2005; Meyer & Fourie, 2004; Truter, 2008), including being reliable, honest, acting with integrity within a group, having emotional intelligence and the ability to understand others, possessing strong social skills, and is both patient and understanding. In addition, mentors are expected to possess skills on the specific subject/s that they are providing mentoring for (Meyer & Fourie, 2004; Truter, 2008). Some of these skills are: (a) Having adequate knowledge and experience in the field being mentored; (b) Interested in self-development; (c) Willing to share their own experiences; (d) Having the capability to ask questions in a proper and correct manner; (e) Being a good listener; (f) Being able to help their mentee to manage the knowledge that they gain, and (g) Being an effective role-model so as to help their mentee gain the relevant skills.

The mentee is another important element of the mentoring process, and is defined as a person who is in receipt of mentoring from a mentor. Meyer and Fourie (2004) defined a mentee as a person who is responsible for their own self-development, whilst aware of the importance of learning from others. In the literature, researchers have underlined certain qualities that a mentee should possess in order to benefit from the mentoring process, with Clutterbuck (2005) listing them as good communication skills, being open to new information and experiences, having a sense of trust, willing to communicate with their mentor, being committed to learning, being responsible to both their fellow group members and their mentor, and being open to criticism.

There are certain factors that make for an effective mentoring process, apart from the qualities of both the mentor and mentee/s, including determining the criteria of how mentors and mentees are selected (Conolly & Blunt, 2006; Meyer & Fourie, 2004), describing the goals and desired outcomes expected to be realized during the mentoring process (Conolly & Blunt, 2006), clearly sharing the content and activities related to the mentoring process with the mentees (Meyer & Fourie, 2004), and organizing the setting where process take place. Similarly, the literature revealed that shared decision-making processes (Gabriel & Kaufield, 2008; Kopcha, 2010; Roth & Tobin, 2002), forming equal relationships (Wilkinson et al., 2014), defining mutual expectations and goals (Baran, 2016; Barker, 2006), practicing together as a form of sharing responsibility (Baran, 2016), feedback-oriented communication (Baran, 2016), and being adaptive to mentees' needs (Rajuan, Beijaard, & Verloop, 2010; van Ginkel, Oolbekkink, Meijer, & Verloop, 2016) are all critical elements to the success of mentoring programs. These criteria ensure that both mentee and mentor grow together personally (Baran, 2016; Huizing, 2012; Simpson et al., 2007), professionally (Hudson, 2004, 2013; Lopez-Real & Kwan, 2005; Lumpkin, 2011; Mena et al., 2017; Zachary, 2009), and pedagogically (Baran, 2016; Hudson, 2004). In the study of Baran (2016), it was stated that mentoring programs contribute to the professional growth of mentors by affording them the opportunity to build professional friendships and to improve their communication skills.

As a system, inclusive education has become pervasive, adding additional responsibilities for teachers independent of their specific domain. Professional development of inservice teachers is necessary considering the requirements of special education (SE). In this regard, mentoring has been attributed to being an effective means for inservice teachers to gain the necessary skills for conducting effective education (Hairon et al., 2019; Hobson, Ashby, Malderez, & Tomlinson, 2009; Ingersoll & Strong, 2011; Mathur, Gehrke, & Kim, 2013; S´anchez-García, Marcos, GuanLin, & Escribano, 2013), which also includes the practices of inclusive education (Falvey, Coots, Bishop, & Grenot-Scheyer, 1989; Hobbs & Westling, 2010; Rudiyati, 2014). According to the literature, mentoring practices in SE are found to be effective with teachers (Akçamete, Aslan, & Dinçer, 2010; Boyer & Lee, 2001; Heirdsfield, Walker, Walsh, & Wilss, 2008; Patton et al., 2005; White & Mason, 2006), students (Pleiss & Feldhusen, 1995), as well as for other stakeholders (Moon & Callahan, 2001; Ota & Austin, 2013). In a mentoring study conducted by Rudiyati (2014) with teachers of students with VI, it was shown that mentoring can be effective in increasing teachers' inclusive skills.

The Purpose of the Study

The current study attempts to overcome the problems that students with VI face in science classes, and the challenges that inservice teachers are required to confront in serving to inclusive students, by improving the inclusive teaching practices of inservice teachers. For this purpose, an Accessible Science Project for Students with Visual Impairment (ASVI) mentoring program was conducted. The main purpose of the program was to develop innovative teaching methods for the effective teaching of science to third-grade and fourth-grade students with VI.

In order to achieve this, the ASVI program aimed to guide classroom and science teachers do develop or adapt instructional materials based on the objectives of science classes and the needs of students with VI. Despite the project targeting students with VI, the adopted materials can be applied to students with different types and levels of disability. Thus, the program has the potential to contribute to teachers in preparing materials for students with a variety of disabilities, which in turn affects their inclusive teaching practices. The focus of the current study was generally on the mentors, since the literature emphasizes that many aspects of the mentoring process impacts on the mentors rather than the mentees, which is based on mentors' self-reflections of their own practices (Lopez-Real & Kwan, 2005; Simpson, Hastings, & Hill, 2007).

Method

The main purpose of the current study was to explore inservice teachers' perceptions, as mentees, about faculty members' mentoring roles, and also to understand the mentors' self-perceptions. The study was designed as an exploratory case study that guided an in-depth analysis of both the mentors' and mentees' perceptions of an Accessible Science for Students with Visual Impairment (ASVI) mentoring program. In the context of an exploratory case study, the purpose is to extend understanding of a complex social phenomena in real-life contexts (Ogawa & Malen, 1991). Exploratory case studies are qualitative in nature. Baxter and Jack (2008) indicated that qualitative case study methodology provides tools for researchers to learn about complex phenomena within their specific context. The case in the current study focused on the hands-on material design for a science course designed specifically for VI students, with teachers from different schools participating in the study. Implementing an exploratory case study approach enabled a detailed investigation of the mentoring process through a multiple-case mentor-mentee context (Yıldırım & Şimşek, 2011).

The following research questions have driven the current study:

(1) What are the mentors' perceptions of the ASVI mentoring program?
(2) What are the mentees' perceptions of their mentors in terms of fulfilment of their mentoring roles?

Participants

The participant group of the current study was comprised of both mentors and mentees.

The first group of participants in the current study were the mentors, who were either faculty members of Special Education (n = 8) or Science Education (n = 2) departments at Aksaray University, Turkey. Sampling method was implemented in two steps. First, the mentors were recruited based on purposeful sampling, then, based on the convenience sampling method, mentors who were available to participate were recruited. The two faculty members from the Science Education department were Professors. Three of the faculty members from the Special Education department were Assistant Professors, while the remaining five faculty members were Research Assistants studying for doctoral degrees in SE at various universities in Turkey. The mentors' role was twofold; to mentor inservice teachers in the development and adaptation of science education instructional materials for students with VI, and also to guide their mentees in the construction of knowledge on how to educate students with VI. Also, the mentors introduced assistive technologies to their mentees for use within the project, including the use of Braille printers, 3D printers, 3D pens, and a tactile-copying machine.

The second group of participants were inservice classroom and science teachers working in primary education. At the outset of the study, an announcement was undertaken through the Aksaray Provincial Directorate of National Education to all schools in the Aksaray province. The announcement included information about the project, its procedures, and that the project was looking to recruit volunteer participants for the 1-year research project. Mentees were recruited based on the purposeful sampling method. Based on the selection criteria, teachers who had one or more inclusive students in their class or in their school (i.e., not specifically in their own class) were welcomed to apply to join the project in order to increase their inclusive teaching skills. Considering the increasing inclusive practices in Turkey, there is a high possibility of teachers having an inclusive student in their class in the future. A total of 23 teachers (10 science education and 13 classroom teachers) attended all stages of the project.

The matching of mentors and mentees was applied on a random basis, with five groups formed consisting of two mentors and four or five mentees.

Accessible Science for Students with Visual Impairment (ASVI)

The current study was conducted as part of the "Accessible Science for Students with Visual Impairment" (ASVI) project, which was supported by the Sabancı Grant Foundation 2019 Program for a period of 1 year. The project was conducted at the Aksaray University, in Turkey, with cooperation of the Aksaray Provincial Directorate of National Education.

The current study focused on the ASVI mentoring program which was conducted in two stages, and with four panels and six workshops, however, the current study specifically focuses on the workshops. Four panels were conducted with the teachers at the beginning of the program, during which information about the workshops was provided, the needs and individual characteristics of students with VI were discussed, and the project's details were clarified. The workshops were then conducted following on from the panels. During the workshops, five groups of inservice teachers developed materials under the guidance of two mentors. The output of the workshops was a set of 22 science education materials that had been developed and adapted based on the individual needs of VI students.

Workshops

The workshops were conducted during November 2019 and January 2020, and over a total of 6 days Each workshop started at 8:30 in the morning and lasted until the groups had completed their assigned materials. For groups who could not complete their materials during one day, they continued working with the same material the next day. The critical issue during the material development was consideration of the needs of students with visual impairment, and what kind of modifications they needed in order to access these materials.

Prior to the workshops, the mentors examined the science textbooks and prepared an objective list where students with VI required specific materials. Then, before each workshop, the mentors assigned objectives for each of the groups, with each group responsible for developing different sets of materials. All the required equipment necessary for developing the materials were prepared in advance of each workshop by the mentors, and distributed to the tables where each group would work. All of the workshops were conducted at the university's premises. Depending on the process, some groups were assigned more than one objective and were responsible for the development of more than one material. According to the difficulty associated with creating the materials, some groups continued with the same activity over two days of their workshops. Table 1 presents information about the dates and objectives of the each workshop, whilst Figure 1 presents a visual of one of the materials that was prepared.

Table 1. Workshop dates and objectives for material development

Date	Objective	Group
November 18, 2019	Compares areas covered by land and water on earth's surface on the model.	Group 1
November 18, 2019	Understands the world is made up of layers.	Group 1
November 18, 2019	Prepares a model of the earth.	Group 2
November 18, 2019	Discovers the forces of push and pull.	Group 3
November 18, 2019	Presents observational results of the lifecycle of a plant.	Group 4
November 18, 2019	Explains the basic functions of sense organs.	Group 5
November 20, 2019	Prepares a model of the earth.	Group 2 (continued with same material)
November 20, 2019	Explains what should be done to protect sense organ health.	Group 1
November 20, 2019	Hardness/softness, flexibility, brittleness, color, odor, taste, roughness, and smoothness.	Group 3
November 20, 2019	Conducts experiments to understand magnets.	Group 4
November 20, 2019	Conducts experiments to understand push and pull forces.	Group 5
November 22, 2019	Hardness/softness, flexibility, brittleness, color, odor, taste, roughness, and smoothness.	Group 3 (continued with same material)
November 22, 2019	Conducts experiments to understand push and pull forces.	Group 4 (continued with same material)
November 22, 2019	Conducts experiments to understand push and pull forces.	Group 5 (continued with same material)

Date	Objective	Group
November 22, 2019	Recognizes the elements that constitute a simple electrical circuit and their functions.	Group 1
November 22, 2019	Discovers how light is necessary for vision.	Group 2
January 13, 2020	Conducts experiments to understand push and pull forces.	Group 5 (continued with same material)
January 13, 2020	Classifies sound sources as natural or artificial.	Groups 1 & 2 (two different materials)
January 13, 2020	Relates rocks with minerals and discusses the importance of rocks as raw materials.	Group 3
January 13, 2020	Discusses types of rock and minerals found in Turkey (e.g., gold, boron, marble, lignite, copper, hard coal, silver).	Group 4
January 14, 2020	Classifies sound sources as natural or artificial.	Groups 1 & 2 (two different materials) (continued with same material)
January 14, 2020	Discusses types of rock and minerals found in Turkey (e.g., gold, boron, marble, lignite, copper, hard coal, silver).	Group 4 (continued with same material)
January 14, 2020	Explains the formation of fossils.	Group 3
January 14, 2020	Constructs a working electrical circuit.	Group 5
January 14, 2020	Explains the formation of fossils.	Group 3 (continued with same material)
January 15, 2020	Explains the relationship between sound intensity and distance.	Group 1
January 15, 2020	Discovers that that every sound has a source, and that sounds spread in all directions.	Group 4
January 15, 2020	Using the sense of hearing, makes inferences about the approach and distance of a sound source and its location.	Group 5

Figure 1. A visual of Earth model

Each workshop included three modes of mentoring as "discussion," "material development," and "feedback and reflection." At the beginning of the workshops, discussions were held with regards to the objectives and the potential materials that could be developed based on the characteristics of the target group. Each group undertook a brainstorming activity and then decided on the specific features of the materials they would develop prior to commencing the material development process. The groups each worked collaboratively throughout the process. After completion of each item of instructional material, feedback and reflection sessions were then conducted. The mentors visited each of their groups and provided them with feedback, and discussed the materials that they had developed.

Data Sources

Three data sources were used in the current study, with focus group interview schedules for teachers, semi-structured interview schedules for mentors, and observation forms completed by the mentors.

Focus Group Interview Schedules: At the end of the study, five focus group interviews were conducted with the teachers in order to gather their insights about the mentors. Each focus group interview lasted between 45 and 60 minutes. The focus group interviews were conducted by the current study's researchers, who each possessed prior experience with qualitative studies. The focus group interview guidelines included questions about the participants' general perceptions related to the project, their perceptions about their mentors; specifically, what they thought about their mentors' contribution to the project, their mentoring strategies, and how the mentors affected the effectiveness of the workshops.

Semi-structured interview schedules for mentors: Also, at the end of the study two of the researchers conducted interviews with the faculty members in order to understand their mentoring experiences from a mentors' perspective. Interviews lasted around 30 minutes with each mentor. Interview guidelines included questions about their general perceptions related to the project, what they gained from the mentoring process, and in what specific factors they contributed to their mentees.

Observation Forms: The current study included prolonged observation of the mentees during their workshops. Holistic description of the events was undertaken by two researchers, who completed observation forms and also took field notes during the collaborative activities.

Data Analysis

For the qualitative data of the study, thematic analysis was performed in order to describe the phenomenon. Thematic analysis is a flexible qualitative data analysis technique that provides a rich and detailed presentation of the collected data (Braun & Clarke, 2006). First, the interview recordings were transcribed into written form. For data analysis, the step-by-step approach as proposed by Nowell, Norris, White, and Moules (2017) was applied. As the first step, the researchers familiarized themselves with the data (Braun & Clarke, 2006) through "careful reading and re-reading of the data" (Rice & Ezzy, 1999, p. 258). Next, the researchers separately generated initial codes for the data. During their coding, the researchers identified the significant elements of the data and attached labels to index them in relation to themes (King, 2004). The data were organized based on the interview questions, and each of the respondents' answers were categorized based on their consistencies and differences. During the third step, the researchers looked for themes in the data, and then sorted and collated the coded data (Braun & Clarke, 2006) according to a data-driven inductive approach in order to generate the themes (Boyatzis, 1998). During the fourth step, the researchers reviewed the themes so as to explore whether or not they appeared to form a coherent pattern. The researchers then met to examine the themes and subthemes, and worked towards reaching a consensus. For example, some of the themes that had little support were changed. During the fifth step, the themes were defined and named by the researchers, having met to discuss the themes in detail. The themes were the organized repeatedly until consensus was reached. Finally, the researchers wrote the findings of the data. These six steps were implemented for both types of data collected; interviews and observation forms.

Trustworthiness

In order to strengthen the research design, the researcher took issues of the study's trustworthiness (Lincoln & Guba, 1985) into consideration during both the data collection and analysis process. In order to ensure credibility triangulation, prolonged field engagement and peer debriefing were employed. Data triangulation was achieved through having multiple data sources, with interviews and observation used together to validate the results through comparison of the observational data with the interviews. Investigator triangulation was also performed in order to minimize potential researcher bias. For that purpose, both researchers conducted the data collection process. Additionally, the qualitative data were analyzed independently by two researchers and the findings then compared. Prolonged field engagement is another credibility strategy concerning the investment of adequate time within a research site in order to learn the culture, testing for possible misinformation, and for the establishment of trust (Lincoln & Guba, 1985). The researchers in the current study conducted repeated observations and held in-depth interviews; staying for a prolonged period of time at the research site. Peer debriefing was also ensured since the two researchers worked separately at the beginning and then jointly examined the themes until a consensus was reached.

In order to ensure dependability (which is termed as reliability in quantitative studies), both audit trail and intercoder agreement (Silverman, 2000) were applied. Audit trail in the current study was achieved by recording and documenting the whole process in order that an external observer could trace the research on a step-by-step basis. Intercoder agreement or intercoder reliability is a strategy that requires different coders to analyze transcribed data (Creswell, 2008; Miles & Huberman, 1994). The current study did not aim for an exact match in the coding, but for consistency among the codes and categories. For that purpose, both of the researchers worked throughout the data analysis phase.

Results

The findings of the study are presented under subheadings of "Mentors' perceptions" and "Mentees' perceptions" according to the relevant themes.

1) Mentors' Perceptions

During the interview, the mentors were asked to explain their perceptions related to the mentoring program. First, all of the mentors indicated that the program was able to reach and achieve its aims and in line with the goals, and that the mentees achieved the necessary gains in the design and adaptation of instructional materials aimed at inclusive students. In addition, the mentors strongly emphasized that besides designing and adapting materials, the mentees' awareness about disabled students increased. Digital assistive technologies used for students with VI were introduced to the teachers during the process and they were informed about these technologies. Regarding this issue, one the mentors stated that:

> *The most important outcome of the program was the materials themselves since these materials can be used for all students with different disabilities, even with those without a disability. [Mentor 1]*

The mentors viewpoints regarding the mentoring program were clustered around two main themes. First, they expressed their perceptions of their own personal gains, and then explained their perceptions related to the mentoring procedure. Table 2 represents the mentors' perceptions during the process.

Table 2. Mentors' Perceptions

Themes	Sub-themes	Categories
Personal Perceptions		
	Improved Professional Practice	Sharing knowledge and vision, Gaining new academic perspectives
	Effective Collaborative Work	Collaboration, Shared decision-making, Practicing together and reflection
	Building Collegiality	Different profession, Working for a common goal, Relevance of the topic
Procedurel Perceptions		
	Identifying Mutual Expectations	-
	Discussion	-
	Material Development	-
	Feedback and Reflection	-

Personal Perceptions

The first issue mentioned by the mentors was with regards to their personal perceptions toward the mentoring program. They evaluated it as an invaluable opportunity to learn new things and as a means to improving themselves. Having the opportunity to work with teachers and colleagues from two different discipline areas promoted positive outcomes of the study. The mentors reported that it was a reciprocal process in which both

sides of the relationship had learned. The mentors' perceived benefits of the mentoring experience in various practices as "improved professional practice," "effective collaborative work," and "building collegiality."

Improved Professional Practice. The mentors perceived the greatest benefit of the mentoring exercise to be the opportunity to improve their professional practice in terms of sharing their knowledge, gaining new academic perspectives, and improving themselves with regards to project management issues. They indicated that they had the chance to share their knowledge and vision related to inclusive education practices with classroom teachers. As the mentors were mostly new to their academic career, they first considered the opportunity to work with teachers as the actual practitioners, and from whom they learned new things and then shared these ideas with each other. As they significantly emphasized, their experience with the teachers also contributed to them considering to undertake more academic studies in the future. Similarly, conducting the mentoring program as two mentors together provided opportunities to develop abilities to strengthen their capacities professionally, since the mentors were from different academic disciplines. One of the mentors indicated that:

> *Specifically working with the teachers helped to gain new insights, especially about the studies I had conducted in the field. Since the teachers, as actual practitioners, explained the problems they faced related to special education issues, we had the chance to learn more about the actual problems they faced, which in turn helped me to think about future academic studies with the target group. [Mentor 2]*

Effective Collaborative Work. This professional improvement in turn enhanced their ability to nurture collaborative working skills. The mentoring relationship included the factors of collaboration, shared decision-making, practicing together, and also of reflection, which together promoted effective collaborative working. These common shared experiences promoted their effectiveness as a team. Working for a common and important goal was a key motivating factor for the mentors, which in turn increased their keenness to work and collaborate effectively.

Building Collegiality. The other category mentioned was about building collegiality, as during the process the mentors conducted the workshops partnered alongside a second mentor. This was the first such experience for eight of the mentors in the study, which provided them with an invaluable experience. The mentors from different disciplines reflected their knowledge effectively during the mentoring program. A collegial relationship was effectively fostered which was related to the relevance of the study's goals. They indicated that the mentoring program was conducted based on a real need, which increased their motivation as they were working for a common goal. This level of keenness and effectiveness is exemplified in the following quote:

> *As a team, we established collegiality since we believed in the necessity of the program's focus. There was a real solidarity between the mentors, and also between the mentors and mentees. Thus, we created a network of support, and learned a great deal from each other. [Mentor 6]*

To summarize, these positive experiences promoted corresponding changes in the mentors' self-perception of their own personal development and understanding of their value within the workplace.

Procedural Perceptions

The second theme of the mentors' perceptions related to procedural issues. As the mentors evaluated the mentoring program as being a successful and productive process, they mainly related this success to the structure of the mentoring program itself. Thus, their answers to the interview questions were categorized according to the procedural steps as they evaluated the mentoring program based on these stages.

Identifying mutual expectations. Prior to the workshops, the mentors held discussions with their mentees during the panels. These discussions primarily concerned the topics of SE and the problems faced by and needs of students with VI. During these panels, the mentors and mentees also discussed potential solutions that could be enhanced for the education of students with VI. This ensured that both sides' expectations related to the outcomes of the workshops were defined, which corresponded to the main theme of the project being that the process should be personally relevant to both parties. The participants raised similar ideas related to the mentoring program, as well as how to conduct the process, which also ensured active participation of the mentors.

> *From the beginning of the material development process, discussing problems related to the education of the VI as well as potential solutions helped to increase the sense of belonging of us all. In this way, we had the chance to work towards a common goal that was relevant to real-life experience. As I observed, we all worked with a high level of motivation in order to successfully reach our determined goals. [Mentor 2]*

Discussion. Before each workshop, the groups discussed the objectives that had been assigned to them, and then brainstormed about what could be designed and how it should be designed in order to teach that specific subject to students with VI. These discussion sessions ensured that both the mentees and their mentors took part in the research together, which helped promote the establishment of a co-equal working relationship. Also, these discussions held prior to the actual practical material development provided the teachers with an opportunity to learn how to think and what to think while considering the individual differences of students with VI and their individualized needs for the learning of a new subject. Thus, the teachers learned to think of ways to produce better outcomes for their students. Conducting effective discussions with the involvement of all the participants was also found to be essential to the effectiveness of the process.

> *During the discussion sessions, the teachers' awareness of students with VI and their specific needs increased. The teachers held discussions in groups and spent effort to find solutions for an effective material development. These discussions contributed to both their personal development and to the material development process. [Mentor 4]*

Material Development. The mentors strongly emphasized that the hands-on material development process was conducted very effectively, having brainstormed every step as a team, and reached common ideas within the group. The mentors attributed the success of the material development process to the relevance of the study, and to its applicability to a real-life application based on real experiences. The applicability of acquired knowledge and practices was articulated.

Feedback and Reflection. After each material development process had concluded, each mentor pair visited all of their groups and provided process feedback at the group level. Process feedback included task-related behaviors and focused on the convenience of the materials to students with VI. Additionally, during this session, the mentees were required to reflect on their opinions as related to the feedback and what they had learned following the materials development process.

2) Mentees' Perceptions

During the interviews, the mentees were asked to evaluate their perceptions in terms of their mentors. As the mentees highlighted, the mentors, as faculty members, had a powerful role in influencing and shaping the inclusive teaching practices of inservice teachers. Similarly, the mentees stated that the mentoring program provided an opportunity for their professional renewal, and helped to advance their knowledge with regards to inclusive teaching practices. Thus, the mentees found the mentorship experience to be beneficial in increasing their knowledge related to the education of students with VI. Moreover, they added that they might require ongoing support during their actual teaching practices in the field. The mentees' perceptions clustered around four themes; namely "endorsing the setting up of co-equal relationships," "adapting the mentoring program to the mentees' needs and suggestions," "being tech-savvy," and "being proactive." Table 3 presents an overview of the mentees' perceptions.

Table 3. Mentees' Perceptions

Themes	Sub-themes
Endorsing the setting up of co-equal relationships	Mutual respect, Trust, Equal balance of power
Adapting the mentoring program to the mentees' needs and suggestions	Construction of practical knowledge, Workshop days and duration
Being tech-savvy	Effective assistive technology use, Creating assistive technology awareness and familiarity
Being proactive	Highly focusing on the big picture of inclusive education, Foreseeing possible problems

Endorsing the setting up of co-equal relationships

The most strongly emphasized issue by the mentees was the mentors' effectiveness in promoting the establishment of a co-equal relationship with their mentees, which was characterized by mutual respect, trust and equal balance of power. The mentees, as teachers, indicated that at first they were biased against the university faculty members, having had concerns that the mentors would try to direct them and manage the whole process based on their own preferences. However, all of the focus groups stated that the mentors were very effective in managing the process with respect and trust, and were able to promote a co-equal relationship with the teachers they were mentoring. The mentees stated that they felt that they were treated as equals with the mentors, which ensured their keenness to join in throughout the program. The mentors both began with and maintained an equality-centered paradigm in which all parties established a partnership which nurtured a reciprocal relationship. This equal balance of power ensured the development of a true partnership. On this, one of the groups stated that:

> *The way the mentors valued the teachers and the way they looked into the eyes of the teachers was so nice. It is the mentors who brought us to this point, through their warm behavior and cordiality. To be honest, if the mentors had not worked with us as equals, we would have stepped back. The mentors' efforts during the process motivated us; otherwise, we would have thought that we had got out of school voluntarily to spend time on the project, but question why the mentors were not working with us. However, the mentors worked equally and enthusiastically with us. [Focus Group 3]*

Similarly, a member of the same group also mentioned that:

> *When I came here, there was a very warm environment and everyone's opinions were taken as equal. My opinion towards academics has changed in a positive way, I can really say that. Before, I was very prejudiced against academicians, but here the mentors valued all of us. [Focus Group 3]*

Adapting the mentoring program to the mentees' needs and suggestions

Adapting the mentoring program was the second-most mentioned theme regarding the mentors. The mentees indicated that the mentors' being adaptive to their needs ensured an effective mentoring process. Especially, the mentors were reported as being able to adapt the process concerning the construction of practical knowledge in materials development for students with VI. They shaped the process depending on the mentees' needs and perceptions. When the mentees had different ideas, the mentors tried to implement them. The second issue mentioned was about adopting the workshop schedule and duration based on the convenience of the mentees. Each time when they were planning the process, the views and opinions of the mentees were given due consideration. Thus, aligning the mentoring program with the expectations and opinions of the mentees ensured the mentees desire to be active participants in the mentoring program, and thereby to exhibit a positive attitude with the mentors.

Being tech-savvy

The mentees also reflected their perceptions regarding assistive technology usage during the study. They indicated that the mentors were technologically competent, or "tech-savvy," and were able to effectively integrate assistive technologies during the materials development process, which provided additional insights for the mentees. As the mentees strongly emphasized, building technology awareness and familiarity during the hands-on material development process proved to be an invaluable experience. Thus, the mentees admired the mentors' tech-savviness which contributed to the mentees reconsidering their own teaching practices so as to produce better outcomes for students with VI in their classes, as exemplified in the following quote:

> *We really appreciate the technological assistance provided by the mentors. We had the chance to become familiar with a variety of technologies, and noticed how practical they were. For example, the Braille writer was really easy to use. Even though we did not know about the existence of such kinds of technology beforehand, we noticed how easy it was to use. The 3D pens we learned can be used in classes to make the learning environment more attractive and motivating. The mentors were really keen on being portrayed as tech-savvy and modern instructors. [Focus Group 1]*

Being proactive

The proactivity of the mentors was also articulated by the mentees as one of the things they admired about the mentors. The mentees stated that the mentors were highly focused on the bigger picture throughout the program of delivering inclusive education, which increased their proactivity in foreseeing the problems that we might encounter. This issue positively affected the efficiency of the mentoring process.

Discussion and Implications

The current study aimed to contribute to the literature on inclusive education by conducting an Accessible Science for Students with Visual Impairment (ASVI) mentoring program, highlighting both the self-perceptions of mentors and of the procedure itself; as well as the mentees' perceptions related to their mentors. The study's findings suggest that such kinds of mentoring studies that promote partnerships between universities and central education authorities should be conducted more, which in turn could result in positive outcomes for both faculty members and teachers alike. As is frequently noted in the literature, effective mentoring studies within an inservice context (Mathur et al., 2013; S´anchez-García et al., 2013) should ensure that both mentors and mentees learn together (Haggarty, 1995; Halai, 2006; Hobson et al., 2009) since the tendency is to aim for a sense of communal development (Forde & O'Brien, 2011).

The findings of the current study focused more on the "mentors," presenting the participants' perspectives related to the mentors and their mentoring aspects. The mentors first reflected on their personal gains, indicating that the mentoring program was personally rewarding (Huizing, 2012; Simpson et al., 2007) since such positive gains would benefit them throughout their academic careers. The mentors reported that mentoring was perceived as beneficial in terms of professional practice, collaborative working, and in the building of collegiality.

Immersing faculty members with inservice teachers within the ASVI program helped the faculty members to improve their professional practices in terms of sharing knowledge and their vision related to inclusive education, as well as gaining new academic perspectives. These current findings contribute to the literature, which also shows that mentoring has been widely recognized as an important aspect of gaining in professional practice for mentors (Hagger & McIntyre, 2006; Hudson, 2013; Lopez-Real & Kwan, 2005; Zachary, 2009). One critical output of the current study relates to the mentors' experience in sharing knowledge and vision reciprocally with teachers. Prior to the study, the faculty members had no experience of personal reflection from working with teachers as part of a long-term project. This was also found to have been discussed in the literature, since mentoring affords significant opportunities for the sharing of experiences, perspectives, and enthusiasm about various themes based on expertise (St-Jean & Audet, 2009). During this sharing process, the mentors benefitted from the experiences of their mentees and by learning more about the problems that they faced in the classroom. The faculty members also had the chance to gain new perspectives, which might help shape their future academic studies. However, this finding might be considered as specific to the current study, since the mentors were mostly new to their academic careers and therefore more open to new perspectives. Similarly, some studies in the literature highlighted certain benefits of mentoring for mentors in terms of

advancing their pedagogical knowledge (Hudson, 2013), which can be associated with the current study's results.

The other positive gain for the mentors was improving their effective collaborative working skills. The study's findings showed that improving the effective collaborative working skills within a mentoring program especially conducted between faculty members and teachers requires shared decision-making and joint practicing. Research studies have indicated that shared decision-making processes (Baran, 2016; Gabriel & Kaufield, 2008; Kopcha, 2010; Roth &Tobin, 2002) and practicing together as a form of sharing responsibility (Baran, 2016) are critical to the success of mentoring programs, and which is also an indicator of effective collaborative working. The current study's findings also support that effective collaborative working increases the motivation of both stakeholder parties due to working towards a common and relevant goal.

The findings also showed the power of such a mentoring program in building collegiality. As the mentors hailed from various academic departments, they were afforded the opportunity to develop ideas from different perspectives throughout the study. In addition to brainstorming with the teachers, the mentees also discussed and brainstormed as a group of academicians, which enlarged their academic vision and facilitated the establishment of a trust-based collegiality. This finding is also in line with the literature, emphasizing that mentors can benefit from mentoring programs in gaining new perspectives (Simpson et al., 2007), whilst at the same time improving the working relationships with their colleagues (Baran, 2016; Davies, Brady, Rodger, & Wall, 1999). In addition, the current study's findings suggest that building a collegial relationship ensured a meaningful growth was experienced by both mentors and mentees during the study.

The second issue that the mentors mainly discussed related to the workshop procedures and the steps that were implemented. The findings implicated that understanding the issues related to inclusive education and being able to implement effective inclusive practices requires activities to be conducted based on practice, feedback, and reflection. However, before conducting these steps, identifying mutual expectations based on effective discussion is key (Baran, 2016), and which positively affects the overall procedure, especially in ensuring the active participation of the mentees. As emphasized in the literature, when the expectations are mutually agreed upfront, the mentees' personal commitment increases (Barker, 2006), and this also indicates that the mentors will likely conduct the process based on mutual respect (Huskins et al., 2011). Thus, identifying and aligning mutual expectations with regards to the mentoring process can be considered as an essential element of effective mentoring (Baran, 2016; Rajuan, Beijaard, & Verloop, 2010).

Another element critical to the improvement of inclusive education practices requires inservice teachers to experience hands-on materials development in order to better understand the needs of the target group and thereby to aim to meet those needs. The findings support that hands-on material development is sine qua non for the professional development of effective inclusive practices by inservice teachers. The current research also revealed that ending the process with feedback and reflection can lead to increased communication between groups, as well as the satisfaction and motivation of the mentees. All of these procedural steps affected the overall quality of the mentoring program. Other mentoring studies conducted on inclusive education practices also suggest that such studies involve collaboration among different stakeholders, which contributes to teachers more effectively solving problems associated with inclusive education (Falvey, et al., 1989).

The current study also examined the perceptions of the mentees related to their mentors, which were found to be positive. The first topic was endorsing the setting-up of co-equal relationships, which was closely related to the aforementioned factor of there being shared decision-making and practicing together. Parallel to this, the study showed that the mentees found that being treated equally was considered to be a positive example of their interaction with the mentors. These findings imply that the establishment of equal relationships opened the way for mutual respect, trust, and an equal balance of power. This also corroborates with the literature, which emphasizes the significance of the equal relationship (Driscoll, Parkes, Tilley-Lubbs, Brill, & Bannister, 2009; Wilkinson et al., 2014), since mentoring is a bidirectional undertaking in its current form.

The second item that the mentees significantly commented on related to their mentors' being adaptive to their (the mentees') needs and suggestions. The mentees saw this as valuable since it motivated them to both continue with the process, and also to volunteer more within their group. As indicated in the literature, adapting materials based on the mentees' needs was considered to be a precondition for effective mentoring (Rajuan et al., 2010; van Ginkel et al.,2016).

Additionally, the mentees favored two important aspects related to the mentors in the current study, which were their mentors being both tech-savvy and proactive. These items were specific to the current study and to the

context, and which thereby positively affected the mentees' perspectives and the overall effectiveness of the process.

Conclusion and Suggestions

Mentoring, as a bidirectional undertaking, is a form of concurring with the best practices of inclusive education. With the mentors as the focus of the current study, they had the opportunity to implement procedures based on effective mentoring, and were thereby able to help the inservice teachers develop professionally in their preparations. Mentoring was highlighted as a means of overcoming some of the problems that the inservice teachers' faced in their educating of students with special needs.

Since the current study was a long-term research and required full involvement of its participants, only 23 teachers joined the study through to the end of the workshops. For this reason, it is suggested that these types of mentoring studies, aimed at the improvement of inclusive practices, could be conducted with a greater number of participants. Similarly, experimental studies could be conducted to compare different mentoring strategies. Moreover, the current study was limited to the development and adaption of science education materials for third-year and fourth-year VI students. In the future, studies that focus on the upper grades and also different disciplines could be conducted. Thus, inclusive practices will benefit further through a process of continuous improvement and thereby become more pervasive and seen less as being exclusive, distinctive, or uncommon. Another limitation of the current study was that the mentors did not observe the mentees' actual classroom practices. Conducting a mentoring study that included an element of coaching would provide a more sound and comprehensive interpretation of the teachers' understanding of inclusive education practices.

Acknowledgements or Notes

References

Abner, G. H., & Lahm, E. A. (2002). Implementation of assistive technology with students who are visually impaired: Teachers' readiness. *Journal of Visual Impairment & Blindness*, *96*(2), 98-105.

Akçamete, G., Aslan, B., & Dinçer, Ç. (2010, May). *Uygulama öğretmenlerinin mentörlük becerilerinin değerlendirilmesi* [Evaluation of practice teachers' mentoring skills]. Paper presented at the Uluslararası Öğretmen Yetiştirme Politikaları ve Sorunları Sempozyumu II. Hacettepe University, Ankara, Turkey.

Altunay Arslantekin, B. (2012). Engel türüne göre materyal uyarlama, hazırlama ve kullanımı [add translation here]. In A. Ataman (Ed.), *Temel eğitim öğretmenleri için kaynaştırma uygulamaları ve özel eğitim* [Inclusive education practices for classroom teachers and special education] (pp. xx-xx). Ankara, Turkey: Vize.

Ambrosetti, A., & Dekkers, J. (2010). The Interconnectedness of the Roles of Mentors and Mentees in Preservice Teacher Education Mentoring Relationships. *Australian Journal of Teacher Education, 35(*6), 42-55. http://dx.doi.org/10.14221/ajte.2010v35n6.3

Aslan, C. (2016). Görme engelliler için yardımcı teknolojiler [Assistive technology for visually impaired]. In S. Çakmak (Ed.), *Özel eğitim ve yardımcı teknolojiler* [Special education and assistive technologies] (pp. 56-92). Ankara, Turkey: Vize.

Atika, I. N., Ediyanto, E., & Kawai, N. (2018). Improving Deaf and Hard of Hearing Students' Achievements Using STS Approach: A Literature Review. *International Journal of Pedagogy and Teacher Education*, *2*, 3-19.

Aykut, Ç., & Özmen-Güzel, R. (2010). Hayat bilgisi ve sosyal bilgiler öğretimi [Life studies and social studies teaching]. In İ. H. Diken (Ed.), *İlköğretimde kaynaştırma* [Inclusive education practices in primary school] (pp. 480-515). Ankara, Turkey: Pegem Akademi.

Baran, E. (2016). Examining the Benefits of a Faculty Technology Mentoring Program on Graduate Students' Professional Development. *Journal of Digital Learning in Teacher Education, 32*(3), 95-104. https://doi.org/10.1080/21532974.2016.1169958

Bardin, J. A., & Lewis, S. (2008). A survey of the academic engagement of students with visual impairments in general education classes. *Journal of Visual Impairment & Blindness*, *102*(8), 472-483.

Barker, E. R. (2006). Mentoring – a complex relationship. *Journal of the American Academy of Nurse Practitioners, 18*(2), 56-61.

Batu, S., & Kırcaali-İftar, G. (2005). *Kaynastirma* [Inclusion]. Ankara, Turkey: KOK.

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report*, *13*(4), 544-559.

Betts, L. R., & Cross, A. (2010). Reflections on supporting a visually impaired student complete a biological psychology module. *Psychology Teaching Review*, *16*(1), 58-60.

Bogner, D., Hurd, D., Wentworth, B., Ristvey J., & Arens, S A. (2011). *Visualizing science with adapted curriculum enhancements*. *ACE Resource Manual*. McREL. Available online at https://sites.google.com/site/mcrelace/.

Bogner, D., Wentworth, B. L., Ristvey, J., Yanow, G., & Wiens, R. (2006). Our place in the spongy universe. *The Science Teacher*, *73*(3), 38-43.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development.* Sage

Boyer, L., & Lee, C. (2001). Converting challenges to success: supporting a new teacher of students with autism. *Journal of Special Education, 35*(2), 75-83.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oaClutterbuck, D. (2005). Mentoring and the legal profession: Developing professionals for real. *Development and Learning in Organizations, 19*(1), 10-12.

Conolly, J., & Blunt, R. J. S. (2006). Perceptions of mentoring: Expectations of a key resource for higher education. *South African Journal of Higher Education*, *20*(2), 195-208.

Creswell, J. W. (2008). *Educational research planning, conducting and evaluating quantitative and qualitative research.* Prentice Hall.

da Mota Silveira, H., & Martini, L. C. (2017). How the New Approaches on Cloud Computer Vision can Contribute to Growth of Assistive Technologies to Visually Impaired in the Following Years. *Journal of Information Systems Engineering & Management*, *2*(2), Article 9.

Darrah, M. A. (2013). Computer haptics: A new way of increasing access and understanding of math and science for students who are blind and visually impaired. *Journal of Blindness Innovation and Research*, *3*(2), 3-47.

Davies, M. A., Brady, M., Rodger, E., & Wall, P. (1999). Mentors and school-based partnership: ingredients for professional growth. *Action in Teacher Education, 21*(1), 85-96.

Driscoll, L. G., Parkes, K. A., Tilley-Lubbs, G. A., Brill, J. M., & Bannister, V. R. (2009). Navigating the lonely sea: Peer mentoring and collaboration among aspiring women scholars. *Mentoring and Tutoring: Partnership in Learning, 17*(1)*,* 5-21.

Ediyanto, E., Atika, I., Hayashida, M., & Kawai, N. (2017). A Literature Study of Science Process Skill toward Deaf and Hard of Hearing Students. In *1st Annual International Conference on Mathematics, Science, and Education (ICoMSE 2017)* (pp. 190-195). Atlantis Press. https://dx.doi.org/10.2991/icomse-17.2018.23

Fairchild, T. N., & Henson, F. O. (1993). *Engelli Çocuklarda Kaynaştırma Eğitimi* [Inclusive Education in Children with Disabilities] (S. Ünlü Trans./Ed.). Eskişehir, Turkey: Anadolu University.

Falvey, M., Coots, J., Bishop, K., & Grenot-Scheyer, M. (1989). Educational and curricular adaptations. In S. Stainback, W. Stainback, & M. Forest (Eds.), *Educating all students in the mainstream of regular education* (pp. 143-1158). Baltimore: Brookes.

Forde, C. & O'Brien, J. (2011). Taking continuing teacher education forward. In C. Forde, & J. O'Brien (Eds.), *Coaching and mentoring: Developing teachers and leaders* (pp. 82-90). Edinburgh, United Kingdom: Dunedin Academic.

Fraser, W. J., & Maguvhe, M. O. (2008). Teaching life sciences to blind and visually impaired learners. *Journal of Biological Education*, *42*(2), 84-89.

Gabriel, M. A., & Kaufield, K. J. (2008). Reciprocal mentorship: An effective support for online instructors. *Mentoring and Tutoring: Partnership in Learning, 16*(3), 311-327.

Godshalk, V. M., & Sosik, J. J. (2003). Aiming for career success: The role of learning goal orientation in mentoring relationships. *Journal of Vocational Behavior*, *63*(3), 417-437.

Gray, C. (2009). A qualitatively different experience: Mainstreaming pupils with a visual impairment in Northern Ireland. *European Journal of Special Needs Education*, *24*(2), 169-182.

Haggarty, L. (1995). The use of content analysis to explore conversations between school teacher mentors and student teachers. *British Educational Research Journal, 21*(2), 183-197.

Hagger, H., & McIntyre, D. (2006). *Learning teaching from teachers: Realising the potential of school-based teacher education*. Maidenhead, United Kingdom: Open University.

Hairon, S., Loh, S. H., Lim, S. P., Govindani, S. N., Tan, J. K. T., & Tay, E. C. J. (2019). Structured mentoring: Principles for effective mentoring. *Educational Research for Policy and Practice.* Advance online publication. https://doi.org/10.1007/s10671-019-09251-8

Halai, A. (2006). Mentoring in-service teachers: issues of role diversity. *Teaching and Teacher Education, 22*(6), 700-710.

Hasper, E., Windhorst, R. A., Hedgpeth, T., Van Tuyl, L., Gonzales, A., Martinez, B.,...Baluch, D. P. (2015). Methods for creating and evaluating 3D tactile images to teach STEM courses to the visually impaired. *Journal of College Science Teaching, 44*(6), 92-99.

Hatlen, P. (1996). The core curriculum for blind and visually impaired students, including those with additional disabilities. *RE:view Online, 28*(1), 25-32.

Heirdsfield, A. M., Walker, S., Walsh, K., & Wilss, L. (2008). Peer mentoring for first-year teacher education students: The mentors' experience. *Mentoring & Tutoring: Partnership in Learning*, *16*(2), 109-124.

Hobbs, T., & Westling, D. L. (2002). Mentoring for inclusion: A model class for special and general educators. *Teacher Educator, 37*(3), 186-201.

Hobson, A. J., Ashby, P., Malderez, A., & Tomlinson, P. D. (2009). Mentoring beginning teachers: What we know and what we don't. *Teaching and Teacher Education*, *25*(1), 207-216.

Hudson, P. (2004). Specific mentoring: a theory and model for developing primary science teaching practices. *European Journal of Teacher Education, 27*(2), 139-146. https://doi.org/10.1080/0261976042000223015

Hudson, P. (2013). Mentoring as professional development: 'growth for both' mentor and mentee. *Professional Development in Education, 39*(5), 771-783. https://doi.org/10.1080/19415257.2012.749415

Huizing, R. L. (2012). Mentoring together: A literature review of group mentoring. *Mentoring and Tutoring: Partnership in Learning, 20*(1), 27-55

Huskins, W. C., Silet, K., Weber-Main, A. M., Begg, M. D., Fowler, V. G., Hamilton, J. & Fleming, M. (2011). Identifying and Aligning Expectations in a Mentoring Relationship. *Clinical and Translational Science Journal, 4*(6), 439-447.

Ingersoll, R. M., & Strong, M. (2011). The impact of induction and mentoring programs for beginning teachers: A critical review of the research. *Review of Educational Research*, *81*(2), 201-233.

Janse Van Rensburg, K., & Roodt, G. (2005). The perceptions of employment equity and black economic empowerment as predictors of union commitment. *South Africa Journal of Industrial Psychology*, *31*(1), 55-64.

Jones, M. G., & Broadwell, B. (2008). Visualization without vision: students with visual. In J. K. Gilbert, M. Reiner, & M. Nakhleh (Eds.), *Visualization: Theory and practice in science education* (pp. 283-294). Dordrecht, Germany: Springer.

Jones, M. G., Minogue, J., Oppewal, T., Cook, M. P., & Broadwell, B. (2006). Visualizing without vision at the microscale: Students with visual impairments explore cells with touch. *Journal of Science Education and Technology, 15*(5-6), 345-351.

Kamali Arslantas, T., Yıldırım, S., & Altunay Arslantekin, B. (2019). Educational affordances of a specific web-based assistive technology for students with visual impairment. *Interactive Learning Environments* (Advance online publication). https://doi.org/10.1080/10494820.2019.1619587.

Kızılaslan, A., Zorluoğlu, S. L., & Sözbilir, M. (2020). A hands-on classroom activity to teach science concepts for students with visual impairment. *Science Activities, 56*(4), 130-138.

King, N. (2004). *Using templates in the thematic analysis of text. In* C. Cassell & G. Symon (Eds.), Essential guide to qualitative methods in organizational research (pp. 257–270). Sage

Klasen, N., & Clutterbuck, D. (2002). *Implementing Mentoring Schemes: A Practical Guide*. Oxford, United Kingdom: Elsevier.

Koehler, K. E., & Wild, T. A. (2019). Students with Visual Impairments Access and Participation in the Science Curriculum: Views of Teachers of Students with Visual Impairments. *Journal of Science Education for Students with Disabilities*, *22*(1), 8

Koehler, K. E., Wild, T. A., & Tikkun, S. (2018). Implications of 3-D Printing for Teaching Geoscience Concepts to Students with Visual Impairments. *Journal of Science Education for Students with Disabilities, 21*(1), 49-81.

Koenig, A. J., & Holbrook, M. C. (2000). *Foundations of education, Volume II: Instructional strategies for teaching children and youths with visual impairments*. Sewickley, PA: AFB Press.

Kolitsky, M. A. (2014). Reshaping teaching and learning with 3D printing technologies. *e-mentor*, *56*(4), 84-94.

Kopcha, T. (2010). A systems-based approach to technology integration using mentoring and communities of practice. *Educational Technology Research and Development, 58*(2), 175-190

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: SAGE Publications.

Lohmeier, K. L. (2009). Aligning state standards and the expanded core curriculum: Balancing the impact of the No Child Left Behind Act. *Journal of Visual Impairment & Blindness*, *103*(1), 44-47.

Lopez-Real, F., & Kwan, T. (2005). Mentors' perceptions of their own professional development during mentoring. *Journal of Education for Teaching, 31*(1), 15-24.

Ludi, S., Canter, A., Ellis, L., & Shrestha, A. (2012). Requirements gathering for assistive technology that includes low vision and sighted users. In *2012 First International Workshop on Usability and Accessibility Focused Requirements Engineering (UsARE)* (pp. 25-31). IEEE.

Luecke, R. (2004). *Coaching and Mentoring: How to Develop Top Talent and Achieve Stronger Performance.* Boston, MA: Harvard Business Press.

Lumpkin, A. (2011). A Model for Mentoring University Faculty. *The Educational Forum, 75* (4), 357-368

Lunney, D. (1995). Assistive technology in the science laboratory: A talking laboratory work station for visually impaired science students. *Information Technology and Disabilities*, *2*(1).

Mathur, S. R., Gehrke, R., & Kim, S. H. (2013). Impact of a teacher mentorship program on mentors' and mentees' perceptions of classroom practices and the mentoring experience. *Assessment for Effective Intervention, 38*(3), 154-162. https://doi.org/10.1177/1534508412457873

Mena, J., Hennissen, P., & Loughran, J. (2017). Developing pre-service teachers' professional knowledge of teaching: The influence of mentoring. *Teaching & Teacher Education, 66*, 47-59. https://doi.org/10.1016/j.tate.2017.03.024

Metatla, O. (2017, July). Uncovering challenges and opportunities of including children with visual impairments in mainstream schools. In *Proceedings of the 31st British Computer Society Human Computer Interaction Conference* (Article 102). Swindon, United Kingdom: British Computer Society.

Meyer, M., & Fourie, L. (2004). *Mentoring and coaching: tools and techniques for implementation.* Johannesburg, South Africa: Knowres.

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

Milli Eğitim Bakanlığı. (2018). *Özel Eğitim Hizmetleri Yönetmeliği* [Special Education Services Regulation]. Resmî Gazete 07.07.2018 Sayılı 30471. Ankara, Turkey: Milli Eğitim Bakanlığı. Retrieved from https://orgm.meb.gov.tr/meb_iys_dosyalar/2018_07/09101900_ozel_egitim_hizmetleri_yonetmeligi_0 7072018.pdf.

Milli Eğitim Bakanlığı. (2020). *Millî eğitim istatistikleri örgün eğitim 2018-2019* [National education statistics formal education]. Resmi İstatistik Programı. Ankara, Turkey: Milli Eğitim Bakanlığı. Retrieved from https://sgb.meb.gov.tr/meb_iys_dosyalar/2018_09/06123056_meb_istatistikleri_orgun_egitim_2017_2 018.pdf.

Moon, T. R., & Callahan, C. M. (2001). Curricular modifications, family outreach, and a mentoring program: Impacts on achievement and gifted identification in high-risk primary students. *Journal for the Education of the Gifted*, *24*(4), 305-321.

Moreland, L. M. (2015). *Science for Visually Impaired Students and Accessible Technology* (Doctoral dissertation). Marshall University, Charleston, WV. Retrieved from http://mds.marshall.edu/cgi/viewcontent.cgi?article=1983&context=etd

Morelle, M. (2016). *Challenges experienced by learners with visual impairment in two mainstream primary schools in Klerksdorp, Dr. Kenneth Kaunda District*. (Doctoral dissertation). University of South Africa.

Morelle, M., & Tabane, R. (2019). Challenges experienced by learners with visual impairments in South African township mainstream primary schools. *South African Journal of Education*, *39*(3), Article 1615.

Mundilarto, M. (2002). Posisi dan peranan strategis pengajaran IPA (Fisika) di SLTP [Position and Strategic Role of Teaching Natural Sciences (Physics) at SLTP]. *Cakrawala Pendidikan, Jurnal Ilmiah Pendidikan, 2002*(3), 339-356.

Nees, M. A., & Berry, L. F. (2013). Audio assistive technology and accommodations for students with visual impairments: Potentials and problems for delivering curricula and educational assessments. *Performance Enhancement & Health*, *2*(3), 101-109.

Negrete, O., Lisboa, A., Peña, F. J., Dib, C. O., & Vargas, P. (2020). Teaching labs for blind students: equipment to measure the thermal expansion coefficient of a metal. *European Journal of Physics, 41*(3), 035704.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods, 16*(1) https://doi.org/10.1177/1609406917733847

Ogawa, R. T., & Malen, B. (1991). Towards rigor in reviews of multivocal literatures: Applying the exploratory case study method. *Review of Educational Research, 61*(3), 265-286.

Okonkwo, H. C., Fajonyomi, M. G., Omotosho, J. A., Esere, M. O., & Olawuyi, B. O. (2017). Challenges, counselling needs, and coping strategies of students with visual impairment in regular secondary schools in Nigeria. *Human and Social Studies*, *6*(1), 111-137.

Ota, C. L., & Austin, A. M. B. (2013). Training and mentoring: Family child care providers' use of linguistic inputs in conversations with children. *Early Childhood Research Quarterly*, *28*(4), 972-983.

Patton, K., Griffin, L. L., Sheehy, D., Arnold, R., Gallo, A. M., Richardson, K.,...James, A. (2005). Navigating the mentoring process in a research-based teacher development project: A situated learning perspective. *Journal of Teaching in Physical Education, 24,* 302-325.

Pleiss, M. K., & Feldhusen, J. F. (1995). Mentors, role models, and heroes in the lives of gifted children. *Educational Psychologist*, *30*(3), 159-169.

Porter, J., & Lacey, P. (2008). Safeguarding the needs of children with a visual impairment in non-VI special schools. *British Journal of Visual Impairment*, *26*(1), 50-62.

Rajuan, M., Beijaard, D., & Verloop, N. (2010). The match and mismatch between expectations of student teachers and cooperating teachers: Exploring different opportunities for learning to teach in the mentor relationship. *Research Papers in Education, 25*(2)*,* 201-223.

Ramrathan, L., & Mzimela, J. (2016). Teaching reading in a multi-grade class: Teachers' adaptive skills and teacher agency in teaching across grade R and grade 1. *South African Journal of Childhood Education*, *6*(2), Article 448.

Rice, P. L., & Ezzy, D. (1999). *Qualitative research methods: A health focus* (Vol. 720). Victoria, Australia: Oxford University.

Rosenblum, L. P., Ristvey, J., & Hospitál, L. (2019). Supporting Elementary School Students with Visual Impairments in Science Classes. *Journal of Visual Impairment & Blindness, 113*(1), 81-88.

Roth, W.-M., & Tobin, K. (2002). *At the elbows of another: Learning to teach through co-teaching.* New York, NY: Lang.

Rudiyati, S. (2014). Improving Skills of Candidate Teachers of Children with Visual Impairment as Sighted Guide. *Dewantara, 2*(1), 24-33.

Rule, A. C. (2011). Tactile Earth and space science materials for students with visual impairments: Contours, craters, asteroids, and features of Mars. *Journal of Geoscience Education, 59*(4), 205-218.

Rule, A. C., Stefanich, G. P., Boody, R. M., & Peiffer, B. (2011). Impact of adaptive materials on teachers and their students with visual impairments in secondary science and mathematics classes. International *Journal of Science Education, 33*(6), 865-887.

Şahin, M., & Yorek, N. (2009). Teaching Science to Visually Impaired Students: A Small-Scale Qualitative Study. *Online Submission*, *6*(4), 19-26.

S´anchez-García, A. B., Marcos, J. J. M., GuanLin, H., & Escribano, J. P. (2013). Teacher development and ICT: The effectiveness of a training program for in-service school teachers. *Procedia-Social and Behavioral Sciences,* 92, 529-534. https://doi.org/10.1016/j.sbspro.2013.08.713

Sapp, W., & Hatlen, P. (2010). The expanded core curriculum: Where we have been, where we are going, and how we can get there. *Journal of Visual Impairment & Blindness*, *104*(6), 338-348.

Silverman, D. (2000). *Doing qualitative research: A practical handbook*. Thousand Oaks, CA: Sage.Simpson, T., Hastings, W., & Hill, B. (2007). 'I knew that she was watching me': the professional benefits of mentoring. *Teachers & Teaching, 13*(5), 481-498.

Smith, D. W., Kelley, P., Maushak, N. J., Griffin-Shirley, N., & Lan, W. Y. (2009). Assistive technology competencies for teachers of students with visual impairments. *Journal of Visual Impairment & Blindness*, *103*(8), 457-469.

St-Jean, E., & Audet, J. (2009). Factors leading to satisfaction in a mentoring scheme for novice entrepreneurs. *International Journal of Evidence Based Coaching and Mentoring, 7*(1), 148-161.

Supalo, C. A., Humphrey, J. R., Mallouk, T. E., Wohlers, H. D., & Carlsen, W. S. (2016). Examining the use of adaptive technologies to increase the hands-on participation of students with blindness or low vision in secondary-school chemistry and physics. *Chemistry toenderEducation Research and Practice, 17*(4), 1174-1189.

Toenders, F. G., de Putter-Smits, L. G., Sanders, W. T., & den Brok, P. (2017). Improving physics teaching materials on sound for visually impaired students in high school. *Physics Education, 52(*5).

Truter, I. (2008). Responsibilities of a leader as mentor and coach:: management. *South African Pharmaceutical Journal*, *75*(5), 58-61.

Urquhart, V. (2012). How One Teacher, Two Students with Visual Impairments, and a Three-Year R&D Project Could Change How All Students Learn Science. *Journal of the American Academy of Special Education Professionals, 2012*(Spring-Summer), 185-196.

van Dijk, H. G. (2008). The talent management approach to human resource management: attracting and retaining the right people. *Journal of Public Administration*, *43*(SI 1), 385-395.

van Ginkel, G., Oolbekkink, H., Meijer, P. C., & Verloop, N. (2016). Adapting mentoring to individual differences in novice teacher learning: the mentor's viewpoint. *Teachers and Teaching: Theory and Practice, 22*(2), 198-218.

Wang, J. (2002). Contexts of mentoring and opportunities for learning to teach: A comparative study of mentoring practice. *Teaching and Teacher Education, 17(*1), 51–73. https://doi.org/10.1016/S0742-051X(00)00038-X

White, M., & Mason, C. Y. (2006). Components of a successful mentoring program for beginning special education teachers: Perspectives from new teachers and mentors. *Teacher Education and Special Education*, *29*(3), 191-201.

Wilkinson, C., Pennington, T. R., Whiting, E., Newberry, M., & Feinauer, E. (2014). Triad relationships and member satisfaction with paired placement of student teachers. *Australian Journal of Teacher Education, 39*(8), 162-184.

Yalcin, G. (2020). Görme yetersizliği olan öğrencilere yönelik ortam, materyal ve program/öğretimsel düzenlemeler [Setting, materials and program / educational arrangements for students with visual impairments]. In P. Piştav Akmeşe & B. Altunay (Eds.), *İşitme ve görme yetersizliği olan çocuklar ve eğitimleri* [Children with hearing and vision impairments and their education] (pp. 271-293). Ankara, Turkey: Nobel.

Yalcin, G., & Altunay Arslantekin, B. (2019). Görme yetersizliği olan öğrenciler için genişletilmiş çekirdek müfredat ve dinleme becerileri [Extended core curriculum and listening skills for visually impaired students]. *Aksaray Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 3*(2), 298-323.

Yıldırım, A., & Şimşek, H. (2011). *Sosyal bilimlerde araştırma yöntemleri* [Research methods in social sciences]. Ankara, Turkey: Seçkin Yayınları.

Zachary, L. (2009). Examining and expanding mentoring practice. *Adult Learning, 20*(1-2), 43-45.

A Scale Development Study to Determine Disciplined Mind Features of 4th Grade Students

İjlal OCAK[1], Kerem İÇEL[2]
[1] Afyon Kocatepe University
[2] Republic of Turkey Ministry of National Education

To cite this article:

A Scale Development Study to Determine Disciplined Mind Features of 4th Grade Students

İjlal Ocak[1], Kerem İçel[2*]

[1] Afyon Kocatepe University
[2] Republic of Turkey Ministry of National Education

Abstract

The aim of this study is to develop a Disciplined Mind Scale (DMS) in order to determine the disciplined mind features of 4th grade students. Considering that students around the age of 11 can have some scientific thinking skills, it is thought that the disciplined mind features of 4th grade students should be determined. The sample of this research, in which the survey method was used, consists of 400 students studying in Afyonkarahisar-Turkey, in the 2018-2019 academic year. 23 items were removed from the item pool consisting of 50 items and a final form of 27 items, 7 negative and 20 positive, was obtained. KMO (Kaiser-Meyer Olkin) value was calculated as .862 and Cronbach alpha value as .826. As a result of the factor analysis, it was seen that the factor loading values of the items were between .426 and .786. It was determined that the scale consists of 5 factors (thinking like a scientist, making interdisciplinary connections, motivation to live with discipline, deep learning, connection with daily life). As a result of validity and reliability analyzes, it was seen that it can be accepted as a valid and reliable measurement tool consisting of 5 sub-dimensions that measures disciplined mind traits of 4th grade students.

Key words: Five minds, Disciplined mind, Disciplined mind scale.

Introduction

Just as in the past, learning is of great importance today. Learning is a need that exists in the creation of the individual. In the world of the 21st century, there have been some changes in the expectations of the society from individuals, as well as changes and developments in learning needs and teaching styles. Individuals are expected not only to have knowledge, but also to be able to use their knowledge skillfully and adapt it to new situations. The globalizing world order expects individuals to have creativity and innovation skills.

Individuals will be expected to develop some types of mind in the future (Nofsinger & Young, 2010). These mind types are "disciplined mind", "synthesizing mind", "creative mind", "respectful mind" and "ethical mind" (Gardner, 2006). Three of these mind types, called five mind types, consist of cognitive mind types, and the remaining two are relational mind types (Stork, Wodilla, Brown, Ogilvie, Rutter & Trefry, 2010). Individuals who can only have five mind types will be able to produce unique products. Educators will aim to develop five minds in individuals. When the characteristics of the type of individual that educators want to train in the future with the five mind areas put forward; It envisages a human model that has disciplined thinking skills, has been able to develop the synthesizing mind structure and thus acquires creative thinking skills, can demonstrate the skills to respect the rules of ethics and ethical principles while exhibiting these skills, and that can be beneficial to the society, the environment and the world. It is stated that it is of great importance to develop a disciplined mind in order to develop the synthesizing mind and creative mind features in the five mind areas (Gardner, 2006).

Disciplined mind means having a mind developed specific to the discipline. One of the main goals in the disciplined mind is that individuals are mastered in at least one discipline (Chang & Lee, 2008). This discipline can be a branch of art, profession, history or sociology from the humanities. In one discipline, the process of qualification can take up to ten years (Stork et al, 2010). It is thought that developments such as Newton's

---

disclosure of the law of gravity and Wilhelm C. Röntgen's finding of x-rays, which have a great importance in our lives, did not come into existence all of a sudden.

The disciplined mind begins only during adolescence and continues for the rest of a person's life (Pava, 2008). Therefore, it is thought that the disciplined mental characteristics of the students should be developed starting from the primary school period. Gardner states that although he studied psychology, it took him ten years to learn to think like a psychologist. Several years of intense engagement with the subject area are required to internalize a thinking style. Disciplined mind refers to the ability to adopt an academic discipline's thinking style (Schneider, 2014). Considering that the process of competence in a discipline can last up to ten years, the importance of developing a disciplined mind from primary school becomes apparent. All professional groups, whether they are lawyers or engineers, should have the basic principles and knowledge of their profession enough to deserve to be a member of their profession. An individual with a disciplined mind should have the ability to think specific to his profession (Sawyer, 2008).

It is believed that the five minds must be developed for the future (Nofsinger & Young, 2010). It is thought that learners with a disciplined mind will thus become lifelong learners. Individuals who cannot master one or more disciplines will not be able to succeed in any challenging workplace and will be limited to mundane tasks that are not of great importance (Chang & Lee, 2008). Who believes that current formal education prepares students for the possible worlds of the future, but primarily for the past world (Essig, 2012), education for five minds is challenging in all contexts. Developing a disciplined mind requires constant effort over a long period of time. In a context in which test scores guide more and more educational decisions, it is questionable whether training is possible for five minds (Davis & Gardner, 2012).

The aim of this research is to develop a Disciplined Mind Scale (DMS) to examine the disciplined mind features of 4th grade students. It is stated that children around the age of 11 have scientific thinking skills such as observing facts, recording data, and determining the effects of independent variables on dependent variables (Keys & Bryan, 2001). It is thought that it is important to examine disciplined mind features, since the 4th grade students, which constitute the universe of the study, are also in the 10-11 age group. It is stated that it may take approximately ten years to specialize in one or more disciplines. Individuals who specialize in disciplines will be the wanted individuals of the future (Gardner, 2006). Individuals with a disciplined mind are those who use and master "major scientific disciplines and ways of thinking about major professions" (Pava, 2008).

In primary schools in Turkey, the students are gaining 1st grade reading and writing skills. In addition, it is aimed to gain listening and speaking skills (Turkish Republic Ministry of Education, 2017). In Turkey, Mathematics and Turkish courses are started in the 1st grade primary school and continues throughout the primary school. Science course starts in the 3rd grade of primary school. It is thought that students need some pre-learning and positive attitude about these disciplines in order to teach disciplines in the secondary school and high school period.

As a result of the literature review, it was observed that five themes related to disciplined mind were formed (Can Aran, 2014). These themes appear as thinking like a scientist, making interdisciplinary connections, motivating to live with discipline, learning in depth, and connecting with daily life.

"Thinking like a scientist" enables students to focus on real-world issues that they find relevant to their own lives (Williams, Papierno, Makel & Ceci, 2004). The scientist must have an enlightened personality with a universal thinking structure (Ortaş, 2004). Scientists should be curious, open-minded, free, resourceful, and have the ability to communicate effectively (Jarrard, 2001). Scientists should not hesitate to reveal the facts and should have high character (Yıldırım, 2006). Scientists often try to explore their surroundings by asking the question of why (Chiappetta & Koballa, 2006) and they are people who research what is, not what should be (Abruscato, 2000). With the education program "Thinking Like a Scientist" developed by the Cornell Children's Research Institute (CIRC), it is aimed to reach students with innovative materials that train thinking and reasoning in scientific method about the problems in daily life (Williams, Papierno, Makel & Ceci, 2004). Science is a human activity and it is important to give students a comprehensive view of the nature of science (Bybee, 2006).

"Interdisciplinary connection" is the holistic approach of a concept, theme or problem using the method and language of more than one discipline (Jacobs, 1989; Erickson, 1995). According to the concept of interdisciplinary connection, disciplines are interconnected and real life problems do not always have one true (Perkins, 1994). By establishing interdisciplinary connections, it is possible to view information from different

angles (Şahbaz & Çekici, 2012). In order to achieve permanent learning, a connection must be established between the learned information (Bruner, 1999).

A student who is "motivated to live with discipline" will not need external reinforcements thanks to the pleasure of internal motivation (Kelecioğlu, 1992). Motivation is an inner force that drives the individual. If students find the information they learn meaningful and worth learning, it will enable them to be motivated against learning (Chiappetta & Koballa, 2006).

There are internal movements in the process of "deep learning". The process of students searching for meaning between concepts or disciplines and creating a meaningful link occurs (Ekinci, 2009). A student who strives for in-depth learning creates a purposeful and orderly study order by truly engaging with the subject area. He tends to investigate the reasons behind what is presented to him (Biggs & Kirby, 1983).

As individuals can find the opportunity to make connections between what is learned through daily life, permanent learning will take place (Bruner, 1999), individuals' preparation for life and being able to attribute meaning to the events in their daily lives are among the main objectives of education (Coştu, Ünal, & Ayas, 2007). Considering that students may have difficulties in determining how to relate what they learn in school and classroom environment in daily life (Doruk & Umay, 2010), purposeful teaching environments that can guide the use of what is learned in daily life should be designed instead of a random design in the arrangement of learning environments (Dewey, 2010).

Method

Research Design

In this study, survey method was used, which aims to describe the views and characteristics of large masses (Büyüköztürk, 2015) and can be generalized towards the represented universe in line with the data obtained from the sample (Cohen, Manion, & Morrison, 2007; Özdemir, 2015). Research data were collected from 4th grade students through a questionnaire.

Population and Sample

The population of the study consists of 4558 4th grade students studying in the 2018-2019 academic year. The sample of the study consists of 400 students, 207 male (%) and 193 female (%) students attending 4th grade. Simple random sampling technique was used in the study. It is assumed that the number of participants can represent the universe (Büyüköztürk, 2015; Çıngı, 1994).

Item Pool Phase

In this study, first of all, the relevant literature was researched and disciplined mind features were examined under 5 themes (Can Aran, 2014). Open-ended questions were created according to the determined themes and expert opinion was sought. Then, a form consisting of open-ended questions was distributed to 138 4th grade students and they were asked to answer in writing in the classroom. In the light of the answers obtained, a trial form consisting of 50 items was obtained by referring to the opinions of two experts who are faculty members in the field of Curriculum and Instruction and an expert who is a faculty member in the field of Science Education.

Statistics

Following the creation of the item pool, the trial form was administered to 400 students attending 4th grade. By applying KMO test and Bartlett's test of sphericity on the obtained data, the suitability of the data to the exploratory factor analysis was determined.

KMO and Barlett test values were examined in order to determine the suitability of the data obtained from the DMS trial form to exploratory factor analysis. A new structure can be revealed by questioning the relationships between variables with the exploratory factor analysis (Can, 2017). According to the result of exploratory factor analysis conducted to examine the construct validity, the KMO (Kaiser Meyer Olkin) test value calculated as .84 shows that the sample size is sufficient. The value of .000 found as a result of Barlett's test of sphericity shows that the data meet the multiple normality assumption (p < .01). The fact that the KMO coefficient is greater than .60 and the result of the Barlett test is significant indicates that the data set is suitable for principal component

analysis and is sufficient in terms of sample size (Can, 2017). In this case, it was concluded that the data obtained from the trial application of the scale was suitable for factor analysis.

Factor analysis is a statistical technique that aims to explain the measurement of variables that measure the same structure or quality by collecting them together with a small number of factors (Büyüköztürk, 2011). The factor loads of the items in the DMS were calculated, and the factors under which the items were found were determined with the transformed components matrix. In order to examine validity in item analysis, the mean scores of the end groups were compared using the t test. In order to determine whether there is a significant relationship between the dimensions of the DMS, Pearson correlation analysis has been performed. Confirmatory factor analysis was conducted to test the compatibility of the 5-factor scale, which was formed as a result of the exploratory factor analysis, with the model. Confirmatory factor analysis is a type of factor analysis that tests specific hypotheses about the structure and relationships between hidden variables underlying the data (Field, 2013). The reliability of the resulting factors was analyzed by calculating Cronbach's Alpha values. Alpha coefficient, which is also accepted as the lowest limit of the reliability coefficient, can be accepted as an internal consistency index (Tekindal, 2015).

As a result of the analyzes performed on the trial form, which initially consisted of 50 items, 23 items were removed from the scale. The analyzes made in the development of the Disciplined Mind Scale are given below.

Results

Exploratory Factor Analysis

By starting the analysis of the scale consisting of 50 items with factor analysis, a total of 14 factors were obtained at the first stage, and it was determined that the 14 factors that were obtained explain 56% of the total variance. As a result of the factor analysis, 23 items were removed from the scale. During factor analysis, items no 2, 25, 8, 7, 22, 38, 9, 45, 40, 33, 21, 32, 30, 35, 36, 13, 27, 20, 46, 34, 37, 42, 48 were Since it was seen that the difference between matrix values was less than .10, they were excluded from the scale. The common variance value explained by each item must be at least .10 (Seçer, 2015). However, due to the large number of factors, a line chart was used to determine the actual number of factors. One of the two statistical techniques generally used in determining the number of factors is eigenvalue plot and the other is scree plot (Salkind ve Green, 2005). In order to determine the number of factors, it was observed that there were 5 main break points in the line graph and the slope started to disappear after these break points. The number of components indicated by the point where the slope starts to disappear in the graph is taken as the number of factors to be calculated. The interval between two points on the line chart indicates a factor. During factor analysis, it can be said that it would be a much more accurate approach to consider more than one technique together instead of using a single technique during the factor determination or factor number decision (Seçer, 2015). Based on the main breaking points, the scale is limited to 5 factors.

The KMO value obtained as a result of factor analysis has increased to .862, and 5 factors explain 47.99% of the total variance. Items with a KMO value of .862 and an initial eigenvalue greater than 1.00 were included in the scale. The values obtained can be accepted as an indicator that factor analysis can be applied (Cureton and D'Agostino, 1983). The variance values explained for the total factor analysis of the disciplined mind scale are given in Table 1.

Table 1. Total Variance Explained Table of DMS

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.72	21.186	21.186	5.72	21.186	21.186	3.135	11.61	11.61
2	3.07	11.372	32.557	3.07	11.372	32.557	2.98	11.036	22.647
3	1.677	6.212	38.769	1.677	6.212	38.769	2.609	9.663	32.31
4	1.352	5.009	43.778	1.352	5.009	43.778	2.39	8.851	41.161
5	1.139	4.217	47.995	1.139	4.217	47.995	1.845	6.834	47.995

As seen in Table 1, it is seen that a structure with 5 factors has been formed. According to Seçer (2015), the concept of eigenvalue in factor analysis is a condition that shows the variance explained by a factor alone, and in factor analysis, the eigenvalue of a sub-dimension is expected to be at least 1. In addition, each of the sub-factors is expected to explain at least 5% of the total variance in the scale. In this sense, when determining the number of factors in a scale, dimensions with an eigenvalue above 1 and the variance value explained above 5% should be determined as sub-dimensions. According to Tekindal (2015), it is not appropriate to make a decision about the items just by looking at the common variance. It is desirable that the variance explained should be high. It is desirable to achieve the highest variance with the least factor. In social sciences, variances explained between 45 and 60 are generally encountered. It is seen that the obtained scale explains the total variance rate by 47.99%, the lowest sub-dimension eigenvalue is 1.84 and the sub-factors explain the total variance at the lowest 6.83%.

As a result of the evaluation made by considering all the criteria, 23 items of the 50-item trial scale were removed from the scale, and in this case, 27 items remained in the scale, 7 of which were negative and 20 of which were positive.

The factors under which the items belonging to the Disciplined Mind Scale are collected and the factor loads of the items are given in Table 2.

Table 2. Rotated Component Matrix[a]

DMS	Component				
	1	2	3	4	5
Factor 1. Thinking like a scientist (α=.767)					
DMS_8	.661				
DMS_13	.653				
DMS_12	.638				
DMS_16	.625				
DMS_3	.608				
DMS_6	.546				
DMS_9	.426				
Factor 2. Interdisciplinary connection (α=.769)					
DMS_22		.746			
DMS_21		.736			
DMS_26		.713			
DMS_18		.650			
DMS_15		.584			
DMS_23		.559			
DMS_25		.457			
Factor 3. Motivation to live with discipline (α=.708)					
DMS_4			.671		
DMS_11			.597		
DMS_14			.595		
DMS_5			.589		
DMS_2			.584		
Factor 4. Deep learning (α=.734)					
DMS_7				.709	
DMS_10				.684	
DMS_1				.576	
DMS_17				.565	
DMS_19				.532	
Factor 5. Connection with Daily life (α=.565)					
DMS_24					.786
DMS_27					.700
DMS_20					.536

As seen in Table 2, as a result of the analysis, it is seen that the factor load values for all five factors vary between .426 and .786. 7 items in the 1st factor (Thinking like a scientist), 7 items on the 2nd factor (Interdisciplinary connection), 5 items on the 3rd factor (Motivation to live with discipline), 5 items on the 4th factor (Deep learning), and the 5th factor (Connection with daily life). establishment) It has been determined that there are 3 items.

Another internal criterion of examining validity in item analysis is the comparison of the mean scores of the end groups (higher group-lower group) at the item level (Can, 2017). Accordingly, the answers given by the 27%

upper group and 27% sub group (108 people with the highest and lowest scores) were compared with the unrelated t test. The results of the t test made are given in Table 3.

Table 3. Comparison of Responses Given on Final Form in terms of Upper Group-Subgroup

Item Number	Groups	N	X	sd	t	p
DMS_1	Higher Group	108	3.40	1.223	-8.478	.00
	Lower Group	108	4.58	.762		
DMS_2	Higher Group	108	4.08	.892	-9.065	.00
	Lower Group	108	4.90	.321		
DMS_3	Higher Group	108	3.71	1.051	-8.642	.00
	Lower Group	108	4.68	.505		
DMS_4	Higher Group	108	3.83	1.195	-8.666	.00
	Lower Group	108	4.87	.379		
DMS_5	Higher Group	108	4.24	.935	-7.125	.00
	Lower Group	108	4.91	.309		
DMS_6	Higher Group	108	3.49	1.264	-9.299	.00
	Lower Group	108	4.72	.544		
DMS_7	Higher Group	108	2.93	1.331	-8.888	.00
	Lower Group	108	4.30	.901		
DMS_8	Higher Group	108	3.13	1.370	-11.401	.00
	Lower Group	108	4.76	.573		
DMS_9	Higher Group	108	3.14	1.398	-8.282	.00
	Lower Group	108	4.49	.942		
DMS_10	Higher Group	108	2.92	1.418	-7.553	.00
	Lower Group	108	4.21	1.059		
DMS_11	Higher Group	108	4.00	1.10199	-8.597	.00
	Lower Group	108	4.94	.26768		
DMS_12	Higher Group	108	3.34	1.161	-10.967	.00
	Lower Group	108	4.71	.580		
DMS_13	Higher Group	108	3.15	1.382	-11.511	.00
	Lower Group	108	4.77	.479		
DMS_14	Higher Group	108	3.91	1.184	-8.954	.00
	Lower Group	108	4.95	.211		
DMS_15	Higher Group	108	3.48	1.562	-9.018	.00
	Lower Group	108	4.90	.503		
DMS_16	Higher Group	108	3.48	1.392	-10.333	.00
	Lower Group	108	4.90	.321		
DMS_17	Higher Group	108	3.09	1.226	-10.298	.00
	Lower Group	108	4.51	.754		
DMS_18	Higher Group	108	2.84	1.511	-6.339	.00
	Lower Group	108	4.09	1.384		
DMS_19	Higher Group	108	3.05	1.205	-9.882	.00
	Lower Group	108	4.41	.774		
DMS_20	Higher Group	108	3.89	1.400	-6.230	.00
	Lower Group	108	4.81	.613		
DMS_21	Higher Group	108	2.87	1.512	-10.283	.00
	Lower Group	108	4.61	.884		
DMS_22	Higher Group	108	2.36	1.342	-9.932	.00
	Lower Group	108	4.16	1.329		
DMS_23	Higher Group	108	3.30	1.475	-9.571	.00
	Lower Group	108	4.81	.712		
DMS_24	Higher Group	108	4.31	1.007	-5.305	.00
	Lower Group	108	4.90	.572		
DMS_25	Higher Group	108	2.23	1.287	-8.447	.00
	Lower Group	108	3.83	1.481		
DMS_26	Higher Group	108	2.60	1.459	-11.102	.00
	Lower Group	108	4.58	1.144		
DMS_27	Higher Group	108	4.30	1.155	-6.037	.00
	Lower Group	108	4.98	.135		

According to Table 3, as a result of the comparison made, it was determined that the difference between lower group and higher group of the items to be included in the final scale was significant.

As a result of the analyzes made, it was thought that the scale had five factors, and this situation reflected five basic structures (thinking like a scientist, interdisciplinary connection, motivation to live in discipline, deep learning, connecting with daily life). The results of the Pearson correlation analysis made to determine whether there is a relationship between the factors are given in Table 4.

Table 4. Relationship Between Factors

	N	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	X	SD
Factor 1	400	-	.138**	.533**	.545**	.279**	4.0853	.73594
Factor 2	400	.138**	-	.168**	.005	.235**	3.5529	1.00611
Factor 3	400	.533**	.168**	-	.415**	.397**	4.5315	.56029
Factor 4	400	.545**	.005	.415**	-	.200**	3.7453	.85455
Factor 5	400	.279*	.235**	.397**	.200**	-	4.6529	.62090

(**p<.01)

According to Table 4, as a result of the Pearson correlation analysis performed to determine whether there is a significant relationship between the DMS dimensions, it was found that there is a statistically positive significant relationship between them at .01 significance level. The highest positive relationship is between the 1st factor and the 4th factor (r = .545 **, p<.01), the lowest positive relationship is between the 1st factor and the 2nd factor (r = .138 **, p< .01).

Confirmatory Factor Analysis

Confirmatory factor analysis was conducted to test the compatibility of the 5-factor scale, which was formed as a result of the exploratory factor analysis, to the model. After the analysis, the program suggested a modification between 6 items in order to have a significant decrease in the Chi-square value and a better level of fit indices. Modifications were made between items numbered "11 to 15", "17 to 4", "18 to 12", "24 to 10", "28 to 23" and "50 to 44". After the modification, the fit indices were analyzed and it was found that the fit indices were acceptable and some were at excellent levels. Confirmatory factor analysis data of DMS are given in Table 5.

Table 5. Acceptable Fit Indexes and Values Found

Fit Index	Acceptable Values	Found Value
χ^2 / sd	$0 \leq \chi^2/sd \leq 2$ perfect fit $2 < \chi^2/sd \leq 3$ acceptable fit	1.38
GFI	$.90 \leq GFI \leq 1.00$ perfect fit $.85 \leq GFI < .90$ acceptable fit	0.93
AGFI	$.90 \leq AGFI \leq 1.00$ perfect fit $.85 \leq AGFI < .90$ acceptable fit	0.91
NFI	$.95 \leq NFI \leq 1.00$ perfect fit $.90 \leq NFI < .95$ acceptable fit	0.93
NNFI	$.97 \leq NFI \leq 1.00$ perfect fit $.95 \leq NFI < .97$ acceptable fit	0.98
CFI	$.97 \leq CFI \leq 1.00$ perfect fit $.95 \leq CFI < .97$ acceptable fit	0.98
RMSEA	$0 \leq RMSEA \leq .05$ perfect fit $.05 < RMSEA \leq .08$ acceptable fit	0.031
RMR	$0 \leq RMR \leq .05$ perfect fit $.05 < RMR \leq .08$ acceptable fit	0.06
IFI	$0.95 \leq IFI \leq 1.00$ perfect fit $0.90 \leq IFI < 0.95$ acceptable fit	0.98

GFI: Goodness of Fit Index, **AGFI:** Adjusted Goodness Fit Index, **NFI:** Normed Fit Index, **NNFI:** Unscaled Fit Index, **CFI:** Comparative Fit Index, **RMSEA:** Root Mean Square Error of Approximation, **RMR:** Root Residual Mean of Squares, **IFI:** Incremental Fit Index.

According to Table 5, the fit indices of the scale show values in the range of excellent to acceptable levels. First, the ratio of Chi-Square value (427.71) to degrees of freedom (308) (X2 / sd = 1.38) shows a perfect fit. Moreover, Unscaled Fit Index (NNFI = .98), Comparative Fit Index (CFI = .98), and Incremental Fit Index (IFI = .98) Goodness of Fit Index (GFI = .93), Adjusted Goodness Fit Index (AGFI = .91).) and Root Mean Square Error of Approximation (RMSEA = .031) and perfect fit index values. On the other hand, Normed Fit Index

(NFI = .93) and Root Residual Mean of Squares (RMR = .20) were found to have acceptable fit index values. All these values show that the data set has a good and acceptable fit index and the scale is suitable for the model. In summary, the model formed as a result of exploratory factor analysis was verified by confirmatory factor analysis.

Reliability

The data on the variance rates and Cronbach α coefficients explained by the sub-factors of the disciplined mind scale are given in Table 6.

Table 6. Variance Ratios and Alpha Coefficients Explained by DMS's Sub-Factors

Factor	Variance Explained (%)	Alpha (α) Coefficient
1. Thinking Like a Scientist	11.610	.767
2. Interdisciplinary Connection	11.036	.769
3. Motivation to Live with Discipline	9.663	.708
4. Deep Learning	8.851	.734
5. Connection with Daily Life	6.834	.565
TOTAL	47.995	

When the Cronbach α coefficients in Table 6 are examined; It was calculated as .767 in the 1st factor, .769 on the 2nd factor, .708 on the 3rd factor, .734 on the 4th factor and .565 on the 5th factor. The total alpha value of the scale is .826. When the evaluation criteria used in the evaluation of the alpha coefficient are examined, if $0.00 \leq \alpha \leq .40$, the scale is not reliable. If $.40 \leq \alpha \leq .60$, the scale has low reliability. $.60 \leq \alpha \leq .80$ is highly reliable. If $.80 \leq \alpha \leq 1.00$, the scale is a highly reliable scale (Kalaycı, 2006). In this case, it can be said that the Disciplined Mind Scale with Cronbach α coefficient ,82, has a very high reliability.
DMS final form expressions and factor load values are given in Table 7.

Table 7. DMS Final Form and Factor Loadings

Item Number	Factor Load	Items in DMS
1	.576	I question the accuracy of what I learned in the lessons.
2	.584	I work regularly to be successful in my lessons.
3	.608	I use what I learned at school in my daily life.
4	.671	I would like to learn more from what I have learned in my lessons.
5	.589	I enjoy learning new information.
6	.546	I will definitely investigate a topic that is caught in my mind.
7	.709	I can scientifically explain the causes of natural events (rain, snow, etc.).
8	.661	I enjoy telling what I have learned to people around me.
9	.426	I'm interested in studying the development of living things (e.g. the growth process of a cat).
10	.684	I can scientifically explain how a plant grows.
11	.597	I will be happy when I have new information..
12	.638	I apply what I have just learned in my life.
13	.653	I like to share what I've learned with my friends.
14	.595	The more I learn, the happier I feel.
15	.584	**If I know math well, it would be okay if I don't know the other subjects.**
16	.625	I enjoy researching information that I am curious about.
17	.565	I can explain the reasons for events in daily life with what I learned in lessons.
18	.650	**Knowing math very well is enough to reach my dream job.**
19	.532	I can explain the reasons for events in daily life with the information I learned at school.
20	.536	I must continue to learn after I have a profession.
21	.736	**A science course is enough to be an inventor.**
22	.746	**It is sufficient to learn the courses in the field I want to specialize in.**
23	.559	**I don't want to make a new invention.**
24	.786	All the lessons I have learned are necessary for me.

25	.457	**I can't get the same pleasure from all the lessons.**
26	.713	**It is enough to study enough to pass the lessons.**
27	.700	All the lessons I have learned are important to me.

Items written in bold are negative items. The items in the final form consisting of 20 positive items and 7 negative items, totaling 27 items, were renumbered and made ready for implementation. Items numbered 3, 6, 8, 9, 12, 13, 16 constitute the "thinking like a scientist" factor, items 15, 18, 21, 22, 23, 25, 26 constitute the "interdisciplinary connection" factor, items numbered 2, 4, 5, 11, 14 constitute the "motivation to live with discipline" factor, items numbered 1, 7, 10, 17, 19 constitute the "deep learning" factor, and items numbered 20, 24, 27 constitute the "connection with daily life" factor.

Scoring Phase

There are 7 negative items (15, 18, 21, 22, 23, 25, 26) in the scale that must be reverse coded. The lowest score that can be obtained from DMS is 27 and the highest score is 135. DMS level 27-63 point range is considered as low level, 64-99 point range as medium level, 100-135 point range as high level.

Disscussion

One of the main goals in the disciplined mind is that individuals have mastered at least one discipline. Considering that the process of competence in a discipline can last up to ten years, it is thought that the disciplined mind should be developed from primary school. Disciplined mind features were examined under the themes of " thinking like a scientist, interdisciplinary connection, motivation to live with discipline, deep learning, connection with daily life".

In the theme of "thinking like a scientist" of the disciplined mind, it has been observed that the characteristics of a scientist include courage, determination to work, curiosity, impartiality, desire to share what they know with other people, tolerance, patience, and a creative personality. Science includes more than scientific knowledge. Science; It is the process of people revealing the unknown through creativity, calculation skills, curiosity, courage and patience (Bybee, Powell, & Trowbridge, 2008).

In the theme of "interdisciplinary connection" of the disciplined mind, it is seen that it is necessary to specialize in thinking skills specific to several disciplines (Gardner, 2006), since a single-disciplinary thinking style will not be sufficient. By establishing interdisciplinary connections, it is possible to view information from different angles (Şahbaz & Çekici, 2012). In order to reach permanent information, a connection must be established between the learned information (Bruner, 1999).

In the "motivation to live with discipline" theme of the disciplined mind; the characteristics of students' enjoyment of the learning process, the transformation of learning into a passion, learning after formal education, being willing to show knowledge to other people, and constantly striving to train and develop itself are seen. The concept of motivation includes complex behaviors that lead to an internal movement consistent with personality and cognitive characteristics (Krathwohl, Bloom, & Masia, 1964).

In the theme of "deep learning" of the disciplined mind, the features of being in an effort to understand and interpret instead of reading by rote, to establish a connection between knowledge, to understand the subject of learning in depth, to apply the learned information to a situation that it has not encountered before. A student who strives for in-depth learning creates a purposeful and orderly study order by truly engaging with the subject area. Students tend to investigate the reasons behind what is presented to them (Biggs & Kirby, 1983).

In the "connection with daily life" theme of the disciplined mind, we encounter the features of making a conscious and deep view about the facts about the time in which we live, and the effort to understand the world, in an effort to learn information in a meaningful way. Permanent learning takes place, as individuals can find the opportunity to make connections between what is learned by making connections with daily life (Bruner, 1999). Purposeful teaching environments that can guide how to use the learned in daily life should be designed (Dewey, 2010). What is learned at school will be of little importance unless it is used in problem solving in daily life (Entwistle, 2009).

When other studies on disciplined mind features are researched, no study was found at primary school level. Erik-Soussi (2008) investigated how the achievements of those working at the administrative level of

universities are affected by five types of mind. Miller (2011) examined the forums created for adolescent books in terms of characteristics of five areas of mind. It was seen by Yılmaz (2012) that secondary school 6th grade textbooks were examined in terms of five mind types, including the disciplined mind type. Bowen (2013) examined the four-year colleges, official colleges and university level assistant managers in the USA in terms of their institutional development according to their ability to use five areas of mind. It has been observed that these studies, which have been carried out in five areas of the mind, are generally carried out on adults and are limited in number.

When the scale development studies on the disciplined mind features were researched, it was seen that the disciplined mind scale consisting of 22 items was developed by Can Aran (2014) at the 7th grade level of secondary school. As the world turns into a more global nature, this will require a more modern way of learning and thinking (Nofsinger & Young, 2010). Since it is necessary to develop a "disciplined mind" for lifelong learning (Chang & Lee, 2008), which is a requirement of the 21st century, it is necessary to do research on what affects the disciplined mind.

No disciplined mind scale developed for different school or grade levels was found. Developed in the disciplined mind features, DMS is a 28-item scale developed at the 4th grade level of primary school. "Disciplined mind" is defined as the characteristic of thinking styles in major disciplines (Wrenn, 2010), and it takes several years of intense work with the subject area to internalize a thinking style (Schneider, 2014). By presenting the disciplined mind features of the students at the 4th grade of primary school; The effects of teaching processes or various variables in schools on the disciplined mind can be revealed.

Conclusion

In the beginning, after the trial form, which consisted of 50 items, was applied to 400 students, as a result of the analysis, 23 items were removed and a total of 27 items remained, 7 of which were negative and 20 of which were positive. It was observed that these items reflect the five basic structures taken into account in the preparation of the scale (thinking like a scientist, interdisciplinary connection, motivation tol ive with discipline, deep learning, connection with daily life). As a result of the analysis, it is seen that the factor loading values for all five factors vary between .426 and .786 at a high level. When the items in Factor I (3, 6, 8, 9, 12, 13, 16) are considered; It has been found that it measures the level of "thinking like a scientist". When the items in the factor II, (15, 18, 21, 22, 23, 25, 26) are considered; It has been determined that it measures the level of "interdisciplinary connection". When the items (2, 4, 5, 11, 14) in Factor III are examined; It has been determined that it measures the "motivation to live with discipline" level. When the items (1, 7, 10, 17, 19) in Factor IV are examined; It has been determined that it measures the level of "deep learning". When the items in Factor V (20, 24, 27) are examined; It has been found that it measures the level of "connection with daily life". Considering the reliability of the determined factors, $\alpha = .767$ for Factor I (thinking like a scientist), $\alpha = .769$ for Factor II (interdisciplinary connection), $\alpha = .708$ for Factor III (motivation to live with discipline), Factor IV (deep learning) and $\alpha = .734$ for Factor V (connection with daily life). The Cronbach Alpha value, which is the reliability coefficient for the whole scale, was determined as .826 ($\alpha > .70$). This indicates that the scale is highly reliable. As a result of the validity and reliability analysis of the DMS, it can be accepted as a valid and reliable measurement tool consisting of 5 sub-dimensions that measures the disciplined mind traits of primary school 4th grade students.

Recommendations

This study has a quantitative design aimed at developing a Disciplined Mind Scale (DMS). By applying DMS on different samples, DMS levels of 4th grade students can be examined in terms of various variables. Factors affecting DMS levels of 4th grade students can be investigated through experimental studies.

Notes

This scale development study is part of the second author's master's thesis titled "Investigation of the relationship between the disciplined mind attributes and STEM attitudes of elementary school 4th grade students (Afyonkarahisar sampling)". In addition, the abstract of this study was presented as an oral presentation at the UBEK-ICSE (2019) International Science and Education Congress with the title of "Primary School 4th Grade Disciplined Mind Features Scale Development Study".

References

Abruscato, J. (2000). *Teaching Children Science: A discovery approach.* Usa: Allyn and Bacon.

Biggs, J., & Kirby, J. (1983). *Approaches to Learning in Universities and CAEs.* 26(2), 3-9.

Bowen, M.T. (2013). *An examination of ınstitutional advancement vice presidents' reports at four-year public and private historically Black Colleges and Universities regarding their use of the five minds to promote stewardship*, 96 (Unpublished doctoral dissertation). University of Hartford.

Bruner, J. S. (1999). *Process of education (25th edition).* Cambridge: Harvard University Press.

Büyüköztürk, Ş. (2011). *Sosyal bilimler için veri analizi el kitabı [Manual of data analysis for social sciences].* Ankara: Pegem.

Büyüköztürk, Ş., Kılıç Çakmak, E. Akgün, Ö. E., Karadeniz, Ş. ve Demirel, F. (2015). *Bilimsel araştırma yöntemleri [Scientific research methods].* Ankara: Pegem.

Bybee, R. W. (2006). Scientific Inquiry and Science Teaching. L. B. Flick, & N. G. Lederman içinde, *Scientific Inquiry and Nature of Science* (s. 1-14). Netherlands: Springer.

Bybee, R. W., Powell, J. C., & Trowbridge, L. W. (2008). *Teaching Secondary School Science.* New Jersey: Pearson.

Can, A. (2017). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi (5. basım) [Quantitative data analysis in the scientific research process with SPSS (5th edition)].* Ankara: Pegem Akademi.

Can Aran, Ö. (2014). *Disiplinli zihin özellikleri açısından fen ve teknoloji eğitimi ve öğrenci düzeylerinin incelenmesi [An investigation of students' disciplined mind characteristics, and science and technology education].* (Unpublished doctoral dissertation). Hacettepe Üniversitesi, Ankara.

Chang, A. S. C., & Lee, N. H. (2008). Developing global citizens through Gardner's" Five minds for the future".*Asia-Pacific Education Research Association Conference 2008, Singapore.*

Chiappetta, E. L., & Koballa, T. R. (2006). *Science Instruction in the Middler and Secondary Schools.* Pearson.

Cohen, L., Manion, L. & Morrison, K. (2007). *Research methods in education.* London: Routhledge.

Coştu, B., Ünal, S., Ayas, A., ve Alipaşa, A. Y. A. S. (2007). Günlük yaşamdaki olayların Fen Bilimleri öğretiminde kullanılması [Using the events in daily life in Science teaching]. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, 8(1), 197-207.

Cureton, E. E.& D'Agostino, R.B. (1983). *Factor analysis an applied approach.* Broadway: Lawrence Erlbaum Associates, Inc. Publishers.

Davis K. & Gardner H. (2012). Five minds our children deserve: Why they're needed, how to nurture them. *Journal of Educational Controversy, 6(1),* 1-9.

Dewey, J. (2010). *Okul ve toplum [The school and society].* (H. A. Başman, Çev.) Ankara: Pegem.

Doruk, B. K., ve Umay, A. (2010). Matematiği günlük yaşama transfer etmede matematiksel modellemenin etkisi [The effect of mathematical modeling on transferring mathematics to daily life]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *41*(41), 124-135.

Ekinci, N. (2009). Üniversite öğrencilerinin öğrenme yaklaşımları [University students' learning approaches]. *Eğitim ve Bilim,* 34(151), 74-86.

Entwistle, N. (2009). *Teaching for Understanding at University.* Newyork: Palgrave Macmillan.

Erik-Soussi, C. (2008). *Leading in the Liberal Arts and Post- Secondaty Institutions: How Five Cognitive Minds Aid The Good Work Of College Presidents.* University of Hartford.

Essig L. (2012). Frameworks for educating the artist of the future: Teaching habits of mind for arts entrepreneurship. *Artivate, 1(2),* 65-77.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics.* Sage publications.

Gardner, H. (2006). *Five minds for the future.* Boston: Harvard Business School Press.

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

Jarrard, R. (2001). *Scientific Methods*. An online book. University of Utah, Salt Lake City.

Kelecioğlu, H. (1992). Güdülenme [Motivation]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 7, 175-181.

Keys, C. W. & Bryan, L. A. (2001). Co-constructing ınquiry-based science with teachers: Essential research for lasting reform. *Journal Of Research In Science Teaching ,* 38(6), 631-645.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of educational objectives: Handbook 2: Affective domain*. Longman.

Miller, D. L. (2011). *An Analysis for Habits of Minds in Blogs about Young Adult Books.* Arizona State University.

Milli Eğitim Bakanlığı [Ministry of Education] (2017). Fen Bilimleri öğretim programı [Science curriculum]. Retrieved from http://mufredat.meb.gov.tr/ ProgramDetay. aspx?PID=325

Nofsinger, C. & Young, A. (2010). Review of 5 Minds for the Future by Howard Gardner. *Education Review, 13*.

Ortaş, İ. (2004). Öğretim üyesi ya da bilim insanı kimdir [Who is a lecturer or scientist]. *Pivolka,* 3(12), 11-16.

Özdemir, E. (2015). Tarama yöntemi [Survey method]. M. Metin (Ed.), *Eğitimde bilimsel araştırma yöntemleri* (pp. 77-97) [*Scientific research methods in education* (pp. 77-97)]. Ankara: Pegema.

Pava M. L. (2008). Loving the distance between them:'Thinking beyond howard gardner's "five minds for the future". *Journal of Business Ethics, 83(2),* 285-296.

Perkins, D.N. (1994). *The Intelligen eye*. Sanat Monica, CA: The Getty center for education in the arts.

Salkind, N. J. & Green, S. B. (2005). *Using SPSS for windows and macintosh: Analyzing and understanding data (fourth edition)*. Pearson Prentice Hall.

Sawyer, R. K. (2008). Howard Gardner's Recommendations for the Future. *The American Journal of Psychology, 121(4), 671-675.*

Seçer İ. (2015). *SPSS ve LISREL ile pratik veri analizi, analiz ve raporlaştırma [Practical data analysis, analysis and reporting with SPSS and LISREL]*. Ankara: Anı.

Schneider, N. (2014). Five minds for adult education: Howard Gardner kritisch gewürdigt. *DIE Zeitschrift für Erwachsenenbildung*, (1), 30-32.

Stork, D., Woodilla, J., Brown, S. M., Ogilvie, J., Rutter, R., & Trefry, M. (2010). Starting with Howard Gardner's five minds, adding Elliott Jaques's responsibility time span: Implications for undergraduate management education. *Organization Management Journal*, 7(1), 28-38.

Şahbaz, N. K., ve Çekici, Y. E. (2012). Disiplinler arası bir disiplin olarak Türkçe eğitimi [Turkish education as an interdisciplinary discipline]. *Electronic Turkish Studies*, 7(3), 2367-2382.

Tekindal, S. (2015). *Duyuşsal özelliklerin ölçülmesi için araç oluşturma (3. Baskı) [Creating a tool for measuring affective characteristics (3rd Edition)]*. Ankara: Pegem Akademi.

Williams, W. M., Papierno, P. B., Makel, M. C., & Ceci, S. J. (2004). Thinking like a scientist about real-world problems: The Cornell Institute for Research on Children Science Education Program. *Journal of Applied Developmental Psychology*, 25(1), 107-126.

Yılmaz, Ü. (2012). *İlköğretim Fen ve Teknoloji dersi ders, çalışma ve kılavuz kitaplarının beş zihin modeli açısından değerlendirilmesi [Evaluation of elementary science and technology lesson, study and guide books in terms of five mind models]*. (Unpublished master's thesis). Akdeniz University, Antalya.

An Action Research Study on Teaching the Landform Concepts in a Fifth Grade Social Studies Course

Yusuf İNEL[1], Emrah URHAN[2]
[1] Uşak University
[2] Republic of Turkey Ministry of National Education

To cite this article:

An Action Research Study on Teaching the Landform Concepts in a Fifth Grade Social Studies Course

Yusuf İNEL[1], EmrahURHAN[2*]
[1]Uşak University
[2]Republic of Turkey Ministry of National Education

Abstract

This study aims to find out the effect of activities and practices on concept education to help carry out activities to teach these concepts more effectively and permanently, and determine student preconceptions about the concepts related to the landforms in the 5th-grade social studies curriculum in the learning subdomain called "People, Places and Environments." It was conducted as an action research study, employing a qualitative research design. The participants of the study were 16 fifth graders who were studying at a state secondary school in Bekilli District of the province of Denizli, Turkey. Mind maps, worksheet activities, models and mock-up, focus group interviews and field notes were used as the data collection tools, concept development tools, and process support tools. Collected in four class hours, the data were subjected to statement-based content analysis, and then compared with the criteria developed for the related concept and converted into findings and conclusions. The findings indicate that the student group has a partial preliminary knowledge and perception of the distinctive landforms they see around them. However, the student group's preliminary knowledge and perception of the landforms they do not observe around them is insufficient. The results obtained at the end of the instructional intervention process indicate that the student group is confused about the definitions of the bay, gulf, cape, peninsula, island, strait, and ocean. With the activities implemented after the additional action plan, the student group made progress by effectively learning the basic and sub-dimensional landforms they were confused about. The student group found it useful to carry out mind maps and application activities and achieved effective learning with models and mock-ups. Based on research result, use of sightseeing –observation techniques can be recommended to social studies teachers.

Keywords: Social Studies, Concept Teaching, Landforms, Action Research

Introduction

The modern formal education approach views the type of teaching based solely on factual knowledge as inappropriate because of its immense volume and insufficiency in meeting all learner needs by itself. Instead, it views concept teaching based on factual associations resulting from drawing parallels among similar objects, events, and processes more worthwhile to apply. The concept is an information structure that is expressed in a word, which exists in the mind of the person as a general meaning, and which is used to represent the variable points of different kinds of objects and different phenomena (Ülgen, 2004, p.107). Concepts are tools for making categorical distinctions and associating different facts. Objects, people, events with similar characteristics can be grouped into ideas and processes through concepts (Tokcan, 2015, p.7). In this respect, concepts are special tools used by the human mind in their thinking processes. In understanding and perceiving the world, an individual needs concepts for self-expression and using highly comprehensive and diverse information (Senemoğlu, 2004).

Considering the context and environment learned, concepts can be learned through socialization in school or in a wide variety of social networks such as family, work environment, and social environment. This type of socialization-based concept learning process is simpler and more random, which consists mostly of concrete concepts. However, the concept teaching process based on school and curriculum is more difficult, because concepts can remain abstract and can consist of complex patterns. Supporting this view, Gagne divides concepts into abstract and concrete, stating that concrete concepts are perceived through sense organs, but abstract concepts cannot. Concrete concepts can be learned in informal ways from an early age, but abstract concepts require a systematic teaching process, i.e. formal education (as cited in Alkaş, 2014, p.70).

The most important tools of formal education are schools and curriculum employed in schools. The social studies curriculum was updated in 2018, placing a stronger emphasis on concept teaching, value teaching, and skill teaching. As an interdisciplinary field of study, the field of social studies primarily focuses on the social sphere and considers the facts and problems related to this sphere. Taking into account previous definitions of social studies, Öztürk (2011, p.24) made the following definition *"Social studies is a curriculum that merges the knowledge and methods obtained from the social and humanities with the aim of educating active citizens who can take decisions based on knowledge and can solve problems in the national and international conditions which are changing in every way."*

Human and human lives are the main subjects in social studies education. Understanding human and human life is only possible by understanding and knowing phenomena and the concepts, and performing relational analyses of them. This phenomena, events, concepts and generalizations are selected from various disciplines.

Concept teaching has always been viewed as crucial in social studies education, which is a multidisciplinary field and has also been highlighted in the updated 2018 national-level curriculum. This is because the social studies course should be taught based on understanding instead of direct transmission of information, and thus helping students understand, use and learn to produce new kinds of information (Bart and Demirtaş, 1997, p.10). In addition, concept teaching enables problem-solving, reasoning, distinguishing between right and wrong, simplifying learning and remembering, and increasing academic achievement (Doğanay, 2005, p.277). For these reasons, the national-level social studies program prepared in 2005 stresses concept teaching by using classifications in teaching concepts, underlines the use of various concept teaching approaches, and advises practitioners to eliminate the confusion of meanings and concepts and to prevent forming misconceptions.

In the social studies course, students are observed to have difficulties in understanding the concepts they encounter for the first time, who may confuse them and form misconceptions (Yazıcı and Samancı, 2003, p.83). Indeed, a review of the related literature reveals that Avcı (2015) investigated 6th- grade students' levels of understanding geography concepts and their misconceptions related to these concepts in a social studies lesson, and found that these concepts were not adequately learned and that the students had many misconceptions regarding these concepts. Gültekin (2016) conducted a thesis study by referring to the opinions of teachers in teaching some concepts in the fifth-grade social studies unit "*Turkey Step by Step*", and reached the conclusion that some concepts in the unit could not be taught, that these concepts were not appropriate for the age level of the students, these abstract concepts were not encountered by the students in daily life, and that teachers struggled to teach these concepts by traditional methods. In his doctoral dissertation research, Akşit (2016) first determined the background knowledge and perceptions of students about some concepts in the unit "the Journey through the Turkish History" in the seventh-grade social studies curriculum, and prepared and implemented some activities to ensure that they learned the concepts in the unit more effectively and permanently. He also observed and evaluated the effects of the prepared activities on the process of concept teaching. Subaşı (2013) examined the learning levels of the concepts in the fifth grade "Turkey Step by Step" unit in his master's thesis and completed his research by administering a 20-item multiple-choice test to 361 students in the schools of Çankaya central district of Ankara. He concluded that the concepts of reign, secularism, cultural element, chronology, place and value in the unit were not fully learned. Doğrukök (2004) selected 45 concepts for his study on the acquisition levels of 6th-grade social studies course concepts and concluded that the level of learning these geography concepts was low. Conducting a study with 6th-grade students, Yükselir (2006) compared the traditional method used in teaching concepts with the concept analysis method experimentally, reaching the conclusion that the concepts related to geography and the Earth are learned more permanently and substantially by applying the concept analysis method. Süer (2010) used the action research method in a 6th-grade social studies course to determine the degree of learning achieved regarding some geography concepts and the readiness levels of students about learning these geography concepts, by using a 20-question achievement test as the pre-test and giving a post-test at the end of the action activities process, and concluded that the targeted concepts and topics were adequately remembered. In addition, the findings demonstrated that the level of concept learning by the student group increased and that the use of multiple methods and techniques in

teaching geographical concepts led to more effective learning. To reveal any misconceptions regarding the concepts taught in the 6th grade "Life on Earth" unit, Akdağ (2010) gave a 30-question achievement test to 197 students in Eskişehir province of Turkey and found that the students had some misconceptions. He concluded that the concepts that the students had the highest number of misconceptions about Sphere versus geoid, meridian versus parallel, meridian versus equator, mathematical position versus special position, map versus sketch, and special map versus physical map. Studying the learner acquisition level of the concepts of life on earth unit in the 6th-grade social studies course, Talay (2011) selected some human and physical geography concepts in the textbook as the research subject and concluded that the course outcomes related to these abstract concepts, especially the concepts related to the physical geography subject were not adequately achieved by the students. In a study of 6th-grade students conducted by Avcı (2015) to determine the level of their understanding geographical concepts and whether they form misconceptions, it was found that the research group did not adequately understand many concepts of geography and that many students had misconceptions.

Dere and Aktaşlı (2019) conducted an action research about teaching the conceptsof production, distribution and consuption. In their research they prepared activities and aimed to reveal the effectiveness of their activities. As a result of research, they concluded that students learned the concepts of production, distribution and consuption better, and overcome misconceptions.In their study with24 middle school students, Dere and Aktaş (2020) gave lecture to students with activities they prepared before to change students conscious consumerism and increase their awereness. They determined that activities prepared as a result of their study in the form of action research, enabled students to perceive "conscious consuption"as a responsible, sustainable and thougtfulsonsuption by positively changing students' conscious consumerism perceptions and awereness.

Our review of the related literature reveals no one-on-one study that matches our research, but the concepts examined in the current study have been the subject of some other studies conducted at different grade levels. It was observed that there is some confusion and misconceptions regarding the geography subjects that are analyzed in our study, the use of traditional methods are insufficient to ensure learning retention of the geographical concepts, and diversifying the teaching process is a must. In line with the renewed curriculum and syllabi, the concept of the *unit* has now been removed from the content of social studies, instead, replaced by the *learningarea* approach. With this new design, grade-based changes were made in the course subjects. Some 6th-grade subjects have been moved to 5th grade. Along with this change, the fact that there has been no study dealing with a similar subject at the grade levels where "People, Places and Environments" learning area is taught reveals the originality of the present study. Therefore, with its particular focus and content, our study can be expected to help researchers working on the teaching of geography subjects in social studies by eliminating the existing gap in the field.

The starting point that attracts the researcher to the study is the following question: "*Are the concepts included in the social studies curriculum learned by the students?*" As such, the aim of this study is to prepare and implement activities for more effective and permanent learning the *"landforms and sub-concepts of landforms"* in the *"People, Places and Environments"* learning domain in the 5th grade social studies program, and to observe the effects of the prepared activities on the process of concept teaching.

Derived from the main problem, the following sub-problems were aimed to be answered in the study:

1. What is the pre-knowledge and perceptions of the students about the concepts of landforms that are covered by the learning domain?
2. What activities can be prepared for the meaningful, effective and lasting teaching of the concepts of landforms covered within the scope of the study
3. How did the activities developed for the study affect or change the concept teaching process?

Method

Research Design

The action research, a qualitative research design, was used in the study. Considering the teacher responsibility of producing solutions to the practical problems of the teaching field, Yıldırım and Şimşek (2005, p.306) use the term "teacher research" for action research. Action research is a type of research that is based on classroom and in-school practice aimed at improving the quality of teaching (Johnson, 2014, p.19). Viewed as such, it is evident that action research is not only about determining a situation and problem but also about application and

process. With a holistic nature, action research has a productive nature, characterized by reflection, participation and development (Yıldırım and Şimşek, 2005, p.78). Action research improves a teacher's reflective thinking skills as well. Action research in the field of practice supports teacher professional development, offers new experiences, and helps apply theory to the field (Johnson, 2014, p.26).

In this study Practise/mutual cooperation/disscussion-oriented action research approach atype of action research has been adopted.In this approach the researcher and practitioner identify possible problems that may arise in practise and the factors that may reveal these problems and possible solitions by working together. Since this approach is aimed at improving the application it is also called practise-oriented action research.

Study Group

The study group of this study was selected by criterion sampling (Yıldırım ve Şimşek, 2005), which is a purposive sampling technique. The study group consists of 16 students studying in the 5th grade of a public secondary school in Bekilli town located in Denizli province of Turkey. Due to research ethics, the real names of the students were kept anonymous and each student was assigned a code as their pseudonym.

The characteristics of the students who constitute the study group were classified, organized and then presented. Using the e-School application data, these characteristics were categorized by parental professional and educational status and the average student scores from the previous school year were explained through subjective values indicating the level of achievement. In the tables, the pseudonyms (codes) are used for the students forming the study group. To determine the financial situation of the participants, the data obtained by the classroom counselor and school counselor were used. The table below explains the demographic characteristics of the study group.

Table 1. Demographic characteristics of the study group

F	Code Name	Gender	Paternal profession	Maternal profession	Paternal Education	Maternal Education	Revenue	Academic level of success
1	Irmak	Girl	Teacher	Teacher.	BA	BA	High	Very High
2	Taha	Boy	Security P.	Housewife	High School Degree	High School Degree	Average	High
3	Yasin	Boy	Farmer	Farmer	High School Degree	High School Degree	Very High	Average
4	İlhami	Boy	Driver	Housewife	High School Degree	Elementary School	Average	Average
5	Ela	Girl	Worker	Baby-sitter	BA	High School Degree	Low	Average
6	Ecrin	Girl	Ironmonger	Housewife	Elementary School	Elementary School	Average	Low
7	Havva	Girl	Farmer	Housewife	High School Degree	Secondary School	High	Average
8	Muham.	Boy	Employee	Baby-sitter	High School Degree	Elementary School	Low	Very High
9	Ayşegül	Girl	Medical Officer	Housewife	BA	High School Degree	Average	Average
10	Muhsin	Boy	Ironmonger	Security P.	BA	BA	Very High	High
11	Fatma	Girl	Employee	Housewife	Secondary School	Elementary School	Average	Average

12	Asunur	Girl	Tradesman	Housewife	High School Degree	High School Degree	Very High	Average
13	Nisa	Girl	Shepherd	Housewife	Elementary School	Elementary School	Very Low	High
14	Ramazan	Boy	Tradesman	Housewife	High School Degree	Elementary School	High	Average
15	Mustafa	Boy	Worker	Housewife	High School Degree	Elementary School	Low	Very High
16	Öykü	Girl	Teacher.	Accountant	BA	BA	High	Average

As the subject of the study is landforms, the characteristics of the particular geographical area where the study group lives are also important. These can be described as follows:

Bekilli town is administratively a part of the province of Denizli. The town is located in the north of Denizli. It is 86 kilometers away from Denizli city center and is located 83 km from the neighboring province of Uşak (an administrative center) in the north. Bekilli's neighboring border from the north is the town of Karahallı, a part of Uşak, bordered by Çal town of Denizli in the south, Ulubey town of Uşak in the west, and Çivril town of Denizli in the east.

The land area of Bekilli is 247 km², with an altitude of 850 meters. Its Gömce neighborhood has the highest altitude. The town is surrounded by hills: the Asar (Hisar) Hill (957 m.) in the south direction, Aslankara Hill in the north, Hocalı Hill (932 m.) in the west, Uzunçalı and Zıntı Hill in the east (894 m.), and Tatar Hill (908 m.) in the southwest.

The central and rural parts of Bekilli are characterized by hilly terrain. The land that can be considered as flatland is the "Medele Plain" located in the neighborhoods of Çokaşlı, Yeşiloba and Kutlubey which lie in the west and northwest of the town center. The Büyük Menderes River runs 5 km south of the district, which is home to some projects of power generation and agricultural irrigation that offer no benefits for Bekilli. Because the valley where the river flows is too deep and steep, personal opportunities for irrigation are rather limited as well.

Data Collection

In this study, various techniques and methods were used under the main headings of pre-implementation data collection tools, data collection tools used during the implementation and data collection tools that support the process. The mind map was used to determine the learning status of the participants in the pre-implementation and post-implementation stages, worksheets and extracurricular activities were used to teach the targeted concepts, and field notes and focus group interviews were used to support the data collection process. In the first hour of the class, each student prepared a mind map so that their existing knowledge and perceptions of the landform concepts could be revealed, during the second hour of class concept acquisition activities and worksheets were used, and during the third hour of the class, to determine the final status of student learning, a mind map regarding the concepts of landforms was prepared again, the gaps in learning were determined and additional activities were implemented in line with the additional action plan. To increase the quality of the data collected, descriptive field notes were taken throughout the process, and focus group meetings were held at the end. The data collection phase of the study was carried out by the researchers themselves.

Data Analysis

Data analysis can be defined as the process of exporting the data obtained by various data collection tools. During the data analysis phase of this research, the data analysis methods proposed by Wolcott (1994) were followed, the authenticity of the data was preserved, the data were subjected to systematic analysis, and the experience gained by the researcher in the process was reflected in the study.

The main data collection tool of this study is the mind maps created on the concepts of landforms. The analysis of these mind maps was performed as follows:

1. All mind map forms related to landform concepts were read at once.
2. All the words that students use in their mind maps were transferred verbatim to the related tables.
3. The frequency of the words that the students expressed was determined.
4. Based on the definitions of landform concepts and the concept analysis proposed by Martorella (Doğanay, 2005, p.277- 278) and the 5th grade Social Studies textbook, a set of criteria was developed for the concept and the secondary dimensions created from students' statements associated with the concept were also added to these criteria.
5. The words expressed by the students were classified as "related" or "unrelated" according to the criteria.
6. The codes and themes, as implied by the words expressed by the participants were obtained.
7. Based on the codes and themes a common mind map was created for the holistic and visual presentation of the data. The valid student statements were indicated on the criterion map with a frequency value.
8. In the final stage, the students' mind maps were described, compared and interpreted in light of the criteria.
9. The worksheets were analyzed by scoring them, and the field notes and focus group interviews were reflected in the findings as they were.

Implementation Process

In this action research study, the data on the concepts of landforms were collected in a week during four class hours. Additionally, it was supported by focus group interviews at the end of the data collection process. The action research cycle proposed by Johnson (2014, p.20) was followed in conducting the study.
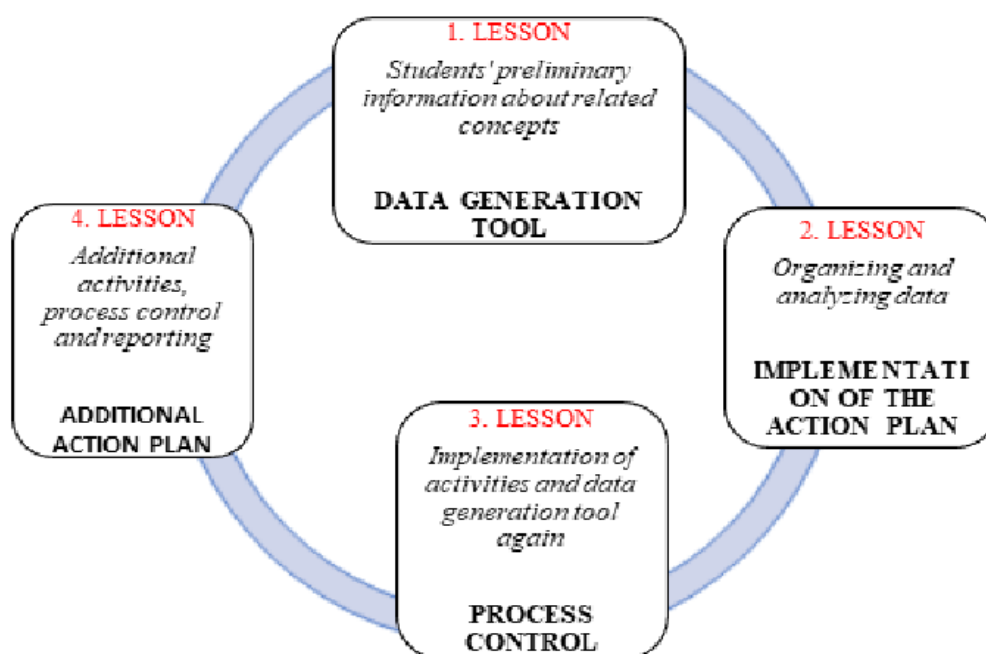
Figure 1. Research cycle proposed by Johnson (2014, p.20)

During the research phase of the study, in the first lesson of the week, the researcher conducted mind map work on the concepts of landforms to reveal the background knowledge and perceptions of the participants about the relevant concept. Following the determination of the background knowledge and perceptions, the concept teaching was carried out with various methods and techniques according to the action plan prepared in the second and third hours, and reinforcement was given by using additional activities and worksheets. In the second half of the third hour, the final assessment of the participants' learning status of the concepts of

landforms was carried out. The same mind mapping tool used in the first assessment was used for the final assessment as well. After the final results were analyzed, if the learning was at a lower level and limited, the additional action plan was used and the additional activities were implemented during the fourth-course hour in the same week. During the research process, field notes were kept during every practice hour, and after the completion of the process, focus group interviews were conducted on the last day of the week during the recesses and lunchbreaks for general evaluation purposes.

According to the action plan created to teach the landform concepts, the steps were taken as described in the table below.

Table 2. Implementation process of the Landform Concepts

	Asessment of the Current Status (1st Hour)	Action Activities and Final Assessment (2nd and 3rd Hours)	Additional Action Plan (Lesson 4)
WEEK 2 (Landforms)	1.Determination of the existing status 2. Review of the 1st week subject 3. Introduction to landforms	1.Teaching the subject of *landforms* 2. Teaching the "Different Landforms" activity 3. Determinationj of the final status	1. Teaching to eliminate any learning gaps. 2. Construction of 3D landforms material
	4. Field notes	5. Field notes	5. Field notes 6. Focus group interviews

A benchmark mind map for the concepts of landforms was developed by the researcher to be able to rigorously analyze the data obtained from the preliminary and final assessments. Martorella's (1998, as cited in Doğanay, 2005, p.278) "*Concept Analysis Strategy*" was used to develop the benchmark mind map. The mind map developed as a benchmark for the concepts of landforms is given below. Initial and final assessment findings were reached by taking into account the above criteria.

Figure 2. A benchmark mind map for the landform concepts

Results and Discussion

In determining the preliminary knowledge and perceptions of the student group about the landform concepts, 10 students related to the mountain concept, 6 students related to the plain concept, 4 students related to the river concept, 4 students related to the sea concept, 3 students related to the plateau concept, 3 students, 2 students related to the concept of the valley, 1 student related to the concept of ocean left a statement, made a visual

drawing and brought a short definition. No expression was found in the first case determination regarding the concepts of gulf, peninsula, cape, and island, strait, bay. It was understood that the students were familiar with the concepts of island, peninsula, but did not evaluate these concepts in the earth category, so they did not leave an expression to these concepts in the first case study. It is understood from the interviews that the bay, headland, strait and bay concepts are not known as a geographical term. In the first case determination, it is understood from the field notes, focus group interviews and the mind map, which is the tool of data production, that the students do not have information in a structured format about the landforms. In accordance with the action plan for the concepts of landforms, "different landforms" activity was applied and the results of the activity were analyzed after the subject transfer with the support of plain expression and smartboard applications.

Accordingly, 14 students made correct pairings for the island, lowland, ocean and river concepts in the activity based on matching the titles and visual drawings of the landforms with the definitions; 13 students made correct matches for the lake concept, and 12 students made correct matches for the mountain and plateau concepts.
In the second part of the event, 16 students made correct matches for the sea concept, 13 students for the valley concept, 13 students for the peninsula concept, 8 students for the promontory concept, 6 students for the strait concept, and 5 students for the bay and inlet concepts. The student code-named Ela confused the island and lake concepts both in visual drawing and definition, which can be attributed to haste and carelessness because the researcher's field notes indicate that Ela was the first to complete the activity and hand it over. The student code-named Taha confused the visual drawing and short description of the concepts of plateau and mountain. However, he was able to distinguish these two concepts from each other in the focus group interview. Examining the activity paper he submitted, the reason for his confusion about these two concepts appears to be more of haste than lack of learning. The student code-named Ayşegül confused the concepts of lowland and plateau, the concepts of mountain and hill, and the concepts of lake and island. The expression of "wide flatland" under the concepts of plain and plateau, the expression of "mound" under the definition of mountain-hill concepts, the similarity of the images of the lake and island drawings, and the carelessness of the student are seen as the main reasons for making wrong matches. The student code-named Fatma code incorrectly matched the concepts of the mountain - plateau, river - ocean, and plain-lake.

Similar expressions found in the definitions of the concepts, the lack of adequate guidance and the fact that this student was easily distracted can be viewed as the reasons for this failure. The student codenamed Nisa confused the mountain - plateau concepts and river - ocean concepts. The reason is thought to stem from the similarity of definitions and the effect of the inability to concentrate on the subject.

While the concepts of sea, valley and peninsula were learned at a sufficient level in the second part of the activity, the promontory and the bay were confused by the majority of the group. The mind map form was used again and analyzed for final assessment after the activities related to the concepts of landforms were completed. According to the final assessment, 13 students provided a statement on the mind map, made a drawing, and gave a short definition for the mountain concept, 13 students for the plain concept, 10 students for the sea concept, 9 students for the plateau concept, 8 students for the lake concept, 7 for the river concept, 6 students for the valley concept, 2 students for the ocean, promontory, and gulf concepts, and 1 student for the bay, island, and peninsula concepts. Considering the final assessment, it was observed that students made progress in all concepts compared to the initial assessment. However, according to the results of the implemented activity, fewer expressions were found for the concepts in the final assessment of learning than provided for the initial assessment. According to the focus group interview conducted to find out the cause of this problem, the students knew the concepts of landforms but could not express them because they were not guided in the mind map. However, there are instructions, shapes and definitions in the activity. Since the students were asked only to make matches, the number of correct statements was higher than the final assessment. The final assessment showed that the concepts of river, valley, ocean, promontory, gulf, bay, island, and peninsula were inadequately learned, which made the researchers implement an additional action plan to improve learning on these concepts. Comparison of the initial (10) and final (13) assessments revealed that the mountain was the earth formation that the student group provided the highest number of expressions in both activities. The concept with the highest number of student statements in the final assessment (13) compared to the initial assessment (4) is plain. The strait emerged as the landform that was mentioned in neither assessment. As the potential reasons for this, it might be argued that the teacher was incapable of making this concept more tangible and that the student group had difficulty in creating a scheme related to the concept of strait since they had not encountered this land formation in their immediate environment. In addition, the reason why very few expressions were provided for the concepts of the bay, gulf, peninsula, promontory, island, and the ocean could be because they are very similar landforms whose definitions and images make distinguishing them difficult, and these are not part of students' daily lives.

The analysis of the initial and final assessments of learning revealed that all the students submitted at least one statement in the initial assessment and that the student who submitted the highest number of statements was Ela (5). The student who submitted the lowest number of expression was Asunur (1). Irmak (8), Havva (8), and Muhsin (8) were the students who submitted the largest number of statements in the final assessment. It is noteworthy that Havva was one of the most articulate students who provided the highest number of statements in the final assessment. Havva was made sure to get engaged in the lesson while this concept was being taught, in-class work was performed to increase the level of readiness, and thus Havva's concept learning skill was enhanced. Ela (5) who submitted the highest number of statements in the initial assessment, submitted the same number of statements in the final assessment, which shows that Ela did not have a meaningful learning experience through the activities implemented. In addition, Asunur (1), Ayşegül (1), Nisa (1) were the students who provided the lowest number of expression in the final assessment. Surprisingly, Nisa and Ayşegül submitted a fewer number of expressions than they did for the initial assessment. Students were asked about this during the focus group interviews. Nisa said that she did not like nor understood the social studies lesson, and Ayşegül said that she was careless, had family problems, could not pull herself together and she asked the researcher for help during another assessment work.

According to the action plan, an additional action plan was introduced during the fourth-course hour. Three-dimensional landforms included in the additional action plan were studied with the student group. Accordingly, the student group was aimed to see all the landforms in a holistic and interrelated way, some mortar was prepared by blending straw and mud with the participation of all students, a 3-D model containing the landforms was made under the guidance of the teacher on a suitable platform, and after the mud (the main material of the model) was dry, the landforms were painted appropriately.

The presentation of the subject was repeated with the interaction of the student group on the model that emerged as a result of the activity. As can be inferred from the field notes, the student group worked as a team while performing this activity, did a division of work, and learned by having fun. The focus group interviews with the students about the activity showed that they were happy with their work, that they were proud to produce an original educational material that they were satisfied to contribute to their own learning with their own product. The students' statements during the focus group interviews made it clear that thanks to this activity, they were able to learn the concepts of promontory, bay, peninsula, island, and strait. Considering the geographical context in which the research was carried out, these landforms were among the concepts that the student group had never encountered before. The students interacted with landforms such as mountains, plains, plateaus, streams, and valleys and they were made aware of their interactions through the examples provided by the teacher. After explanations were made when needed, the associations with daily life were made, and the principle of teaching by going from the immediate environment to the distant environment was applied, a significant improvement in learning was demonstrated by all the students in the group and the subject became clear for all of them. Therefore, utmost attention should be paid to the principle of "moving from the closer towards the farther away" in the teaching of geography subjects in the secondary school social studies course. Using the prompts in the immediate environment is the key point in following this principle. In addition, when applying the principle of closer-to-farther, abundant examples should be given, and due attention should be paid to make sure these examples are directly from students' life.

Fifth graders need teacher support and guidance in making sense of the concepts, information and shapes they encounter for the first time and creating new diagrams. A maximum effort needs to be spent on embodying the topics to be learned as much as possible, to add information from daily life, to show concepts three dimensionally. In addition, the ability to distinguish between similar concepts seems to have not sufficiently developed for student groups at this age level since almost all students had problems with these distinctions. In the course of the research, particular attention was paid to the observations reflected in the field notes as stated above. The researcher made a point of updating personal knowledge of the context and subject in each step of the study and making progress by drawing lessons from previous practices.

Conclusion

As regards the landform concepts, the student group had knowledge about the concepts of mountains, plains and streams, as indicated by the initial assessment. Since these landforms were the type they could observe directly around them, it was normal for them to have such preliminary knowledge. However, they had very little knowledge of the concepts of plateau, lake, sea, valley, and ocean. They had no knowledge of the concepts of

gulf, peninsula, promontory, island, strait, and bay, which could be attributed to the fact that they do not come across any of these landforms in the geographical location they live in.

When the student group's recent knowledge and perceptions of landforms considered, it can be said that they made progress in the concepts of mountains, plains, sea, plateau, lake, stream and valley and that the "Various Landforms" action activity worked in teaching these sub-dimensions. In addition, the landforms these students learned best were those that they frequently encountered in their daily lives, which may have helped them enhance their learning. On the other hand, they were not able to make adequate progress in the concepts of promontory, bay, island, and peninsula, and they also confused the visuals and definitions of these concepts. The action activity did not adequately serve the learning of these sub-dimensions. Another reason that may have played a role in this could be that these landforms are too rare for students to encounter in the geographical area they live in. Talay (2011) found that the subjects of physical geography were not adequately understood by the students and that they differed in their acquisition of these concepts. Avcı (2015) also concluded that middle school students confuse the concepts related to geography. While the action activity in the relevant area of our research worked to teach the concepts that were observable in the immediate environment of the group, it failed to teach the concepts that were not around them and remained abstract to the students, leading to conceptual confusion. This may indicate that there is a need for embodiment, abundant sampling, models and materials in the teaching of landforms. This was taken into account in the additional action activities.

The additional activity of "3-D Landforms Material Construction" was carried out to teach all of the aforementioned concepts, especially those that turned out to be challenging for the students, in an effort to ensure holistic learning. The activity was based on making Earth shapes as models from some mortar prepared with soil, straw and water. After the practice, the student group corrected their learning about the Earth formations, about which they experienced a conceptual confusion, and had the opportunity to reinforce what they had already known accurately. In their studies Avşar (2010), Meydan (2001), Çelik (2018) highlighted the importance of using materials in teaching concepts specific to social studies and concluded that they contribute positively to learning, which is a conclusion that supports the findings of the current study. Beyond simply using the given materials, the students participating in the study prepared their own materials as well. The additional activity can, therefore, be considered more effective in teaching landform concepts and in correcting misconceptions regarding landforms.

It was observed that the students in the study group had the utmost pleasure in producing new things and doing the activities where they reached new knowledge through their own observations and experiences. In particular, the additional action activities were based on this type of practice, and the experimental and observational practices were the practices that the student group acquired new learning and addressed their knowledge gaps best. During the research, active learning practices that went beyond the traditional approach made a great contribution to the overall progress of the group. Furthermore, the student group needed plenty of examples, connections to their lives, and achieved meaningful learning to the extent they were able to meet these needs. This is in line with the learning outcomes that the constructivist learning approach aims to achieve as well.

Recommendations

While the students learned the landforms that were visible in their own environment more easily through the worksheets; they had difficulty in learning the landforms that did not exist in the environment they live in, with which they had no personal experiences. Regarding the landforms in the content of the learning domain, practitioners are recommended to teach these concepts by following the principle of "teach from closer-to-farther," presenting plenty of examples, using tangible instructional materials, making students prepare their own course materials, and using a 3-D model.

It was also found that the students who had participated in field trip activities were more successful and responded with a higher number of statements these concepts in their initial and final mind maps. In their study with primary school teacher candidates, Akkuş and Meydan (2013) found that if teacher candidates used field trips in their teaching, student interest in the lesson would increase. Çiftçi and Dikmenli (2016), in their study with geography teachers, found that extracurricular activities increase the retention of learned knowledge. Therefore, using excursions, field trips, and other observation activities are highly recommended in teaching and reinforcing landform concepts.

References

Akdağ, Ş. (2010). *İlköğretim 6. sınıf öğrencilerinin sosyal bilgiler dersi yeryüzünde yaşam ünitesindeki kavram yanılgıları [The misunderstanding of the expressions about the life on the, the expressions about the life on the land,the six grade students]* (Unpublished master's thesis). Afyokkarahisar Kocatepe University, Insitute of Social Siences, Afyonkarahisar.

Akkuş, A. & Meydan, A. (2013). Sosyal bilgiler öğretiminde tarihi ve coğrafi mekân uygulamalarının değerlendirilmesi [Evaluatıon of hıstorıcal and geographıcal place ımplementatıons ın socıal scıences teachıng]. *Uluslararası Avrasya Sosyal Bilimler Dergisi, 4*(13), 14-30.

Akşit, İ. (2016). *7. Sınıf sosyal bilgiler dersi Türk tarihinde yolculuk ünitesinde geçen bazı kavramların öğrenilmesi sürecinde karşılaşılan sorunların belirlenmesi ve çözümüne yönelik bir eylem araştırması [An action research intended to solve and detect the encountered issues that are used during the process of studying of some concepts on the 7th grade social studies corse in the unit 'voyage in the history of Turkish]* (Unpublished master's thesis). Uşak University, Insitute of Social Siences, Uşak.

Alkış, S. (2014). Sosyal bilgilerde kavram öğretimi [Concept teaching in social studies] M. S. (Ed.) içinde, *Sosyal Bilgiler Öğretimi [Social studies teaching] (s.67-90)*. Ankara: Pegem Akademi.

Avcı, G. (2015). *Altıncı sınıflarda sosyal bilgiler dersinde öğrencilerin coğrafi kavramları anlama düzeyleri ve kavram yanılgılarının belirlenmesi [The determination of students' levels of understanding geographical concepts and geographical misconceptions in sixth grade social studies course]* (Unpublished master's thesis). Dokuz Eylül University, Insitute of Education Sciences. İzmir.

Avşar, S. (2010). *İlköğretim ikinci kademe sosyal bilgiler derslerinde coğrafya konularının öğretiminde materyal kullanımının öğrenci başarısına etkisi [Second stage geography elementary social issues course of their teaching materials using the effect of student achievement]* (Unpublished master's thesis*)*. Selçuk University, Insitute of Education Sciences, Konya.

Barth, J. L., & Demirtaş, A. (1997). *İlköğretim sosyal bilgiler öğretimi* [Elementary social studies education]. Ankara: YÖK/Dünya Bankası Milli Eğitimi Geliştirme Projesi Öğretmen Eğitimi Dizisi.

Çelik, K. (2018). *Sosyal bilgiler dersi yeryüzünde yaşam ünitesi iklim ve insan ilişkisi konularının materyal destekli öğretiminin bazı değişkenler açısından öğrencilerin akademik başarılarına etkisi [The effects of teaching life on earth unit climate and human relationship issues with in terms of some variables material assisted instruction on the academic achievement social studies course]* (Unpublished master's thesis). Mehmet Akif Ersoy University, Insitute of Education Sciences, Burdur.

Çifci, T. & Dikmenli, Y. (2016). Coğrafya öğretmenlerinin okul dışı coğrafya öğretimine ilişkin görüşleri [Geography Teachers Views with Regard to Teaching Geography Outdoors]. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi (KEFAD), 17*(1), 363-382.

Dere, İ, & Aktaşlı, İ. (2019). Üretim, dağıtım ve tüketim kavramlarının öğretimine ilişkin bir eylem araştırması [An Action Research on the Teaching of Production, Distribution and Consumption Concepts]. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 23*(3), 1389-1406.

Dere, İ. & Aktaşlı, İ. (2020). Conscious consumerism education in social studies courses: Reflections of an action research. *International Journal of Contemporary Educational Research, 7*(1), 215-227. DOI: https://doi.org/10.33200/ijcer.666525

Doğanay, A. (2005). Öğretimde kavram ve genellemelerin geliştirilmesi [Development of concepts and generalizations in teaching]. C. Ö. (Ed) içinde, *Hayat Bilgisi ve Sosyal Bilgiler Öğretimi* [Life Science and Social Studies Teaching] (s. 227-255). Ankara: Pegem Akademi Yayınevi.

Doğrukök, A. (2004). *İlköğretim 6. sınıf sosyal bilgiler dersinde yer alan kavramların kazandırılma düzeyi: Konya örneği [The to be enabled level of concept in 6'th plas social studies course: Simple Konya]* (Unpublished master's thesis). Gazi University, Institute of Education Sciences, Ankara.

Gültekin, E. (2016). *Ortaokul 5. sınıf sosyal bilgiler dersi "adım adım türkiye" ünitesinde bulunan kavramların kazandırılmasına ilişkin öğretmen görüşleri [Teachers' views on teaching the concepts of the 'Step by step Turkey' unit of secondary school 5th grade Social Science lesson]* (Unpublished master's thesis*)*. Adnan Menderes University, Insitute of Social Siences, Aydın.

Johnson, A. (2014). *Eylem araştırması el kitabı. [Action research handbook] (Çev. Y. Uzuner. M. Özten Anay).* Ankara: Anı Yayıncılık.

Martorella, P. (1996). *Teaching social studies the elemantary school / the basics og citizenship.* New Jersy: Prentice-Hall, İnc.

MEB. (2005). *İlköğretim sosyal bilgiler dersi 6-7. sınıf öğretim program [Primary social studies course 6-7. classroom curriculum].* Ankara: MEB.

Meydan, A. (2001). *İlköğretim birinci kademe sosyal bilgiler öğretimi coğrafya ünitelerinin işlenişinde laboratuvar ve görsel-işitsel materyal kullanımının öğrencilerin niteliksel gelişimine etkisinin değerlendirilmesi [Evaluation of the effect of the usage of laboratories and audio-visual materials, in teaching first stage primary education teaching social science geography units, on the development in quality]* (Unpublished master's thesis). Selçuk University, Insitute of Education Sciences, Konya.

Senemoğlu, N. (2004). *Kuramdan uygulamaya gelişim, öğrenme ve öğretim [Development, learning and teaching from theory to practice].* Ankara: Gazi Kitapevi.

Süer, S. (2010). *6. sınıf sosyal bilgiler dersinde coğrafi kavramlar bakımından öğrencilerin hazırbulunuşluk düzeyleri ve bu kavamların geliştirlmesi [Social sciences course of 6th grade and teaching of these terms via different methods and techniques]* (Unpublished master's thesis). Marmara University, Insitute of Education Sciences, İstanbul.

Talay, F. (2011). *İlköğretim 6. sınıf sosyal bilgiler dersi kapsamında işlenen yeryüzünde yaşam ünitesinde geçen coğrafya kavramlarının kazanılmışlık düzeyi. [For the students of the 6. grade students in social sciences, the assessment of acquisition level of geographical concepts in life unit]* (Unpublished master's thesis). Kastamonu University, Insitute of Education Sciences, Kastamonu.

Tokcan, H. (2015). *Sosyal bilgilerde kavram öğretimi [Concept teaching in social studies].* Ankara: Pegem Akademi.

Ülgen, G. (2004). *Kavram geliştirme kuramlar ve uygulamalar (4.Baskı)*, [Concept development theories and applications (4th Edition)] Ankara: Nobel Yayın Dağıtım.

Wolcott, H. (1994). *Transforming qualitative data:.* Newbury Park CA: Sage.

Yazıcı, H., & Samancı, O. (2003). İlköğretim öğrencilerinin sosyal bilgiler ders konuları ile ilgili bazı kavramları anlama düzeyleri [The levels of elementary school students' understanding of some concepts related to social studies course subjects]. *Milli Eğitim Dergisi*, (158), 83-90.

Yıldırım, A., & Şimşek, H. (2005). *Sosyal bilimlerde nitel araştırma yöntemleri (5.Baskı) [Qualitative research methods in social sciences (5th Edition)].* Ankara: Seçkin Yayınevi.

Yükselir, A. (2006). *İlköğretim altıncı sınıf sosyal bilgiler programında geçen kavramların kazanımı ve kalıcılığında kavram analizi yönteminin etkisi.[The Effect of the concept analysis method on the learning and permanance of the concepts of social sciences programme in the sixth ear of the primary education]* (Unpublished master's thesis). Çukurova University, Insitute of Social Sciences, Adana.

Preservice Middle School Mathematics Teachers' Definitions of Algebraic Expression and Equation

Pınar Yıldız[1], Azime Atay[2], Ş. Koza Çiftçi[3]
[1]Çanakkale Onsekiz Mart University
[2]Erciyes University
[3]Akdeniz University

To cite this article:

Preservice Middle School Mathematics Teachers' Definitions of Algebraic Expression and Equation

Pınar Yıldız[1*], Azime Atay[2], Ş. Koza Çiftçi[3]
[1]Çanakkale Onsekiz Mart University
[2]Erciyes University
[3]Akdeniz University

Abstract

Using correct definitions of the mathematical concepts is crucial for learning and teaching of any mathematical content. Being able to make mathematically correct definition of the concepts is an indicator of teachers' content knowledge. The purpose of this study is to determine how preservice middle school mathematics teachers define the concept of algebraic expression and equation. The participants of this case study were 35 preservice middle school mathematics teachers. The data were collected through written exam and semi-structured interviews. Written exam includes two questions asking preservice teachers to define equation and algebraic expression and write an example of each. Only 9 of participants correctly defined algebraic expression. Preservice teachers' definitions of algebraic expression were classified under three themes which are expressions containing unknown, expressions containing equality, and mathematical expressions. Two themes arose from preservice teachers' definition of equation: expressions with unknown, and expressions with equality.

Key words: Preservice mathematics teachers, mathematics teacher education, algebraic expression, equation.

Introduction

Teachers are the most important factors affecting teaching and learning process of mathematics (Ball, 1990; Charalambous, Hill, Chin & McGinn, 2019; Shulman, 1986, 1987). Teachers' knowledge about the subject that they will teach is the most important factor affecting student success. Content knowledge is the knowledge of teachers about definitions of concepts, relation between concepts and mathematical notation of subject they teach (Ball, Tames &d Phelps, 2008). For students to learn mathematics topic and concept correctly, teachers should have sufficient content knowledge about the subject (Ball, 1990, Ball, Thames & Phelps, 2008; Fennema, Sowder & Carpenter, 1999). Researches show that teachers' content knowledge has effect on their teaching practice and thus students learning (Ball, Lubienski & Mevborn, 2001; Brizuela, 2016; Charalambous et al., 2019; Copur-Gencturk, 2015; Hill & Ball, 2004; Hill, Rowan & Ball, 2005; Tchoshanov, Cruz, Huereca, Shakirova, Shakirova & Ibragimova, 2017). It cannot be expected from a teacher with lack of knowledge on subject he teaches to perform a successful teaching practice (Ball, 1990; Ball et al., 2008; Dreher, Lindmeier, Heinze & Niemand, 2018). To strengthen students' mathematical knowledge, teachers should be strengthened first (Ma, 2010).

Teachers should know the concepts included in content area and different representations of the relationships between these concepts (Ball, 1990). One of the most important dimensions of content knowledge is to be able to make a mathematically correct definition (Ball et al., 2008). Teachers should be able to define the concept correctly first. Making correct definition of concepts is the basis of concept knowledge about the subject that, they will teach (Ball, 1990, Ball et al., 2008).

Learning algebra is challenging for students and teachers' content knowledge related to mathematical concepts determine how students make sense of algebraic concepts (Stephens, Ellis, Blanton & Brizuela, 2017). Just focusing on operational properties of algebraic concepts prevents students from correctly structuring their mathematical meaning. Students develop many different misconceptions as they deal with concepts of algebra

---

because they focus on rules and operation (Akkan, Baki & Çakıroğlu, 2012; Dede & Argün, 2003; Weinberg, Dresen & Slater, 2016). Students have difficulty to understand letter symbols used in algebra. They find it difficult to make sense of letter symbols as a numerical value that represents a variable (Brizuela, 2016; Christou & Vosniadou, 2012). Students also learn algebraic concepts incorrectly due to the erroneous meanings they attribute to letter symbols (placeholder, abbreviation of an object, etc.). For example, they define algebraic expression and equations as they include a letter and an operation symbol because of wrong learning related to letter symbol (Stephens et al., 2017).

The most difficult concepts in teaching algebra are the concepts of algebraic expression and equation. The concept of algebraic expression and equation form the basis of teaching advanced algebra (Wasserman, 2016). In order for these concepts to be learned correctly by students, teachers' content knowledge must be strong. Attorps (2003) stated that teachers' understanding of equation concept is incorrect. Teachers explained the meaning of the equation through operational processes. Another study examining conceptions of algebra conducted by Stephens (2008) included 30 preservice elementary teachers. Result of the study indicated that preservice teachers' content knowledge related to algebra is quite limited. Most of participants equated algebra just with manipulation of symbols and implied that solution strategies including traditional symbol manipulation may be more valuable for algebra learning than strategies demonstrating conceptual understanding. Stump and Bishop (2002) found that 30 prospective teachers who participated in their studies defined algebra with a focus on symbol, letter and problem solving, and they did not mention algebraic relations. Zuya (2017) worked with 54 preservice secondary school mathematics teachers to investigate preservice teachers' knowledge and skills related to algebra. Finding of the study showed that the preservice teachers' performances were very low in tasks demanding conceptual knowledge for solution while they performed above average in tasks requiring procedural knowledge. What do teachers view as algebra determines what will worth to teach in the classroom (Stephens, 2008). That is, teachers' definitions of algebra play an important role during organization of teaching process (Stephens, 2008; Stump & Bishop, 2002).

Teachers and pre-service teachers who view algebra not as a way of thinking but as a solution process in which a series of operations are performed do not use algebraic relations in operations involving basic concepts of algebra, such as algebraic expression and equation. They mostly focus on procedural information and cannot make conceptual explanations regarding these procedural processes (Agarwal, 2006; Black, 2007; Hohensee, 2017; Odumosu & Fisayi, 2018; Welder, 2007; Welder & Simonsen, 2011; Zuya, 2017). Therefore, it is important to determine how pre-service teachers define the two basic concepts of algebra; algebraic expression and equation. In this context, this study aims to determine how middle school mathematics teacher candidates define the concept of algebraic expression and equation, which form the basis of algebra learning.

Method

Case study design was adopted for this study. Case study is a qualitative inquiry which aims to explore a program, an activity, process or one or more cases in depth through multiple data sources (Creswell, 2007). Written exams and interviews were used as data sources. Case studies help researcher to investigate a problem in detail. Therefore, this study was conducted as a case study in order to investigate preservice middle school mathematics teachers' knowledge related to basic algebra concepts.

Participants

Participants of this study were 35 second grade preservice middle school mathematics teachers studying in a state university in Turkey. For selection of participants, convenience sampling and criterion-sampling methods were used. A convenience sample is a group of people who are chosen for a study because they are available (Fraenkel, Wallen & Hyun, 2012). , In convenience sampling, participants are selected from a group of people easy to reach in order to prevent loss of time, money and labour. (Büyüköztürk, Çakmak, Akgün, Karadeniz & Demirel, 2013). The criteria about determining the participant was to take "Introduction to Algebra" and "Linear Algebra" courses, which are important for improvement of preservice teachers' content knowledge of algebra. Taking and succeeding these course were taken as criteria because contents of these courses includes basic algebra concepts and proof of algebraic theorems. Preservice teacher learn background theorem of basic and simple algebra rules. So, at the end of these courses, preservice teachers should have deep understanding of algebraic concepts, theorems and logical explanation of operation with symbols. Also, preservice teachers' algebraic thinking skills are expected to develop after taking these courses.

Data Collection

Research data were collected through written exam and semi-structured interviews. In this context, the research was carried out in two stages. First, written exam that includes two open-ended questions was administered to preservice teachers to determine the definitions of the equation and algebraic expression. To ensure intelligibility of questions, written exam questions were checked and revised in terms of content and language by two experts and one in-service mathematics teacher. Preservice teachers were asked to define both equation and algebraic expression and write an example of each. To ensure reliability and validity, two experts checked content and language of questions. The written exam lasted 15-20 minutes. Second, clarification interviews were conducted with 7 preservice teachers. Interviewing is an effective way of checking accuracy of inferences that the researcher makes through written exam or observations (Fraenkel, Wallen & Hyun, 2012). Interviews help the researcher to understand what participants think. Interviewees were selected based on diversity of definitions. The interviews took on average 5 to 7 minutes. Interviews were recorded with an audio recorder.

Data Analysis

Content and discourse analysis were utilised to discover hidden meaning in preservice teachers' written and spoken expression. For content analysis, the processes of coding data, creating codes and themes, defining and interpreting the findings were followed (Patton, 2002). First, written exam papers of preservice teachers were examined and a summary of their answers to all questions were written up. Then, codes are established based on definition of preservice teachers. Lastly, codes were reviewed and general themes were created. Also, interviews were transcribed and data were analysed. To ensure reliability, all data were coded separately by two researchers. Then, codes and themes were discussed with the participation of third researcher and final version of codes and themes were created. The reliability rate between two coding was determined as 92%. Preservice teachers' definitions of algebraic expression led to creation of three major themes, which are expressions containing unknown, expressions containing equality, and mathematical expressions Also, preservice teachers' definitions of equation collected under two themes: expressions with equality and expressions with unknown.

Findings

Preservice Teachers' Definitions of Algebraic Expression

Preservice teachers' definitions of algebraic equation were collected under 3 different themes. Preservice teachers defined algebraic expressions as expressions containing unknown, expressions containing equality, and mathematical expressions (Table 1).

Table1. Themes related to definition of algebraic expression and number of participants

Theme	Categories	Number of participants
Expression containing unknown	Expression with unknown	12
	Expression with unknown, number and operation	9
	Expression with unknown and without equality	7
Expressions containing equality		4
Mathematical expression		3

Expression containing unknown: 28 of participating preservice middle school mathematics teachers defined algebraic expression as expression containing unknown. Three different codes were established: Expression with unknown, expression with unknown, number and operation, and expression with unknown and without equality. 12 of preservice teachers who defined algebraic equation as expression, which contain unknown refereed only unknown inclusion in their definition. These preservice teachers' definitions were erroneous. Some examples of preservice teachers' definitions are given below.

> *"It is an expression containing an unknown number."(7x-2)(2x+3y)* (PT9)
> *"It is an expression of any number as unknown." (2x+y)* (PT19)

Preservice teachers who defined algebraic expression as expression that contain unknown gave correct examples for algebraic expression. Although preservice teachers' examples of algebraic expression contain fixed terms and operations, they did not mention these concepts in their definitions. In the interviews, preservice teachers elaborated their definitions and stated that algebraic expression contains numbers and operations. During interview, participant 14 made following statement.

> Researcher: *How would you explain algebraic expression?*
> PT14: *For an expression to be an algebraic expression, it must contain an unknown. Actually, x, in 3x, is an algebraic expression own its own.*
> Researcher: *What else can an algebraic expression contain?*
> PT14: *Besides, there is another number in addition to unknown, 3+x is also an algebraic expression.*
> Researcher: *So, what is important?*
> PT14: *The important thing is that something performs an operation with the unknown.*

Nine of preservice teachers who defined algebraic expression as expressions containing unknown mentioned concepts of unknown, number and operation in their definitions. These preservice teachers defined algebraic expression correctly and gave correct examples. They also explained their definitions in detail during interview. Seven of preservice teachers who defined the algebraic expression as expressions containing unknown stated that algebraic expression does not contain equality. Preservice teachers defined algebraic expression to a limited extent without equality. Explanation of Participant 18 during interview is given below as an example.

> PT18: *It is an expression that involves at least one unknown. It is not expected this expression to be equal to a number (His gave (3a+4) as example of algebraic equation in written exam.).*

In the interviews, Participant 10 and Participant 18 emphasized that algebraic expression does not involve equality. It can be said that this situation is a result of preservice teachers' associating algebraic expression with the equation. The statements of Participant 18 reflect this situation.

> Researcher: *You define algebraic expression as expressions with at least one unknown and it is not expected this expression to be equal to a number.*
> PT18: *Yes, that is, there is equality in the equation. When equality is added to the algebraic expression, it becomes an equation.*
> Researcher: *What is the difference between algebraic expression and equation?*
> PT18: *As I said, there is no equality in algebraic expression. There is no value equal to.*

Expression containing equality: 4 of participants defined the algebraic expression as "expressions that contain equality" which is wrong. Below are some examples of their answers to the question in the written exam.

> *"It is expressions that we establish equality between numbers." (3x = 7) (PT22)*
> *"They are numerical expressions containing equality." (2 + 3 = 5) (PT33)*

The preservice teachers' examples for algebraic expression are also incorrect. Two preservice teachers gave an equation as an example of algebraic expression, and other two of them gave number sentence example. In the semi-structured interview conducted with Participant 6, the preservice teacher made the following explanation.

> Researcher: *You gave 2x + 5 = 11 as an example for algebraic expression. Is 2x + 5 algebraic expression in this expression, or is this entire expression (2x+5=11) an algebraic expression?*
> PT6: *Yes, whole of it is an algebraic expression.*

Mathematical expression: 3 of preservice teachers who participated in the study defined algebraic expression as *"mathematical expressions"* in written exams. 2 pre-service teachers defined algebraic expressions as they are *"mathematical expression",* while 1 preservice teacher defined algebraic expression as they are *"mathematical expression, something like rational and root numbers"*. For algebraic expression, they gave 7/3, $\sqrt{5}$ and sin30 = 1/2 as examples. It was determined that algebraic expression definitions and examples of these pre-service teachers were wrong.

Preservice Teachers' Definitions of Equation

Equation definitions of the preservice teachers who participated in the study were collected under two different themes. In this context, preservice teachers defined the equation as expressions with equality and expressions with unknown. Name of themes and distribution of number of preservice teachers are given in Table 2.

Table 2. Themes related to definition of equation and number of participants

Theme	Category	Number of participants
Expression with equality	Expression with unknown	17
	Expression with algebraic expression	6
	Equality of two expression	4
Expression with unknown		8

Expression with equality: 27 preservice teachers defined the equation as "expressions that contain equality". Three different categories were obtained from definitions of preservice teachers: expression with unknown, expression with algebraic expressions and equality of two expressions.

17 of the teacher candidates who defined the equation as expressions with equality defined the equation as expressions with unknown. Preservice teachers used only concepts of equality and unknown in their definitions. In their definitions, they did not explain mathematical meaning of unknown inclusion in an equation. Examples of preservice teachers' definitions are given below.

> *"There is equality in the equation. It contains one or more unknowns."* $2x + 3 = 11$ *(PT12)*
> *"It is the equality of the unknown to a number or to an unknown."* $x + y = 3$, $x-y = 2$ *(PT1)*

In the interviews, preservice teachers could not elaborate their definitions. On the other hand, examples given by preservice teachers in wren exam were correct examples of the equation. However, the preservice teachers gave examples of equations where the unknown is on the left side of the equation. In the interviews, it was seen that preservice teachers thought that the left side of the equal sign is the problem and it should be equal to a result. A part of interview conducted with Participant 16 in relation to this situation is presented as an example.

> Researcher: *You explained the equation as an expression consisting of one or more unknowns, which is equal to a value. You have given examples of $3x + 6 = 20$ and $x = 2$. Should there be a known value on the right side of equality?*
> PT16: *Yes, so it has to be equal to a result. However, I am thinking now, there may be an unknown expression on the right side of equality.*

6 of the 27 preservice teachers who defined the equation as expressions with equality used the concept of algebraic expression in their definitions. 4 preservice teachers defined equation as equality of algebraic expression to a number and 2 pre-service teachers defined equation as equality of two algebraic expression.

> *"The equation is equality. It is equality that an algebraic expression is equal to a known value."* $4x + 5 = 9$ *(PT27)*
> *"It is the equality of two algebraic expressions."* $5x + 48 = 12$ *(PT35)*

In the interviews, it was seen that preservice teachers who defined the equation as the equality of two algebraic expressions, wrote examples, which is not congruent to their definitions. A part of the interview with Participant 35, which shows that the definition of the equation in the minds of preservice teachers is not clear, is presented as an example.

> Researcher: *You defined the equation as the equality of two algebraic expressions. You wrote the expression, $5x + 48 = 12$, as an example of equation. Can you explain?*
> PT35: *There can be an unknown on the right side or equality or it can contain no unknown. It is important to have equality.*
> Researcher: *Is $1 + 2 = 3 + 0$ an equation?*
> PT35: *Of course, it is an equation.*

27 of preservice teachers who defined the equation as expressions with equality, 4 used only the concept of equality in their definitions. Examples of preservice teachers' definitions are presented below.

> *"They are expressions that are equal to each other."* $2 + 3 = 1 + 4$ *(PT23)*
> *"It is the equality of two different expressions."* $3x-5 = 10$ *(PT26)*

2 of preservice teachers gave correct examples for the equation. In interviews conducted with these preservice teachers, they stated that the equation contains unknown. This shows that preservice teachers' definition of equations is missing. The other 2 preservice teachers wrote down number sentences as examples of equation. It is clarified in the interview that these preservice teachers thought that it was sufficient to have equality for the equation. The dialog below reflects this situation.

> Researcher: *You have explained that the equation shows expressions that are equal to each other. You gave $2 + 3 = 1 + 4$ as example. Can you explain?*
> PT: *There is something known and proven. In my opinion, the equation is equality.*
> Researcher: *Is $2x = 6$ an equation?*
> PT: *Yes, it has equality. If there is equality, it is the equation.*
> Researcher: *Is it sufficient to have equality to be an equation?*
> PT: *Yes.*

Expressions with unknown: 8 preservice teachers defined equation as "expressions that contain unknown". Although these preservice teachers did not use the concept of equality in their definitions, examples they gave contain equality and they gave the correct examples for the equation. In interviews conducted with 2 preservice teachers, they stated that the equation contains equality and the unknown is found thanks to equality. Examples of preservice teachers' definitions are presented below.

> *"Expressions that contain at least one unknown."* $2x + 3 = 2y$ *(PT24)*
> *"Expressions with one or more unknowns."* $2x + 13 = 47$ *(PT21)*

Discussion and Conclusion

Mathematical definitions have an important role in mathematics teaching process. Defining mathematical concepts correctly is important for development of conceptual understanding (Edwards & Ward, 2004; Morgan, 2005; Mosvold & Fauskanger, 2013). Conceptual understanding is important for students' learning as well as for defining teachers' content knowledge competence (Morgan, 2005). Teachers' content knowledge is a determining factor of their teaching practices (Ball et al.,, 2008). Thus, it is very important to define basic algebra concepts correctly during algebra teaching process. In this context, this research focuses on how middle school mathematics teacher candidates define the concept of algebraic expression and equation as a dimension of content knowledge.

In the middle school mathematics curriculum (MEB, 2018), algebraic expression is defined as "expressions that contain at least one unknown and one operation". The algebraic expression definitions of the participating 35 preservice teachers classified under three themes: expression containing unknown (28 preservice teachers), expression containing equality (4 preservice teachers) and mathematical expression (3 preservice teachers). It is seen that definitions of preservice teachers (28 preservice teachers) who define the algebraic expression as expressions that contain unknown are insufficient. Most of these preservice teachers (12 preservice teachers) used only expressions containing unknown statements in their definitions. Chalouh and Herscovics (1988) stated that algebraic expression is often defined as expressions that involve variables. It was emphasized that such formal definition is not sufficient and meaningful for students to understand the concept. Defining algebraic expression only to a limited extent (defining algebraic expressions limited to variables) can make it difficult for students to understand other mathematical concepts (coefficient, term, constant term, and similar term) (MEB, 2018) contained in algebraic expression. Failure to understand these concepts correctly by students will make it difficult for them to learn advanced algebra topics (Kieran, Pang, Schifter & Ng, 2016; Stephens, Ellis, Blanton & Brizuela, 2017; Tirosh, Even & Robinson, 1988). For example, it is necessary to understand terms, coefficients and similar term concepts in order to learn the process of operations with algebraic expressions. For this reason, it is important to correctly express these concepts when defining algebraic expression. On the other hand, defining algebraic expression only as expressions containing unknown may cause students to develop different misconceptions that other mathematical concepts, which contain unknown, are also algebraic expressions.

Participants who defined algebraic expression as expressions containing unknown stated that algebraic expression also contains numbers and operations in their definitions (9 preservice teachers). This definition coincides with the definition given in the middle school mathematics curriculum (MEB, 2018). These preservice teachers gave correct examples for algebraic expression. 7 preservice teachers stated that algebraic expression does not involve equality. The reason why the preservice teachers emphasized that algebraic expression does not involve equality is due to their understanding that relation between algebraic expression and equation is operation-oriented. Interview data also support this situation. Two preservice teachers stated that there was no equality in algebraic expression, but it turns out to an equation when equality was added to algebraic expression. These explanations showed that preservice teachers explained the difference between algebraic expression and equation as operation-oriented (there is no equality in algebraic expression). They did not make any conceptual explanations regarding the mathematical meaning of why algebraic expression does not involve equality. Researches also show that preservice teachers explain basic algebra concepts with a focus on letters and symbols and focus on the operational process (Attorps, 2005; Black, 2007; Welder & Simonsen, 2011). Understanding the underlying mathematical meaning of operations is important for meaningful learning (Van de Walle, Karp, Bay-Williams, 2013).

4 pre-service teachers incorrectly defined algebraic expression as expressions with equality. These preservice teachers mentioned only the concept of equality in their definitions. 2 teacher candidates wrote down equation (e.g. $3x = 7$) and two teacher candidates wrote down number sentence (eg $2 + 3 = 5$) as an example of algebraic expression. 3 preservice teachers defined algebraic expression as mathematical expressions. These definitions are not an explanatory definition and examples given by the preservice teachers for algebraic expression ($7/3$, $\sqrt{5}$ and $\sin 30 = 1/2$) are incorrect. Based on these findings, it is understood that preservice teachers have wrong knowledge about algebraic expression.

In the middle school mathematics curriculum (MEB, 2018), the equation is defined as "equality that involve unknown and true for some values of the unknown." The majority of participants (28 preservice teachers) defined the equation as expressions with equality. 17 of these pre-service teachers defined equation as equality, which contain unknown. Preservice teachers gave correct examples for the equation. On the other hand, the teacher candidates' definitions do not exactly match up to the definition given in the middle school mathematics curriculum (MEB, 2018). In addition, since not every expression (eg identity) containing equality and unknown is an equation, preservice teachers' definitions are incorrect. These findings show that pre-service teachers define the concept of equation with a focus on letters and symbols. In the interviews, teacher candidates could

not elaborate their definitions. They could not make conceptual explanations about the meaning of the equation concept. The examples given by preservice teachers for the equation also support these findings. Preservice teachers think that the right side of equal sign is result while right side is the problem (Van de Walle, Karp, Bay-Williams, 2013). For this reason, they gave examples of equations where the unknown is on the left side of the equal sign. This shows that pre-service teachers consider the equation as computing the amounts and do not have knowledge about the relational meaning of equality.

4 pre-service teachers defined equation as the equality of algebraic expression to a number and 2 pre-service teachers defined as the equality of two algebraic expressions. The definitions of teacher candidates do not match the definition of the equation (MEB, 2018). These definitions do not meet the mathematical meaning of the equation concept. This shows that pre-service teachers think that the difference between equation and algebraic expression is existence of equality. The statements of prospective teachers that there is no equality in algebraic expression support this situation. 4 pre-service teachers defined the equation, using only the concept of equality, as equality of expressions. 8 of preservice teachers defined the equation incorrectly as expressions with unknown. These findings show that pre-service teachers focus on operational processes while defining the equation. Similar findings were obtained in other studies in the literature (Attorps, 2003; Tanisli & Köse, 2013; Yıldız, 2016). In the study carried out by Yıldız (2016) with three secondary school mathematics teachers, teachers defined the concepts of algebraic expression and equation with a focus on letters, operations and equality. They stated that the difference of the equation from the algebraic expression is the equation contains equal sign.

Recommendations

The results obtained from this research show that majority of participating preservice middle school mathematics teachers defines algebraic expression and equation wrongly. An important result of the study is that preservice teachers define algebraic expression and equation based on letters, operations and symbols. Similarly, teacher candidates did not make any explanations about the mathematical meaning of the concepts in the interviews and they focused on the operational processes. Similar results were obtained in the studies conducted in the literature (Attorps, 2003; Stephens, 2008; Strumps & Bishops, 2002; Yıldız, 2016). Ball (1990) pointed out that teachers should know the underlying meanings of operational processes, and that concept knowledge is an important part of teachers' content knowledge competencies. These results indicate that preservice education program should include applications that will enable preservice teachers to learn the definition and meaning of mathematical concepts. In the mathematics teaching courses, a teaching environment should be created in which preservice teachers can examine and discuss mathematical concepts not only theoretically but also practically.

In this study, the data were obtained from the written answers given by the preservice teachers to the questionnaire. One-to-one interviews were held with only some of the teacher candidates. In future studies, more in-depth findings can be obtained by conducting extensive interviews with fewer participants and using different data collection tools (focus-group interviews, observation of practice and internship courses, etc.).

References

Agarwal, S. (2006). *The nature of pre-service secondary mathematics teachers' knowledge of mathematics for teaching of functions* (Unpublished doctoral dissertation). State University of New York, Buffalo, United States.

Akkan, Y., Baki, A. & Çakıroğlu, Ü. (2012). 5-8. sınıf öğrencilerinin aritmetikten cebire geçiş süreçlerinin problem çözme bağlamında incelenmesi [Examination of the 5th – 8th grade students' transition process from arithmetic to algebra with regard problem solving]. *Hacettepe University Journal of Education, 43*, 1-13.

Attorps, I. (2003). Teachers' images of the 'equation' concept. *European Research in Mathematics Education*, *3*, 1-8.

Ball, D. L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *Elementary School Journal, 90*(4), 449-446.

Ball, D.L., Lubienski, S. & Mewborn, D. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), Handbook of Research on Teaching (pp. 433-456). New York: Macmillan.

Ball, D. B., Thames, M. H. & Phelps, G. (2008). Content knowledge for teaching: What makes it special?. *Journal of Teacher Education, 59*(5), 389-407.

Black, D. J. W. (2007). The relationship of teachers' content knowledge and pedagogical content knowledge in algebra, and changes in both types of knowledge as a professional development (Unpublished doctoral dissertation). Auburn University, Auburn.

Brizuela, B. M. (2016). Variables in elementary mathematics education. *The Elementary School Journal*, *117*(1), 46-71.

Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2013). Bilimsel araştırma yöntemleri (15. Baskı) [Scientific research methods (15th ed.)]. Ankara: Pegem Akademi.

Charalambous, C. Y., Hill, H. C., Chin, M. J. & McGinn, D. (2019). Mathematical content knowledge and knowledge for teaching: exploring their distinguishability and contribution to student learning. *Journal of Mathematics Teacher Education*, 1-35.

Christou, K. P. & Vosniadou, S. (2012). What kinds of numbers do students assign to literal symbols? Aspects of the transition from arithmetic to algebra. *Mathematical Thinking and Learning*, *14*(1), 1-27.

Copur-Gencturk, Y. (2015). The effects of changes in mathematical knowledge on teaching: A longitudinal study of teachers' knowledge and instruction. *Journal for Research in Mathematics Education*, *46*(3), 280-330.

Chalouh, L. & Herscovics, N. (1988). Teaching algebraic expressions in a meaningful way. In A. F. Coxford (Eds.), *The ideas of algebra, K-12.* (pp. 33-42). Reston, VA: National Council of Teachers of Mathematics.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). London: Sage.

Dede, Y. & Argün, Z. (2003). Cebir, öğrencilere niçin zor gelmektedir? [Why do students have difficulty with algebra?]. *Hacettepe University Journal of Education, 24,* 180-185.

Dreher, A., Lindmeier, A., Heinze, A. & Niemand, C. (2018). What kind of content knowledge do secondary mathematics teachers need?. *Journal für Mathematik-Didaktik*, *39*(2), 319-341.

Edwards, B. S. & Ward, M. B. (2004). Surprises from mathematics education research: Student (mis) use of mathematical definitions. *The American Mathematical Monthly*, *111*(5), 411-424.

Fennema, E., Sowder, J. & Carpenter, T. P. (1999). Creating classrooms that promote understanding. In E. Fennema, and T. A. Romberg (Eds.). *Mathematics classrooms that promote understanding* (pp. 185-199). Hilsdale, NJ: Lawrence Erlbaum Associates.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2013). *How to design and evaluate research in education* (8th ed.). New York: McGraw-Hill.

Hill, H. C. & Ball, D. L. (2004). Learning mathematics for teaching: Results from California's Mathematics Professional Development Institutes. *Journal for Research in Mathematics Education, 35*(5), 330-351.

Hill, H. C., Rowan, B. & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371-406.

Hohensee, C. (2017). Preparing elementary prospective teachers to teach early algebra. *Journal of Mathematics Teacher Education*, *20*(3), 231-257.

Kieran, C., Pang, J., Schifter, D. & Ng, S. F. (2016). *Early Algebra. Research into its nature, its learning, its teaching.* Hamburg: Springer Open.

Ma, L. (2010). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Routledge.

Milli Eğitim Bakanlığı (2018*). Matematik dersi (1-8. sınıflar) öğretim programı [Mathematics curriculum (1st - 8th grade)].* Ankara: Milli Eğitim Bakanlığı.

Morgan, C. (2005). Word, definitions and concepts in discourses of mathematics, teaching and learning. *Language and Education*, *19*(2), 102-116.

Mosvold, R. & Fauskanger, J. (2013). Teachers' beliefs about mathematical knowledge for teaching definitions. *International Electronic Journal of Mathematics Education*, *8*(2-3), 43-61.

Odumosu, O. & Fisayi, A. (2018). Teachers' content and pedagogical knowledge on students' achievement in algebra. *International Journal of Education and Research*, *6*(3), 83-94.

Patton, M. Q. (2002). Qualitative research and evaluation methods. Thousand Oaks, CA: Sage Publications.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harward Educational Review, 57*(1), 1-21.

Stephens, A. C. (2008). What "counts" as algebra in the eyes of preservice elementary teachers?. *Journal of Mathematical Behavior, 27,* 33-47.

Stephens, A. C., Ellis, A. B., Blanton, M. & Brizuela, B. M. (2017). Algebraic thinking in the elementary and middle grades. *Compendium for Research in Mathematics Education*, 386-420.

Stump, S. L. & Bishop, J. (2002). Preservice elementary and middle school teachers' conceptions of algebra revealed through the use of exemplary curriculum materials. In D. S. Mewborn, P. Sztajn, D.Y. White,

H. G. Wiegel, R. L. Bryant, and K. Nooney (Eds.). *Proceedings of the twenty-fourth annual meeting of the international group for the psychology of mathematics education* (pp. 1903–1914). Columbus, OH: ERIC.

Tanisli, D. & Kose, N. Y. (2013). Preservice Mathematics Teachers' Knowledge of Students about the Algebraic Concepts. *Australian Journal of Teacher Education*, *38*(2), 1-18.

Tchoshanov, M., Cruz, M. D., Huereca, K., Shakirova, K., Shakirova, L. & Ibragimova, E. N. (2017). Examination of lower secondary mathematics teachers' content knowledge and its connection to students' performance. *International Journal of Science and Mathematics Education*, *15*(4), 683-702.

Tirosh, D., Even, R. & Robinson, N. (1998). Simplifying algebraic expressions: Teacher awareness and teaching approaches. *Educational Studies in Mathematics, 35*(1), 51-64.

Van de Walle, J. A., Karp, K. S. & Bay-Williams, J. M. (2013). *Elementary and middle school mathematics: Teaching developmentally* (8th ed.). New Jersey: Pearson Education.

Weinberg, A., Dresen, J. & Slater, T. (2016). Students' understanding of algebraic notation: A semiotic systems perspective. *The Journal of Mathematical Behavior*, *43*, 70-88.

Wasserman, N. H. (2016). Abstract algebra for algebra teaching: Influencing school mathematics instruction. *Canadian Journal of Science, Mathematics and Technology Education*, *16*(1), 28-47.

Welder, R. M. (2007). *Preservice elementary teachers' pedagogical content knowledge of prerequisite algebra concepts* (Unpublished doctoral dissertation). Montana State University, Montana.

Welder, R. M. & Simonsen, L. M. (2011). Elementary teachers' mathematical knowledge for teaching prerequisite algebra concepts. *Issues in the Undergraduate Mathematics Preparation of School Teachers, 1*.

Yıldız, P. (2016). Middle mathematics teachers' knowledge for teaching algebra: A multiple case study (Unpublished doctoral dissertation). Hacettepe University, Ankara.

Zuya, H. E. (2017). Prospective teachers' conceptual and procedural knowledge in mathematics: The case of algebra. *American Journal of Educational Research*, *5*(3), 310-315.

The Alignment between the Official Curriculum and the Taught Curriculum: An Analysis of Primary School English Curriculum

Eda Karabacak[1], Dilruba Kürüm-Yapıcıoğlu[2]
[1] Kütahya Dumlupınar University
[2] Anadolu University

The Alignment between the Official Curriculum and the Taught Curriculum: An Analysis of Primary School English Curriculum[*]

Eda Karabacak[1][**], Dilruba Kürüm-Yapıcıoğlu[2]
[1]Kütahya Dumlupınar University
[2]Anadolu University

Abstract

The aim of this research is to analyze the alignment between the official primary school English curriculum and the taught curriculum. In this mixed method research, concurrent equal status design was adopted. Questionnaires, semi-structured interviews, teachers' reflection diaries, and research diaries were used to collect the data. The participants were the teachers who were teaching English in primary schools in Kütahya in the 2016-2017 academic year. The quantitative data of the study were analyzed by using descriptive statistics analysis and the qualitative data were analyzed with inductive analysis. The result of this research showed that the English teachers can align the curriculum in terms of its learning outcomes and content. However, they cannot completely align the curriculum in terms of its teaching-learning process and evaluation. Moreover, the study found out the classroom teachers can partially align the curriculum and the multigrade classroom teachers cannot ensure the alignment between the official and the taught curriculum to a considerable extent. The lack of technological equipment and materials, challenges of teaching in a multigrade classroom, teachers' professional competencies, and beliefs are among the factors affecting the alignment.

Key words: Curriculum alignment, Curriculum fidelity, Multigrade classrooms, Taught curriculum, Written curriculum

Introduction

Education, which is one of the most discussed concepts in the literature, can be defined as the process of providing experiences to students that will change their behavior in the desired direction, and help them discover and improve themselves. Varış (1994) states that in order to achieve the desired development in the behavior of the individual, the purpose of education, the teaching-learning process and the outcomes should be determined, and these can be achieved through the curriculum. This is to say that the curriculum is crucial to achieve what is intended with education. According to Bobbitt (1918, p. 43), curriculum is the learning experiences that the school consciously plans to complete the abilities of individuals. According to Ertürk, the curriculum guides the teachers by presenting the information they need to know on how to teach the students and to carry out the teaching-learning process effectively (1979, p. 22-23). On the other hand, Wiles and Bondi (2011, p. 1) state that the curriculum includes plans, objectives, activities or learning outcomes presented in different environments and in different ways, but they also relate the curriculum to the textbook, syllabus, teacher guide or learning package.

The wide scope of the concept of the curriculum and the diversity of the meanings attributed to it have led to the emergence of different types of curriculum. There are several classifications of curriculum types in the literature, one of these belongs to Glatthorn (2000), who indicates that there is a big difference between the written curriculum and the taught curriculum. Glatthorn (2000, p. 83-84) refers to seven different types of curricula. The first one is the recommended curriculum which is recommended by scholars. The second one is the written curriculum which can be defined as the curriculum prepared by the government to be used in state

schools. This type of curriculum is also referred as the official curriculum in the literature. He defines the taught curriculum as the curriculum the teacher actually teaches in the classroom and the learned curriculum as the curriculum that students actually learn. The next type of curriculum he mentions is the assessed curriculum. This includes what is assessed in the tests prepared by teachers and the government. The supported curriculum includes the textbooks and other resources that support the curriculum. As the last type of curriculum, he mentions the hidden curriculum which is the non-intended curriculum that students learn from their environment, or the policies of the school.

Glatthorn and Jailall (2009, p. 111) state that there are gaps between different types of curriculaum, and this highlights the process of the curriculum alignment. Curriculum alignment not only refers to the alignment between the dimensions of the curriculum but also refers to the alignment between different curriculum types (Squires, 2009, p. 4-5). English (1992, p. 63) defines curriculum alignment as the alignment between the curriculum or the content of a source, such as coursebook used instead of the curriculum and the content of the exams applied. Glatthorn and Jailall (2009, p. 111) state that teachers usually use the written curriculum to look at what they will teach at the beginning of the year and they care about the tested curriculum more. For this reason, they state that the alignment between the written curriculum and the taught curriculum is weak, but the alignment between these two types of curricula is important for the success of the whole curriculum. In order to ensure the success of the curriculum, it is important to find out whether the curriculum is changing in the process of implementing it in the classroom, or in other words whether there is an alignment between the written curriculum and the taught curriculum. While Glatthorn and Jailall (2009) use the term written curriculum, some researchers use the term official curriculum to define the curriculum prepared by the government to be used in state schools. In this study, the official curriculum was adopted as it is thought to reflect the concept of the curriculum analyzed in this study better.

Curriculum fidelity is one of the terms used to analyze if there is an alignment between the official and the taught curriculum (Furtak et al., 2008, p. 362; Dusenbury et al., 2003, p. 240). Curriculum fidelity can be defined as the implementation of the official curriculum by the teacher or the practitioner of the curriculum by staying faithful to its principals (Bümen, Çakar & Yıldız, 2014, p. 205). Penuel, Phillips and Harris (2014), on the other hand, see the alignment between the official curriculum and the taught curriculum from two different perspectives, actor-oriented and integrity. They express the integrity perspective as teachers' degree of alignment of the materials with the objectives and principles of the curriculum. The actor-oriented perspective explains how teachers interpret the directions on how to use the materials and how they adapt the materials to their classroom. Furthermore, they argue that curriculum alignment is the process of determining the difficulties teachers experience during the curriculum implementation and the reasons behind the changes they make.

Curriculum explains the aims, content, activities to be used in the teaching-learning process and how they are organized. However, although there is a detailed official curriculum, the different perspectives of teachers might cause differences in implementation even among the teachers who are using the same curriculum (Walker & Soltis, 2004, p. 1-2). Each teacher implements the curriculum by considering the characteristics of their students, the physical conditions of the classroom, the materials they own, and their own personal preferences (Livingstone et al., 1986, p. 2-7). This causes teachers to make changes in the official curriculum as they implement it, and that affects the success of the curriculum. Morgan and Xu (2011, p. 3-4) state that most studies on the curriculum reveal that the intended curriculum is different from the implemented curriculum. The reasons for this are that teachers are not renovative, they have difficulty in making changes in their teaching styles, and they think the characteristics of their students and the curriculum do not overlap. This causes the curriculum, which is prepared by the state to be applied jointly in all schools, to transform into the taught curriculum due to the changes teachers make while implementing it. The transformation of teachers' official curriculum into the taught curriculum can adversely affect the success of the curriculum.

The changes teachers make while implementing the curriculum affect the success of it. For this reason, finding out what teachers change when they are implementing the curriculum and why they make certain changes is important to make sure the desired outcomes are being reached. In Turkey, there are some challenges faced while implementing and ensuring the success of some curricula, and the primary school English curriculum is one of them. As a result, a lot of studies on teaching English and the English curricula have been carried out over the years. Teaching English in Turkey has gained importance since the Second World War (Cem, 1998, in Demirel, 2003, p. 7) and with the foundation of TED college in the Republic's first years, English was taught partially or completely (Çelebi, 2006, p. 287- 289). With the transition to eight years of compulsory education in 1977, it was decided to start teaching English from the 4th grade for five years; student-centered English curricula were also introduced in 1997 (Demirel, 1999, p. 27). The Ministry of National Education launched a new primary school English curriculum in 2006, which adopted a student and teacher-centered approach

(MoNE, 2006, p. 2). In 2013, with the transition to the 4 + 4 + 4 education system, a new primary school English curriculum was introduced, and it was decided to teach English starting from the 2nd grade (MoNE, 2013). This curriculum emphasizes language learning as a process that not only consists of grammatical structures, but also that enables constant communication and interaction through the use of the language. Communicative approach, learner autonomy and intercultural awareness are defined as the basis of the curriculum. The primary focus of the curriculum at the primary school level is listening and speaking skills, reading and writing skills, on the other hand, are given a very limited focus (MoNE, 2013). For each grade level, there are ten suggested themes which reflect students' lives. In some of these themes, there are elements of traditions of different countries in order to create intercultural awareness. In the teaching-learning process, it is suggested that teachers use different types of materials instead of coursebooks. As for the evaluation, the use of self-assessment, written and oral exams, homework, projects, and the European Language Portfolio is suggested (MoNE, 2013).

Developing reading and writing skills is not an aim in the primary school English curriculum; therefore, students should not have notebooks, and reading and writing activities should not exceed ten words in one lesson hour. The focus is on communication at primary school level, so students should be encouraged to keep the conversation going even if they make mistakes. Therefore, when students make mistakes, the teachers are advised not to correct their mistakes directly but to use the structures correctly to help students learn. Places such as classrooms, parks, and school gardens in students' immediate surroundings were chosen as the context of the teaching-learning process (MoNE, 2013).

The primary school English curriculum, which was enacted in 2013, was updated by taking the opinions of the stakeholders into account. With the Law No. 15 of 2018 of the Ministry of National Education, it was decided to start implementing the new primary school English curriculum as of the 2018-2019 academic year. Although the general structure and approach of the curriculum stay the same, the learning outcomes related to intercultural awareness are omitted in the new curriculum. One other prominent difference between the new curriculum and the previous one is that the new one emphasizes teaching values such as friendship, honesty, and patriotism as part of the language teaching process (MoNE, 2018).

The English language curriculum development studies in Turkey started in 1997, and continued in 2006, 2013 and 2018. Over the years, the philosophy of the curriculum, its objectives, content, the teaching-learning process, the assessment methods, and the roles of the teachers have changed. To illustrate, while evaluation activities used to be limited to pen and paper exams, alternative evaluation techniques have gained importance. The teacher-centered curriculum has become student-centered. In order to ensure the success of the changes in the curriculum, it is important to determine how the official curriculum is implemented in the classroom. Therefore, this study aims to reveal the level of the alignment between the official curriculum and the taught curriculum. The questions to be answered in this research are as follows:

1. What are the opinions of the teachers regarding the implementation of the official primary school English curriculum?
2. What are the problems that teachers experience while implementing the official primary school English curriculum and their suggestions to eliminate these problems?

Methodology

Research Model

In this study, a mixed research method was used to collect comprehensive data by combining the strengths of qualitative and quantitative methods. Creswell (2009, p. 14) states that all methods have their strengths and weaknesses, so collecting both quantitative and qualitative data by applying the mixed method will balance the weaknesses of the two methods. Johnson and Christensen (2014, p. 434-435) state that timing and paradigm should be considered when deciding on the pattern of a mixed method research. Whether the qualitative and quantitative methods will be used concurrently or sequentially is about the timing and whether the qualitative and quantitative methods will be used equally, or one will be more dominant than the other is related to the paradigm. In this study, qualitative and quantitative methods were used equally, and qualitative and quantitative data collection tools were used concurrently. As a result, this study is based on concurrent equal status design.

The Research Process

The research process started in June 2016 with the literature review to form the basis of the research and prepare the data collection tools. The number of primary schools and teachers were taken by contacting the Kütahya Provincial Directorate of National Education to choose the participants of the research. Then, these schools were called, and the exact number of the teachers was finalized. During the calls, it was found that classroom teachers carry out English lessons in some schools and there are multigrade classrooms in some of these schools. Later, data collection tools were prepared, presented to the experts' opinions, and piloted.

The collection of research data started with the reflection diaries. At the end of the first term, questionnaires were conducted, and semi-structured interviews with volunteer teachers were made. To collect data on the implementation of the curriculum during the second term, the teachers started writing the reflection diaries at the beginning of the second term. At the end of the term, first the questionnaires, then the semi-structured interviews were conducted. After collecting all the data, first the questionnaires, then the semi-structured interviews and then the reflection diaries were analyzed, and the findings were revealed. Afterwards, findings which were written separately for each data collection tool were organized and combined under the themes.

Participants

As it was necessary to collect data for a long time and the researcher had a full-time job, contacting the schools and teachers easily was prioritized. As a result, the participants were determined as English teachers and classroom teachers who were teaching English in primary schools in the city center of Kütahya. The participants are presented in Table 1. After deciding on the schools from which to collect the data, the schools were visited, the principals and the teachers were informed, and the teachers were asked to participate in the research voluntarily.

Maximum diversity was considered when determining the teachers who would write the reflection diaries. The criteria to choose them were that they were English teachers, the years of their professional experience were different, and the school they worked in were at different socioeconomic levels. In the first term, fifteen teachers volunteered to write the reflection diary, and ten teachers completed their diaries at the end of the term. At the end of the fall term, three teachers decided to leave the study but the other seven continued to write their diaries along with six new teachers. At the end of the second term, ten of the thirteen teachers completed their diaries.

The teachers to answer the questionnaires were determined by criterion sampling. The criteria to choose them were that the teachers were carrying out primary school English lessons and were not the participants of the reflection diaries. Outside of the teachers who wrote the reflection diaries in the first term, there were 31 English teachers and 28 of them answered the questionnaires. In addition, four classroom teachers in two schools who were teaching English answered the questionnaires. In the second term, there were 32 English teachers except the teachers who wrote the reflection diaries and 27 of these teachers answered the questionnaires. Four classroom teachers who answered the first term questionnaires also answered the second term questionnaires.

Criterion sampling was used when determining the participants of the semi-structured interviews. One of the criteria to choose them was that there were teachers from the three groups; English teachers, classroom teachers and multigrade classroom teachers who were teaching English in primary schools. The other criterion was that the schools where the teachers worked were at different socioeconomic levels. During the school visits, voluntary participation of teachers was requested. From the teachers who agreed to do an interview, three English teachers, two 4th grade classroom teachers and three multigrade classroom teachers were chosen based on the criteria. Semi-structured interviews were held with these eight teachers at the end of the first and the second term. The details about the participants can be seen in Table 1.

Table 1: *Information about the participants*

		Participants of the reflection diary		Participants of the questionnaire		Participants of the semi-structured interview	
		First term	Second term	First term	Second term	First term	Second term
Branch	English teacher	10	10	28	27	3	3
	Class teacher	-	-	4	4	5	5
Number		10	10	32	31	8	8

Data Collection

In this study, reflection diaries, questionnaires, semi-structured interviews, and research diaries were used to collect data. Data collection tools were used simultaneously in accordance with the design of the research. In Figure 1, information about the order in which the data collection tools were used is presented.
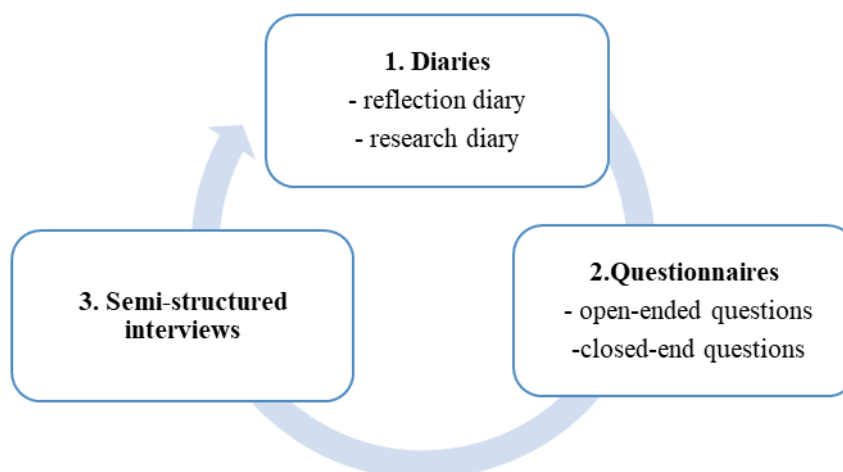

Figure 1: *Data collection process*

As can be seen in Figure 1, the data collection process was circular. It started with the reflection diaries in the first term. Then it continued with questionnaires and semi-structured interviews. Throughout this process, the data were also collected through research diaries. In the second term, the data were collected in the order followed during the first term.

Reflection Diaries

The reflection diaries were used to determine the opinions of the teachers about the implementation of the primary school English curriculum, and problems and suggestions regarding its implementation process. First, the teachers were asked to write their experiences by considering the learning outcomes, content, teaching-learning process, and evaluation dimensions of the curriculum. However, because the data obtained were not sufficient, a reflection diary form was prepared to guide the teachers and receive more detailed information. This form started with the learning outcomes of each unit followed by six questions. The questions were about the learning outcomes, the content, text and activity types, the teaching-learning process, the projects and the activities in the assessment section, the problems faced in the process and the suggestions for solutions. The reflection diary forms for the first term were delivered in a file to the teachers who wanted hard copies of them and sent by email to those who asked for soft copies in October and November. The diaries completed were collected from the teachers in January and February. In the second term, the reflection diary forms were delivered to the teachers in February and the completed diaries were collected in June.

Questionnaires

In order to find answers to the research questions, a questionnaire was prepared for each grade level by taking the 2nd, 3rd, and 4th grade curriculum into consideration (Karabacak, 2018, p. 225-233). There are ten units for each grade level in the curriculum, and the teachers stated that they planned five of these units in the first term and five in the second term in their annual plans. There were a lot of questions in the questionnaire and in the case that they were answered at the end of the year, the teachers might have forgotten the details about what they did during the first term. As a result, two separate questionnaires were developed. The first term questionnaires include five sections which have items about the learning outcomes of the first five units, suggestions for the teaching-learning process and evaluation dimensions, suggested text and activity types, and personal information. The second term questionnaires include a section about the general objectives of each grade level in addition to the sections in the first term questionnaires. After preparing the questionnaires, the opinions of five faculty members working in two different universities in the field of Curriculum Development and Instruction were obtained on the questionnaire forms. Then, they were piloted by ten volunteer English

teachers working in different cities and primary schools. As the questionnaires included items related to all the units to be covered during a term, they were conducted in the last week of each term in order not to miss any data. The fall term questionnaires were conducted in January, in the last week of the 2016-2017 fall term. The spring term questionnaires were conducted in June, in the last week of the second term.

Semi-structured Interviews

For the semi-structured interviews separate forms were created for English teachers, classroom teachers and multigrade classroom teachers. The reason for this is that while English teachers are teaching in all of the 2nd, 3rd and 4th grades, classroom teachers are teaching in one class and the multigrade classroom teachers can teach in two or four different grades at the same time. The interview form was prepared in five sections considering the curriculum. The form includes questions such as teachers' experiences of teaching English, how they use the curriculum, questions about the learning outcomes, content, teaching-learning process and evaluation dimensions of the curriculum, problems faced during the implementation of the curriculum and their suggestions for solutions. The same questions were used for English teachers and classroom teachers, but the questions of three grade levels were included in the same form for English teachers considering the possibility of English teachers teaching three grade levels. As classroom teachers conduct classes in only one grade level, a separate form has been prepared for each grade level. Before preparing the interview form for multigrade classroom teachers, the researcher did an interview with a volunteer multigrade classroom teacher to have a clear picture of how the English lessons are conducted in multigrade classrooms. After that, the interview form was prepared based on the information taken from the teacher in addition to the principals of the curriculum. Next, the forms prepared were presented to the opinion of a primary school English teacher, who was writing her dissertation in the field of Curriculum Development and Instruction and had taken a qualitative research methods course. Then, the interview form was piloted with a primary school English teacher.

Semi-structured interviews on the implementation of the curriculum during the 2016-2017 fall term were held in February and March. Interviews for the spring term were conducted in June. Some of the teachers who had already been interviewed were interviewed again during the term to get additional information. These interviews were held between February and June. Interviews lasted a minimum of 5 minutes 13 seconds and a maximum of 49 minutes 56 seconds.

Research Diaries

The researcher who collected the data kept diaries to follow her progress on the research and plan her next steps from the time she started her preparations for this research until the research was completed. The diaries also include impressions gained before and after interviews with teachers, additional information given by the teachers, and information obtained from telephone interviews. The data in the diaries were used as supporting data when necessary.

Data Analysis

The analysis of the data collected during the research started with quantitative data, then qualitative data were analyzed. Quantitative data obtained from the questionnaires were analyzed by descriptive statistical analysis. The purpose of descriptive statistics is to describe data with techniques such as central tendency measurements, percentage, and frequency distributions (Büyüköztürk, 2018, p. 5). Quantitative data obtained from this research were analyzed by the researcher using SPSS, percentage and frequency values were taken for each item. To make the findings reader-friendly, the findings of the quantitative data were presented in four categories: a small minority of the participants (≤25), a minority of the participants (26%-49%), a majority of the participants (51%-75%), a great majority of the participants (≥76).

After the analysis of quantitative data, the qualitative data obtained from the reflection diaries and the semi-structured interviews were analyzed by thematic analysis. Thematic analysis is the coding of data and creating categories and themes by trying to find patterns between these codes (Glesne, 2013, p. 259-260). In this study, one of the researchers first transferred the semi-structured interview data, the reflection diaries data, and the qualitative data in the questionnaires to Nvivo. Then, the semi-structured interview data were coded by the researchers using Nvivo. Expert opinion on the congruence of the codes was received from a Ph.D. student who was working as an instructor of English and was writing her dissertation in the field of Curriculum Development

and Instruction. After necessary changes were made in line with the expert opinion, themes were created. For the congruence of the themes, expert opinion was received from a faculty member working in the field of Curriculum Development and Instruction. Following the expert opinion, the necessary changes were made, and the analysis of the semi-structured interviews was completed.

The data obtained from the reflection diaries were coded and themed by one of the researchers using Nvivo. The congruence of the codes and themes was checked by the other researcher. Then, the necessary changes were made, and the analysis of the reflection diaries was completed. Finally, as the qualitative data received from the questionnaires were not dense, they were analyzed manually by one of the researchers.

Credibility and Ethics in Research

Different data collection tools such as semi-structured interviews, reflection diaries, questionnaires, and research diaries were used to support each other for the credibility of the study. The questionnaire, semi-structured interview and reflection diary forms were prepared based on the primary school English curriculum. Throughout the process, establishing positive relations with the participants by keeping constant communication with school visits, telephone calls and social media was given importance. Thanks to the positive relationships established, additional interviews were held to get detailed information throughout the process or to eliminate any conflicts between the data.

During the analysis of the data, expert opinions were obtained, and the reliability of the analysis was ensured. In addition, it was ensured that all experts were working in the field of Curriculum Development and Instruction and in areas related to teaching English. All findings were described in detail and presented by establishing relations between the data obtained from different data collection tools. In addition, the findings were presented with quotations of the participants.

Ethical and research permits have been obtained from the relevant institutions to provide research ethics. All the administrators and participant teachers in the schools where the research would be conducted were informed in detail about the purpose and the process of the research. Teachers were asked to participate in the study voluntarily, and they were informed that they could withdraw from the research at any time and that their personal rights would be protected. During the semi-structured interviews, both verbal and written approvals of the participants were obtained. Based on the principle of fair gain, printed and electronic materials such as educational games, flash cards, power point presentations that could be used in their lessons were shared with teachers throughout the process. While presenting the research findings, nicknames were used instead of the participants' names.

Findings

Findings Regarding the Learning Outcomes of the Curriculum

According to the reflection diaries used to monitor the process regarding the achievement of the learning outcomes of the curriculum, the teachers taught the learning outcomes in the first, second, third, fourth and fifth units of the 2nd grade curriculum. However, there were teachers who stated that they experienced problems while teaching the learning outcomes in these units. For example, one teacher stated that students could not learn the alphabet in the 1st unit because of the lack of time. In the 2nd unit, two teachers stated they had problems in teaching one of the learning outcomes which is stated as "Students will be able to greet people others in other languages in addition to Turkish and English." in the curriculum. One of them stated that the students had difficulty in learning the words related to greeting in different languages, and the other stated they had problems with pronouncing these words. When the questionnaire data were analyzed, it was found that three teachers stated this learning outcome was not achieved, and five teachers did not teach it. Three of these teachers stated that they did not find this learning outcome necessary and one stated they did not want to cause confusion. Another stated that they did not spend time on it because of time constraints and as they did not want students to be confused. A teacher interviewed stated that she taught all the learning outcomes except for this one because she thought students could feel confused. Hazal explained the reason for not teaching this learning outcome by saying:

> "They learned how to read and write in Turkish the previous year, I thought teaching them words in different languages might make them confused. That's why I didn't want to teach them." (Interview 1, 05'56"-06'16")

While five of the teachers stated that the 3rd unit of the 2nd grade curriculum was difficult compared to the students' levels, one teacher stated that there was content density in the 4th unit. In the 5th unit, one of the teachers stated that they experienced a shortage of time. In the 6th unit, one of the teachers stated that she taught most of the learning outcomes, but she did not state the learning outcomes she did not teach. The other teachers stated that they taught all the learning outcomes in the 6th unit. In the 7th, 8th, 9th and 10th units, all the teachers taught all the learning outcomes. However, in these units, there were teachers who stated that the students had difficulty in learning and pronouncing the words.

According to the questionnaire data, one teacher did not teach the learning outcomes in the 3rd unit. The teacher said they will teach this unit during the second term since the unit is hard and confusing for the student level. Ayla, one of the five teachers who stated that the 3rd unit was difficult for the level of the students in the reflection diaries, said that this unit should be at the end of the curriculum. She said "For the 3rd unit, it was a little heavy for my students. These topics should have taken place in the 9th or 10th unit, but we have mostly achieved our goals." Zehra, one of the teachers interviewed on the subject, explained that this unit is above the student level, and expressed her opinions by saying:

> "Because there are verbal directives that consist of direct verbal and sentence patterns of two or three words, and I have tried and failed in previous years for children to perceive, accept, practice and speak them. So, I left it to the second semester because their level of English and their vocabulary will be a bit better, and they'll be familiar with the language." (Interview 1, 9'45 " - 10'05 ").

According to the questionnaire data, it was found that teachers taught all the learning outcomes except for the ones in the 2nd and 3rd units mentioned above. On the other hand, in the semi-structured interviews, Zehra stated that she did not teach the learning outcomes in the 3rd unit in the first term, but she taught them in the second term. The other two English teachers stated that they taught all the learning outcomes. One of the two multigrade classroom teachers, who had the 2nd grade students in her class, stated that she taught the words and sentence patterns she chose from the textbooks, and the other teacher said he chose topics from the 2nd and the 3rd grade coursebooks and taught them to both grades.

When analyzing the reflection diaries for the 3rd grade units, it was seen that there are not any learning outcomes that the teachers did not teach. However, there were some teachers stating that they had problems in some units. For example, Özge said the 1st unit was difficult for her students by saying: "It takes a while to teach the Wheel of Fortune unit. It is not possible to spend a lot of time on teaching numbers. Due to this, the teaching hours must be longer." In the semi-structured interviews, a teacher stated that this unit was an unnecessary unit because it is not related to students' daily lives. Three of the teachers who wrote the reflection diaries stated that they had a lack of time in the 3rd unit while two teachers stated that the present continuous tense was difficult for the students in the 4th unit.

According to the questionnaire data for the 3rd grade learning outcomes, one teacher could not reach the learning outcomes stated as "Students will be able to ask and answer questions about the quantity of things." in the 5th unit. The data show there are no other learning outcomes which were not taught or reached. In the semi-structured interviews, the English teachers stated that they taught all the learning outcomes. One of the multigrade classroom teachers stated in the second term interview that he was not sure whether he taught all the learning outcomes because he did not do the listening activities, so he thought he might have skipped some learning outcomes related to the listening skills. On the other hand, two multigrade classroom teachers stated that they taught subjects, words and sentence patterns that were in the coursebooks.

The opinions of the teachers who wrote the reflection diaries about the 4th grade units show that all the teachers except for one taught all the learning outcomes in 1st unit. However, three teachers stated that there were too many learning outcomes, and two emphasized the lack of time. In addition, one of the teachers did not teach one learning outcome which is stated as "Students will be able to say 'thank you' in different languages." because it was hard. The questionnaire data show that there are two other teachers who did not teach the same learning outcome. One of them did not teach it because they thought it was not necessary for the students. The other stated that they did not find it necessary and it was not related to English. In the interviews, Hazal stated that she taught all the learning outcomes of the 4th grade curriculum except for this one because she thought other learning outcomes were more important. She expressed her opinions as follows:

"Well, I didn't spend much time on this learning outcome and set it aside because I wanted to focus on teaching other learning outcomes fully. This was necessary for me to focus on the others. I thought it wasn't really important to teach this learning outcome before I could finish teaching the others." (Interview 1, 07'58- 08'13")

The reflection diaries show that the teachers taught all the learning outcomes of the 2nd unit. However, one teacher mentioned that she had a lack of time. In the 6th unit, all the teachers except for one stated that they taught all the learning outcomes. Kemal, who stated that he did not teach some learning outcomes in the 6th unit because they were not in the book or on the internet, explained his opinion by saying "I couldn't teach some of the learning outcomes above because not all of them are in the book or on the internet. I taught only three or four of them." The teachers said that they taught all the learning outcomes in the 7th, 8th, 9th, and 10th units. However, one teacher said that there were too many learning outcomes in the 7th unit, one teacher said there were too many learning outcomes in the 8th unit, and one said the allocated time was not enough in the same unit.

According to the questionnaire data, while there are two teachers who said they did not teach the learning outcome stated as "Students will be able to extract the gist and relevant specific information in short, recorded passages about predictable everyday routines which are spoken slowly and clearly," one teacher stated that this learning outcome was not achieved. A teacher who stated that they only taught this learning outcome in the 5th unit due to the density of the content expressed his opinions by saying, "Due to lack of time and learning outcomes density in each unit, we started unit 5 the last week." There are three teachers who stated that they did not teach the other four learning outcomes in this unit. One teacher stated that they were behind the schedule and could not start the 5th unit, and one teacher said that they did not start doing this unit.

In the questionnaires, one teacher stated that they could not achieve the learning outcome stated as "Students will be able to start and continue a conversation by using simple phrases and sentences, along with mimics and gestures, to make their meaning clear." One teacher who stated that none of the learning outcomes in the 9th and 10th units were taught expressed why by saying, "We had fallen behind the schedule." While there is one teacher who could not achieve the learning outcome stated as "Students will be able to express how they feel and what they want in simple and short sentences and phrases," in the 10th unit, there are three teachers who stated they could not achieve the learning outcome stated as "Students will be able to identify popular food across cultures." Apart from these learning outcomes, there are no others which were not taught or achieved.

In the interviews, two of the English teachers said they taught all the learning outcomes. One of the classroom teachers said he taught all the learning outcomes in the first term, but not all of them during the second term. The other classroom teacher, Ferhat, said that he mostly taught words and he did not teach all the learning outcomes since he did not know much English. One of the multigrade classroom teachers said he was not sure if he taught all the learning outcomes because he did not spend time on listening exercises so he might have not taught some learning outcomes related to listening skills.

Findings Regarding the Content of the Curriculum

Findings in the reflection diaries of the 2nd grade units indicate that most of the teachers focused on teaching the alphabet and similar words in Turkish and English in the 1st unit. In the 2nd unit, teachers stated that they spent time on teaching words, sentences, and questions about greetings and meeting. The subjects taught in the 3rd unit were the directions, orders, and in-class instructions. While the numbers and classroom objects were focused on by the teachers in the 4th unit, likes and dislikes, objects, their colors, and numbers were taught in the 5th unit. The teachers said that they taught toys, games, the games students like and dislike, instructions, and the simple present tense in the 6th unit. In the 7th unit, clothes, the parts of the body, requests and pronunciation were taught by the teachers. In the 8th unit, they focused on animals and where they live, pronunciation and prepositions. They taught fruit, prepositions, pronunciation in the 9th unit while they focused on teaching animals, abilities of animals and students, and likes and dislikes in the 10th unit.

The teachers who wrote the reflection diaries for the 2nd grade mentioned various problems. For example, they said that there was a lexical density in the 3rd, 4th, 7th, 9th, and 10th units. Two teachers said the words in the 6th unit were difficult for the students' level. Özge explained her opinions on the lexical density by saying "There are too many words, 18 words about fruit and other words were too many for the students and they had difficulty in memorizing them."

The 3rd grade reflection diaries show that the teachers taught numbers, verbs, directions, instructions, and words about the wheel of fortune game in the 1st unit. The teachers focused on the members of family in the 2nd unit while they focused on adjectives, abilities, physical appearance, and objects in the 3rd unit. In the 4th unit, they taught feelings and the present continuous tense. They taught shapes, toys, the numbers, and colors of toys in the 5th unit whereas they taught prepositions, rooms, objects and where they are in the 6th unit. Buildings, where buildings and people are, prepositions and simple instructions were their focus in the 7th unit. In the 8th unit, they taught vehicles, asking for and giving directions, asking and saying where vehicles are. Weather conditions and the present continuous tense were their focus in the 9th unit while they focused on animals, their colors, numbers, abilities, likes and dislikes in the 10th unit. The teachers addressed the problems they experienced about the content of the 3rd grade units. For example, two of the teachers stated that the content of the 1st unit was dense while one teacher mentioned that the 5th unit was dense in terms of the content. One teacher mentioned that they did not have enough time to complete the content in the 6th unit.

The 4th grade reflection diaries show that the teachers taught numbers, classroom objects, words, sentences and questions for in-class instructions, requests, asking for permission, and accepting or refusing an offer in the 1st unit. In the 2nd unit, it is seen that the teachers' focus was on the subjects such as countries, nations, where people are from and their nationality, introducing yourself while they focused on leisure activities, verbs, activities they like and dislike in the 3rd unit. They taught abilities, activities, possessive adjectives, possessions in the 4th unit and telling the time, days, daily routines, and prepositions in the 5th unit. In the 6th unit, they focused on saying where objects and people are, prepositions, laboratory equipment, imperatives and orders related to doing experiments. In the 7th unit, they taught occupation and where people work, daily routines, the simple present tense, likes and dislikes while they taught clothes, weather conditions, seasons, possessions, requests, likes and dislikes in the 8th unit. In the 9th unit, personality traits and physical appearance were the topics they focused while their focus in the 10th unit was on food, drinks and their amount, feelings, requests, offers, accepting and refusing.

The problems raised by the teachers regarding the content of 4th grade units are subject density and some subjects' being difficult for students. Five teachers mentioned that there were too many topics to teach in the 1st unit. For example, Ayla said "There are too many topics in one unit. To be honest, I had a lot of difficulty, I was able to finish the unit in time, though. I spent quite a lot of time on each topic." In addition, some teachers stated that the content of the 2nd, 5th and 8th units were dense. In the 5th unit, four teachers said that the students had difficulty in learning how to tell the time. One teacher said the words in the 6th unit were hard for the students and two teachers said the students found the simple present tense in the 7th unit hard.

In the reflection diaries and interviews, the teachers mentioned that they had problems because the content of the English and other courses were not parallel. For example, a teacher who wrote the reflection diary stated that students had difficulty in learning how to tell the time in English because they did not know how to do it in Turkish. Two teachers stated that students had difficulty in learning terms such as nation, country, and continent in the 2nd unit of the 4th grade curriculum as they did not know them in Turkish. Pınar, one of the teachers who touched on this problem, said: "The learning outcome on recognizing the flags does not coincide with the Social Studies course, which means they have not learned most of them. They did not even know what the country meant, what the nation meant Turkish." On the other hand, Nehir suggested some courses should be parallel to English by saying "First of all, the lessons must be connected. The students must first recognize these concepts in their own language." Metin, one of the two teachers who were interviewed and touched on the problem of lessons not being parallel, said, "I'll teach adjectives, well, in the 3rd grade, but first I have to teach what an adjective is. If they did not learn this in Turkish lesson, I cannot explain it, we had such problems." (Interview 1, 06'26 "- 06'39")

Findings Regarding the Teaching-Learning Process of the Curriculum

In the interviews, the teachers were asked how they used the curriculum and what other resources they benefited from while they were preparing for their lessons. While the two classroom teachers stated that they benefited from the curriculum in the process, an English teacher emphasized that she benefited from the curriculum when choosing the content to be taught. Ferhat, one of the classroom teachers, stated in his interview that he took advantage of the curriculum which was implemented in 2008 by saying "The current curriculum is not guiding me. I am following the old curriculum." (Interview 1, 06'33"- 06'40"). In the second term interview, Ferhat said that he was using the curriculum which was implemented in 2013. On the other hand, he said he did not take a look at this new curriculum. Six of the teachers said they did not use the curriculum. There are six teachers who used the yearly lesson plans. One of these teachers used them only to fill in the classroom notebook, in which

the teachers write the learning outcomes and topics they cover during the lesson. Another teacher said she used the yearly lesson plans to follow the dates of the exams. In addition to these, all the teachers use the students' coursebooks and seven of them use the teacher's books while getting ready for their lessons. Metin, who used the teacher's books to choose the content to be taught, expressed his opinion in these words:

> "I used the teacher's books, but only partially. The reason we use them is that you do not have to make a lesson plan when you use them. If I do not use them, then I will have to prepare a daily lesson plan, I used them because of this reason." (Interview 2, 06'10"-06'27")

One of the English teachers, Hazal said she used the students' coursebooks by saying:

> "I follow the students' coursebooks. I conduct my lessons as the same way the book is organized. I mean, I look at the words in the book. I try to teach the words, the highlighted words, sentence structures, the sentences in the listening and speaking parts in the books." (Interview 1, 11'01"-11'28")

Two of the classroom teachers stated that they got help from their colleagues while they were getting prepared for the lessons. They explained they were doing this as they were concerned with not knowing English and teaching something wrong. In addition to this, Ferhat, who is a classroom teacher, emphasized his professional competencies in determining the subjects to be taught while planning his lessons. He explained his opinions by these words:

> "In fact, I decide what to teach by thinking over what I can teach because I mean, I don't try to teach something I don't know… First, I take a look at the topics I can teach a few days before the lesson, then I teach them. I don't try to teach a topic that is beyond my abilities, to be honest." (Interview 1, 05'30"-05'38"; Interview 06'22"-06'31").

The teachers were asked about which text and activity types they used in the teaching and learning process. Twelve teachers, who wrote the reflection diaries about the 2nd grade curriculum, stated that they used "Songs, Games, Listening, Matching, Drawing and Coloring, Dialogues, TPR, Speaking, Filling the blanks, Questions and Answers, Puzzles, Drama/Miming, Pair Work, Flashcards, Puppets, Reading, Cartoons, Roleplay, Video/Film, Writing". 12 teachers, who wrote the reflection diaries about the 3rd grade curriculum, said "Listening, Matching, Games, Drawing and Coloring, Songs, Questions and Answers, Drama/Miming, Arts and Crafts, Visuals, Speaking, Dialogues, Reading, Animations, Videos, Puzzles, Flashcards, Cut and Paste, Finger Puppets, TPR, Writing" were the text and activity types they used. The text and activity types that the 4th grade teachers used were "Matching, Questions and Answers, Listening, Games, Dialogues, Drawing and Coloring, Songs, Speaking, Writing, Roleplay, Drama/Miming, Puppets, Flashcards, Reading, Film/Video, Stories, Pictures."

In the interviews, the teachers said that they used "Songs, Cartoons, Illustrations, Stories/Fables, Picture Dictionaries, Lists, Posters, Tables, Charts, Dialogues, Rhymes, Menus, Dictionaries" as the activity types. On the other hand, two classroom teachers and two multigrade classroom teachers stated that they rarely used text types during the class, and they did not use activity types. Two of them said that the reason they did not use text types was because of their lack of professional competencies. Ferhat, who did not use activity types, explained his reason by saying "Because I don't know English, I have incompetency." (Interview 1, 09'43"- 09'45"). Mehmet, who is a multigrade classroom teacher, stated that he did not use activity types so as not to cause distraction in the class where there were students from three different grade levels.

In the questionnaires, teachers were asked about the frequency of following one of the suggestions for the teaching-learning process in the curriculum, which is stated as "Using fun visual, audio, and audiovisual tools and materials in learning." It is seen in the 2nd grade level questionnaire data that a majority of the teachers in the first term and a great majority of them in the second term said they "always" used them. According to the 3rd grade level data, a great majority of the teachers "always" used them both terms. The 4th grade data show that a majority of the teachers both terms said they "always" used them. In the semi-structured interviews, English teachers stated that they used audio, visual and audio-visual materials. Two of the multigrade classroom teachers stated that they could not use audio materials because they did not have the technological equipment to use them whereas one of them said he did not use them in order not to distract students. One of the classroom teachers said that he used visual and audio materials while the other said that he used only visuals in the first term, but he did not use either visual or audio materials in the second term.

In the questionnaire, teachers were asked about the frequency of spending time on teaching listening skills which is one of the main focuses of the primary school curriculum. Looking at the questionnaire data of the 2nd grade, it is seen that a minority of the teachers "always", another minority of them "generally," and a small minority of them "sometimes" taught it in the first term. In the second term, a minority of them "always," another minority of them "generally," and a small minority of them "sometimes" spent time on the listening skills. When the 3rd grade questionnaire data were analyzed, it was seen that a minority of the teachers "always," another minority of them "generally," and a small minority of them "rarely" taught listening skills in the first term. In the second term, a majority of the teachers "always," a minority of them "generally," a small minority of them "sometimes" and another small minority of them "rarely" did listening activities to improve listening skills. Considering the frequency of teaching listening skills in the 4th grade, a minority of the teachers "always," another minority of them "generally," and a small minority of them "sometimes" spent time on improving listening skills in the first term. In the second term, it is seen that a minority of the teachers "always," another minority of them "generally," a small minority of them "sometimes" and another minority of them "rarely" taught listening skills.

In the interviews, the teachers were asked the kinds of activities they used to develop the four basic skills and how often they used these activities. Six of the teachers stated that they used the listening activities in the book or the ones they found on the internet to help students to improve their listening skills. One of the multigrade classroom teachers stated that he did not do any listening activities since they would distract the students. When the teachers were asked how much time they spent doing listening activities, one of the English teachers said in each lesson, one said once a week and one said rarely. Zehra, who rarely did listening activities, said the following statement:

> "I spend time on listening activities rarely. I can honestly confess that doing listening activities has only twenty or twenty five percent of importance to me. It has to be this way because the education system is mostly test-based. While we are teaching for the exams and we are trying to make them have more true answers on tests, listening exercises always remain in the background." (Interview 1, 24'28"-24'47")

One of the multigrade classroom teachers stated that he did listening activities once a month, one of them spent one hour a week on listening activities, and the other one did not focus much on listening due to the lack of technological equipment. One of the classroom teachers indicated that he did not do any listening activities because of the lack of technological equipment whereas the other teacher said he did not spend much time on listening during the first term and he did not do listening activities at all during the second term. In addition to these, one of the classroom teachers said that he did not do listening activities due to the lack of materials and technological equipment.

In the questionnaires, teachers were asked how much time they spent on teaching speaking skills which is given a primary focus in the curriculum. The 2nd grade questionnaires show that a minority of the teachers "always," a minority of them "generally" and another minority "sometimes" spent time on teaching speaking skills in the first term. In the second term, it is indicated that a minority of the teachers "always," another minority of them "generally" and a small minority of them "sometimes" taught speaking skills. In the 3rd grade, it is seen that a minority of the teachers "always," a small minority of them "generally," and another small minority "sometimes" taught speaking skills in the first semester. In the second term, a minority of the teachers "always," another minority of them "generally," and a small minority of them "sometimes" spent time on teaching speaking skills. When 4th grade questionnaire data were analyzed, it was understood that a minority of the teachers "always," another minority of them "generally," and a small minority of them "sometimes" spent time on speaking skills in the first term. In the second term, a minority of the teachers "always," another minority of them "generally," and a small minority of them "sometimes" and another small minority "rarely" taught speaking skills.

In the interviews, the teachers stated that they used pair work and group work activities, dialogues with puppets, questions and answers, role play and drama in order to improve students' speaking skills. Three of the teachers said they had speaking activities with their students every week. One of the teachers did speaking activities in each lesson during the first term, but twice a month during the second term. One of the teachers did speaking activities once a week during the first term and once in two weeks during the second term. Moreover, one of the English teachers said she did not focus much on improving their speaking skills.

During the interviews, five of the teachers said they spent time on reading and writing activities in the classroom. Zehra, who is one of the English teachers, said she spent more time on reading and writing than listening and speaking. She expressed her opinion by saying "Yes, reading and writing is more often because of

the importance of the state exams, just like I said before. Reading and writing are thought to be more useful, so I spend more time on these in my classes." (Interview 1, 30'23"- 30'39"). Two of the teachers said they used reading and writing activities rarely, one of them used them in each lesson, and one teacher used reading activities generally but used writing activities sometimes. One of the teachers said he always used these activities during the first term, but he said he used them less during the second term. In addition to these, one classroom teacher and three multigrade classroom teachers stated that they focused teaching vocabulary a lot in their classes.

In the interviews, the teachers were asked questions about the suggestions for the implementation of the curriculum. For example, it is recommended that students should not use notebooks in the primary school level. Based on this, the teachers were asked whether their students had notebooks. All the teachers, except for one, stated that the students had notebooks. When asked why they prefer to use notebooks; for example, two of the classroom teachers and an English teacher stated that they made students use notebooks because they thought that they learned more permanently when they wrote.

The primary school English curriculum advises teachers not to correct students' mistakes directly. In the interviews, two classroom teachers stated that they made immediate corrections whereas the English teachers stated that they said the correct pronunciation of the words themselves in order to make the students realize their mistakes and learn the correct pronunciation of the words.

The 2nd and 3rd grade curricula propose to take advantage of the extracurricular environments in the teaching-learning process. When the questionnaire data for the 2nd grade were analyzed, it was seen that a small minority of the teachers "always," another small minority of them "generally," a minority of them "sometimes," and a small minority of them "rarely" did activities out of the classroom, and a small minority of the teachers did not do any activities outside the classroom. In the second term, it is seen that a small minority of the teachers "always," a minority of them "generally," another minority "sometimes" and a small minority "rarely" did activities in different places outside the classroom while a small minority "never" did activities outside the classroom. When the 3rd grade questionnaire data for the first term were analyzed, it was found that a small minority of the teachers "generally," a minority of them "sometimes," a small minority "rarely," and another small minority "never" did activities outside the classroom. The second term questionnaire data show that a small minority of the teachers "always," a small minority of them "generally," another small minority of them "sometimes," a small minority of them "rarely," and another small minority of them "never" did activities outside the classroom. In the interviews, when the teachers were asked if they followed this recommendation of the curriculum, two of them stated that they benefited from environments such as the playground and school garden, one of them stated that they did not do any activities outside the classroom due to lack of time and two of them said they did not do activities outside the classroom due to inconvenient weather conditions.

The curriculum recommends that teachers conduct their lessons in English. During the interviews, the teachers were asked which language they prefer as the language of instruction in their classes. A multigrade classroom teacher stated that she was teaching in Turkish because she could not speak English. Two classroom teachers and two multigrade classroom teachers stated that they mostly used Turkish because of their low proficiency level of English while one English teacher stated that he spoke Turkish more so as not to create negative attitudes towards the lesson. One of the English teachers said that she used English more as the students' English levels increased while one explained that she used English but explained complicated instructions in Turkish.

Another recommendation of the curriculum to teachers is to enable students to develop positive attitudes towards language learning. In order to achieve this, some teachers stated that they tried to make the lessons fun, gave awards, and used different materials. There are also teachers who stated that they gave additional exercises to highly motivated students, gave more space to the activities that students enjoyed, and established positive communication with students.

Findings Regarding the Evaluation Process of the Curriculum

There are projects in the assessment section of each unit of the primary school English curriculum. For example, in the 1st unit of the 2nd grade curriculum, there is a project stated as "Students prepare a visual dictionary to show the words they know in English." In the reflection diaries, questionnaires and interviews, the teachers were asked if they did these projects, and if they did not, why they did not do them. When the opinions of the teachers who wrote the reflection diaries for the 2nd grade curriculum were analyzed, it was understood that four

teachers did the projects in the 1ˢᵗ and the 2ⁿᵈ units while six teachers did the projects in the 3ʳᵈ, 4ᵗʰ and 5ᵗʰ units. Seven teachers stated they spent time on the projects in the 6ᵗʰ unit and there were no teachers stating that they did not do the projects in the 7ᵗʰ unit. Six teachers in the 8ᵗʰ unit and eight teachers in the 9ᵗʰ and 10ᵗʰ units stated that they completed the projects. In addition to these, there are teachers who stated that they had the projects done as homework because there was not enough time to cover them in the classroom. However, Ülkü stated that the activities were not beneficial for her student when they were assigned as homework because the activities were done by the families. She expresses her opinion by saying "We did not do the projects. Families do homework assignments and students are passive while doing them. When we want to do them in the classroom, we cannot have enough time."

When the questionnaire data regarding the unit projects were analyzed, it was seen that a majority of the teachers completed the project in the first unit. In the second unit, there are three projects and the data show that a majority of the teachers did not do the first and the third projects while a majority of them did the second project. There is one project in the 3ʳᵈ unit, and it is seen that the majority of the teachers did this project. There are two projects in the 4ᵗʰ unit, and it is seen that a minority of the teachers did the first project, and another minority partly did it. The second project was partly done by a minority of the teachers while it was not done by another minority. There are two projects in the 5ᵗʰ unit, and it is seen that the majority of the teachers did both. In the 6ᵗʰ unit, there are two projects; the data show that the first of these projects was done by a majority of teachers, and the second one was done by a great majority of the teachers. There are two projects in the 7ᵗʰ unit, and it is seen that a great majority of the teachers did both projects. The data show a majority of the teachers did both projects in the 8ᵗʰ unit and a majority of the teachers did both projects in the 9ᵗʰ unit. There are two projects in the 10ᵗʰ unit, and it is seen that a majority of the teachers did both projects.

When the opinions of the teachers who explained the reason for not doing the projects in the first five units of the 2ⁿᵈ grade curriculum were analyzed, it was understood that the lack of time, economic problems and lack of technological equipment were the reasons for not doing the projects. For example, one of the teachers who stated that the first project in the second unit was not done due to the lack of technological equipment said, "Because most of the students do not have the necessary equipment to record sound." A teacher who stated that they did not do the second project in the 4ᵗʰ unit due to the lack of time said "There was no time for them to prepare within 2 lesson hours, but I brought them a puzzle and we played with it." A teacher who said that he did not do any projects in the last five units of the 2ⁿᵈ grade curriculum said, "I could not catch up with the pacing of the curriculum because I am not an English teacher." Other teachers who explained the reasons for not doing the projects stated that they did not do them because of the lack of time or because the projects were not suitable for the level of the students.

There are also teachers who stated that they did different activities instead of doing projects in the 2ⁿᵈ grade units. For example, a teacher who stated that she did not do the second project in the 2ⁿᵈ unit said, "They did not prepare masks, I took puppies to the class." Another teacher who stated that she did not do the second project in the fourth unit said, "The making puzzle activity was not done. Instead, they learned a song about the numbers, and they danced to the song with the moves of it." A teacher who stated that she implemented the second project in the 9ᵗʰ unit by changing it said "The second project about Fruits in the 9ᵗʰ unit was applied differently. The students prepared a poster about the fruit."

When the reflection diaries of the 3ʳᵈ grade units were analyzed, it was found that three teachers in the 1ˢᵗ unit, six teachers in the 2ⁿᵈ unit and four teachers in the 3ʳᵈ unit did projects. While there were five teachers in the 4ᵗʰ unit stating that they did the project, six teachers in the 5ᵗʰ unit, seven teachers in the 6ᵗʰ, 7ᵗʰ and 8ᵗʰ units did projects. While the number of teachers who did the projects in the 9ᵗʰ unit was five, eight teachers in the 10ᵗʰ unit did the projects. There were teachers who stated that they did not do the projects in the 3ʳᵈ grade units and assigned them as homework due to the lack of time. For example, one of these teachers, Gözde, expressed that she assigned the project in the 6ᵗʰ unit as homework by saying "Project: They did "dream home" project as homework. We didn't do it in the classroom so as not to waste time and not to let some students lose their attention." In addition, Nilüfer stated that she did the project of the 9ᵗʰ unit by making changes to it and explained the reason as follows:

> "In the project part, I asked them to make pictures about weather conditions and prepare questions and answers instead of the weather conditions in different cities. I did it this way because they said they wouldn't be able to find and draw pictures about different cities."

When the 3ʳᵈ grade questionnaire data were analyzed, it was seen that there is only one project in the 1ˢᵗ unit and a great majority of the teachers did this project. There are three projects in the 2ⁿᵈ unit, and the first and the

second projects were done by a great majority of teachers. However, a majority of the teachers did not do the third project. There are two projects in the 3rd unit, and it is understood that a minority of the teachers did the first of these projects, and another minority partly did it. It is seen that a majority of the teachers partly did the second project and a minority of them did it. The data show that a majority of the teachers did the first project in the 4th unit while a minority of them did the second project. A majority of the teachers did both projects in the 5th unit. In the 6th unit, there are two projects, and a majority of the teachers did both of them. There is one project in the 7th unit and a great majority of the teachers did this project. There are two projects in the 8th unit. It is seen that a majority of the teachers did the first of these projects while a minority of the teachers did the second project and another minority partly did the second project. It is understood that the first of the two projects in the 9th unit was done by a majority of the teachers and the second was done by a great majority of the teachers. It is seen that a minority of the teachers did the first project in the 10th unit, another minority partly did it, and a small minority did not. It is seen that the second project in this unit was done by a majority of the teachers.

When the statements of the teachers explaining the reasons for not doing the projects in the 3rd grade units were analyzed, it was understood that they did not do the projects due to the lack of time, the lack of technological equipment, economic problems and teachers' lack of professional competency. For example, a teacher explained the reason for not doing the third project in the 2nd unit by saying "Students do not have the tools to record sound." On the other hand, a teacher explained the reason for not doing the second project in the 8th unit by saying "Lack of time and knowledge." In addition, there are teachers who stated that they had different activities done instead of the projects in the units. For example, one teacher who said they did not do the projects in the 4th unit stated that "The projects of the Feelings unit were made in different ways. Instead of making a poster, games that students guess what their friends did in the classroom were played."

When the 4th grade reflection diaries were analyzed, it was observed that six teachers in the 1st unit, five teachers in the 2nd unit, and eight teachers in the 3rd and 4th units did the projects. Four teachers in the 5th unit, six teachers in the 6th unit, seven teachers in the 7th unit, and eight teachers in the 8th, 9th and 10th units stated that they did the projects. The teachers who stated that they did not do the projects in the 4th grade curriculum said that they did not do the projects and assigned them as homework due to the lack of time. On the other hand, a teacher said that the project was difficult in the 3rd unit, so he had another activity done instead. In the 6th unit, a teacher stated that she did not do the project because it was difficult. In the 7th and 9th units, a teacher stated that she did not have the projects done because she did not find them useful for the students. In addition, Kemal, who stated that he wrote the reflection diaries based on the student's coursebook and not the curriculum, said that he did not do any projects in the 1st unit since there were no assessment activities in the book. He explained his opinion by saying "Unfortunately, there are not many activities in the assessment sections of the book. I also did not evaluate them with something else."

According to the data obtained from the 4th grade questionnaires, it is seen that the first project in the 1st unit was done by a majority of the teachers and the second project was done by a minority of them. There are two projects in the 2nd unit, it is seen that a majority of the teachers did the first one of these projects. On the other hand, a minority of the teachers did the second project and another minority partly did it. There are two projects in the 3rd unit and a majority of the teachers did both of them. There are two projects in the 4th unit, and it is seen that a majority of the teachers did the first project. The second project was done by a minority of the teachers and another minority of them partly did it. There are three projects in the 5th unit, the first of these was done by a minority of the teachers, and another minority partly did it. It is seen that a minority of teachers did the second project while another minority partly did it. The third project was done by a majority of the teachers. There is one project in the 6th unit and a majority of the teachers did this project. There are three projects in the 7th unit, and it is seen that a majority of the teachers did all three of them. There are three projects in the 8th unit, it is seen that a majority of the teachers did the first and second projects. A minority of the teachers did the third project in this unit, another minority partly did it, and a small minority did not. It is seen that a majority of the teachers did both projects in the 9th unit. While a majority of the teachers did the first of the two projects in the 10th unit, a minority of the teachers did the second project, and another minority partly did it.

When the statements of the teachers explaining the reasons for not doing the projects were analyzed, it was understood that the reasons for the teachers not doing the projects were lack of time, crowded classes, lack of opportunities and teachers' lack of knowledge in English. For example, a teacher explained their reason for not doing the first project in the 3rd unit by saying "I am not good at English and there's time constraint." A teacher who stated that they did not do both projects in the 3rd unit due to the crowded classes said "There are 6 class hours for one unit and there are 35 students in the class. Unfortunately, I cannot spare time because such activities take a lot of time in the classroom." A teacher who did not do both projects in the 4th unit stated that

they cannot do the projects due to the heavy load of the curriculum. One of the teachers who did not do the project: "Students prepare a puppet with seasonal clothes and describe him / her (video recording is suggested)" said that they did not do this project due to the limited opportunities.

In the interviews, English teachers stated that they did not do all of the projects. Two teachers stated that they assigned the projects as homework. One of them stated that she assigned them as homework due to the lack of time, and the other teacher said she did this because the students did not bring their materials to the class. The teachers also mentioned that when they assigned the projects as homework, parents complained, or some students did not do their homework. In addition, two of the teachers stated that they sometimes changed the projects to make them more suitable for their students and easier to do on their own. On the other hand, classroom teachers stated that they did not do the projects due to their low proficiency level of English and the lack of physical equipment.

In addition to the projects in each unit of the curriculum, there is the European Language Portfolio which starts in the 1st unit and ends in the 10th unit. The teachers were asked if they spent any time doing the European Language Portfolio and if not, why they did not. The questionnaire data show that in the 1st unit of the 2nd grade curriculum, a small minority of the teachers did the European Language Portfolio, a minority of them partially did it and another minority of the teachers did not do it. In the 10th unit, a minority of the teachers spent time on it, another minority partially did it and a minority of the teachers did not spend any time on it. The 3rd grade questionnaire data show that a small minority of the teachers started the European Language Portfolio in the 1st unit, a minority of them partially did it and another minority of them did not do it. In the 10th unit, a minority of the teachers did the European Language Portfolio, another minority of them partially did it and a small minority of them did not do it. The 4th grade questionnaire data show that in the 1st unit, a small minority of the teachers did the European Language Portfolio, a majority of the teachers partially did it and a minority of the teachers did not do it. In the 10th unit, a minority of the teachers spent time on completing it, another minority of them partially spent time on it, and a small minority did not do it.

When the reasons why the teachers did not do the European Language Portfolio were analyzed, it was understood that they did not spend time on them due to the lack of time and their low proficiency level of English. Moreover, they did not do it because they did not find them necessary and they think it was not suitable for primary school students' level. For example, a teacher who did not do the European Language Portfolio, which starts in the 1st unit of the 2nd grade curriculum, said "We do not do this portfolio since the application of (CEF) for foreign language has not been implemented in the primary school curriculum yet." Another teacher who did not do the European Language Portfolio in the 10th unit of the 3rd grade said, "Their level was not suitable." A teacher who stated that they did not do the European Language Portfolio in the 1st unit of the 4th grade explained the reason for not doing it by saying "My low level of English and time limitation."

Five of the teachers who were interviewed, expressed their opinions on the European Language Portfolio. They all stated that they did not spend time doing it. Two of the teachers stated that they did not have any information about it while one stated that they did not do it because it was difficult to follow the students. The two classroom teachers stated that they did not do it because of their low proficiency level of English.

The curriculum recommends teachers use alternative assessment techniques such as peer assessment, self-assessment, and portfolio. Based on this, during the interviews the teachers were asked if they used such activities. Six out of the eight teachers stated that they did not use peer assessment or self-assessment. Only two of the teachers stated that they benefited from the portfolio while evaluating the students. Seven teachers said that they graded students based on their performance in class. In addition, two of the teachers stated that they used observation forms. The teachers stated that they used pencil and paper exams only for the 4th grade students. However, a multigrade classroom teacher stated that she gave quizzes to the 2nd and the 3rd grade students to determine whether they learned the words she taught.

Conclusion, Discussion and Suggestions

In this study, the alignment between the official primary school English curriculum and the taught curriculum was analyzed. By working with English teachers, classroom teachers and multigrade classroom teachers who carry out primary school English lessons, the alignment was investigated by trying to reveal how faithful the teachers were to the curriculum during the implementation of the official curriculum, the reasons for not following the curriculum, the problems they had and their suggestions. In the study, it was observed that English teachers ensured the alignment in terms of the objectives and the content of the curriculum, but they had

problems in implementing the recommendations of the curriculum regarding the teaching-learning process and evaluation process. Therefore, they were not able to fully ensure the alignment between the official curriculum and the taught curriculum. The study showed that the classroom teachers were not able to fully achieve the curriculum alignment due to their professional competencies, the lack of technological equipment and materials. On the other hand, multigrade classroom teachers were found to be unable to ensure the alignment as a result of the problems they experienced due to the unique structure of multigrade classes and the same problems experienced by the classroom teachers.

When the alignment of the objectives between the official curriculum and the taught curriculum was analyzed, it was seen that English teachers taught all the learning outcomes in the 3rd grade curriculum. This shows that the curriculum alignment was achieved in terms of the learning outcomes of the 3rd grade curriculum. However, there are learning outcomes that some teachers did not teach in the 2nd and 4th grade curriculum. One of these teachers stated that the reason for not teaching all the learning outcomes in the 6th unit of the 4th grade curriculum was that those learning outcomes were not in the book or on the Internet. This can be considered as an indication that the teacher was not following the curriculum but the coursebook. Other learning outcomes not taught are the two learning outcomes under the heading of intercultural awareness in the 2nd and 4th grade curricula. They did not teach these learning outcomes because they thought these goals were not related to teaching English, they were challenging for students and teachers found it more important to focus on teaching English words than teaching words in different languages. The fact that the teachers did not teach these learning outcomes can be interpreted as an indication that the curriculum alignment was not fully achieved in terms of the learning outcomes of the 2nd and 4th grade curricula. In the study where the relationship between the secondary school science technology curriculum and the taught curriculum was analyzed, Ntoi (2007) found that the objectives of the curriculum could not be achieved. The results of this research also support our study as it showed that the lack of curriculum alignment was related to the deficiencies in teacher training, the teachers' implementation of the curriculum based on their own perspectives and the readiness level of the students. While the process of analyzing the data of this research was continuing, the new curricula were introduced in 2018. When the 2018 curriculum was analyzed, it was seen that the learning outcomes related to intercultural awareness were omitted, and this change in the curricula can be seen as a result of curriculum development studies conducted with the feedback received from the teachers.

The findings obtained from the classroom teachers and the multigrade classroom teachers show that the curriculum alignment could not be ensured in terms of the learning outcomes of the curriculum. According to the teachers' opinions, this is because the teachers have problems in terms of their proficiency level of English and their professional competencies in teaching English. In addition, the fact that these teachers attach more importance and devote more time to teaching core lessons such as Turkish and Mathematics causes the lack of alignment. The fact that there are students from different grade levels in a multigrade classroom also makes it difficult to implement the curriculum designed for single-grade classes and to teach all learning outcomes. For example, one of the multigrade classroom teachers stated that he did not do listening activities since that would cause distraction in the classroom; therefore, he was not sure that he taught all the learning outcomes in the curriculum. Sidekli, Coşkun and Aydın (2015) also reached a similar conclusion in the study they conducted with the multigrade classroom teachers in order to identify the problems they had and possible solutions. Summak, Summak and Gelebek (2011) express that it is very difficult for the teacher to implement the curriculum developed for single-grade classes in multigrade classrooms, and all the learning outcomes cannot be reached. In their study, they also found out that the teacher could sometimes accomplish one learning outcome of one grade level in a classroom with three different grade levels. Based on these results, it can be said that it is important to develop a separate English curriculum for multigrade classes.

In terms of the content of the curriculum, it was found that English teachers taught subjects not included in the curriculum in four units at three grade levels. Apart from this, the content that teachers taught in their lessons is aligned with the curriculum. One of the subjects which is not in the curriculum, but the teachers taught in their lessons was "clothes" in the 7th unit of the 2nd grade curriculum. This unit is related to the parts of the body, and it is thought that the teachers taught clothes in this unit because they were related to the parts of the body. It can also be said that based on the flexibility of the curriculum, the teachers shape the curriculum to fit the needs of their students. The study conducted by Ziebell (2010) on the alignment of the elementary mathematics curriculum and the implemented curriculum support our research as it also found that the teachers changed the curriculum to meet the students' needs. It is seen that the teachers taught "likes and dislikes" in the 8th unit of the 4th grade level that is not included in the curriculum. This topic is in the 7th unit and it is thought that teachers spent time on this topic in the 8th unit to revise the previous unit. As a matter of fact, as the primary school English curriculum is spiral, it recommends teachers repeat the previous subjects. When the content that

the teachers taught is considered as a whole, it can be said that they were able to ensure the curriculum alignment between the official and taught curriculum in terms of the content of the curriculum.

Considering the alignment of the text and activity types that the curriculum suggests being used in the teaching-learning process and what is actually used in the classroom, it is seen that English teachers use the text and activity types suggested by the curriculum. The study conducted by Küçüktepe, Eminoğlu-Küçüktepe and Baykın (2014) on the 2nd grade curriculum also shows that teachers use the activities suggested by the curriculum. The fact that the multigrade classroom teachers and classroom teachers used the text types in a limited way and did not use the activity types is not only related to the professional competencies of the teachers, but also related to the unique structure of the multigrade classes. For example, one of the multigrade classroom teachers stated that he did not use the activity types to avoid distraction in the classroom. This shows that the teacher had problems in ensuring the curriculum alignment due to the features of the multigrade classroom. While English teachers always use audio, visual and audiovisual materials, multigrade classroom teachers and classroom teachers rarely use them due to the lack of technological equipment in their schools and the coexistence of two or four different grade levels in one classroom. Kaya and Ok (2016) carried out research on the implementation of the 2nd grade English curriculum of 2013 and most of their participants were classroom teachers. Similarly to our research, their study revealed that teachers did not use the visual and audio materials sufficiently due to the lack of equipment in schools.

The primary school English curriculum focuses on listening and speaking skills. However, this study shows that teachers do not spend much time on improving these skills, and the activities they use to develop these skills are limited. The reasons for these are related to the lack of technological equipment in schools, the unique structure of the multigrade classes and professional competencies of the teachers. Similarly, Kaya and Ok (2016) revealed that the activities teachers spent the least time on were listening and speaking activities, which is due to the professional competencies of teachers and the low proficiency level of students to perform these activities. Primary school English curriculum states that reading and writing activities in the 2nd, 3rd and 4th grades should be included as extra-curricular activities, and that they can be used in the 3rd and 4th grades without exceeding ten words (MoNE, 2013, p. VI). However, the interviews show that two classroom teachers and one English teacher often did reading and writing activities within the lesson while other English teachers rarely did. These show that teachers could not fully ensure the curriculum alignment in terms of implementing the suggestions of the curriculum on four basic language skills.

When looked at how the suggestions of the curriculum for the teaching-learning process are implemented in the classroom, it is seen that none of the teachers were able to fully create the alignment. While the curriculum states that students should not have notebooks in English lessons, the students of all the teachers, except for one, had notebooks. The curriculum recommends that students' mistakes should not be corrected immediately. English teachers carry out their classes in accordance with this recommendation of the curriculum, but classroom teachers correct students' mistakes immediately. In their study, in which most of the participants were classroom teachers, Kaya and Ok (2016) also reached the conclusion that students' mistakes were corrected during communication. Based on these, it can be said that this is related to the professional competencies of the classroom teachers on teaching English. The curriculum advises teachers to benefit from out-of-class environments during the teaching-learning process, but it seems that teachers do not follow this suggestion of the curriculum due to adverse weather conditions and lack of time. Another suggestion of the curriculum is that the courses are conducted in English. While a multigrade classroom teacher does not speak English and conducts her lessons entirely in Turkish, other teachers use both Turkish and English. However, multigrade classroom teachers and classroom teachers use Turkish more. Similarly to the result of this study, in the study they carried out Şad and Karaova (2015) found that the classroom teacher conducted his lessons mostly by speaking Turkish. The study revealed that all the teachers acted in accordance with the curriculum's suggestion of encouraging students to have a positive attitude towards learning English.

There are projects in the assessment section of each unit in the curriculum; moreover, it suggests alternative assessment techniques such as peer assessment and self- assessment. The findings of the study show that the alignment was not fully ensured in terms of the evaluation dimension of the curriculum. English teachers do not do all of the suggested projects in each unit of the curriculum. The greatest reason for this is the lack of time. Similarly, Arı's study (2014) with teachers teaching the 6th grade and Yörü's study (2012) with teachers teaching the 8th grade show that teachers do not do the project and performance assignments because they take a lot of time. Sometimes teachers assign projects as homework due to the lack of time, and some teachers make changes on the projects to make them more suitable for their students. This can be interpreted as teachers trying to make the curriculum suitable for their students by considering their competencies and needs. The classroom teachers, on the other hand, do not do projects due to the lack of equipment in schools and their professional competencies. In his reflection diaries, one of the English teachers explained the reason for not doing the

projects in some units as there were no evaluation activities in the book. It can be said that this shows the teacher sees the coursebook as the curriculum and does not follow the suggestions of the curriculum in terms of the evaluation dimension.

One of the evaluation components of the curriculum is the European Language Portfolio, and the study shows that the teachers did not spend time on it. The reasons for this are the teachers' lack of knowledge on how to do it and the classroom teachers' professional competencies. When the curriculum that was put in action in 2018 is examined, it is seen that the European Language Portfolio is omitted. This change is also thought to have occurred as a result of the curriculum development studies. It is also seen that most of the teachers do not follow the curriculum's suggestion of using peer assessment, self-assessment, and portfolio to evaluate students. Except for one multigrade classroom teacher, the others used paper and pencil tests only in the 4th grade level. The teachers also evaluate students by considering their in-classroom performance. When the findings of the evaluation dimension of the curriculum analyzed as a whole, it can be said that the alignment could not be fully ensured. The results of the research carried out by Özüdoğru (2016) on the 2nd grade curriculum also showed that the implementation level of the items related to the assessment was not sufficient.

Teachers were asked how they used the curriculum while getting ready for their lessons. While an English teacher stated that she benefited from the curriculum when deciding what content to teach, both classroom teachers stated that they used it while preparing for their lessons. However, one of the classroom teachers stated that he used the book prepared in accordance with the curriculum that was put in action in 2008 and ended in 2013. The second term he stated that he used the 2013 curriculum. When asked if the teacher read over the curriculum, he said he did not. Verhoeven and Verloop's (2010) study supports the findings of our study as they also found that the teachers acted according to the old curriculum in terms of the objectives and evaluation dimensions of the curriculum, not the new curriculum and the reason for this was related to the teachers' beliefs. Other teachers stated that they used student coursebooks and teacher guidebooks while preparing for their lessons. Based on these findings, it can be said that the teachers did not fully benefit from the curriculum while preparing for their lessons and during the teaching-learning process. It is thought that the reason for this might be related to the teachers' beliefs that coursebooks and guidebooks are sufficient to carry out their lessons and their lack of information on the curriculum. In the studies they carried out on the 2nd grade English curriculum, Kandemir (2016) and Özüdoğru (2016) also concluded that the teachers' knowledge of the curriculum was insufficient, which supports the finding of our research. Büyükduman (2005) and İyitoğlu and Alcı (2015) found that teachers perceived the textbook as the curriculum, and this also supports the results of our study. MoNE (2017, p. 13) describes one of the field knowledge competencies of teachers as "They can explain the curriculum of their field with all its elements.". However, this study shows that the teachers' information on the curriculum is not sufficient. This affects the alignment between the official curriculum and the taught curriculum. For this reason, it can be said that in-service trainings are crucial to improve their awareness and information about the curriculum, and thus ensure the curriculum alignment and make the curricula successful.

This research shows that English teachers, classroom teachers and multigrade classroom teachers ensured curriculum alignment in different levels. It is seen that English teachers created curriculum alignment in terms of the objectives and the content of the curriculum. However, considering that these teachers did not benefit from the curriculum except for one teacher, it can be said that the alignment ensured was not the curriculum alignment but the coursebook alignment. Furthermore, the reason teachers could not ensure the curriculum alignment in terms of the teaching-learning process and evaluation is about teacher beliefs and their lack of information about the curriculum. This study is supported by Hannigan's (2015) conclusion that teachers' lack of information about the curriculum caused them to become inadequate in preparing sufficient resources and assessment activities to support students' learning. In addition, the lack of time, and density of the learning outcomes and the content negatively affected teachers to ensure the curriculum alignment. Turan Özpolat (2015) also reached the conclusion that teachers could not fully achieve the curriculum alignment and that various factors related to school, teachers and students affected it. The study conducted by Scheker Mendoza (2011) on the alignment of the reading comprehension curriculum and the taught curriculum also supports this study with its conclusion that the decisions teachers made while implementing the curriculum were related to their previous experiences, beliefs, and the trainings they received. Based on all these, it can be said that it is important to make the curriculum balanced in terms of time, learning outcomes and content by conducting curriculum development studies and carrying out in-service trainings to increase teachers' literacy of the curriculum.

In the study, it was concluded that the classroom teachers and the multigrade classroom teachers were not able to ensure the curriculum alignment. This is due to the teachers' professional competencies, lack of information on the curriculum and the unique structure of the multigrade classrooms. In their study on the curriculum

fidelity, Nevenglosky, Cale and Aguliar (2019) also reached the conclusion that teachers did not have enough information about the expectations of the curriculum and how to implement it successfully, and they needed in-service training for the implementation of the new curriculum. The study carried out by Shah (2015) on the English curricula also shows that along with teacher beliefs, the level of students, lack of resources and problems related to teachers' professional competencies negatively affect the curriculum alignment, and by supporting the teachers with in-service trainings, curriculum alignment can be improved. As a result, teachers should be supported with in-service trainings so that they can improve their proficiency level of English and learn how to teach English. The teachers who have the most difficulty in ensuring the curriculum alignment are multigrade classroom teachers. For this reason, curricula specific to multigrade classes can be prepared to ensure curriculum alignment and success in teaching language. In addition, it was observed that the teachers did not read the curriculum and did not have sufficient information about the concept of the curriculum. Thus, supporting teachers with in-service trainings on the scope of the curriculum and how they can be used in the most effective way is important for the success of the curriculum. This study is limited to the views of the primary school English teachers and classroom teachers in the center of Kütahya. Collecting data from a variety of teachers from different regions and cities can provide a more comprehensive insight of the curriculum alignment. In this study, the teachers could not be observed in the classroom. Thus, the findings are limited to the teachers' views collected through interviews, reflection diaries and questionnaires. In future studies, the alignment between the official and the taught curriculum can be analyzed through observations to get a clearer picture of the alignment between the two types of curriculum.

References

Arı, A. (2014). İlköğretim altıncı sınıf İngilizce dersi öğretim programına ilişkin öğretmen görüşleri [Teacher opinions about evaluation of 6th grade English lesson curriculum in primary schools]. Kuramsal Eğitimbilim Dergisi, 7(2), 172-194.

Bobbitt, F. (1918). The curriculum. Cambridge, Massachusetts: The Riverside.

Bümen, N. T., Çakar, E. & Yıldız, D. G. (2014). Türkiye'de öğretim programına bağlılık ve bağlılığı etkileyen etkenler [Curriculum fidelity and the factors affecting the fidelity in Turkey]. Kuram ve Uygulamada Eğitim Bilimleri, 14(1), 203-228.

Büyükduman, F. İ. (2005). İlköğretim okulları İngilizce öğretmenlerinin birinci kademe İngilizce öğretim programına ilişkin görüşleri [The opinions of elementary school English teachers on the English curriculum for elementary schools]. Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 28, 55-64.

Büyüköztürk, Ş. (2018). Sosyal bilimler için veri analizi el kitabı: istatistik, araştırma deseni, SPSS uygulamaları ve yorum [Data analysis handbook for social sciences: statistics, research design, SPSS applications and interpretation]. Ankara: Pegem.

Çelebi, M. D. (2006). Türkiye'de anadili eğitimi ve yabancı dil öğretimi [Education of native language and teaching foreign language in Turkey]. Erciyes Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 1 (21), 285-307.

Creswell, J. W. (2009). Research design: qualitative, quantitative, and mixed methods approaches (3rd Ed.). Thousand Oaks, California: Sage.

Demirel, Ö. (1999). İlköğretim okullarında yabancı dil öğretimi [Teaching foreign language in primary schools]. İstanbul: MEB.

Demirel, Ö. (2003). Yabancı dil öğretimi [Teaching foreign languages]. Ankara: Pegem.

Dusenbury, L., Brannigan, R., Falco, M. & Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. Health Educatıon Research, 18 (2), 237–256.

English, F. W. (1992). Deciding what to teach and test. Newbury Park, California: Corwin.

Ertürk, S. (1979). Eğitimde program geliştirme (3. Baskı.) [Curriculum Development in Education (3rd Ed.)]. Ankara: Yelkentepe.

Furtak, E. M., Ruiz-Primo, M. A., Shemwell, J. T., Ayala, C. C., Brandon, P. R., Shavelson, R. J. & Yin, Y. (2008). On the fidelity of implementing embedded formative assessments and its relation to student learning. Applied Measurement in Education, 21 (4), 360-389.

Glatthorn, A.A. (2000). The Principals as curriculum leader: shaping what is taught and tested (2nd Ed.). Thousand Oaks, California: Corwin.

Glatthorn, A. A. ve Jailall, J. M. (2009). The principals as curriculum leader: shaping what is taught and tested (3rd Ed.). Thousand Oaks, California: Corwin.

Hannigan, J. E. IV (2015). Investigating the impact of common core integration on early literacy curriculum systems (Unpublished dissertation). California State University, Educational Leadership, Fresno.

İyitoğlu, O. & Alcı, B. (2015). A qualitative research on 2nd grade teachers' opinions about 2nd grade English language teaching curriculum. Elementary Education Online, 14(2), 682-696.

Johnson, B. & Christensen, L. (2014). Eğitim araştırmaları: nicel, nitel ve karma yaklaşımlar [Educational research: quantitative, qualitative, and mixed approaches]. (S. B. Demir, Trans.). Ankara: Eğiten Kitap.

Kandemir, A. (2016). İlkokul 2. sınıf İngilizce öğretim programının katılımcı odaklı program değerlendirme yaklaşımıyla değerlendirilmesi [An evaluation of 2nd grade English curriculum within a participant-oriented program evaluation approach] (Unpublished master's thesis). Pamukkale Üniversitesi, Eğitim Bilimleri Enstitüsü, Denizli.

Karabacak, E. (2018). İlkokul resmi İngilizce öğretim programı ile uygulamadaki program arasındaki uyumun incelenmesi [Analyzing the alignment between the primary school official English curriculum and the taught curriculum] (Unpublished master's thesis). Anadolu Üniversitesi, Eğitim Bilimleri Enstitüsü, Eskisehir.

Kaya, S. & Ok, A. (2016). The second grade English language curriculum: theory-practice congruence. Pegem Eğitim ve Öğretim Dergisi, 6(4), 491-512.

Küçüktepe, C., Eminoğlu Küçüktepe, S. & Baykın, Y. (2014). İkinci sınıf İngilizce dersi ve programına ilişkin öğretmen görüşlerinin incelenmesi [An investigation of teachers' views on the second grade English course and curriculum]. Hasan Ali Yücel Eğitim Fakültesi Dergisi, 11(22), 55-78.

Livingstone, I. D., Postlethwaite, N. T., Travers, K. J. & Suter, L. E (1986). Second international mathematics study. perceptions of the intended and implemented mathematics curriculum. contractor's report 1986. http://files.eric.ed.gov/fulltext/ED274525.pdf

MEB (2006). İlköğretim kurumları İngilizce dersi öğretim programı [Elementary schools English curriculum]. Ankara.

MEB (2013). İlköğretim kurumları (ilkokullar ve ortaokullar) İngilizce dersi (2,3,4,5,6,7 ve 8. sınıflar) öğretim programı [Elementary schools (primary and secondary schools) English (2nd,3rd,4th,5th,6th,7th, and 8th grades) curriculum]. Ankara.

MEB (2017). Öğretmenlik mesleği genel yeterlikleri [General competencies of the teaching profession]. Ankara.

MEB (2018). İngilizce dersi öğretim programı (ilkokul ve ortaokul 2,3,4,5,6,7 ve 8. sınıflar) [English curriculum (primary and secondary school 2nd,3rd,4th,5th,6th,7th, and 8th grades)]. Ankara.

Mendoza, A.S. (2011). Educational reform and teachers' decision making: relationship between the intended and the implemented reading curriculum in Dominican primary schools (Unpublished dissertation). State University of New York, Department of Educational Administration and Policy Studies, Albany.

Morgan, C. & Xu, G.R. (2011). Reconceptualising 'obstacles' to teacher implementation of curriculum reform: beyond beliefs. Mathematics Education and Contemporary Theory Conference. http://www.esri.mmu.ac.uk/mect/papers_11/Morgan.pdf

Nevenglosky, E. A., Cale, C. & Aguliar, S. P. (2019). Barriers to effective curriculum implementation. Research in Higher Education Journal, Volume 36.

Ntoi, L. (2007). Incorporating technology into the Lesotho science curriculum: investigating the gap between the intended and the implemented curriculum (Unpublished dissertation). University of Western Cape, Department of Education, Cape Town.

Özüdoğru, F. (2016). İlkokul 2. sınıf İngilizce öğretim programının diller için Avrupa ortak başvuru metni doğrultusunda aydınlatıcı değerlendirme modeli ile değerlendirilmesi [Evaluation of primary school 2nd grade English language teaching curriculum through illuminative evaluation model in line with the Common European Framework of Reference for languages] (Unpublished dissertation). Anadolu Üniversitesi, Eğitim Bilimleri Enstitüsü, Eskişehir.

Patton, M. Q. (2014). Nitel araştırma ve değerlendirme yöntemleri [Qualitative research and evaluation methods] (M. Bütün & S. B. Demir, Trans.). Ankara: Pegem.

Penuel, W.R., Phillips R.S. & Harris, C.J. (2014). Analyzing teachers' curriculum implementation from integrity and actor-oriented perspectives. Journal of Curriculum Studies, 46(6), 751-777.

Shah, F. H. (2015). An investigation into Lecturers' Beliefs and Implementation of the English Language Curriculum Change at Higher Education Level in Pakistan (Unpublished Dissertation). University of Exeter, Education and lifelong learning, Exeter.

Sidekli, S., Coşkun, İ. & Aydın, Y. (2015). Köyde öğretmen olmak: birleştirilmiş sınıf [Being a teacher in a village: multigrade classes]. Trakya Üniversitesi Sosyal Bilimler Dergisi, 17(1), 311-331.

Squires, D.A. (2009). Curriculum alignment: research-based strategies for increasing student achievement. Thousand Oaks, California: Corwin.

Summak, M. S., Summak, A. E. G. & Gelebek, M. S. (2011). Birleştirilmiş sınıflarda karşılaşılan sorunlar ve öğretmenlerin bakış açısından olası çözüm önerileri (Kilis ili örneği) [Problems encountered in multigrade classes and teachers' suggestions for possible solutions (a case study in Kilis province)]. Gaziantep Üniversitesi Sosyal Bilimler Dergisi, Sayı: 10(3), s. 1221-1238.

Şad, S. N. & Karaova, M. (2015). İlkokul ikinci sınıf İngilizce dersi bağlamında dinleme becerisi öğretimi: bir durum çalışması [Teaching listening skill within the context of primary school 2nd grade English course: a case study]. Eğitimde Nitel Araştırmalar Dergisi, 3(2), 66-95.

Turan Özpolat, E.T. (2015). Öğretmenlerin program uyumluluğu ve program uyumluluğunu etkileyen faktörlerin analizi (ortaokul 5. sınıf fen bilimleri dersi örneği) [Teacher's curriculum alignment and an analysis of the factors affecting curriculum alignment (the case of secondary school grade 5 science lesson)] (Unpublished Dissertation). Gaziantep Üniversitesi, Eğitim Bilimleri Enstitüsü, Gaziantep.

Varış, F. (1994). Eğitimde program geliştirme: teori ve teknikler (5. Baskı) [Curriculum development in education: theory and techniques (5th Ed.)]. Ankara: Alkım.

Verhoeven, P. & Verloop, N. (2002). Identifying changes in teaching practice: Innovative curricular objectives in classical languages and the taught curriculum. Journal of Curriculum Studies, 34:1, 91-102.

Walker, D. F. & Soltis, J. F. (2004). Curriculum and aims (4th Ed.) (J. F. Soltis, Ed.). New York: Teachers College.

Wiles, J. W. & Bondi, J. C. (2011). Curriculum development: a guide to practice. New Jersey: Pearson Education.

Yörü, B. (2012). İlköğretim sekizinci sınıf İngilizce öğretim programına ilişkin öğretmen görüşleri [Teachers' opinions on the 8th grade primary school English curriculum] (Unpublished master's thesis). Eskişehir Osmangazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Eskişehir.

Ziebell, N. (2010). Examining the alignment of the intended curriculum and performed curriculum in primary school mathematics and integrated curriculum (Unpublished master's thesis). The University of Melbourne, Melbourne Graduate School of Education, Melbourne.
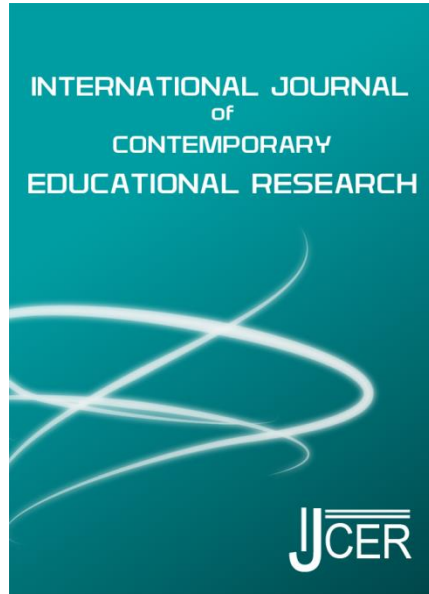
The Effects of Software-Aided Mind and Argument Mapping on Learning in Higher Education

Hülya Yıldızlı[1], İrfan Şimşek[1]
[1]İstanbul University- Cerrahpaşa

To cite this article:

The Effects of Software-Aided Mind and Argument Mapping on Learning in Higher Education

Hülya Yıldızlı[1*], İrfan Şimşek[1]
[1]İstanbul University-Cerrahpaşa

Abstract

By carrying out software-aided mind/argument mapping applications in the "Theories of Learning and Teaching" course offered to post-graduate students completing a master's degree in education, the present research aimed to examine the effects of the applications on academic achievement, as well as explore students' opinions on these applications. The research followed a mixed-method design integrating quantitative and qualitative research methods. Participants consisted of post-graduate students (six students) who were taking the course titled "Theories of Learning and Teaching". The experimental process lasted for 12 weeks. The experimental part of the research has been realized in two stages: In the first stage, the students were given the opportunity to learn about and practice preparing software-aided mind and argument mapping applications on computers and tablets for two weeks. In the second stage, participants were instructed about the theoretical dimension of teaching and learning processes and of the software-aided mind and argument mapping. Theories of Learning and Teaching Course Achievement Test, Reflective Diaries, Student Self-Assessment Forms and Semi-Structured Interview Schedules were utilized as data collection tools. The results suggested that the use of software-aided maps in learning and teaching processes has a considerably positive effect on students' academic achievement. Besides, it has been observed that the students acquired various skills with regards to the use of such tools.

Key words: Academic achievement, higher education, software-aided argument mapping and mind mapping

Introduction

The efficiency of learning and increasing productivity of students in an educational environment are considered as high priority goals in today's paradigms concerning learning and teaching. The significance of these paradigms is on the rise at higher education level. It is a significant question as to how the learning experiences of individuals should be, in order for them to develop knowledge in meaningful structures in mind and to be specialized in their own fields. Because of this importance, the paradigms that prioritize the active participation of students in the learning process emphasize the diversity of learning experiences at higher education level. The increasing enrichment in the variety of instruments in parallel with the rise of computer technology provided a fertile ground to nurture new experiences in learning environments. As a result, new software-aided tools came into prominence in learning environments. In the last 15-20 years, software-aided packages that facilitate the visual presentation of linkages (concept mapping, mind mapping or argument mapping), relations and information among different ideas were produced (Davies, 2010). The increase in information and speed gave rise to the need for using those tools with brain-friendly technologies (Buzan, 2016). Then what are these maps? What are their advantages in learning environments?

What is mind mapping?

Mind mapping, developed by Tony Buzan in 1970's, is a technique that enhances learning skills of individuals, their creativity as well as productivity (Mento et al, 1999). Activating certain parts of the human memory, a mind map is a graphic and network-oriented technique that utilizes keywords, images which promote new associations and ideas in order to store and organize a set of information. Each element of memory is activated

---

by this mind-mapping technique. This is a key aspect to unlock facts, ideas and knowledge as well as to unleash the spectacular potential of human mind (Buzan, 2006). Obviously, mind-mapping is not a simple note taking method. These maps are the non-linear visual representations of the ideas and interrelations of them (Biktimirov & Nilson, 2006) and a network of interrelated concepts. These free-forms are at the same time a means for finding out the creative relations among the ideas (Davies, 2010).

Since there are certain criteria as to how each technique, method and strategy is used, there are also certain rules to be followed in the case of the efficiency and productivity of mind-mapping. According to Buzan (2006; s.144-145), *the mind map has four essential characteristics:*
(1) The subject of attention is crystallized in a central image.
(2) The main themes of the subject radiate from the central image as branches.
(3) Branches comprise a key image or key word printed on an associated line. Topics of lesser importance are also represented as branches attached to higher level branches.
(4) The branches form a connected nodal structure.
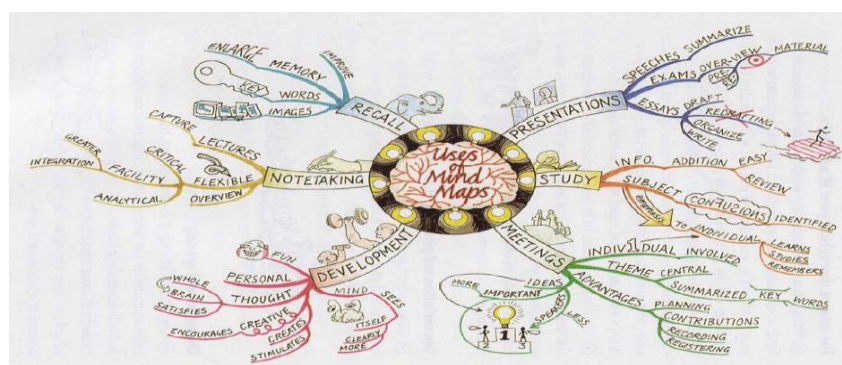
Figure 1: Mind map, source: Buzan (2006)

As can be seen in Figure 1, mind maps can be enriched by images, colours and words in the process of mapping. This dynamic structure also presents a kind of individual-specific learning map.

Mind maps can be utilized in almost any sort of activity that contains ideas, planning, or creativity (Buzan, 1989). These maps facilitate the formation of visual images to enhance the learning potential of the students and function as a reflective tool for creating wide associations among the learning materials (Budd, 2004); they also enable learners to evaluate their knowledge (Wandersee, 1987) and provide a better approach for the teachers in order to build up more efficient dynamics with those students who adopt varying ways of learning (Nesbit & Adesope, 2006). Having been utilized in all stages of a course (preparation, introduction, progression, evaluation), mind-mapping can also be useful in processes such as note-taking, problem solving, brainstorming, studying, planning, research, collection of data acquired by various sources, systematic presentation and forming opinions about complicated subjects and so on (Adodo, 2013).

The widespread use of technology in learning environments has enabled these maps to be created using a variety of phones, tablets, PCs or smart boards (see for example, *Elvedere*: Suthers et al.2001; *Digalo*: Schwarzand Glassner 2007; *Reason!Able*: van Gelder 2002, Miro, Mindmeister, Milanote, mind mapping; Figure 2: e-mind mapping). Such software provides a range of features that stimulate mental processes and creativity. Most web-based applications and software:

- Automatically create regular or coloured mind maps: It is easy to create e-mind maps using modern software. There is no space limitation in electronic maps as in handmade maps.
- These maps can be edited and developed at any time: Even after the mind maps have been created, they can easily be adjusted and re-shaped by new ideas.
- They allow for a more convenient analysis and management of data, as they facilitate the use of a range of tools: These maps render interaction with a large quantity of information much easier. As they become a visual and structured interface for accessing information sources, they prevent confusion and provide quicker and easy access to these sources.
- They enable sharing: They offer a wide range of ways to quickly share mind maps with other people (edition, images, web pages, graphics, and pdf).
- They can be transformed into various communication and report formats: it is an effective tool for presenting ideas which are difficult to express on free-hand maps. The software allows for the presentation of maps in several ways, setting up new branches one by one, making interactive presentations or focusing on the specific subjects (Buzan & Buzan, 2016).
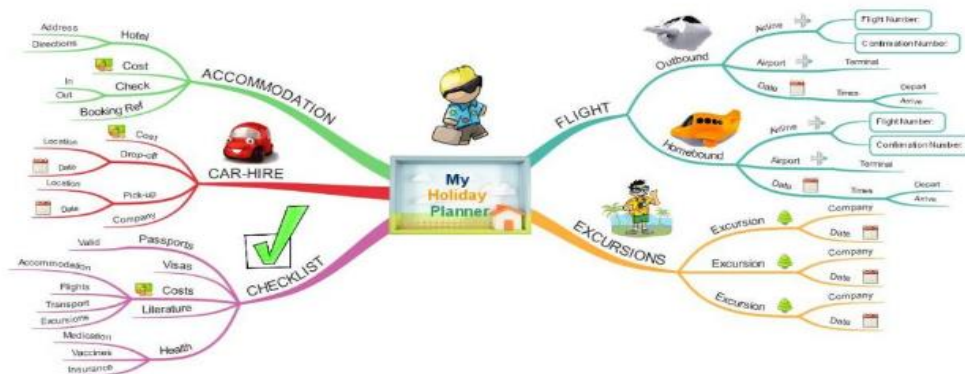
Figure 2. E-mind mapping sample (Source: https://www.toolshero.com/effectiveness/mind-mapping-buzan)

What is argument mapping?

As for the argument-mapping, it has a different format from the mind maps. These maps are important tools in revealing the inferential structure of the arguments. Inter propositional inferences are the main characteristics of argument-mapping (Davies, 2010). Argument mapping is the presentation of an argument in which the inferential structure is usually made clear by graphical techniques and these maps are "box" and "arrow" diagrams where the intersecting points correspond to the claims and highlight the evidentiary relationships of these connections (van Gelder, 2002). In an argument map, the argument is visually represented using a visual flowchart in "box-arrow" form; the boxes are for highlighting propositions whereas the arrows are used to highlight inferential relationships that combine the propositions (Dwyer, Hogan, & Stewart, 2012; van Gelder, 2003). On these maps, "boxes-arrows" in different colours can be used to support or object to (because, thus, but, however etc.) a certain argument.

The use of argument-mapping has increased in learning-teaching environments thanks to its ability to provision supportive settings for critical thinking of students (van Gelder et al, 2004; van Gelder, 2015). Enabling students to consider the informational content more critically, argument-mapping also allows students to control their own information structures and increase their meta-cognitive awareness towards learning (Dwyer, Hogan, & Stewart, 2012). By argument-mapping, students (1) develop a better critical attitude towards arguments (2) evaluate any structure of knowledge better (3) become more open-minded in their thinking processes (4) and get considerably better in argumentative writing (Rider & Thomason, 2014).

The use of software-supported argument-mapping in learning environments, similar to mind mapping, has recently become more widespread. The use of diagrams on ready-made software applications makes the learning process more efficient (Davies, 2012; Davies, Barnett, & van Gelder, 2019; van Gelder et al. 2004; van Gelder, 2007; van Gelder, 2015), functional (van Gelder, 2002) and saves on time by providing ready to use templates at hand (see Figure 3). These software applications enable students to quickly represent the reasoning by using box and line diagrams. These maps make the learning process fun, and allow students to unravel how different arguments function, provide them with opportunity to compare with different logical structures, and enable them to acquire a deeper and practical understanding of the structure of the arguments. Software-aided argument maps do not analyse or check the validity of arguments, but they help students to practically analyse and evaluate the arguments in a better way, as they encourage them to construct arguments openly and meticulously (Davies, 2010; Davies et al, 2019).
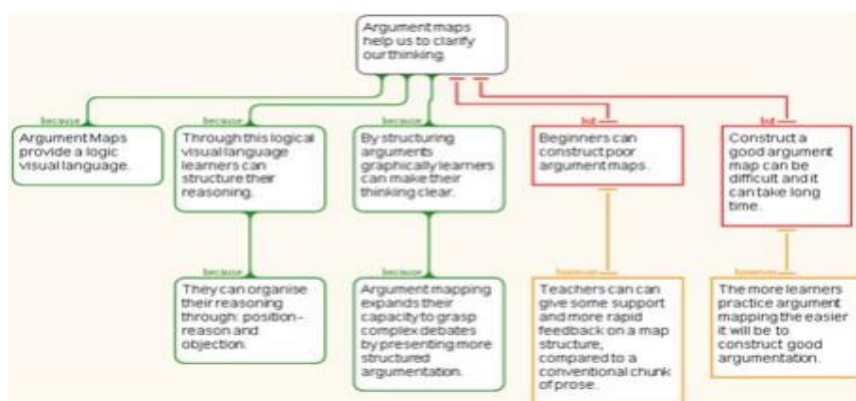
Figure 3: Argument mapping, (Source: Jasem, 2014, books.kmi.open.ac.uk)

The Purpose and Significance of the Research

Mind and argument maps, as mentioned above, are not limited to pen and paper. They are rather popular as software (Tergan, 2006; Zumbach, 2009). Considering the literature; there is a certain amount of research available on these maps and software-aided maps. These studies have focused on the teaching process of different disciplines at different grade levels (including higher education level; Fu, Lin, Hwang, & Zhang, 2019; Israel, Zipp, D'Abundo & Deluca, 2020; Kunsch, Schnarr, & van Tyle, 2014; Lin, Shadiev, Hwang, & Shen, 2020; Luo, 2019, Wu & Wu, 2020). A number of studies also focused on developing higher-order thinking skills (reflective thinking, questioning, reasoning, critical thinking, problem solving etc.) as well as academic achievement (Aljaser, 2017; Butchart et al. 2009; Bütüner & Gür, 2008; Carrington, Chen, Davies, Kaur, & Neville, 2011; Davies, Barnett, & van Gelder, 2019; Eftekhari, Sotoudehnama & Marandi, 2016; Evrekli & Balım, 2010; Jeong, 2020; Mento et al, 1999; Rider & Thomason, 2014; van Gelder & Rizzo 2001; van Gelder et al. 2004; van Gelder, 2007).

The present research attempted to find out whether such software-aided maps have any influence on the academic achievement of graduate students studying in faculties of education. The graduate programs in these faculties have very important goals. In the context of National Qualifications Framework for Higher Education in Turkey (NQF-HETR, 2019), the students at this level are obliged to develop and deepen their knowledge in their area to an expertise level; be able to utilize their theoretical and practical knowledge at an expertise level; be able to integrate and interpret the interdisciplinary knowledge and develop new knowledge by doing so; they are also expected to critically evaluate their expertise level knowledge of their own area and regulate their own learning processes, have advanced level skills on information/communication technologies and software in proportion to the requirements in their field; have ability to make use of their acquired knowledge in problem solving and use their practical skills in interdisciplinary areas. It is clear that post-graduate students are expected to gain meaningful in-depth knowledge of their field, as well as other skills, attitudes and habits. Therefore, it is necessary to provide them with rich learning experiences that support meaningful learning at the graduate level. The assessment of the preliminary knowledge of students, the use of materials which clearly reflect the content of students' basic knowledge, the active participation of students in order for the meaningful development of knowledge and determining as to how the process of learning takes place (or not) for them are all quite important elements (Hay, Kinchin & Lygo-Baker, 2008). This situation also leads to the need for applied studies in the graduate education. This research, thus, originated out of this significance and need. The teachers or prospective teachers who have graduated from the faculties of education are expected to learn various methods and techniques in their teaching practices and how they should be used. These knowledge and skills can only be built in the environments that support learning by practice and experience.

This research has been carried out within "Theories of Learning-Teaching" course that is offered at the faculties of education. The course is very significant in those faculties. By virtue of the knowledge and skills that are obtained in the course, the teacher or prospective teacher can make a plan of teaching in an easier way, he/she can have a better judgement about the educational applications that emerge from certain theories, and he/she can better analyse how/where these theories should be used in the process of learning-teaching (Gredler, 1987). Because of the significance of the course, it is necessary for graduate students to have enriched learning experiences about this course. In addition, students in higher education deal with arguments all the time. Indeed, a major purpose of a higher education is to teach students how to read, understand and respond to complex arguments. The ability to do this makes for highly employable, adaptable and reflectively critical individuals (Davies, 2011). Based on this idea, in this research, certain software-aided mind mapping and argument

mapping techniques were utilized in order to provide the students with the appropriate way of meaningful learning and it was intended to answer the following questions concerning the research:

1) Is there any significant difference between the pre-test and post-test mean scores of the experimental group using software-aided mind mapping and argument mapping?
2) What are the opinions of the students who participated in the experimental research about the application process?

Method

Having focused on the mind and argument mapping techniques and their influence on the process of learning, the research has been designed following a mixed methods research strategy synthesizing quantitative and qualitative methods. Qualitative and quantitative methods were administered simultaneously and given equal priority; however, data collection and analysis were carried out separately and the results were combined while making general interpretation of the data. The quantitative part of the research has been carried out by the experimental research model (one group pre-test post-test design). The qualitative part, on the other hand, was undertaken as a case study. The experimental design of the research is displayed on Table 1. As it is seen on Table 1, the research was carried out with an experimental group. The reason for carrying out the research with a single experimental group is that the number of graduate students and groups was not high enough. This research was completed within the scope of a scientific branch (more specifically Primary Education Master Program) of a state university. The students from the faculties of education apply to study at various scientific branches with their varying points, knowledge and skills. Therefore, it was not possible to find any group that would be the equivalent of the existing one. However, there might be some other groups from other universities who study at the same departments and conduct research in similar areas, which might qualify as the equivalent groups. But in that case, the course would have needed to be taught by different professors. In the literature, many research studies were carried out with one-group experimental design (Öksüz & Coşkun, 2012; Siah, 2019; Yamak, Bulut & Dündar, 2014). In line with the literature, the present research was carried out with one group. In addition, the study emphasized the following: the study was conducted in a single group, but more than one qualitative data collection tool was used to analyse the effects on the students in depth.

The measurement tools within the scope of the research were administered to the experimental group as a pre-test. The experimental process lasted for 12 weeks. The students in the experimental group passed through certain processes such as learning about the ways in which mind and argument maps are created, performing certain tasks related with software-aided mind and argument maps, and making evaluations about the process of learning. In addition, the experimental group was given achievement tests and interview forms at the end of the tasks.

Table 1. The Experimental Design of the Research

Groups	Pre-Test	Experimental Process	Post-Test
Experiment	TOLTCAT 1	The applications aiming at the technology-oriented preparation of mind/argument maps	TOLTCAT 2
		The teaching processes supported by mind/argument maps; the teaching evaluation processes	

TOLTCAT: Theories of Learning-Teaching Course Achievement Test

The qualitative part of the research was conducted using a case study design. The study also attempted to illuminate how the students get through the experimental application process and how they are affected by the process. For this reason, the interviews and document analysis techniques were used so as to collect qualitative data. Three types of qualitative data were collected:

a) First, at the beginning of the experiment, focus group interviews were held with the experimental group to collect information regarding the extent to which they used software-aided mind and argumentation maps during/outside the course as well as their perceived awareness and importance of using these techniques. The reasons for the increase in their awareness: Pointing out to the practices to be done, inspiring their interest, providing motivation for the endurance of learning and the use of these tools in learning.

b) Secondly, the reflective diaries and self-assessment forms written by the students during the experiment were continuously reviewed by the researchers. These forms were collected regularly.

With these data collection tools, the students were provided with feedback about their learning processes.

c) Finally, at the end of the experiment, one-to-one interviews were conducted with the students about the study.

Study Group

The study group of the research is composed of the graduate students at the Department of Elementary Education. The work group consisted of 6 students in total: 2 men, 4 women.

Data Collection Tools

For the research, the following data collection tools have been used:

The Theories of Learning-Teaching Course Achievement Test

Certain achievement tests that consist of open-ended questions regarding the content of the course were created in the study. Open-ended questions were composed of the questions that measured such skills of the students as thinking, reasoning, comparison, generalization, application, analysis, evaluation and so on. Following the preparation of the questions, experts were consulted, and revisions to the questions were made in the measurement tool in line with the feedback received from experts. There were 25 questions in the final test. Because of the high number of questions, the exam took place over three sessions both before and after the experimental procedure. It is quite difficult to grade the answers to the open-ended questions. Therefore, the following principles were taken into consideration in order to ensure the reliability in grading the student answers: (1) It was made clear in the answer key, which outlines the students should mention. (2) Holistic grading method was preferred. (3) While the students' answers were being graded, it was endeavoured to grade the certain answers of all students simultaneously and all at once. (4) In the process of answer grading, the confidentiality of the student names was ensured.

Student Self-Assessment Form

The Student Self-Assessment Form was utilized to evaluate the mind and argumentation maps that the students prepared after the course**.** In these assessment forms, it has been endeavoured to evaluate whether the students had sufficient factual, conceptual and operational knowledge that they acquired in the course about the teaching-learning theories. (1) *Do I know the significant concepts in …. Theory? What are they?* (2) *Do I know the basic assumptions of …Theory? What are they?* (3) *Could I explain the basic theoretical differences between … and … Theories?* Besides, the maps that were prepared by the students were re-evaluated in the form of feedbacks offered to students by the teachers. In accordance with the feedbacks given by the teacher, students were asked to evaluate themselves in these self-assessment forms. For example: (1) *Are the maps that I drew theoretically reliable?* (2) *Have I accurately reflected the details of the theory on my mind maps?* (3) *Have I formed the connections between concepts in my map in accordance with the theoretical framework*?

Reflective Diaries

The reflective diaries are the ones that students had written at the end of each lesson. These were the diaries that covered the feelings and views of students about what they learned in the course, what they found interesting, what they should examine, and how they generally evaluated the maps that they prepared. The researchers had an informative meeting about the writing of the diaries with the students prior to the experimental process. The diaries were collected from the students regularly (each week) and a general evaluation was made.

Semi-Structured Interview Forms (One-to-one Interviews and Focus Group Interviews)

At the beginning of the experimental process, the focus group interviews were conducted with the experimental group about the extent to which the students make use of the software-oriented mind and argumentation maps and the significance of using these techniques and raising their self-awareness of such tools. All participants were present in this meeting. This interview process lasted two hours.

At the end of the research, semi-structured interview forms were utilized to make interviews with the students

about the experimental study. The questions of this interview schedule were re-evaluated in accordance with the expert views. One-to-one interviews with the students lasted on average for 45 minutes. To provide examples of the questions in the interview form: (1) *What do you think about the use of mind and argument maps in learning?* (2) *What do you think about the theories being discussed and learned on the map?* (3) *What do you think about mind and argument maps?* (4) *Do you find it enjoyable to use these maps during and after class? And why?*

Procedures

The experimental phase of the research took place in the following order:

1) The imindmap software was utilized for mind mapping, and Rationale^{TM} program for argument mapping in the present study.
2) The experimental process was 12 weeks long.
   In the first phase of the experiment, informative lessons were held with the students over two weeks about what these maps are and how they are used. During this period, they were also informed about the use of the above software applications. Software-based mind and argument mapping practices were carried out with the students. Sample titles related to the content of the two-week process include: conceptual framework of maps, their usage purposes, advantages, types of software related to maps, software's that were used in the research, how they were used, making applications on maps, and so on.
3) The application process in the research after the two-week introduction lessons of mapping is as follows: Mind mapping activities took place in two ways. First of all, prior to the relevant subject and before the beginning of lessons, the students were asked to create maps that indicated their prior knowledge about the subject. Secondly, following the end of the course, the students were asked to re-create mind maps on the basis of the relevant sources in the course and course notes outside the class. These re-created mind maps were evaluated together with teacher and the students. In this evaluation process, the students introduced their maps and made comparisons between the maps they created before/after the classes. Thus the students were ensured to make a self-assessment by comparing the two mind maps. At the same time, with the mind-mapping software, the students were provided with the opportunity to uncover the conceptual structure of the theory via the mind-mapping software.
4) On the other hand, the students integrated argument mapping into their learning processes as follows: By taking into account the assumptions and principles of the learned theory, they clarified their reflections on the learning-teaching process through in-class discussions and demonstrated these points with argument maps. For example: They were provided with opportunities to enrich the maps using questions such as: How do we structure the learning-teaching process when we consider the basic assumptions and principles of behavioural theory?
5) Students were asked to share the maps they created with one another. The aim here was to enable students to see how information can be organized in different minds, and enrich the learning activity by asking questions to one other.
6) During the experiment, students were given the opportunity to write their reflective diaries and self-assessment forms. It was assumed that these data collection tools would contribute to students' learning processes.

Data Analysis

Both quantitative and qualitative data analysis techniques were employed in the study. The data of the achievement test was analysed by using a statistics program. The data in the scope of the student self-assessment form, reflective diaries and semi-structured interviews, on the other hand, were analysed descriptively.

And also, in order to ensure validity in the study, (a) Expert opinions were obtained in the arrangement of the data collection tools used, (b) Necessary participant confirmation (consent) was obtained after the interviews, (c) The research process was explained in detail, (d) The researchers engaged in a long-term interaction with the study group (prolonged engagement). Different data collection techniques were used together (triangulation) in order to increase the reliability of the research. The methodological process of the research has been expressed in depth. The necessary expert opinions were obtained for the analysis results of the research data.

Results

Within this section, quantitative and qualitative results will be evaluated separately. The data is presented under the categories; (1) the data from achievement tests, (2) the data from self-assessment forms – reflective diaries, and (3) the data from semi-structured interviews.

The Results Obtained from Achievement Tests

The data presented in Table 2 and 2 show meaningful differences between the pre-test and post-test average points of the students who were in the experimental group. As expected, this meaningful difference was in favour of the post-test points. (z=2,05, p<.05). The results indicate the positive effect of the software-aided maps on the students' academic achievement.

Table 2. The results of pre-experimental and post-experimental administration of achievement tests

Post-test-pre-test	N	Mean Rank	Sum of Ranks	z	p
Negative Ranks	0	,00	,00	-2,214[*]	,027
Positive Ranks	6	3,50	21,00		

[*]Based on negative ranks

Table 3. Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Pre-test	6	28,83	1,941	26	31
Post-test	6	105,00	3,347	101	110

The Results Obtained from Self-Assessment Forms – Reflective Diaries

The results obtained from self-assessment forms and reflective diaries are presented in Table 4.

Table 4. The Results Obtained from Self-Assessment Forms

Reflective Diaries	f
A deeper comprehension of the subject	6
Enabling self-assessment	6
Providing a more enjoyable way for in-class/out-class learning	6
Learning to ask questions to the self	6
Providing individual feedbacks effectively	5
Enabling the evaluation of one's prior knowledge	5
Enabling the self to discover learning strategies	6
Enabling the self to form stronger associations among the sets of information	6
Enabling the self to be more cautious in out-class studies	5
Making the remembering easier in the exams	5
Providing internalized hints related to the subject	4
Enabling the self to be more effective in in-class discussions	5
Enabling the self to make more effective presentations of one's ideas in the class	5
Using them effectively in one's professional development	3
It is easier to draw such maps with pen and paper	1
Developing higher-order thinking skills	5
Increasing the interest in/ motivation for learning	4
Increasing academic self-efficacy in the course	3
The philosophical and historical backgrounds are quite important in the formation of argument maps. Otherwise, it leads to the formation of weaker argument maps.	3

Table 4 suggests that self-assessment forms and reflective diaries were used to increase students' activity in the learning process. The majority of the students stated that the use of mind mapping and argument mapping enabled them to learn the subject in-depth, make self-evaluation, render the in/out-class learning environments more enjoyable, and evaluate the prior knowledge about the subject. In relation to this, for example, S2

explained: *At the end of the lesson, I created maps at home again. I searched different sources about how to enrich the map. Then I noticed the increase in my interest in learning and research*. Besides, S6 claimed: *"While evaluating myself with these maps, I also realized that I was actually questioning how to get the right information"*. S3 reported: *"Comparing my first map with the last one, I was quite happy to see how much progress I made"*.

In addition, the students stated that learning with these maps increases attention in extra-curricular activities, facilitate easier recollection of knowledge in exams, can be used as an effective tool in class discussions, enable them to develop high-level thinking skills, and increase interest, motivation and self-efficacy in learning. To illustrate, S5 commented: *"Questioning these theories thoroughly in classroom discussions allowed me to ask myself 'how can one ask different (/genuine) questions?'*. And S2 noted: *"The discussions in the classroom and the examples that came forward during these discussions made it easier for me to recall the information I learned"*.

Moreover, some of the students stated that the effective utilization of these maps in the learning processes had positive contributions to their professional development. For example, S1 put forward the following argument: *"Having studied and mapped the relevant subjects, I wrote questions to myself and answered them all. This is the way I administered them to my own students... They and I have realized that we concentrated our attention on the important and cornerstone elements of the subject in this way"*. In addition, S4 reported that: *"Learning how to learn also teaches one how to teach. For this reason, it led me into questioning my own teaching process"*.

It is an important finding that one of the participant students claimed that it was in fact easier to draw such maps with paper and pencil. Some students also claimed that it was necessary to know the philosophical and historical backgrounds of theories in creating argument maps; otherwise these maps cannot be effective as a learning method. To provide an example, S3 noted: *"While analysing the theories, we also realized that these theories were also affected by the results of a variety of disciplines. Therefore, we need more philosophical and historical background readings"*.

The Results Obtained from the Semi-Structured Interviews

The results obtained from the semi-structured interviews and focus group interview with the students are presented in Table 5.

Table 5. The Findings Obtained from the Semi-Structured Interviews

(a) The results that were obtained from the focus group interview prior to the experiment	f
Not having used software-aided mind maps at all	6
Having heard about what mind maps are during undergraduate classes	5
Not having heard about/not having used argument maps in learning process at all	6
Having a limited knowledge about the purposes and benefits of these maps	6
(b) The results that were obtained from one-to-one interviews after the experiment	
1) The Usage Purposes of Mind and Argument Maps	
Personal planning	4
Note-taking and summarizing whilst studying	6
Searching for alternative solutions to the problems	3
Summarizing books/articles (searching new sources of information)	4
Producing new ideas	5
Making self-assessment	6
Providing permanence (of knowledge) in learning	6
Making learning processes enjoyable	6
Providing associations of the acquired knowledge with the other dimensions of life	6
Discovering the personal ways of learning	6
Providing an easier means to remember information	6
For teachers, learning how to use them in the processes of teaching	5
Discovering alternative ways of thinking (Critical, reflective thinking etc.)	6
Saving time for ourselves	5
In-class assessment and evaluation activities	6
Providing students with alternative perspectives on their current knowledge	6
Making it easier to notice the connections among the factual, conceptual and operational knowledge	6
2) The students' perspectives on the mind and argument-mapping	

Concrete schematization of the current associations, likening's and connections

Summarizing a subject	6
Providing a holistic view on a subject	6
A sensible way of note-taking and noting ideas as map forms	5
Organized, colourful, mnemonic schematization of any long and/or monotonous set of information lists in parallel with the natural stream of consciousness	3
A note-taking method	3
Putting an internalized learning method on paper/computer	2
Visualization of a subject	5
A practical technique that is used to associate different subjects and encapsulate subjects	2
A method that is utilized to make an effective use of knowledge via facilitating a better flow of knowledge in the brain	2
Transformation of a list consisting of monotonous information into a mnemonic, colourful, organized schema that is compatible with the natural processes of the brain	5
Like a city map	1
A mirror of your own natural thinking	3
Being able to compile ideas	

3) The advantages of using the mind and argument-mapping

It makes learning enjoyable by pictures and shaped note-taking methods.	6
It provides a better assimilation of information and ideas in the process of creating maps.	6
Increases attention in the learning process.	3
Unlike classical and long notes, it facilitates the repetition and recall of acquired knowledge.	6
It provides an effective and efficient use of time in learning.	6
It supports creative thinking.	
It allows seeing the main points and sub-headings related to the acquired knowledge.	5
It enables the visuals, forms and concepts of the subject to be associated and integrated in all aspects.	5
Strengthening the neural connections in the brain, it ensures the long-term storage of the acquired knowledge.	3
It helps to discover one's own learning methods.	6
Mapping of the subject results via more mental associations.	4
It increases the level of comprehension in developing connections with real life through increasing mental associations.	4
It provides a quick command of the subject.	5
It provides an opportunity to evaluate and compare theoretical aspects from a holistic point of view.	6
It is effective in comparing different theories in terms of their own characteristics and noticing the differences between them on a single map.	6
It allows spotting the occurrences of the deficiencies in our leaning process.	6
It raises the effectiveness in the assessment of learning.	6
It helps to overcome the presupposed limitations in learning.	3
It improves argumentative reading, writing, questioning.	5

Analysis of Table 5 in relation to the focus group interview with the students in the experimental group prior to the experiment indicates: none of the students had ever used software-aided mind maps, heard of argument maps before, or used them in learning processes. Students also noted that the first time they heard about mind maps was during their undergraduate studies and added that they had limited knowledge about the purposes and benefits of using these maps.

One-to-one semi-structured interviews were conducted with the students after the experiment. These interviews were conducted to illuminate the effects of using mind-mapping and argument mapping on learning processes on the students. Student views were categorized as the *usage purposes of these maps*, *their subjective perspective*s that are formed in the process of using them and the *advantages of using these maps*. As seen above, the students in the experimental group stated that they used software-based mapping tools in their personal planning, seeking alternative solutions for problems, producing new ideas, making learning enjoyable, and developing different ways of thinking. Considering the students' views on the meaning of these mapping tools; the majority of the students found the maps quite functional in summarizing the subject, grasping the subject from a holistic view, effective note-taking, visualizing the subject, and transformed it into a colourful mnemonic organized scheme. And regarding the advantages of using mind mapping and argument mapping, the majority of students considered that they make learning more enjoyable, they are easy to integrate a set of

information, and they facilitate the efficient and effective use of time in the learning process. They also help to discover personal methods of learning and were quite useful in the comparison of various theories of learning. In addition, the students stated that using these maps was effective in the assessment of learning, in developing different mental connotations, the mastery of the subject in a short time, and overcoming any limitations to the learning.

To provide few examples of direct quotations from the students.

S1: *In my opinion, it provides an enjoyable learning process by making a difference with regard to the narratives in the books where the theoretical knowledge is very intensive; and promoting creative thinking. The interesting and fun part is that our brain makes interesting associations with the objects and events we encounter in our own lives and improves the permanence in mind... I think it is pleasant enough for our own personal learning process.*

S3: *In the mind and argument maps I have created; I think that it is quite supportive for the creative thinking to utilize the connections and connotations I have formed on the subject that I work on. Besides, upon mapping some visuals and connotations, I realized that I could expand the subject further and that the subject becomes more persistent (in the memory). When I mapped the topics separately, I realized that I could make more extensive associations.*

Discussion and Conclusion

In this study, it is clear that the use of software-aided mind-mapping and argument-mapping had a positive effect on the academic achievement of the post-graduate students who studied at the faculties of education. Comparing the academic achievement pre-test and post-test scores, a significant difference was found in favour of the post-test scores of the experimental group. As one can see, the instructional use of these maps helps the development of students' knowledge and skills. National and international literature found that that the academic achievement of the student's increased after the use of such software-aided mapping tools (Adodo, 2013; Akınoğlu & Başar, 2007; Balım, 2013; Bessick, 2008; Bütüner & Gür, 2008; Eftekhari, Sotoudehnama, & Marandi, 2016; Evrekli & Balım, 2010; Israel, Zipp, D'Abundo, & Deluca, 2020; Wu & Wu, 2020; Zumbach, 2008). The results of studies in the related literature are compatible with the results of the present study. Knowledge mapping applications contribute to the effective management of information and lead to increased success (Buzan, 2006, Buzan & Buzan, 2016). Therefore, the mind and argument maps which provide meaningful and permanent learning could be used frequently in learning environments.

In this study, the opinions and reflections of the students who participated in the experimental group were analysed in depth and it was found that quantitative and qualitative research results support each other. The qualitative findings obtained from the self-evaluation forms, reflective diaries and semi-structured interviews (one-to-one interviews and focus group interview) were utilized for in-depth illustration of post-graduate students' views participating in the study. Considering the results obtained from the self-assessment forms and reflective diaries, it can be argued that the use of software-aided mapping tools in the learning process enables graduate students to learn the subject in depth, make self-assessment and evaluate the prior knowledge about the subject. Based on the students' views, the following advantages of these tools were reported in enriching the learning process: (1) more attention and care in extra-curricular studies, (2) studying becomes more enjoyable, (3) it becomes easier to repeat and recall the information, 4) in-class discussions become more effective, (5) high-level cognitive skills are developed, (6) increased interest, motivation and self-efficacy in learning. It is also important to note that these applications have a positive impact on the professional development of the students. The direct quotations from students also support these results. There was also one student who had a negative opinion about the use of these maps in learning processes. The student claimed that it was, in fact, easier to draw such maps with pen and paper. Actually, there are advantages in using these software-aided maps. Because the students can make any adjustments in these maps based on preference, they can add any visual that they want, and expand the information structure as much as possible (Buzan, 2006). This (student opinion) can be interpreted as follows: Affective factors such as attitudes and beliefs about technology should be taken into account in technology supported learning and teaching practices since such factors can influence an individual's actions in teaching-learning processes. Therefore, these technology-related variables may also lead to the formation of certain judgments in the learning process. In the literature, it has been indicated that the individual attitudes and beliefs towards technology can affect the usefulness or ease of technology (Shroff, Deneen & Ng, 2011; Teo, 2009; Wixom & Todd, 2005).

The results obtained from the semi-structured interviews conducted with the students at the end of the experimental application process were in line with the above-mentioned results obtained from the self-assessment forms and reflective diaries. The findings from semi-structured interviews were categorized as "the usage purposes of maps", "the meanings of maps for the students", and "the advantages of using these maps". Participant s views on the intended purposes and advantages of the maps suggested that they can be used for personal planning, studying, seeking for different solutions to problems, searching new sources, producing new ideas, providing permanence in learning, discovering personal ways of learning, forming connection of the acquired knowledge with the other areas of life, developing different thinking ways, in-class measurement and assessment activities, making a better analysis of the types of information related to the subject area, considering the subject in a holistic way, evaluating the theoretical knowledge in various aspects, making comparison of the theories easier, developing argumentative reading, writing and questioning skills and so on. It was clear from student comments that the maps that they had been making use of became very useful and effective tools in their learning processes and other domains of their lives. The results obtained from the focus group interview following the experimental procedure also indicated such a development. According to the findings in the focus group interview, it turned out that the students had never made use of these software-aided maps, they heard about them in the classes for the first time and had very limited prior knowledge concerning the purpose and benefits of the maps. The changes in student views have been remarkable. Therefore, one can argue that these maps, which allow enriched learning experiences, are effective in developing students' cognitive and affective skills. Especially the graduate-level studies aiming at the development of high-level skills should support such enriched learning experiences. The findings obtained from the literature concerning the use of argument and mind maps in learning processes are also compatible with the results of the study. For instance, Zumbach (2008) and Jones et al (2012) found positive effects of the mapping techniques on the motivation of students. Moreover, Butchart (2009), Davies (2009), and Dwyer et al (2011; 2012) concluded that the argument-mapping practices promote the critical thinking of students as well as making the learning interesting and more creative (Koznov & Pliskin, 2008). Similarly, Rider and Thomson (2014) argued that these maps contribute to students' argumentative reading, writing and their progression in the questioning skills. Meta-analysis studies also show that argument-mapping-supported learning environments are highly effective in developing students' critical thinking skills (van Gelder, 2015). Adodo (2013) attempted to show that the use of mind-mapping tools raises students' creative thinking and creative problem-solving skills, facilitates memory recall, and strengthens connections with other disciplines. Similar to these findings, Evrekli et al. (2009), in their study with prospective teachers, came to the conclusion that their use of mind mapping can have certain advantages in the purposes of effective note-taking, in-class assessments/evaluations, promoting the permanence of acquired knowledge, and discovering the effective methods in learning. In addition, in this experimental process, the persistent use of these maps before, during and after the learning procedure enabled students to assess their own learning. It was found out that the student who checked his/her prior knowledge was able to structure the learning process more easily. This leads one to expect that their certain attitude toward learning shows up in such a way that renders it more complex and sophisticated. In order to improve the quality of learning at higher education level, it is recommended that such maps be used as the prime strategy in the assessment/evaluation processes (Hay, 2007). Likewise, the findings of this study support such suggestions. The present research also studied as to what sort of meaning structures the participant students developed with regard to these maps. On the consideration of the student views on the mind and argument maps, it is seen that they have given them the following set of meanings: a note-taking method, a concrete schematization of acquired information in mind, mapping the ideas, visualization of the subject, a practical technique to encapsulate the relevant subjects, (something) like a city map, a natural mirror of one's ideas, being able to compile the ideas and so on. The set of meanings that is given by those students is also compatible with the findings from the quantitative and qualitative data collection tools. It was found that the meanings of the maps the students developed in their minds match up with the definitions of these maps in the literature. It is important for this research to consider the meanings formed by the students. Because it is important to practice through them and to direct the learning process through these structured meanings in educational environments. The formation of these meanings by the students shows that the experimental application process has achieved its prime goal. These maps that were used in the scope of the research are at the same time an active note-taking method. Because the standard note-taking methods lack the kind of elements such as visual rhythm, patterns, colours, images, dimensions, spatial awareness, integrity, connections and so on. Those missing elements, however, are crucial in all brain functions, especially in the process of learning. The standard notes conceal the keywords, cloud the recollection of information, and cause a waste of time (reading and reading over the insignificant notes), and they are unable to activate the creative functions of the brain (Buzan & Buzan, 2016). Therefore, it has been argued that utilizing mind-mapping techniques are quite advantageous in learning processes by virtue of activating the higher-order learning potential of the brain. As it has been clarified throughout this article, this research supported the idea that the use of software-aided mapping tools is quite effective in post-graduate students' learning processes and their development of knowledge, skills, and behaviours (as mentioned above). These can be very convenient

techniques that can be used in different stages of the course. These mapping techniques can also be quite useful for creating modern learning environments which are convenient for today's world. Such stages of learning in which theoretical knowledge must be internalized should support the students' development in different ways of thinking. In particular, undertaking the study with post-graduate students in higher education was very important for the students in the sense of experiencing how the learning processes can be enriched in practice. The data obtained from the students support this interpretation.

Recommendations

It is also a necessity to conduct this research in various courses at undergraduate and graduate levels. Those software-aided applications may provide favourable learning environments for courses such as critical and analytical thinking, philosophy of education, and sociology of education. Future studies can investigate the effectiveness of these software-supported maps in such courses. And those can be used for further studies on solving any problem in the classroom, developing the out-class learning environments of students, solving social problems, raising their awareness about the variables that may affect their learning processes and finding out personal solutions in such situations. In fact, these maps can become useful tools for teachers when distance education and digital education tools are important during the Covid-19 pandemic, For example, these tools can be used effectively by teachers before, during and after the courses. In the limited time period of distance education, these maps will facilitate teachers' evaluation processes to reveal the conceptual structures of students in more depth and how the students analyse the arguments. At the same time, using these learning tools as tools that enable students to assess themselves also ensures meaningful learning experiences. These maps can help students organize their notes outside of the classroom and more easily absorb important information with additional reading. This provides them with the opportunity to develop an in-depth understanding of the subject area (Dwyer, Hogan & Stewart, 2012). The limitation of this study was that it was conducted with one group and within one department. Furthermore, the research was carried out within the scope of only one course. These maps can be used as a tool for the integration of disciplines in a study conducted in the teaching of many disciplines. It is also possible to conduct further studies that will also take various affective variables into the account.

Acknowledgements

References

Adodo, S. O. (2013). Effect of mind-mapping as a self-regulated learning strategy on students' achievement in basic science and technology. *Mediterranean Journal of Social Sciences*, *4*(6), 163-172.

Aljaser, A. M. (2017). The effectiveness of electronic mind maps in developing academic achievement and the attitude towards learning English among primary school students. *International Education Studies*, *10*(12), 80-95.

Balım, A. G. (2013). The effect of mind-mapping applications on upper primary students' success and inquiry-learning skills in science and environment education. *International Research in Geographical and Environmental Education*, *22*(4), 337-352.

Bessick, S. C. (2008). Improved critical thinking skills as a result of direct instruction and their relationship to academic achievement. Doctoral dissertation. Indiana University of Pennsylvania.

Budd, J. W. (2004). Mind maps as classroom exercises. *The Journal of Economic Education*, *35*(1), 35-46.

Butchart, S., Forster, D., Gold, I., Bigelow, J., Korb, K., Oppy, G., & Serrenti, A. (2009). Improving critical thinking using web based argument mapping exercises with automated feedback. *Australasian Journal of Educational Technology*, *25*(2), 268-291.

Bütüner, S. Ö., & Gür, H. (2008). Teaching Of Angles and Triangles by Using Vee Diagrams and Mind Maps. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education, 2*(1), 1-18.

Buzan T. (2006). *The Buzan study skills handbook: The shortcut to success in your studies with mind mapping, speed reading and winning memory techniques*. Harlow: BBC Active.

Buzan, T. (1989). *Use both sides of your brain: New mind-mapping techniques to help you raise all levels of your intelligence and creativity based on the latest discoveries about the human brain*. Penguin Books USA Incorporated.

Buzan, T. & Buzan, B. (2016). *The mind map book: Unlock your creativity, boost your memory, change your life*. (Trans: Tercanlı, G.) England: Pearson Publishing.

Carrington, M., Chen, R., Davies, M., Kaur, J., & Neville, B. (2011). The effectiveness of a single intervention of computer-aided argument mapping in a marketing and a financial accounting subject. *Higher Education Research & Development*, *30*(3), 387-403.

Davies, M. (2010). Concept mapping, mind mapping and argument mapping: what are the differences and do they matter? *Higher education*, *62*(3), 279-301.

Davies, M. (2011). Computer-Aided Argument Mapping: An Educational Revolution. *New Media, New Literacies, And New Forms of Learning; Institute Of Education. London*, 1-7.

Davies, M. (2012). Computer-Aided Mapping and the Teaching of Critical Thinking: Part I. *Inquiry: Critical Thinking Across the Disciplines*, *27*(2), 15-30.

Davies, M., Barnett, A., & van Gelder, T. (2019). Using computer-assisted argument mapping to teach reasoning to students. In *Studies in critical thinking*, (chapter: 10) Windsor Studies in Argumentation publisher.

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2011). The promotion of critical thinking skills through argument mapping. In *Critical thinking*. (Eds: Christopher P. Horvath and James M. Forte). Nova Science Publishers, Inc.

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2012). An evaluation of argument mapping as a method of enhancing critical thinking performance in e-learning environments. *Metacognition and Learning*, *7*(3), 219-244.

Eftekhari, M., Sotoudehnama, E., & Marandi, S. S. (2016). Computer-aided argument mapping in an EFL setting: does technology precede traditional paper and pencil approach in developing critical thinking?. *Educational Technology Research and Development*, *64*(2), 339-357.

Evrekli, E., & Balim, A. G. (2010). The effect of use of mind mapping and concept cartoons in science and technology education on students' academic achievements and inquiry learning skill perceptions *The Western Anatolia Journal of Educational Sciences,* 1(2), 76-98.

Evrekli, E., Balim, A. G., & İnel, D. (2009). Mind mapping applications in special teaching methods courses for science teacher candidates and teacher candidates' opinions concerning the applications. *Procedia-Social and Behavioral Sciences*, *1*(1), 2274-2279.

Fu, Q. K., Lin, C. J., Hwang, G. J., & Zhang, L. (2019). Impacts of a mind mapping-based contextual gaming approach on EFL students' writing performance, learning perceptions and generative uses in an English course. *Computers & Education*, *137*, 59-77.

Gredler, M. E. (1997). *Learning and instruction: Theory into practice*. New Jersey: Prentice-Hall.

Hay, D., B. (2007). Using concept mapping to measure deep, surface and non-learning outcomes. *Studies in Higher Education* 32(1), 39–57.

Hay, D., Kinchin, I., & Lygo-Baker, S. (2008). Making learning visible: the role of concept mapping in higher education. *Studies in Higher Education*, *33*(3), 295-311.

Israel, C., Zipp, G. P., D'Abundo, M., & Deluca, D. (2020). Mind mapping to enhance critical thinking skills in physician assistant education: a randomized controlled study. *Journal of Allied Health*, *49*(2), 135-140.

Jeong, A. C. (2020). Developing computer-aided diagramming tools to mine, model and support students' reasoning processes. *Educational Technology Research and Development*, 1-17. (in press)

Jones, B. D., Ruff, C., Snyder, J. D., Petrich, B., & Koonce, C. (2012). The effects of mind mapping activities on students' motivation. *International Journal for the Scholarship of Teaching and Learning*, *6*(1), 1-23.

Koznov, D., & Pliskin, M. (2008, October). Computer-supported collaborative learning with mind-maps. In *International Symposium on Leveraging Applications of Formal Methods, Verification and Validation* (pp. 478-489). Springer, Berlin, Heidelberg.

Kunsch, D. W., Schnarr, K., & van Tyle, R. (2014). The use of argument mapping to enhance critical thinking skills in business education. *Journal of Education for Business*, *89*(8), 403-410.

Lin, L., Shadiev, R., Hwang, W. Y., & Shen, S. (2020). From knowledge and skills to digital works: an application of design thinking in the information technology course. *Thinking Skills and Creativity*, 100646.

Luo, J. (2019). Teaching mode of thinking development learning based on mind mapping in the course of health fitness education. *International Journal of Emerging Technologies in Learning (iJET)*, *14*(08), 192-205.

Mento, A. J., Martinelli, P., & Jones, R. M. (1999). Mind mapping in executive education: applications and outcomes. *Journal of Management Development*, *18*(4), 390-416.

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*(3), 413-448.

Öksöz, Y. & Coşkun, K. (2012). Impact of the school experience courses on special educational needs teachers' sense of self-efficacy. *Journal of Kirsehir Education Faculty*, *13*(2), 131-155.

Rider, Y., & Thomason, N. (2014). Cognitive and pedagogical benefits of argument mapping: LAMP guides the way to better thinking. In *Knowledge cartography* (pp. 113-134).Springer, London.

Shroff, R. H., Deneen, C. C., & Ng, E. M. W. (2011). Analysis of the technology acceptance model in examining students' behavioural intention to use an e-portfolio system. *Australasian Journal of Educational Technology*, *27*(4), 600-618.

Siah, C. J., Lim, F. P., Lim, A. E., Lau, S. T., & Tam, W. (2019). Efficacy of team-based learning in knowledge integration and attitudes among year-one nursing students: A pre-and post-test study. *Collegian Journal of the Royal College of Nursing Australia*, *26*(5), 556-561.

Teo, T. (2009). Modelling technology acceptance in education: A study of pre-service teachers. *Computers & Education*, *52*(2), 302-312.

Tergan, S.-O. (2006). Individual knowledge and information management with concept maps in web-based learning. In H. Mandl& H. F. Friedrich (Eds.), *Handbuch Lernstrategien [Handbook learning strategies]* (pp. 307–324). Göttingen, DE: Hogrefe.

van Gelder, T. J. (2003). Enhancing deliberation through computer supported argument mapping. In P. Kirschner, S. Buckingham Shum, & C. Carr (Eds.), Visualizing argumentation: Software tools for collaborative and educational sense-making (pp. 97–115). London: Springer.

van Gelder, T. J. (2015). Using argument mapping to improve critical thinking skills. In M. Davies & R. Barnett (Eds.), *The Palgrave Handbook of Critical Thinking in Higher Education* (pp.183-192). New York: Palgrave MacMillan.

van Gelder, T., Bissett, M., & Cumming, G. (2004). Enhancing expertise in informal reasoning. *Canadian Journal of Experimental Psychology*, 58, 142–152.

van Gelder, T.J., & Rizzo, A. (2001). Reason! Able across curriculum, in Is it an Odyssey in Learning? Proceedings of the 2001 Conference of ICT in Education, Victoria, Australia.

Wixom, B. H., & Todd, P. A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information systems research*, *16*(1), 85-102.

Wu, H. Z., & Wu, Q. T. (2020). Impact of mind mapping on the critical thinking ability of clinical nursing students and teaching application. *Journal of International Medical Research*, *48*(3), 1-8

Yamak, H., Bulut, N., & Dündar, S. (2014). The impact of stem activities on 5th grade students' scientific process skills and their attitudes towards science *Gazi University Journal of Gazi Educational Faculty*, *34*(2), 249-265.

Zumbach, J. (2009). The role of graphical and text based argumentation tools in hypermedia learning. *Computers in Human Behaviour*, *25*(4), 811-817.

Teacher Candidates' Experiences with the Flipped Classroom Model: A Phenomenological Approach

Serkan Aslan[1]
[1]Suleyman Demirel University

To cite this article:

Teacher Candidates' Experiences with the Flipped Classroom Model: A Phenomenological Approach

Serkan Aslan[1*]

[1]Suleyman Demirel University

Abstract

Today, technology integration has become an important issue in the teaching-learning environment. Developed countries integrate technology into schools from pre-school education to higher education, adopting a student-centred approach in their curriculums. Turkey has embraced a constructivist approach to its curriculum since the academic year 2005, in which studies of the teaching-learning environment ushered in the integration of new technology. Despite this, few studies have been conducted on technology integration in higher education. One approach to this research explores the integration of technology in Turkey into the flipped classroom model by the higher education institution, a subject upon which there has not been a sufficient amount of qualitative research. For this reason, in this study, the opinions of elementary teacher candidates regarding their experiences with the flipped classroom model were examined. The research was designed according to phenomenological research design, one of the foremost qualitative research patterns. The participants of the research are teacher candidates studying at Süleyman Demirel University as classroom teachers. The research sample was determined according to the criterion sampling method. An opinion questionnaire about the flipped classroom was used as a data collection tool in the research. Content analysis, a qualitative analysis method, was used to analyse the data. From the research, it was concluded that teacher candidates have experience with the flipped classroom model, believe this model provides many contributions to them and have encountered various problems with the model, for which they have suggestions on how these can be fixed.

Keywords: Technology, Technology Integration, Flipped Classroom Model, Teacher Candidate, Phenomenology.

Introduction

With the advancements in technology of the twenty-first century, the characteristics expected from students have shifted. Due to the developments and changes in digital technology of our age, the concept of the digital learner has emerged. Digital learners have grown up with technology and are defined as learners that organise their daily lives with technology (Göksun-Orhan, 2019). The basic features of digital learners include that they communicate with each other via e-mail or instant messaging, use online resources and correspond through virtual channels instead of face-to-face (Andone, Dron, Boyne & Pemberton, 2006). The fact that digital learners use so much digital technology in their daily lives has also affected their education. In order to train digital learners in a qualified way, new educational methods and techniques that take digital technology into account have emerged in the teaching-learning environment (Levin & Wadmany, 2005). The flipped classroom model is one of the newest models to emerge from recent literature and has been employed effectively.

Bishop and Verleger (2013) underscore the fact that the flipped classroom model is a teaching-learning model that provides an opportunity to perform activities and find solutions to the problems of the students in the classroom activities with the support of the teacher individually or in group activities. Demiralay and Karataş (2014) defined the flipped classroom model as a 'blended learning model' realised under the guidance of the teacher, in which the information presented by the teacher in the traditional classroom environment is transferred to the online platform, and the homework expected to be done outside of school is also integrated into the classroom. In the flipped classroom model, the teacher provides students with comprehension-level information at home, in the form of videos, articles, homework and so on. The class understands that they are meant to learn by using these tools and applying the knowledge and skills they acquire at school. In this way,

---

* Corresponding Author: *Serkan Aslan, serkanaslan@sdu.edu.tr*

the model enhances classroom interaction while simultaneously enabling effective usage of classroom time by employing more student-centred active learning strategies (LaFee, 2013; Milman, 2012).

The flipped classroom model has many benefits: it allows all students to act according to their own learning speeds and watch video lessons at any time as often as desired; it helps teachers to decide on the difficulty or ease of the content without experience in the school process; it asks students to do their homework in the classroom, allowing teachers to identify situations in which students face problems, understand their learning styles and obtain a more comprehensive perspective on the classroom; it lets students access class content 24/7 with their smartphones, so those who miss lessons for various reasons have the opportunity to watch again from anywhere, making students active thinkers both in and outside of the classroom; in this model, active approaches, methods and techniques for learning are used, improving students' motivation, academic success, high-level thinking and digital skills (Fulton, 2012; Herreid & Schiller, 2013; Morgan, 2014; Talbert, 2012). These advantages are not unique to students and teachers. Parents and members of school management are also positively affected by this process (Erdoğan, 2018).

As mentioned above, there are many advantages to the use of the flipped classroom model. In this regard, the flipped classroom model has become an important issue in higher education, and teacher training institutions have begun to organise the teaching-learning environment according to new approaches, models, methods and techniques. In the 21st century, approaches that centre student activity in the teaching-learning environment have been used. Developed countries have made changes to their curricula and initiated the effective use of technology in the teaching-learning process. As technology integration becomes increasingly imperative in education, it is becoming increasingly imperative for teacher training institutions to integrate technology. In fact, when studies in the national and international literature at the higher education level are examined, positive findings are observed in terms of the flipped classroom model's provision of time for student-centred education, student achievements, motivations, perceptions, positive participation, changing learning habits, encouragement of self-study and development of communication skills (Betihavas, Bridgman, Kornhaber & Cross, 2016; García-Sánchez & Santos-Espino, 2017; Presti, 2016; Turan, 2015; Yıldız, Kıyıcı & Altıntaş, 2016; Zuber, 2016). Therefore, it is useful to explain lessons by considering the flipped classroom model. Trainee teachers trained in this way, when graduated and assigned to work, can easily integrate technology into their lessons and train students according to the requirements of the digital age.

The effective use of the flipped classroom model in higher education contributes to the educations of students studying in these institutions. An experimental investigation into higher education research in Turkey about the flipped classroom model (Duman, 2019; Erdogan, 2018; Tanners, 2018) reveals that little qualitative research into the subject has been done. This situation was evaluated by the researcher as a deficiency. A flipped classroom model was used in an instructional technology course by the researcher. The teaching-learning process was organised by taking this model into consideration. The opinions of elementary teacher candidates regarding this model – what the model contributes to them, what problems are encountered when applying the model and what solutions they suggest to address these problems – are considered an important research subject, and it was therefore decided that conducting such a study is necessary. This research will provide feedback on the model from both teacher candidates and faculty members working in teacher training institutions. In addition, since there are no studies related to this subject in the existing literature, this study intends to fill the gap in the literature.

The aim of this research is to examine the opinions of elementary teacher candidates about their experiences with the flipped classroom model. In this context, answers to the following questions will be sought:
1. What are the opinions of preservice teachers on the flipped classroom model?
2. What are the opinions of teachers on the flipped classroom model in terms of their contribution?
3. What are the problems that preservice teachers experience in a flipped classroom model?
4. What are teachers' suggestions for solutions to the problems they face regarding the flipped classroom model?

Method
Research Design
Phenomenology, a form of qualitative research design, was used in the study. Qualitative research is defined as "a scientific study based on the discovery, understanding and interpretation of various events and phenomena within their environments as a result of a detailed examination" (Tuna, 2015: 360). According to Creswell (2007), qualitative research is the process of making sense of social questions and problems related to people using one's own methods. According to Creswell (2007), phenomenology is defined as the study of the experiences of a select few with a particular phenomenon or concept. According to Johnson and Christensen

(2012), the aim of phenomenological research is to obtain insight into the living worlds of the participants in the research and to reveal the personal meanings through which they have structured their experiences. In this study, the phenomenological research pattern was used in order to examine the opinions of the elementary teacher candidates on the flipped classroom model by delving into their experiences.

Participants

The participants of the research are 25 candidates studying at in a public university to become elementary school teachers. A criterion sampling method was used to determine the participants of the research. Criterion sampling involves the determination of a set of specific criteria, such as events, features, objects and dates, followed by the selection of the sample within the framework of these criteria (Özbaşı, 2019: 121). The criterion used to determine the scope of this research is that teacher candidates must have had experience with the flipped classroom model and taken lessons about the flipped classroom. The main reason for the determination of this criterion is that it is important to obtain healthy information about the flipped classroom from participants familiar with it for the purposes of this study. The researcher has taught an instructional technologies course for second-year students studying in the department of elementary education according to the flipped classroom model. In this respect, according to phenomenological research design, which is the method used in this research, these students are included in the research. The opinions of the participants are explored in depth through examination of their experiences with the flipped classroom model. 16 of the participants are women and 9 are men.

Data Collection Tool

Within the scope of the research, a questionnaire form was prepared to capture opinions of the elementary teacher candidates about the flipped classroom model. Surveys are available in formats consisting of open-ended or closed-ended questions (Tymms, 2017). In this study, the questions in the questionnaire form consisted of four open-ended questions. While developing open-ended questions, the literature was examined and opinions were received from two faculty members who conducted research on the flipped classroom model. After preparing the questions in the questionnaire form, the pilot was implemented. The form was administered to five preservice teacher candidates and its consistency was checked. Consequently, the form was reconfigured according to feedback from both teacher candidates and faculty members. The final form was administered to teacher candidates as part of the original practice.

Data Collection

The data of the research were collected between 02-06 / 03/2020. While collecting the data, the purpose of the research was explained to the participants. The questionnaire was filled in by teacher candidates in approximately 30 minutes and collected at intervals that would not hinder the teacher candidates' lessons. Before applying the questionnaire form, the ethical committee decision was made and the necessary permission was obtained. The participants gave their consent to participate in the research.

Data Analysis

Content analysis, another qualitative data analysis method, was used to analyse the results of the research. Content analysis is used to "identify the existence of words, concepts, themes, idioms, characters or phrases in one or more texts and to enumerate them" (Kızıltepe, 2015: 253). Berg (2001) defined content analysis as "careful, detailed and systematic examination and interpretation of a particular material in order to identify patterns, themes, prejudices and meanings". Content analysis can be conducted in two ways: deductively and inductively. Inductive content analysis was used in this research. In inductive content analysis, the researcher reads the data repeatedly, determining the dimensions that are important for the purpose of the research, decoding the data into categories and grouping similar categories under broad titles to summarise them (Kızıltepe, 2015). In this study, the questionnaire forms were read one-by-one, and the answers to each question were carefully read and coded. Later, these codes were brought together to create themes (Bogdan & Biklen, 2003; Saladana, 2019). Within the scope of the research, the data were analysed using NVivo 11 and the results obtained are presented in figures.

Validity and Reliability in Research

Various strategies were used in the research to ensure validity and reliability. These include the following.
1. Within the scope of the research, all processes – including data collection, analysis and interpretation by the researcher – were questioned with a critical eye and verified to ensure whether the results obtained reflect reality (Yıldırım & Şimşek, 2016).
2. As a result of the research, the participants were confirmed (Güçlü, 2019).
3. Research findings were presented to field experts and their opinions were received (Sığrı, 2018).

4. The compatibility of the themes and codes was examined by two different experts. In cases where agreement could not be reached, the experts talked to each other and reached a compromise (Güler, Halıcıoğlu & Taşğın, 2014).

5. The opinions of the participants were directly quoted in order to increase the reliability of the study (Güler, Halıcıoğlu & Taşğın, 2014).

Ethical Measures Taken in the Research

Ethical issues were respected during the research process. Attention was paid to ensuring that the participation of participants in the research was voluntary. Before administering the survey form, information was provided and permissions were obtained. Participants were informed that they would not be harmed in any way and that they should not write their names on the questionnaire. Within the scope of the research, the anonymity principle was heeded in direct quotations of teacher candidates, and codes such as TC: 1, W (Teacher Candidates: 1, Women) were used instead. Participants were also informed that the research data would not be used outside the scope of the research and that the data would remain confidential (Johnson & Christensen, 2012; Hammersley & Traianou, 2012; Karagöz, 2017). Within the scope of the research, the consent of the participants was obtained by distributing a participant consent form. In addition, the decision of the ethics committee (SDÜ: 874329561 / 050.991) was taken within the scope of the study.

Findings

Findings are presented in terms of the sub-problems of the research. According to the first sub-problem of the research, the opinions of teacher candidates on the flipped classroom model were examined, the findings of which are presented in Figure 1.
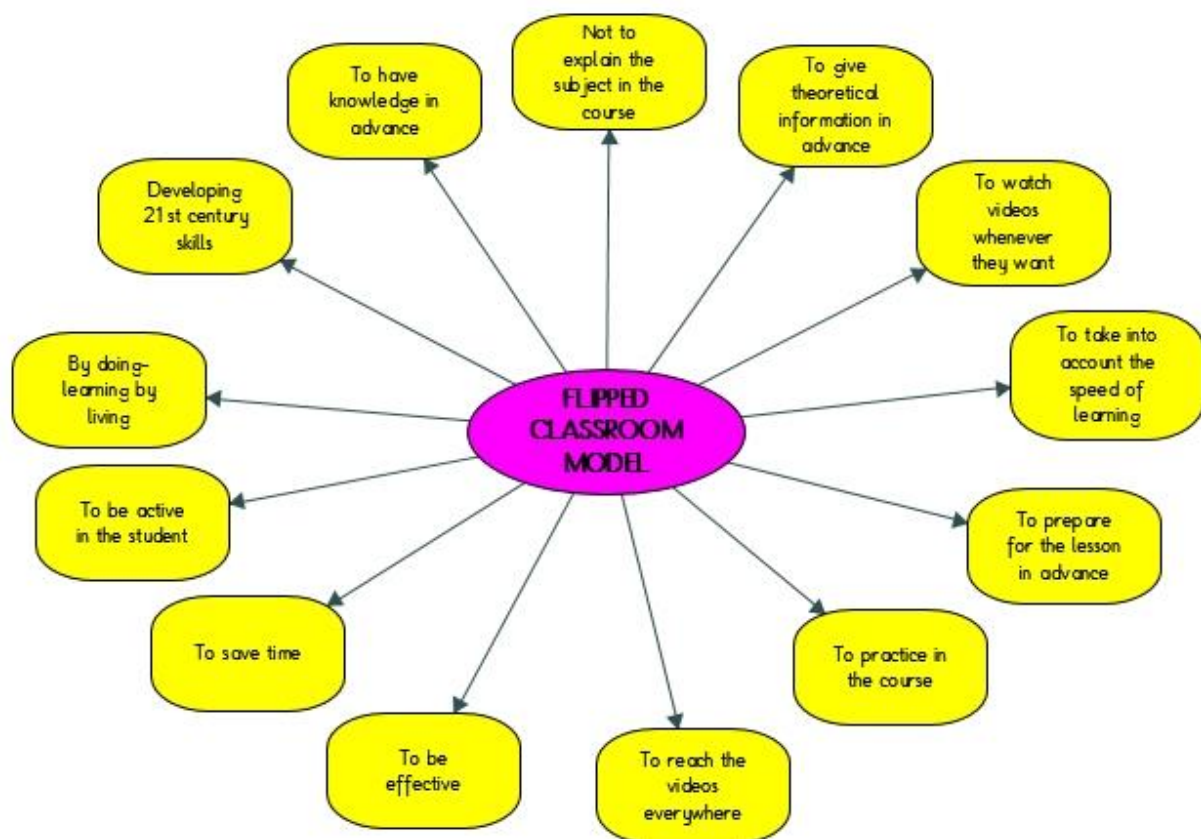
Figure 1. Teacher candidates' opinions about the flipped classroom model

Figure 1 demonstrates that teacher candidates have sufficient information about the flipped classroom model. Teacher candidates, regarding the flipped classroom, had the following to say: it provided them with prior knowledge about the subject to be taught in the lesson, they did not explain the subject of the lesson, they were given theoretical knowledge in advance, they were able to watch the videos whenever they wanted, the model took into account their learning speeds, they had the opportunity to prepare for the lesson in advance, and they had the opportunity to reach the videos anywhere. They stated that the lessons were effective, saved time, resulted in active students, allowed for learning by doing and living and helped them to develop 21st-century skills. The views of some teacher candidates are given below.

"*In the flipped classroom model, before coming to the lesson, we get information about the subject to be covered in the lesson thanks to the videos and articles related to the lesson. In this lesson, the time allocated for the lesson is lessened and allows us to practice to reinforce the subject.*" (TC: 1, M)

"*While doing the definition and memorization part at home on our own, we practice these topics in the classroom. This is a very important situation for us. It provides efficient time in the school. In this model, giving theoretical knowledge in advance makes it easier for us to come prepared for the lesson.*" (TC: 5, M)

Based on the second sub-problem of the research, the opinions of preservice teachers about the contributions of the flipped classroom model to their studies were examined, the findings of which are presented in Figure 2.
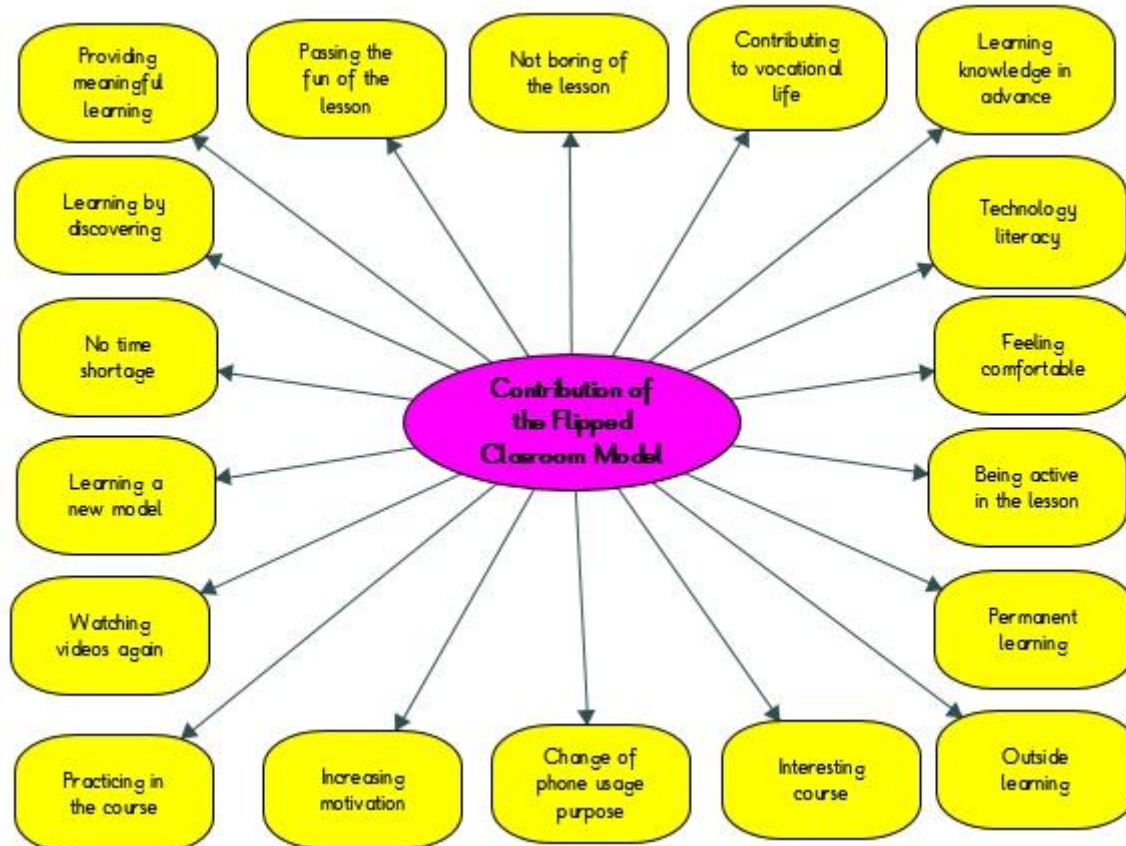
Figure 2. Teacher candidates' opinions about the flipped classroom's contribution

According to Figure 2, teacher candidates stated that the flipped classroom model provides many contributions to them. The teacher candidates stated that the flipped classroom model: interests them to learn by discovering, provides meaningful learning experiences, facilitates the fun of learning, is not boring, contributes to their professional lives, allows them to learn information in advance, helps them develop technology literacy, makes them feel comfortable, encourages active learning, ensures permanent learning and allows for learning outside of the classroom environment. They stated that it was attractive, changed the purpose of using their phones and increased their motivation. Additionally, the applications in the lesson contributed to watching videos, learning a new model and prevented time problems. The views of some teacher candidates are given below.

"*In the flipped classroom model, we learned theoretical knowledge at home and practiced it at school. Thanks to this, I learned the information better and it became permanent. In addition, it was a great advantage for us to give videos, assignments and articles about the subject beforehand. Meaningful learning was also taking place.*" (TC: 6, W)

"*The contribution of this class practice to me was that I did not get bored with 45 minutes of theoretical knowledge in the lesson. It made the lesson more enjoyable and fast. I was constantly actively attending the lesson and my learning was made permanent by reading the homework, videos and articles. It was very advantageous for me to have prior knowledge of the subject of the day. I could answer the questions asked by the teacher.*" (TC: 7, M)

Based on the third sub-problem of the research, the opinions of teacher candidates about the problems they encountered in the flipped classroom model were examined. These findings are presented in Figure 3.
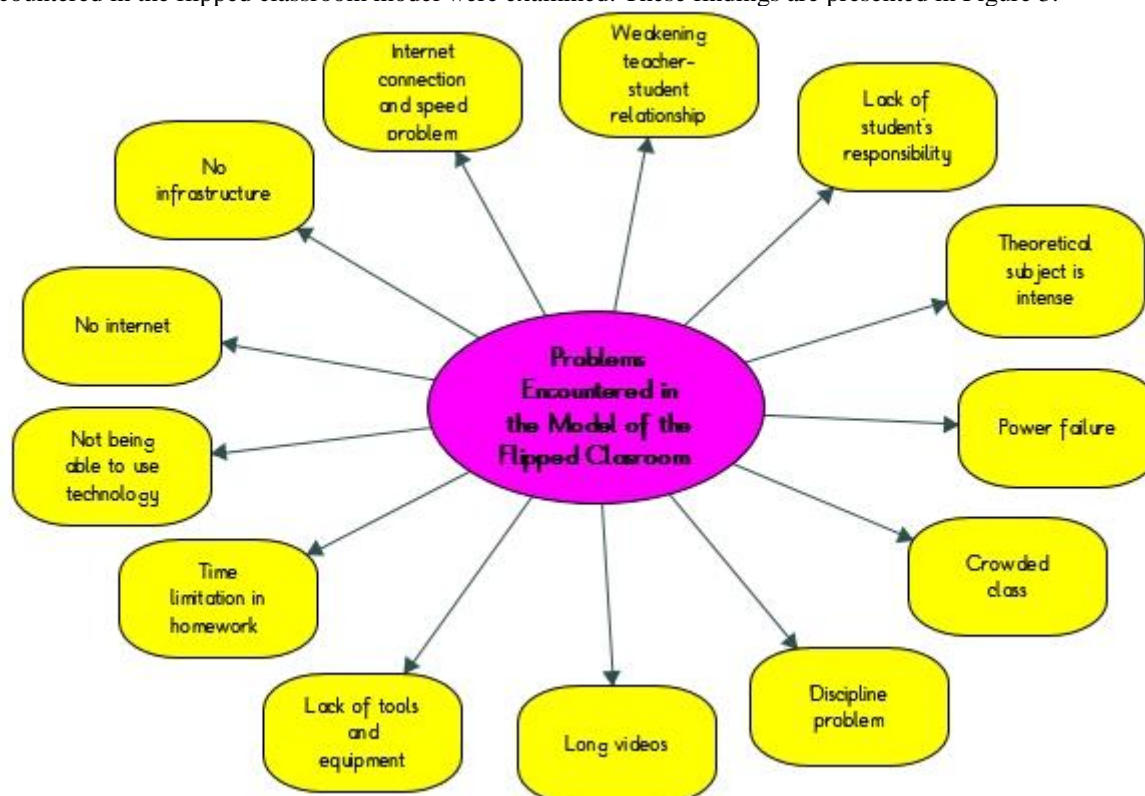
Figure 3. Opinions of teacher candidates about the problems they face in the flipped classroom model

Figure 3 reveals problems encountered by teacher candidates during the implementation of the flipped classroom model: lack of internet, lack of infrastructure, internet connection and speed problems, weakening teacher-student relationships, failure to fulfil students' responsibility, intense theoretical subjects, power failures, crowded classes, disciplinary problems, long videos, lack of tools and equipment and time allowed for assignments. They stated that they had problems such as restriction and not being able to use the technology. The views of some teacher candidates are given below.

"*I think the problem is that there is no connection between the teacher and the student in an emotional sense. The teacher gives the theoretical information, the student watches or reads and comes. There may be questions he wants to ask. I think it weakens the teacher-student relationship.*" (TC: 18, W)

"*I had a problem using technology in the flipped class model. The virtual class implemented in Edmodo was difficult for me. I could not use this program very well. I also had problems on the internet and sometimes I could not connect.*" (TC: 19, M)

Based on the fourth sub-problem of the research, the suggestions of teacher candidates regarding the effective and efficient use of the flipped classroom model were also examined, the findings of which are presented in Figure 4.
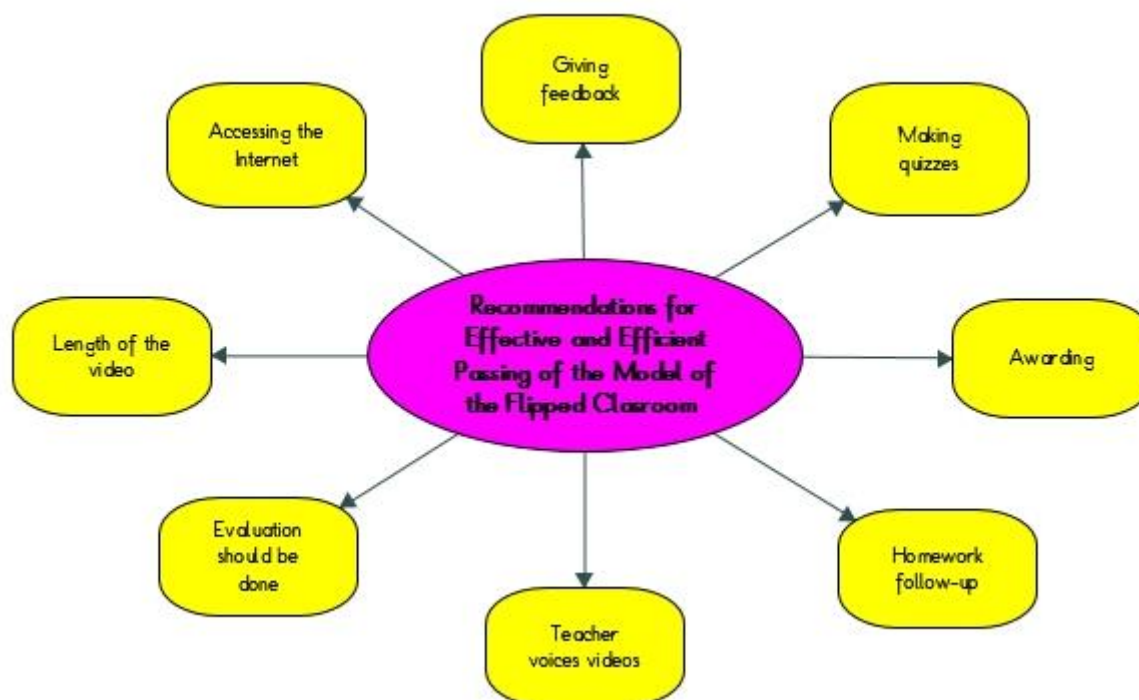
Figure 4. Suggestions of preservice teachers about the effective and efficient use of the flipped classroom model

In Figure 4, participants suggest adjusting the length of the video for effective and efficient implementation of the flipped learning model, providing equal access to the Internet for all students, supplying students with feedback, giving students quizzes, rewarding successful students and groups, following up with homework, videos and suggestions via teacher vocalisation and developing evaluation methods. The views of some teacher candidates are given below.

"*Videos sent to students in the flipped classroom model should definitely not be kept too long. I did not want to watch it when it is long. Short and concise videos should be used instead. In addition, evaluation should be made about the subjects. In this way, it can be determined who has learned the subject.*" (TC: 20, W)

"*The teacher should monitor whether the homework or videos he / she sent are watched. Otherwise, the students come without watching the videos, and this prevents the lesson from being effective. Also, the students must be given feedback by the teacher. Sometimes there is a place he cannot understand.*" (TC: 21, W)

Discussion, Result and Recommendations

In this study, the experiences of the elementary teacher candidates with the flipped classroom model were investigated in depth. As a result of the research, it was observed that teacher candidates are sufficiently informed about the flipped classroom and that their attitudes towards this model are positive. One of the biggest obstacles to technology integration in the teaching-learning process is the negative attitudes and prejudices of students towards technology (Gününç, 2017). In this regard, teacher candidates' knowledge of the flipped classroom model and positive attitudes will ensure that the problems encountered in technology integration will be minimised. This result suggests that teacher candidates will be able to implement the flipped classroom model in their classes when they start working and that they will integrate technology effectively and efficiently. It is not an exaggeration to say that individuals with 21st-century skills will be raised in this way. The fact that teacher candidates say that the flipped classroom helps to develop critical and creative thinking skills and sharpen 21st-century skills supports this interpretation. Studies conducted by Clark (2015), Evseeva and Solozhenko (2015), Muir and Geiger (2016), D'addato and Miller (2016), Özdemir (2019), Szparagowski (2014) and Turan and Göktaş (2015) have concluded that the flipped classroom model improves students' attitudes.

In the study, the opinions of the preservice teachers about the contributions of the flipped classroom model to their own studies were also examined. Research revealed that the flipped learning model helps them to make discoveries, provides meaningful learning, makes lessons fun, contributes to their professional life, allows them to acquire knowledge in advance, improves technological literacy, makes them feel comfortable, facilitates activity in the lesson and makes learning permanent. They noted that it creates out-of-school learning

environments and more interesting lessons, the purpose of using their phones has changed, their motivation has increased, the application of the lesson contributed to watching the videos again, they were able to learn new models and they did not struggle with time problems. These results show that the application of the flipped classroom model in higher education provides many contributions to many students. With this in mind, it will be very useful to apply the flipped classroom model in higher education institutions. In the research carried out by Özdemir (2019), the fact that students felt the flipped classroom model contributed to their studies the opportunity to watch videos, prepare for the lesson, pause the videos, provide permanent learning and supply audiovisual information supports the results of this research. In the results of studies conducted by Chilingaryan and Zvereva (2017), Torun and Darkut (2015) and Lo, Hew and Chen, (2017), the results that overlap with the present research results obtained.

Within the scope of the research, the problems teacher candidates face as a result of applying the flipped classroom model were also examined. During the implementation of the flipped classroom model, teachers encountered a lack of internet, lack of infrastructure, internet connection and speed problems, weakened teacher-student relationships, student failure to fulfil their responsibilities, intense theoretical subjects, power failures, crowded classes, disciplinary problems, long videos, lack of tools and equipment and time in assignments. They stated that they had problems such as restrictions and inability to use the technology required. The national and international literature states that in the application of the flipped classroom model, it is difficult to control whether students watch the videos or not (Arnold-Garza, 2014; Ayçiçek, 2019; Tucker, 2012). In this regard, the problems faced by teacher candidates in the application of the flipped classroom model overlap with those observed in the literature. In studies by Bergmann and Sams (2012), Kara (2016a), Kara (2016b) and Özdemir (2019), similar problems were encountered and identified.

Finally, the suggestions of teacher candidates for administering the flipped classroom model were examined. Findings indicate that for the effective and efficient implementation of the flipped classroom model, video lengths should be properly adjusted, access to the internet should be provided, students should be given feedback, quizzes should be supplied intermittently, successful students and groups should be awarded, teachers should follow up on students' learning with homework, and vocalisation and evaluation should be developed. In the study carried out by Debbağ (2018), teacher candidates developed suggestions such as preventing classroom overcrowding and providing tools and equipment for the flipped classroom to be effective. These suggestions coincide with the suggestions developed as a result of this research.

Based on the research results, the following suggestions have been developed:
1. The flipped classroom model should be used in courses in teacher training institutions, and faculty members should be supported in its implementation. In this way, teacher candidates with high digital self-efficacy and advanced technological literacy can be trained.
2. Measures should be taken by the relevant institutions to eliminate problems such as lack of access to the internet and lack of tools and equipment required by the teachers. In this way, the flipped classroom model can be applied more effectively and efficiently in higher education institutions.
3. The research was carried out on students studying in the department of elementary school teaching. Studies should be conducted to investigate the opinions and experiences of teacher candidates studying in other teacher training programs regarding flipped classroom practices.
4. Conducting research using case models, action research, experimental work, mixed research methods such as case study and the implementation of the flipped classroom model in teacher training programs will contribute further to the literature.

References

Andone, D., Dron, J., Boyne, C. & Pemberton, L. (2006). Are our students digital students? *Association for Leraning Technology*, 82-96.

Arnold-Garza, S. (2014). The flipped classroom teaching model and its use for information literacy instruction. *Communications in Information Literacy, 8*(1), 7-22.

Ayçiçek, B. (2019). Ters yüz sınıf modeli [flipped classroom model]. T. Yanpar-Yelken (Edt.) *Öğretim teknolojileri [Instructional technologies]*. Ankara: Anı Yayıncılık.

Berg, B.L. (2001). *Qualitative research methods for the social sciences.* Boston: Ally and Bacon.

Bergmann, J. & Sams, A. (2012). *Reach every student in every class every day.* United States Of America: Flip Your Classroom.

Betihavas, V., Bridgman, H., Kornhaber, R. & Cross, M. (2016). The evidence for 'flipping out': a systematic review of the flipped classroom in nursing education. *Nurse Education Today, 38,* 15–21.

Bishop, J.L. & Verleger, M.A. (2013,). *The flipped classroom: A survey of the research*. In ASEE National Conference Proceedings, Atlanta, GA.

Bogdan R.C. & Biklen, S.K. (2003). *Qualitative research for education: An introduction to theory and methods*. Boston: Allyn and Bacon.

Chilingaryan, K. & Zvereva, E. (2017). Methodology of flipped classroom as a learning technology in foreign language teaching. *Procedia - Social and Behavioral Sciences, 237*(21), 1500 – 1504.

Clark, K. R. (2015). The effects of the flipped model of instruction on student engagement and performance in the secondary mathematics classroom. *Journal of Educators Online, 12*(1), 91-115.

Cresswell, J.W. (2007). *Qualitative inquiry & research design choosing amang five appproaches*. London*:* SAGE.

D'addato, T., & Miller, L. R. (2016). An inquiry into flipped learnıng in fourth grade math instruction. *Canadian Journal of Action Research, 17*(2), 33-55.

Debbağ, M. (2018). *Öğretim ilke ve yöntemleri dersi öğretim programı için hazırlanan ters-yüz edilmiş sınıf modelinin etkililiği [The effectiveness of the flipped classroom model designed for teaching principles and methods course curriculum]*. Yayınlanmamış doktora tezi. Bolu Abant İzzet Baysal Üniversitesi, Bolu.

Demiralay, R. & Karataş, S. (2014). Evde ders okulda ödev modeli [Homework model at home tutoring school]. *Eğitim ve Öğretim Araştırmaları Dergisi, 3*(3), 333-340.

Duman, İ. (2019). *Etkinlik temelli öğrenmeye dayalı ters-yüz edilmiş sınıf modelinin öğrencilerin akademik başarı ve öğrenme motivasyonları üzerindeki etkisi [The effect of flipped classroom model incorporating activity-based learning on students' academic achievement and learning motivation]*. Yayınlanmamış doktora tezi. Sakarya Üniversitesi, Sakarya.

Erdoğan, E. (2018). *Sosyal bilgiler öğretiminde ters yüz edilmiş sınıf modelinin kullanımı [The use of flipped classroom model in social studies teaching]*. Yayınlanmamış doktora tezi. Gazi Üniversitesi, Ankara.

Evseeva, A. & Solozhenko, A. (2015). Use of flipped classroom technology in language learning. *Procedia - Social and Behavioral Sciences*, *206,* 205-209.

Fulton, K. (2012). Upside down and inside out: Flip your classroom to improve student learning. *Learning & Leading with Technology, 39*(8), 12–17.

García-Sánchez, S. & Santos-Espino, J. M. (2017). Empowering pre-service teachers to produce ubiquitous flipped classes. *Profile Issues in Teachers' Professional Development, 19*(1), 169-185.

Göksun-Orhan, D. (2019). 21. yy. becerileri edinmede dijitalleşen roller [21st century digitized roles in acquiring skills]. Y.K. Türel (Edt.) *Öğretim teknolojileri [Instructional technologies]*. Ankara: ASOS Yayınları.

Güçlü, İ. (2019). *Sosyal bilimlerde nitel araştırma yöntemleri. [Qualitative research methods in the social sciences]*. Ankara: Nobel Yayıncılık.

Güler, A., Halıcıoğlu, M.B. & Taşğın, S. (2014). *Sosyal bilimlerde nitel araştırma yöntemleri [Qualitative research methods in the social sciences]*. Ankara: Seçkin Yayıncılık.

Gününç, S. (2017). *Eğitimde teknoloji entegrasyonun kuramsal temelleri. [Theoretical foundations of technology integration in education]*. Ankara: Anı Yayıncılık.

Hammersley, M. & Traianou, A. (2012). *Ethics in qualitative research: controversies and contexts*. London: SAGE Pub.

Herreid, C.F. & Schiller, N.A. (2013). Case studies and the flipped classroom. *Journal of College Science Teaching*, *42*(5), 62-66.

Johnson, B. & Christensen, L. (2012). *Educational research quantitative, qualitative and mixed approaches*. London: SAGE.

Kara, C.O. (2016a). Ters yüz sınıf. [Flipped classroom].*Tıp Eğitimi Dünyası,* (45), 12-26.

Kara, C.O. (2016b). *Tıp fakültesi klinik eğitiminde "Ters yüz sınıf modeli" kullanılabilir mi? [Can flipped classroom model be used in clinical education program of medical faculty ?]* Yayımlanmamış yüksek lisans tezi. Akdeniz Üniversitesi, Antalya.

Karagöz, Y. (2017). *SPSS ve AMOS uygulamalı nitel-nicel-karma: Bilimsel araştırma yöntemleri ve yayın etiği.[SPSS and AMOS applied qualitative-quantitative-mixed: Scientific research methods and publication ethics]*. Ankara: Nobel Yayıncılık.

Kızıltepe, Z. (2015). İçerik analizi. [Content analysis]. F.N. Seggie & Y. Bayyurt (Edt). *Nitel araştırma yöntem, teknik, analiz ve yaklaşımlar. [Qualitative research methods, techniques, analysis and approaches]*. Ankara: Anı Yayıncılık.

LaFee, S. (2013). Flipped learning. *Education Digest, 79*(3), 13-18.

Levin, T. & Wadmany, R. (2005). Changes in educational beliefs and classroom practices of teachers and students in rich technology-based classrooms. *Technology, Pedagogy and Education, 14*(3), 281-307.

Lo, C.K., Hew, K.F. & Chen, G. (2017). Toward a set of design principles for mathematics flipped classrooms: A synthesis of research in mathematics education. *Educational Research Review, 22*, 50-73.

Meyer, L. (2015, 03 Mart). Research: Active learning more important than flipping the classroom. *The Journal*, https://thejournal.com/articles/2015/03/03/research-usingactive-learning-more-important-than-flipping-the-classroom.aspx Retrived Date: 15.02.2020.

Muir, T. & Geiger, V. (2016). The affordances of using a flipped classroom approach in theteaching of mathematics: A case study of a grade 10 mathematics class. *Mathematics Education Research Journal, 28*(1), 149–171.

Morgan, H. (2014). Focus on technology: Flip your classroom to increase academic achievement. *Childhood Education, 90*(3), 239-241.

Milman, N. (2012). The flipped classroom strategy: What is it and how can it be used? *Distance Learning, 9*(3), 85-87.

Özbaşı, D. (2019). Örneklem belirleme. [Sampling]. K. Yılmaz & R.S. Arık (Edt.) *Eğitimde araştırma yöntemleri. [Research methods in education]*. Ankara: Pegem Akademi Yayıncılık.

Özdemir, M.Ç. (2019). *Ters yüz edilmiş sınıf uygulamalarının geometri öğretiminde kullanılmasının matematik öğretmeni adaylarının geometriye yönelik tutumlarına etkisinin incelenmesi. [The use of flipped classroom in geometry teaching investigation of the effect of the mathematics teacher candidates on the attitudes of geometry]*. Yayınlanmamış yüksek lisans tezi. Bayburt Üniversitesi, Bayburt.

Presti, C.R. (2016). The flipped learning approach in nursing education: a literatüre review. *Journal of Nursing Education, 55*(5), 252–257.

Saladana, J. (2019). *Nitel araştırmacılar için kodlama el kitabı.[Coding handbook for qualitative researchers]*. (Çev. Edt.) A. Tüfekci-Akcan & S.N. Şad Ankara: Pegem Akademi Yayıncılık.

Sığrı, Ü. (2018). *Nitel araştırma yöntemleri.* İstanbul: Beta Yayınları.

Szparagowski, R. (2014). The effectiveness of the flipped classroom. *Honors Projects*, 1-32.

Talbert, R. (2012). Flipped classroom. *Colleagues, 9*(1), 1-3.

Tıkaç, S. (2015). Nitel veri analiz programlarının veri analizinde kullanımı: Nvivo'ya bir bakış. [The use of qualitative data analysis programs in data analysis: a look at Nvivo]. F.N. Seggie & F. Bayyurt (Edt.) *Nitel araştırma: Yöntem, teknik, analiz ve yaklaşımlar. [Qualitative research: Method, technique, analysis and approaches]*. Ankara: Anı Yayıncılık.

Torun, F. & Dargut, T. (2015). Mobil öğrenme ortamlarında ters yüz sınıf modelinin. [A proposal for the applicability of flipped classroom model in mobile learning environments]. *Eğitim Bilimleri Dergisi, 6*(2), 20-29.

Tucker, B. (2012). The flipped classroom. *Education Next, 12(1), 82.*

Turan, Z. & Göktaş, Y. (2015). Yükseköğretimde yeni bir yaklaşım: Öğrencilerin ters yüz sınıf yöntemine ilişkin görüşleri. [A new approach in higher education : the students' views on flipped classroom method]. *Yükseköğretim ve Bilim Dergisi, 5*(2), 156-167.

Turan, Z. (2015). *Ters yüz sınıf yönteminin değerlendirilmesi ve akademik başarı, bilişsel yük ve motivasyona etkisinin incelenmesi.[The evaluation of flipped classroom method and examination of its effects on academic achievement, cognitive load and motivation]*. Yayımlanmamış doktora tezi. Atatürk Üniversitesi, Erzurum.

Tymms, P. (2017). Questionnaires. R. Coe, M. Waring, L.V. Hedges & J. Arthur (Edt.) *Research methods and methodologies in education*. London: SAGE Pub.

Yıldırım, A. & Şimşek, H. (2016). *Sosyal bilimlerde nitel araştırma yöntemleri.[Qualitative research methods in the social sciences]*. Ankara: Seçkin Yayıncılık.

Yıldız, D., Kıyıcı, G. & Altıntaş, G. (2016). Ters-yüz edilmiş sınıf modelinin öğretmen adaylarının erişileri ve görüşleri açısından incelenmesi. [A research into the flipped classroom in terms of the academic achievement, and views of the prospective teachers]. *Sakarya University Journal of Education, 6*(3), 195-210.

Zuber, W.J. (2016). The flipped classroom, a review of the literature. *Industrial and Commercial Training, 48*(2), 97-103.

Integrating L1 into Grammar Teaching as a Remedy for Learners' Unresponsiveness in an ESP Classroom: An Action Research

Handan Çelik[1]
[1]Trabzon University

Integrating L1 into Grammar Teaching as a Remedy for Learners' Unresponsiveness in an ESP Classroom: An Action Research

Handan Çelik[1*]
[1]Trabzon University

Abstract

In this study, the researcher had a group of unresponsive learners taking a year-long English for Specific Purposes (ESP) course in Banking and Finance Department. After the detection of the problem, *unresponsiveness*, one-on-one semi-structured interviews were conducted with the learners (*N*=8) to find the reasons of it, and what they needed accordingly. Data analysis showed they majorly needed Turkish- (L1), which was restricted to the teaching of grammar as they found most beneficial to their learning. Thus, use of L1 for explicit grammar teaching which proceeded schema building plus implicit inferring was examined to see if and how it remedied their unresponsiveness. Systematic data collection included the learners' end-of-course written reflections (*N*=40) and another round of one-on-one semi-structured interviews (*N*=7) to further elicit their views about teaching grammar through L1. The results revealed positive evaluations regarding the benefits of L1 upon more and better input comprehension, vocabulary and grammatical knowledge development, and awareness regarding language learning. They also reported feeling less anxious and stressed, which encouraged their willingness to communicate and participate. Thus, the results verify judicious role of L1 as a scaffold to decrease affective barriers and increase comprehension in language learning.

Key words: action research, English for Specific Purposes, teaching grammar, unresponsive learners, use of L1

Introduction

Teaching ESP, as a specific approach to language teaching relying on its own methodology and content, aims to equip learners with the knowledge and skills for the field of study they aim and need to be ready for (Dudley-Evans & St John, 1998). In ESP classes, the focus is on the identification of sets of transferable language skills, *often through needs analysis*, located within specific contexts (Paltridge & Starfield, 2013). Although ESP teaching relies on its own methodology as an answer to the needs, there is no best way or approach as all methods are a response to a particular group of learners. Therefore, ESP practitioners need to have the ability to assess a situation, then select and adapt their methodology to match learners' needs as "flexibility and willingness to take risks are the name of the game!" (Dudley-Evans & St John, 1998, p.187).

In this regard, despite the promotion of English-only classroom in second, foreign, and specific contexts due to the belief that extensive use of English motivates and improves learners' language skills, L1 which is the ultimate source of their background knowledge of the language and its linguistic features cannot be disregarded (Ellis, 1994). Besides, L1 is a mediator in foreign and second language learning for the teaching of speaking, writing, and grammar as learners are known to rely extensively on it as a cognitive, metacognitive, and pedagogical tool (Butzkamm, 2003; Gass & Selinker, 2008; Giacobbe, 1992; Lida, 2014) in diverse contexts, particularly where they are shy, silent, and reluctant. Moreover, "English only in the classroom is neither conclusive nor pedagogically sound" (Auerbach, 1993, p.5) as it may "result in lengthy, complicated, and incomprehensible explanations that add to teacher talking time" (Szendröi, 2010, p. 41). Thus, allowing L1 in language classroom has many benefits such as affecting later success of learning positively by facilitating transition to English, reducing affective barriers, and integrating learners' authentic and out-of-class experiences into learning (Auerbach, 1993). Furthermore, "not making use of both the L1 and L2 in the classroom is a waste of a valuable resource" (Turnbull, 2018, p.55). Hence, L1 has already been used in English language classes,

---

[*] Corresponding Author: Handan Çelik, handancelik@trabzon.edu.tr

and ESP classroom is not an exception, as it facilitates learners' understanding and provides sense of security, better comprehension and progress, higher achievement, and a positive learning atmosphere (Auerbach, 1993; Bruen & Kelly, 2017; Butzkamm, 2003; Usadiati, 2009). L1 also contributes to language skills development, metalinguistic awareness, negotiation of meaning, reduced cognitive overload, lower affective barriers such as anxiety, and higher self-confidence and self-motivation (Berning, 2016; Boustani, 2019; Bruen & Kelly, 2017; Cheng, 2015; Cuartas Alvarez, 2014; Luchini & Rosello, 2007; Usadiati, 2009). Therefore, although "the use of L1 has been gradually viewed less favorably by second language (L2) teaching theorists after the Grammar-Translation Method lost its prominent role" (Du, 2016, p.359), the interest regarding whether to include learners' L1 in teaching has gained recognition and become a persistent research topic particularly after being neglected for many years within the communicative approach beginning in 1970s.

Literature review: Understanding L1 in teaching English

Evidence regarding the difference L1 made in language classrooms exists in international body of literature. L1 is most frequent to the translation of unknown words and explanations of grammar rules which learners have difficulty in comprehending (Jingxia, 2010). In her study upon the role of using L1 (Indonesian) interchangeably with English in explanations of present perfect tense rules and its effect on writing, Usadiati (2009) reports higher success rate (80 %) compared to when it is not used (45 %). In classes where L1 was periodically used to clarify complex terminology, vocabulary and grammar, and to give instructions, learners' overall exam scores more than doubled and their level of anxiety lessened compared to those where L1 was not permitted (Boustani, 2019; Bruen & Kelly, 2014; Miles, 2004; Teng, 2019). Thus, if L1 is not allowed, it causes insecurities and slow improvement, and when allowed, it facilitates learning as it significantly encourages higher learning performance and sense of confidence in a relaxed classroom atmosphere. Additionally, if allowed to use their L1 (Japanese) when collaborating for writing tasks, learners achieved better written L2 output than those who relied on exclusive use of English during collaboration (Berning, 2016). Besides, as an integral component of learners' identity where prior learning and life experiences are encoded, L1 scaffolds L2 learning (Bismilla, 2011). Thus, it is regarded as "the most important ally a foreign language can have" (Buztkamm, 2003, p.30). Madrinan (2014) arrived at similar results as she saw that learners transferred concepts from their L1. As seen, in any level and context, L1 has various functions ranging from facilitating comprehension and better performance of language skills to mediating the transfer of previous learnings to L2 and foreign language.

When it comes to the Turkish context, use of L1 is a controversial issue as there is no standardization. It is still a matter of concern as teachers face "the dilemma of allowing, limiting, or forbidding it" (Yavuz, 2012, p.4340) since "there has been no absolute research outcome that indicates whether it should be avoided at all costs or not" (Timuçin & Baytar, 2015, p.241). Thus, there is "no concrete agreement among teachers and scholars who are involved" (Yürekli Kaynardağ, 2016, p.5) although research suggested including L1 properly and purposefully (Çelik, 2008). In this regard, teachers suggest using L1 only if it is necessary (Yavuz, 2012) or regard it as an inherent segment of language learning as they think that it fulfills such functions as establishing rapport with students, making clarifications, or giving explanations (Paker & Karaağaç, 2015). Similarly, teachers at primary and secondary levels rely on L1 extensively (48.12 %) causing students to receive inadequate L2 input (İnan, 2016). As for the likely reasons of this extensive L1 use, research shows that in secondary schools, it is mainly to transmit the academic content, manage the classroom and maintain its discipline, and establish rapport with students (Salı, 2014). Similarly, in tertiary level, L1 is used to translate, check comprehension, give instructions, explain grammar, manage the class, or for no obvious reason like random code-switching (Timuçin & Baytar, 2015). Besides, its use in preparatory classes in tertiary level is seen to make a difference in grammar teaching (Uyar, 2012). Similarly, significant evidence of syntactic transfer on the acquisition of verb placement is reported (Mede, Tutal, Ayaz, Çalışır, & Akın, 2014). Lastly, tertiary level EFL students favor the inclusion of L1 due to feeling comfortable, at ease, and less stressed (Debreli & Oyman, 2016). Besides, tertiary level teachers also regard L1 as a facilitating function as they think it can be used for lower-level students for various purposes including teaching grammar, explaining differences between L1 and L2, or solving disciplinary problems (İnal & Turhanlı, 2019).

Consequently, L1 in Turkish context continues to be a persistent issue for various reasons. Therefore, within this study locating itself in an ESP context, whereby use of L1 is clearly under researched, use of L1 was restricted to the teaching of grammar to deal with unresponsive learners. Hence, if and how it remedied their unresponsiveness and benefits, as reported by the learners, are examined.

Research context and the problem

The ESP course, as the context for the study, runs in the Banking and Finance Department of Applied Sciences School of a State University founded in 1982 in northwest Turkey. It is extensive, assessed, and compulsory, and targets the delayed needs (Dudley-Evans & St John, 1998) of the learners (*N*=40) taking the course in their senior year. As there were no programmatic documents regarding the course, *its goals and outcomes, content, methodology, teaching resources and materials, and testing and evaluation issues*, I, as the teacher, adopted an integrated-skills approach, including the four primary *reading, listening, writing, speaking*, and related skills *vocabulary, spelling, pronunciation, grammar, and meaning*. To ease their optimal integration (Oxford, 2001), I selected a coursebook (see Hobbs & Starr Keddle, 2007) to guide the likely content. To this end, each class included *start ups* to warm the learners up and introduce them with key vocabulary and concepts. They were followed by *reading and listening* (depending on the flow of the content) and teaching of *grammar* generated by the topic. Besides, the classes included functional language i.e. *giving opinions, agreeing or disagreeing* etc., and listening tasks asking students to analyze the interaction, and then practicing the language function in meaningful contexts through speaking tasks. Writing tasks, i.e. *writing personal profile, or an application letter* etc., either in-class, self-study, or homework, were also included.

Despite seeming in harmony, in practice it wasn't as the learners were hard to involve although I adjusted my speech through slow talk, simple wording, and repetitions (Gass & Selinker, 2008) and tried to simplify the discussions and used rephrases to encourage them to respond. For instance, in a unit dealing with "Changing World", as a warm up, I simply asked the question, "What is changing?" Although they were supposed to have a lot to say, such as cars, schools, education, technology, communication, media, etc., this was not what happened. I had to prompt them with some more questions like "Is shopping changing? Are you still store shopping or shopping online?, Is travelling changing? Is it becoming faster and more comfortable? etc." However, those responded were very few; only the ones relatively more competent in English. It was the moment that I thought there was a problem, *unresponsiveness*, and I needed to do something. When I made that decision, it had almost been two months since the term started, whereby I experienced many other similar moments. As it was a year-long course, I knew it could not continue this way since it would make no benefit to the learners. Therefore, I knew action research (AR) methodology would suit best for finding a solution to the problem and improving the situation.

Therefore, the following questions were addressed;

1. What do the ESP learners think about the reasons of their unresponsiveness, and what do they suggest solving it?
2. What are the benefits of integrating L1 into grammar teaching, *as reported by the ESP learners*?

Method

Methodology, intervention, and data collection

In AR methodology, there is a cycle working with the combination of some steps, i.e. detecting the problem, developing an action plan to address it, putting the plan into action, collecting data and reflecting to see if and how the action works, and if needed, another round of reflecting and action to improve the plan (Burns, 2009; Kayaoğlu, 2015; Lebak & Tinsley, 2010; Rainey, 2000). Figure below shows the AR cycle (Figure 1).
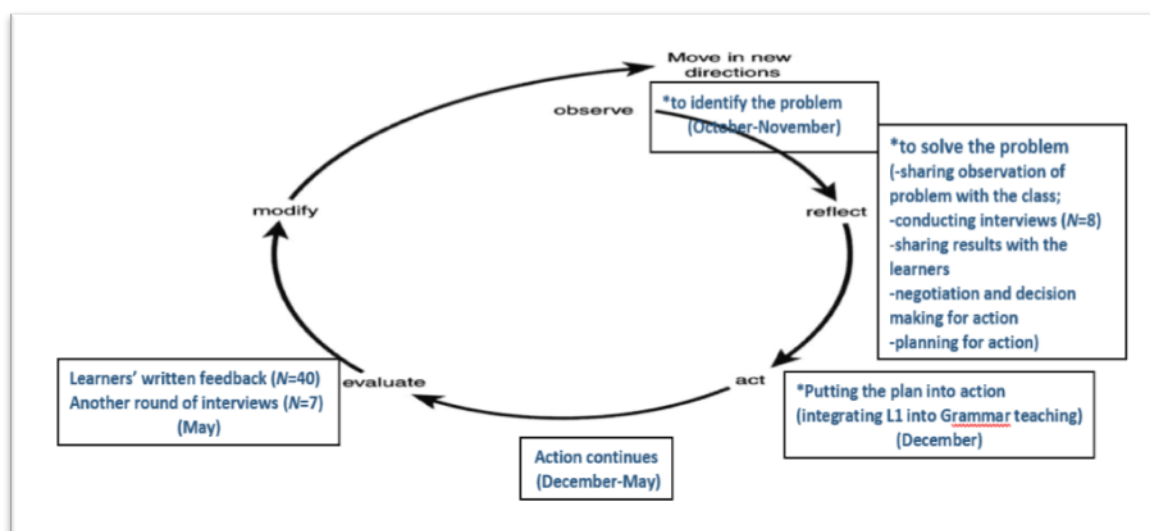
Figure 1. AR cycle in research design and data collection.

The *observation* occurred throughout October-November via deliberate attention to understand what was going on in the class as the learners were hard to involve. Following the detection of the problem, *unresponsiveness,* I did whatever I could, to activate them and become responsive. Seeing that my efforts made no big difference, I kept thinking, *reflection*, as I became sure that I needed to do something. Taking a further step, I shared my observation of the problem with the class and asked them if they wanted to help me find a solution regarding what they needed and wanted. Hence, we planned for interviews, *as preliminary needs analysis*, which were conducted with those (*N*=8) volunteered. For the interviews, a semi-structured interview form developed by the researcher was used. Keeping reflecting on what the learners meant as the interviews continued (end of November), I performed the data analysis. Despite some other issues (which are dealt with in detail in findings) mentioned equally, the learners majorly complained about *use of English as medium of instruction* and suggested that they needed L1 (Turkish).

Having achieved such a finding, I was challenged as I always thought English should be taught in English since being exposed to the language in class, particularly in English as foreign language (EFL) settings and having the opportunities for using the language are key for learners to learn, practice, familiarize themselves with the language, and ultimately become competent in it (Ellis, 1994). However, I also knew that in any context of teaching, learners' needs and wants come first. Thus, I started reviewing if and to what extent L1 should be integrated into teaching, although I knew, it should be to some extent since leaving learners' L1 backgrounds aside was not fair either. Following this inner and mental critic, synthesized with the readings, I shared the findings (Table 2) with the learners and asked them to negotiate. Through the negotiations right at that time, without any interference of me, the learners decided that they needed *L1* particularly *for the teaching of grammar* as they thought that a great extent of their unresponsiveness resulted from their grammatical incompetence which challenged them in comprehending the input and accordingly responding to it. This made sense as not being able to respond indicated lack of communicative competence which includes the "knowledge of rules of phonology, lexis, syntax, and semantics" (Canale & Swain, 1980 as cited in Brown & Abeywickrama, 2010, p.294). The connection between them further refers to grammatical knowledge including grammatical forms plus grammatical and pragmatic meaning (Larsen-Freeman, 1991, 1997 as cited in Brown & Abeywickrama, 2010).

Following this decision making, I took some more time to review the coursebook (see Hobbs & Starr Keddle, 2007) to see how many grammatical points there were ahead to cover, where I would need to switch to Turkish, and how I would do it, either implicitly or explicitly. Thus, the following action plan was formed (see Table 1).

Table 1. Action plan

	WHAT?			HOW?
Actions	*Grammar point;*	*Schema building;*	*Grammar teaching followed;*	*Explicit grammar teaching was followed by;*
A1	Will, be going to, will have to, will be able to	Talking about future predictions	Reading "Future predictions"	"getting the learners to complete a set of given opinions with the grammar points, then to discuss their answers in pairs"
A2	Must, can't, may, might, could	Talking about changing countries and working abroad	Reading "Job adventures"	"getting the learners to match some captions with the rules, and to speculate about given problems"
A3	The Passive	Talking about ethical consuming	Reading "Fairtrade"	"getting the learners to find examples from the reading text".
A4	Conditionals	Talking about diversity at workplace	Reading "Discrimination at work"	"getting the learners to find examples of conditionals in the reading text and to complete the given sentences with the correct form of conditionals"
A5	Past Perfect	Talking about brands and values	Reading "Nike"	"getting the learners to find examples from the reading text and to write explanations for given situations"
A6	Must have/can't have/might-could have +past participle	Talking about workplace, office, and office space	Reading "News from the workplace"	"getting the learners to match the given captions with the given pictures and to complete the given sentences with the grammar point"
A7	Third conditional	Talking about personal qualities, workplace skills	Listening "TEAM building games and activities"	"getting the learners to listen to some situations and imagine the results and to complete the given sentences with the grammar point"
A8	Reported speech: said, told, and questions	Talking about giving presentations	Listening "A presentation delivered to Internet access company"	"getting the learners to listen to more extracts from the programme and correct the errors in given notes; and getting them to report some more extracts from the listening."

As shown, there were eight grammar points. The teaching of each began with schema building through brainstorming. Depending on the presentation of the content, reading or listening was carried out as a means of presenting the learners with the contexts to infer the structure. This implicit activation, which was in English, was followed by explicit teaching of grammar. At that point, I switched to Turkish to present both form and function as the learners needed not only to gain knowledge and skills of the structure but also in which context to use it. The use of Turkish was not simply and only to give a brief explanation of the grammar point, to provide the learners with a deductive rule in other words, rather it was to help them comprehend what the form is, what it does, in which contexts, for which purposes it is used, and to what extent it resembles or differs from the Turkish equivalent. The teaching of each grammar point within each unit took almost an hour which made one-fourth of the weekly hour. The teaching of grammar was then followed by practicing through various tasks (the last column in the table). As there were multiple actions, each also helped me improve those coming next. This *intervention* step, *integration of L1 into grammar teaching*, continued till the end of academic year (May 15th) which made almost 20 weeks.

As the steps in AR are more than teaching, they also included data collection and its analysis to evaluate the effectiveness of the solution (Ellis, 2012). Therefore, evaluation was maintained through the learners' end-of-course written reflections (*N*=40) and another round of interviews (*N*=7) to further open their reflections and encourage them to elaborate on the benefits of the intervention, if there was any. For this second round of interviews, another semi-structured interview form was developed and used. At this point, I need to mention that assessment practices including two mid-term and two end-of-term exams included grammar tasks to see if and

how the use of L1 improved the learners' knowledge and performance. However, they are not included in this paper due to the difficulty of presenting and discussing all findings within the confines of them.

Therefore, qualitative data was appropriate as it would help develop deeper understanding and "capture a sufficient level of detail about the natural context" and "participants' views of the situation being studied" (Dörnyei, 2007, p.38).

Participants

The participants, who were 22 on average, had diverse language learning background and competencies although they were supposed to be true pre-intermediates as they were taking the course. Those interviewed for the preliminary interviews included 5 females and 3 males, and others in the second round of interviews at the end of the term included 4 females and 3 males. Those provided end-of-course written reflections included 14 females and 12 males. All started taking English beginning from the secondary school. As they had been through different high schools, including multi-program and the vocational ones, they had different experiences like an interested and competent teacher in one hand, or a disinterested teacher only using Turkish in the other. They regarded this as a major reason for their lack of language competency. There were also some others who told that another teacher, a physical education for instance, taught since there was no English teacher in their school. Hence, they did not even have the chance to adequately learn English. Some even took the responsibility of their inadequate background of English due to not having self-study skills. Lastly, they were seen to have different motivation and interest in learning which made some of them interested, active, and responsive and many others unresponsive and hard to involve.

Data analysis

The data collected from the interviews and written reflections were qualitatively analyzed and then quantified. The interview data were firstly transformed into texts through transcriptions (Dörnyei, 2007) which were then read several times until familiarization was assured (Creswell, 2009) particularly to "start teasing out the hidden meaning from it" (Dörnyei, 2007, p.242). Analysis of the transcriptions was made through qualitative content analysis including quantification of certain words, phrases, or issues falling into a specific category emerging out of the data (Dörnyei, 2007). For the analysis of the written reflections which were large in sample (N=40), and were comprised of almost a page feedback, in some instances more, data saturation was kept in mind as it would become repetitive (Mason, 2010) which means that there is no more new information, thus no need for further analysis and coding as iterative process of looking back and forth in the data produces no more ideas and categories (Dörnyei, 2007). Therefore, a sample of 26 was achieved through reduction. At this point, I need to clarify that the reduction was not random. All the reflections were read thoroughly a couple of times and coded. However, the coding did not end in the first cycle. I kept coding and recoding (see Saldana, 2009). Through these cycles of coding, I had the chance to see the extent of depth in the reflection. Thus, those which were superficial and only recurrent in that idea (code) without a strong evidence were excluded from the final analysis. For this reason, depth and richness of the reflections, despite indication of the same code, determined the level of saturation as there was no more detailed data (Fush & Ness, 2015). To assure the validation and confidentiality of the findings, another language teacher, who was internal to the research context, but outsider to the study, crosschecked the categories, themes, and verbatim data, and provided confirmation and suggestions as well for the improvisation of the categorizations. As for privacy of the participants and their responses, numbers were given both for the interviews and their written reflections as the findings were reported.

Findings

Learners' reasons regarding their unresponsiveness: findings from preliminary interviews

The first round of interviews elicited the learners' reasons regarding their *unresponsiveness* and negotiation of what they wanted and needed to solve it. Table below presents the findings.

Table 2. Learners' reasons regarding their unresponsiveness

Categories	Themes	f
Teaching methodology	*use of English as medium of instruction	7
Grammatical incompetence	*inadequate syntactic knowledge	7
Listening comprehension incompetence	*inadequate knowledge of sound discrimination/word recognition	7
Previous learning habits	*previous teachers' teaching style & methods	3
	*education system (pass or failure mechanism)	1
	*lack of interest in language learning in early years of it	1
Vocabulary incompetence	*inadequate repository of vocabulary knowledge	5
Lack of motivation in language learning	*coming to the classes unprepared	4
	*external motivation	1
Pronunciation incompetence	*lack of phonology knowledge	3
	*lack of confidence	1
Speaking incompetence	*inability to make full sentences	1
	*inadequate knowledge of mechanics of language (vocabulary, pronunciation, grammar)	2
Speaking anxiety	*exam-oriented system	2
	*inadequate command of English	1
Lack of speaking practice	*limited out-of-class engagement with language	3
Peer/teacher pressure	*fear of being laughed at	2
	*class size	1
Fear of failure	*learned helplessness (due to lack of background knowledge of language)	3
Total		55

As seen, there are various reasons, but some dominate. Firstly, majority of them attached their unresponsiveness to *the teachers' teaching methodology* (*n*=7) relying on use of English as medium of instruction. For instance, Learner1 stated *"Your use of English from the very beginning... We know some of the words, can understand some of them, or go and search for some others. But as the friends told, then the complaints emerged..., 'the teacher always speaks English', at least spread some Turkish around."* This majorly indicates comprehension problem which also suggests inadequacies in their receptive skills. Besides, *grammatical incompetence* (*n*=7) was another reason confirming the solution, *use of L1 in teaching grammar*, they offered. *"We don't know, understand, don't know what to use when to use ... we lack grammar knowledge. ... Some of us speak, but they lack grammar ..., even when you teach grammar in English, I have problems. At least for me, it would be better if it was through Turkish" (Learner2).* As a very basic skill in classroom interaction, *listening comprehension competence which they found themselves incompetent* (*n*=7) was also suggested as a reason since they justified having inadequate knowledge of sound discrimination and word recognition which are key to comprehend the input and respond to it appropriately. Although the learners equally mentioned the three reasons, their primary focus was on the use of English as medium of instruction which they thought to trigger their incomprehension of the input and inability to respond. The presence of grammar and listening incompetence in the background also exacerbated the situation.

Besides, *previous learning habits* (*n*=5), previous teachers' teaching styles and methodologies majorly relying on Grammar Translation Method and adopting Turkish as medium of instruction, were other reasons. Those who did so also thought that lack of interest in their early years of language learning and the education system which they thought to rely on pass or fail mechanism caused them to become unresponsive. *"It is too late to not to understand [to be unresponsive]. It is due to the quality of education in schools, particularly inadequate education in high schools. .... It takes 12 years till the university, but they cannot give directions to a tourist, such a bitter truth. ... The education system in schools relies on memorization, there is no place for practice. There is nothing making you speak English. One just memorizes and passes the exams, and they don't try to learn, but just to pass the course" (Learner4).* Other reasons included; *vocabulary incompetence* (*n*=5), as a natural extension of inadequate vocabulary repository; *lack of motivation* (*n*=5) resulting in coming classes unprepared, just to pass them as an indicator of their external motivation; and *pronunciation incompetence* (*n*=4) as a reflection of inadequate phonology knowledge, production of sounds, sound patterns, and words in speech. Therefore, due to weak knowledge of grammatical forms of language, they regarded *speaking incompetence* (*n*=3) as another reason. Learner7 expressed this nicely; *"Due to inadequacy of both vocabulary and*

pronunciation, it is impossible to move on to speaking. ... Although I love learning new words, due to my language background which I do not see strong, speaking is a problem for me, ... grammar is also a great deficiency..." Similarly, *speaking anxiety* (*n*=3) and *lack of speaking practice* (*n*=3), due to limited out of classroom engagement with language and inadequate command of English, are linked to speaking incompetence too. Lastly, they mentioned *peer/teacher pressure* (*n*=3) resulting from fear of being laughed at when making mistakes, and *fear of failure* (*n*=3) due to weak background knowledge of language.

Benefits of L1 in grammar teaching: findings from end-of-course written reflections and another round of interviews

First and foremost, their written reflections showed that they were good and critical observers of what was being done over the process, as they knew it was a response to their needs. *"What I've liked most about this course is your effort to consider how we can understand better and design accordingly. While I have not understood it at all, I have now become to do so" (written reflection18).* Another said, *"it was extraordinary" (written reflection25).*

Table 3. Benefits of L1 in teaching grammar

Benefits	*f*
More and better input comprehension	15
Vocabulary skills development	12
Improved grammar knowledge	9
Writing skills development	6
Speaking skills development	2
Awareness regarding language learning	2
Breaking down prejudice towards English	1
Awareness regarding lacks	1
Total	*48*

When it comes to the benefits (Table 3), development of *input comprehension* (*n*=15) comes first. *"... Yes, I was challenged to understand, but I have now become to watch most of TV series without subtitles ..." (written reflection4)*. *Vocabulary skills development* (*n*=12) was another benefit as they linked it to the input they received and word building practices regularly dealt with in each topic. Some stated that emphasis put on grammar teaching facilitated their familiarization with new words which is a contribution of learners' L1 and teacher talk to incidental vocabulary learning even when input includes unfamiliar words. For instance, *"Besides use of Turkish in grammar teaching, the reading texts and tasks related to them were very contributory. As we worked on the reading texts, my vocabulary knowledge and awareness regarding grammatical forms developed" (written reflection20)*. Another major benefit is *writing skills development* (*n*=6) as they were regularly involved in writing tasks. In her reflection, one (*written reflection22*) said, *"I already had some grammar knowledge. But, with the contribution of the course, it has become better. This is also due to the teaching of grammar through L1. I've begun writing something in writing sections."* Despite less frequent, few (*n*=2) thought their *speaking skills* developed. Some others emphasized that they started gaining awareness regarding language learning (*n*=2), due to involvement in classroom tasks. Similarly, one (*written reflection16*) told, s*he began breaking down her prejudices towards English thanks to the teacher's teaching methodology*, and her comment shows the need for teaching which is responsive to the contextual factors, realities of the classroom, and its learners' needs and expectations. Lastly, one (*written reflection4*) said special attention paid to the teaching of grammar and use of L1 enabled her to become aware of what grammar forms she lacked and needed to develop.

Another round of interviews was conducted to elicit the learners' further reflections regarding the benefits of teaching grammar through L1. First and foremost, except one, all (*n*=6) thought it worked well. To justify what he thought, Learner3, who was one of those responsive ones, stated *"It was nice. If it had been through English, they [those unresponsive] would have made more mistakes in the exams. It was more comprehensible. Additionally, as the topics went complex, we would have had more difficulty in understanding"*. Similarly, Learner5 thought *"It worked quite a lot. When it was through English, it was hard to comprehend. We got it [what and how to use] better. If it had been through English, it would have made no benefit. Seeing that I got it, I felt motivated"*. Another, who was a retake, also shared similar thoughts as she thought use of L1 made it easier to recall, as she had difficulty in doing so (Learner4). However, Learner6 who was one of those responsive learners thought *"It should be (also have been) through English. Those who are unable to*

understand should have been eliminated somehow in previous years. Not being able to comprehend at this level is not normal. Your use of English as medium of instruction is very nice as it helps us see good use of language, get familiar with new words, their pronunciation, and better listening comprehension." Moreover, they thought their language learning was positively affected as they became to comprehend more and better through "comparing with Turkish (Learner5)", "getting meaning through sentence structure (Learner3)", or "performing better in listening tasks (Learner2)". They also developed awareness regarding how different aspects of language work (metalinguistic awareness) as they "started realizing sentence structure better as it is different from Turkish (Learner3)". They also started making sense of meaning better as they "gained more awareness regarding morphology (Learner1, Learner3)", and "syntax (Learner5)". Besides, they developed self-confidence regarding language comprehension as they "felt better in finding the Turkish equivalent of language forms (Learner4)", particularly in "writing and speaking (Learner1, Learner6)", and "word formation (Learner5)". As seen, they attained self-confidence in various aspects. Similarly, they started feeling less stressed and anxious which previously resulted from incomprehension (Learner1, Learner4). Lastly, to some extent, some stated relating English to Turkish which could show transfer effect of L1, as generally acknowledged. However, for Learner1 it was possible for some Tenses [simple present, present continuous, simple past], but was hard as they continued [she means those that do not coincide e.g. present perfect tense which is hard to comprehend for Turkish learners]. For another, it was hard to do so particularly for if clause Type 2 as he was challenged to comprehend unreal present in Turkish (Learner6). Hence, if the forms in two languages match, learners take the advantage of it, and if vice versa, they are challenged as relating new structures to their already established schemas might be difficult.

Discussion

Despite lack of research on learners' (second, foreign, or ESP) unresponsiveness in language classes, studies investigated the issue from similar perspectives such as silence, reluctance, and speaking anxiety (Cepon, 2016; Iglesias, 2016; Öztürk & Gürbüz, 2014; Savaşçı, 2014; Subaşı, 2010; Tatar, 2005). Similar results such as not being able find correct word(s), not being able to pronounce correctly, lack of speaking practice, lack of communicative competence, lack of grammar knowledge, weak language background, fear of making mistakes, and negative evaluation of peers and teachers are reported among the major causes of it (Cepon, 2016; Öztürk & Gürbüz, 2014; Takkaç Tulgar, 2018; Xie, 2017). Those who are competent are also reported to be unresponsive (Şubası, 2010) which could have a link to low self-confidence (see Akkakoson, 2016; Bailey, 1983, as cited in Ellis, 1994; Savaşçı, 2014; Riasati, 2018) making learners avoid responding. In her study on Turkish international graduate students' silence, Tatar (2005) reveals similar issues i.e. insufficient language skills due to non-nativeness, peers particularly those native and competent dominating the discussions and causing intimidation on those who are relatively less competent, some degree of anxiety due to fear of making mistakes thus losing face and prestigious, and unpreparedness and low self-confidence. Besides, fear of failure and fear of being laughed at impede student success in the classroom (Bledsoe & Baskin, 2014). Hence, it is obvious that responding is a combination of cognitive, intellectual, and affective states.

When it comes to the benefits of the action to solve out unresponsiveness, the learners in this study were positive regarding use of L1 and appreciated it since the classroom atmosphere and their needs were considered. For instance, being able to comprehend the input more and better is an important finding as listening comprehension incompetence was one of the basic reasons for their unresponsiveness (see Table 2). As listening does not simply mean the act of hearing, not being able to comprehend what they listened to indicates their weakness to detect the phonemes, morphemes, grammatical form, intended and implied meaning (see Larsen-Freeman, 1991, 1997 as cited in Brown & Abeywickrama, 2010). Therefore, it seems that the process made a difference on the development of their grammar, vocabulary, and pronunciation knowledge as key issues in language learning for being able to interact and respond since without comprehension of any of these, which does not guarantee but is a prerequisite, learning is almost impossible to occur (Gass & Selinker, 2008; Krashen, 2004). Studies similarly using the L1 (Spanish) selectively also show that L1 functions as an effective tool for cross-linguistic analysis to enable learners to understand and acquire vocabulary in English as they analyze and compare words with their equivalents in L1 (see Cuartas Alvarez, 2014). Besides, the learners' vocabulary development might be linked to the role of L1 in relating new knowledge to the existing one by making semantic and syntactic connections (Yürekli Kaynardağ, 2016) which Butzkamm (2003) regards as "building cross-linguistic networks" (p.35). This could suggest transfer effect (Du, 2016; Madrinan, 2014; Mede et al., 2014) influencing how meaning and content are conveyed through deliberate attention to language forms (Nation, 2003). Moreover, as the rest of the course, the teaching and practice of other skills, were still through English, the input that the learners received through listening tasks, the teacher and peer talks, and practices following grammar teaching could have facilitated vocabulary learning and the ability to comprehend. Despite

considering use of English as medium of instruction as a problem in the beginning, the learners started finding it as a facilitator as the time passed. Therefore, despite the challenges resulting from their language background, motivational, and affective states, this finding could show the difference that persistent use of English, exposure to it, and teacher's talk as a model in language classrooms (Lew, 2015) made since "it serves as the significant, and sometimes only, source of authentic, scaffolded input" (Moeller & Roberts, 2013, p.22).

Writing skills development as another major benefit because of regular involvement in writing tasks supports the contribution of grammar teaching through L1 as it awakens universal grammar (see Butzkamm, 2003), thus results in better performance and active involvement (Lameta-Tufuga, 1994, cited in Nation, 2003; Usadiati, 2009). This could also show that as the learners comprehended the related grammar structure, they became to use it to express their ideas in written language which could have made them perceive that their writing skills developed. Because, it is definite that learners get stuck if they are not able to find the right structure or vocabulary to express their thoughts both written and spoken. Therefore, comprehension of grammatical structures in English and matching them to their equivalents in L1 could improve learners' expressions in written language. Besides, the development of *speaking skill*s could result from the teacher's and peers' talk in the class which provided the learners with the chance of being exposed to the language and practice it, though not directly targeted. Thus, the maximization of classroom input seemed to affect the learners' L2 development (Turnbull, 2001). *Gaining awareness regarding language learning and grammatical knowledge they lack*ed and *breaking down their prejudices towards English* for some learners show that use of L1 in grammar teaching could have activated their awareness regarding if and how both language forms match (Butzkamm, 2003; Sinclair, 1986) as learners could use their L1 as a reference point to make cross-comparison of grammar structures in both languages and to make sense of the input they receive in English (also see Cuartas Alvarez, 2014). Hence, this shows the need for responsive teaching which considers the contextual factors, realities of the classroom and its learners' needs and expectations and which could affect later success of learning (Auerbach, 1993).

When all benefits are critically examined, the process is seen to be effective. This could have a link to L1's potential to lower such affective barriers as speaking anxiety, low motivation, fear of failure, or peer/teacher pressure which the learners included among the reasons of their unresponsiveness. Hence, no matter how learners are challenged as they are exposed to English in class, continuous involvement in it is seen to make a difference, thus as research suggests (Butzkamm, 2003; Cheng, 2015), role of L1 cannot be disregarded to ease the tension in language classroom. Besides, considering that L1 was only integrated to teaching grammar, the benefits emerged out of their reflections could highlight the role of effective use of both L1 and English (Lida, 2014) as learners are known to use their L1 as a cognitive tool (Gass & Selinker, 2008). Moreover, integrating L1 into teaching grammar could have also facilitated the learners' awareness regarding sub-components, i.e. phonology, morphology, syntax, and semantics which the grammatical knowledge and competence includes (Canale & Swain, 1980 as cited in Brown & Abeywickrama, 2010) as the cross-comparison of L1 and English was provided. Lastly, all stated that they became more willing to participate and communicate (also see Uztosun et al., 2014). In this sense, the shift from unresponsiveness to the perception of willingness to participate and communicate could have resulted from developing confidence in their abilities to use English in the class (also see Cuartas Alvarez, 2014). Therefore, in such cases where learners' voices are considered, their motivation and self-esteem are boosted (see Uztosun et al., 2014, 2018).

Therefore, as many others suggest (Awan & Sipra, 2015; Berning, 2016; Çelik, 2008; Nation, 2003) L1 should be used if needed, but over reliance should definitely be avoided as learners may not benefit particularly in EFL contexts where teachers are the only linguistic model and source of input (Turnbull, 2001). However, in cases, like the current, where learners are less proficient and lack adequate knowledge of language, resistance and unresponsiveness might emerge. Therefore, L1 is seen to scaffold cognitive workload and ease comprehension (Bruen & Kelly, 2014).

Consequently, unresponsiveness can be solved through creation of appropriate learning environments and improvement of teaching methods (Bledsoe & Baskin, 2014; Zhouyuan, 2016). Therefore, there needs to be means, teaching methodologies, specific techniques, and classroom research (see Uztosun et al., 2014, 2018) to meet learners' needs to encourage them to become active and responsive.

Conclusion and implications

Firstly, as a response to the learners' unresponsiveness in an ESP context, this study reports the results revealed through the action, *integration of L1 into teaching grammar,* taken with reference to their needs and wants. In particular, the reasons of their unresponsiveness and the benefits of the action were sought.

The study shows that learners' unresponsiveness in language classroom is a combination of cognitive, intellectual, and affective issues (Tatar, 2005). The reason that the learners majorly focused on was the teaching methodology, which adopted English as medium of instruction, but was disliked due to comprehension problems resulting from weaknesses in their language background, thus triggering unresponsiveness. This suggests some certain level of grammar, vocabulary, and phonology knowledge (Gass & Selinker, 2008) is influential on being responsive, but development of which do not emerge overnight. Moreover, motivational concerns together with the social context of the classroom, i.e. peer and teacher pressure, are included among the reasons. Hence, we need competent and interested teachers transforming their classrooms into effective learning environments and giving learners the chances to acquire and develop their language skills as well as their confidence and motivation to become active and responsive. In this case, the role played by responsive teaching is undeniable since acknowledging classroom realities and assessing and addressing learners' needs and wants make a difference. In this regard, action research can be effective to meet learners' needs and empower both learners and teachers (see Ali, 2020).

The results also show that learners' behaviors change when measures are taken. Within this study, the learners held positive perceptions regarding integration of L1 into teaching grammar as it was seen to facilitate language skills development and various aspects of language learning i.e., metalinguistic awareness, input comprehension, and also self-confidence and lower stress and anxiety which were seen to facilitate their willingness to participate and communicate (also see Uztosun et al., 2014). However, it needs to be remembered that L1 had a controlled use as it was only used for teaching grammar. Therefore, use of English as medium of instruction was still seen to contribute as the learners reported development in various aspects of language i.e. input comprehension, awareness regarding their lacks, etc. Therefore, no matter how learners resist as they experience difficulties in comprehension, persistent use of English is seen to change the stance they take towards it. Hence, language teachers should have the awareness and skills to balance L1 and English as leaving L1 aside, only for the sake of English, would have the risk to make no difference as incomprehension would further block learning. Thus, facilitative role of L1 shows that it should and needs to be in language classrooms (Brooks-Lewis, 2009; Bruen & Kelly, 2014; Butzkamm, 2003; Jan et al., 2014) despite the challenges that over reliance on it bring.

Besides the benefits that the learners made, I, as the researcher, should also reflect on my experience as it was a breakthrough to change my mindset. I was challenged in the beginning if I really needed to adopt L1 in teaching. However, what I had found through the learners' reflections showed me that it really worked as it resulted in various benefits. I also observed some learners who were almost completely silent then became responsive. They openly acknowledged the difference, i.e. becoming to comprehend and feeling motivated to join all the classes, that the action made. Hence, in my further teaching practices, I now have a concrete evidence of if and how much L1 works. Therefore, as the existing research suggest, AR has a critical value on teacher learning (Johnson, 2009; Kayaoğlu, 2015).

Lastly, despite all the contributions achieved through the adoption of AR methodology, the study is not free from limitations as it reports the case of a single class, thus findings need to be validated by further research. First and foremost, the data relies on the learners' self-reports and perceptions which might require observation data to validate their perceptions. Besides, due to lack of tests measuring their grammar knowledge before the study, there is no data in hand. Therefore, further studies in similar contexts could begin with a pre-test and re-apply it at the end to see if teaching grammar through L1 makes a difference on grammar knowledge.

Acknowledgements or Notes

Disclosure statement

I have no potential conflict of interest to report.

References

Akkakoson, S. (2016). Speaking anxiety in English conversation classroom among Thai students. *Malaysian Journal of Learning and Instruction, 13*, 63-82.

Ali, A. D. (2020). Implementing action research in EFL/ESL classrooms: a systematic review of literature 2010-2019. *Systematic Practice and Action Research, 33,* 341-362.

Auerbach, E. R. (1993). Reexamining English only in the ESL classroom. *TESOL Quarterly,* 27 (1), 1-18.

Awan, M. A., & Sipra, M. A. (2015). A 'judicious' use of L1 in TL classroom: Socio-political, psychological and functional dimensions. *Advances in Language and Literary Studies, 6* (5), 16-21.

Berning, B. S. (2016). Can't I speak Japanese? The use of students' L1 as a cognitive tool for collaborative writing tasks in a Japanese EFL context. (Unpublished MA Thesis). Hamline University, Saint Paul, Minnesota.

Bismilla, V. (2011). Creating space for students' mother tongues in college classrooms. *Journal of Global Citizenship & Equity Education, 1*(1), 60-95.

Bledsoe, T. S., & Baskin, J. J. (2014). Recognizing student fear: The elephant in the classroom. *College Teaching, 62*, 32-41.

Boustani, K. (2019). The correlation between translation equivalence, as a vocabulary learning strategy, and Tunisian EFL learners' speaking anxiety. *Languages, 4*(19), 1-24.

Brooks-Lewis, K. A. (2009). Adult learners' perceptions of the incorporation of their L1 in foreign language teaching and learning. *Applied Linguistics, 30*(2), 216-235.

Brown, H. D. & Abeywickrama, P. (2010). *Language assessment. Principles and classroom practices* (2nd ed.). New York: Pearson-Longman.

Bruen, J., & Kelly, N. (2014). Using a shared L1 to reduce cognitive overload and anxiety levels in the L2 classroom. *The Language Learning Journal, 45*(3), 368-381.

Burns, A. (2009). Action research in second language teacher education. In A. Burns & J. C. Richards (Eds.), *Second language teacher education* (289-297). New York: Cambridge University Press.

Butzkamm, W. (2003). We only learn language once. The role of the mother tongue in FL classrooms: death of a dogma. *The Language Learning Journal, 28*(1), 29-39.

Cepon, S. (2016). Teachers' and students' perspectives on the reasons for speaking anxiety in English for Specific Purposes. *ESP Today Journal of English for Specific Purposes at Tertiary Level, 4*(2), 184-201.

Cheng, W. (2015). A case study of action research on communicative language teaching. *Journal of Interdisciplinary Mathematics, 18*(6), 705-717.

Creswell, J. W. (2009). The analysis of qualitative data. In K. F. Punch (Ed.), *Introduction to research methods in education* (169-205). Thousand Oaks: Sage.

Cuartas Alvarez, L. F. (2014). Selective use of mother tongue to enhance students' English learning processes … Beyond the same assumptions. *PROFILE: Issues in teachers' professional development, 16*(1), 137-151.

Çelik, S. (2008). Opening the door: An examination of mother tongue use in foreign language classrooms. *Hacettepe University Journal of Education, 34*, 75-85.

Debreli, E. & Oyman, N. (2016). Students' preferences on the use of mother tongue in English as a Foreign language classrooms: is it the time to re-examine English-only policies? *English Language Teaching, 9*(1), 148-162.

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.

Du. Y. (2016). *The use of first and second language in Chinese University EFL classrooms*. Singapore: Springer

Ellis, R. (2012). *Language teaching research & language pedagogy*. West Sussex: Wiley-Blackwell.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Fromkin, V., Rodman, R., & Hyams, N. (2007). *An introduction to language* (8th ed.). Boston: Thomson Wadsworth.

Fush, P. I., & Ness, L. R. (2015). Are we there yet? Data saturation in qualitative research. *The Qualitative Report, 20*(9), 1408-1416.

Gass, S. M., & Selinker, L. (2008). *Second language acquisition. An introductory course.* New York: Routledge.

Giacobbe, J. (1992). A cognitive view of the role of L1 in the L2 acquisition process. *Second Language Research, 8*(3), 232-250.

Hobbs, M., & Starr Keddle, J. (2007). *Commerce 2.* Oxford University Press.

Jan, C., Li, B., & Lin, C. (2014). The use of the mother tongue in Chinese EFL classrooms. *Journal of China University of Science and Technology, 58*, 161-180.

Jingxia, L. (2010). Teachers' code-switching to the L1 in EFL classroom. *The Open Applied Linguistics Journal,* 3, 10-23.

Johnson, K. E. (2009). *Second language teacher education. A sociocultural perspective*. New York: Routledge.

Iglesias, L. M. A. (2016). Students' opinions about the use of L1 in an intermediate level course. (MA dissertation). University of St Mark and St John, Plymouth, UK.

İnal, S., & Turhanlı, I. (2019). Teachers' opinions on the use of L1 in EFL classes. *Journal of Language and Linguistic Studies, 15*(3), 861-875.

Inan, S. (2016). Teachers' use of L1 in young learner EFL classrooms in Turkey. (MA dissertation). Gazi University, Ankara, Turkey.

Kayaoğlu, M. N. (2015). Teacher researchers in action research in a heavily centralized education system. *Educational Action Research, 23*(2), 140-161.

Krashen, S. (2004). Applying the comprehension hypothesis: some suggestions. Paper presented at 13th International Symposium and Book Fair on Language Teaching (English Teachers Association of the Republic of China), Taipei, Taiwan, November 13.

Lebak, K., & Tinsley, R. (2010). Can inquiry and reflection be contagious? Science teachers, students, and action research. *Journal of Science Teacher Education, 21*, 953-970.

Lew, A. W. M. (2015). Input for the second language classroom: Some innovations and insights. *Teachers College, Columbia University Working Papers for TESOL & Applied Linguistics, 15*(2), 54-56.

Lida, T. (2014). A study on the roles of L1 and L2 in foreign language classroom. *Language awareness, language planning, and translanguaging, 31*, 1-15.

Luchini, P., & Rosello, A. (2007). Developing learners' oral communicative language abilities: A collaborative action research project in Argentina. *The Journal of Asia TEFL, 4*(4), 245-271.

Madrinan, M. S. (2014). The use of first language in the second language classroom: A support for second language acquisition. *Gist Education and Learning Research Journal, 9,* 50-66.

Mason, M. (2010). Sample size and saturation in PhD studies using qualitative interviews. *Forum: Qualitative Social Research, 11*(3). Online available at http://www.qualitative-research.net/index.php/fqs/article/view/1428/3027

Mede, E., Tutal, C., Ayaz, D., Çalışır, K. Y., & Akın, Ş. (2014). The effects of language transfer in Turkish EFL learners. *ELT Research Journal, 3*(2), 70-83.

Miles, R. (2004). Evaluating the use of L1 in the English classroom. (MA dissertation). University of Birmingham, Birmingham, UK.

Moeller, A. J., & Roberts, A. (2013). Keeping it in the target language. *Faculty Publications: Department of Teaching, Learning and Teacher Education, 178*, 21-38.

Nation, P. (2003). The role of first language in foreign language learning. *Asian EFL Journal*, *5*(2), 1-8.

Oxford, R. (2001). Integrated skills in the ESL/EFL classroom. Eric Document ED456670. Retrieved from https://eric.ed.gov/?id=ED456670

Öztürk, G., & Gürbüz, N. (2014). Speaking anxiety among Turkish EFL learners: the case at a state university. *Journal of Language and Linguistics Studies, 10*(1), 1-17.

Paker, T., & Karaağaç, Ö. (2015). The use and functions of mother tongue in EFL classes. *Procedia-Social and Behavioral Sciences, 199*, 111-119.

Paltridge, B., & Starfield, S. (2013). *The handbook of English for specific purposes*. (1st Ed.). West Sussex: Wiley-Blackwell Publishing.

Rainey, I. (2000). Action research and the English as a foreign language practitioner: time to take stock. *Educational Action Research, 8*(1), 65-91.

Riasati, M. J. (2018). Willingness to speak English among foreign language learners: A causal model. *Cogent Education, 5*, 1-17.

Saldana, J. (2009). *The coding manual for qualitative researchers.* Thousand Oaks, California: SAGE Publications.

Salı, P. (2014). An analysis of the teachers' use of L1 in Turkish classrooms. *System, 42*, 308-318.

Savaşçı, M. (2014). Why are some students reluctant to use L2 in EFL speaking classes? An action research at tertiary level. *Procedia-Social and Behavioral Sciences, 116*, 2682-2686.

Sinclair, A. (1986). Metalinguistic knowledge and language development. *Advances in Psychology, 39*, 609-627.

Subaşı, G. (2010). What are the main sources of Turkish EFL students' anxiety in oral practice? *Turkish Online Journal of Qualitative Inquiry, 1*(2), 29-49.

Szendröi, I. (2010). Teacher talk in the ESP classroom. The results of a pilot observation study conducted in the Tourism context. *WoPaLP, 4*, 39-59.

Takkaç Tulgar, A. (2018). Speaking anxiety of foreign learners of Turkish in target context. *International Online Journal of Education and Teaching, 5*(2), 313-332.

Tatar, S. (2005). Classroom participation by international students: The case of Turkish graduate students. *Journal of Studies in International Education, 9*(4), 337-355.

Teng, F. (2019). Retention of new words learned incidentally from reading: Word exposure frequency, L1 marginal glosses, and their combination. *Language Teaching Research*, 1-28.

Timuçin, M., & Baytar, İ. (2015). The functions of the use of L1: Insights from an EFL classroom. *Kastamonu Education Journal, 23*(1), 241-252.

Turnbull, M. (2001). There is a role for the L1 in foreign and second language teaching, But …" *Canadian Modern Language Review, 57*(4), 531-540.

Turnbull, B. (2018). The potential impact of cultural and educational background on foreign language teachers' use of the L1. *The Journal of Language Teaching and Learning, 8*(1), 53-70.

Usadiati, W. (2009). Contribution of L1 in EFL teaching. *Kata, 11*(2), 171-184.

Uyar, Y. (2012). Using L1 to enhance grammar learning and having English only policy in EFL classes. Eric Document ED538967 Retrieved from https://files.eric.ed.gov/fulltext/ED538967.pdf

Uztosun, M. S., Skinner, N., & Cadorath, J. (2018). An action research study into the role of student negotiation in enhancing perceived student engagement during English speaking classes at university level in Turkey. *Educational Action Research, 26*(1), 106-126.

Uztosun, M. S., Skinner, N., & Cadorath, J. (2014). An action research study designed to implement student negotiation to improve speaking classroom practice in Turkey. *Educational Action Research, 22*(4), 488-504.

Xie, Q. (2017). Investigating the target language use in and outside business English classrooms for non-English major undergraduates at a Chinese University. *Cogent Education, 4*, 1-16.

Yavuz, F. (2012). The attitudes of English teachers about the use of L1 in the teaching of L2. *Procedia-Social and Behavioral Sciences, 46*, 4339-4344.

Yürekli Kaynardağ, A. (2016). Shall we forget about L1 when teaching English? *Ankara Üniversitesi Tömer Dil Dergisi, 167*(2), 5-14.

Zhouyuan, Y. (2016). The analysis about the factors of silence in college English classroom. *Studies in Literature and Language, 12*(5), 105-110.

Critical Thinking Attitude and Some Other Variables in Predicting Students' Democratic Attitudes

Cüneyit Akar [1], Mehmet Kara[2]
[1]Usak University
[2]Diyarbakır Provincial Directorate of National Education, Turkey

To cite this article:

Critical Thinking Attitude and Some Other Variables in Predicting Students' Democratic Attitudes[*]

Cüneyit Akar [1†], Mehmet Kara[2]
[1]Usak University
[2]Diyarbakır Provincial Directorate of National Education, Turkey

Abstract

This research was conducted to determine the effects of the attitudes towards critical thinking, gender, socioeconomic factors (education level of parents, income level, place of residence) as well as academic success, reading and television viewing habits in predicting the democratic attitudes of the fourth grade students. The research was carried out with the participation of 1066 students studying in the 4th grades of primary schools in the central districts of Diyarbakır in the 2017-2018 academic year. The study is descriptive and has a relational screening model. The "Democratic Attitude Scale" (DAS) developed by Erbil and Kocabaş (2017) was employed to determine the students' democratic attitudes and the "Critical Thinking Attitude Scale" (CTAS) developed by Akar-Vural (2005) was used to measure their critical thinking attitudes. A Personal Information Form was used to get information about the students. As a result of the research, democratic attitudes of the students were found "very good" and their attitudes towards critical thinking were "good". According to the results of the regression analysis, it was determined that all the independent variables of the study predicted the students' democratic attitudes at the level of 17%. It was found that critical thinking attitude scores (2.7%), socioeconomic factors (6.5%), and academic success, television viewing and reading habits (7.5%) significantly predicted the students' democratic attitudes. Considering the regression coefficients, it was seen that gender, maternal education level and reading habits did not contribute significantly to predicting democratic attitudes.

Keywords: *Democratic attitude, Critical thinking attitude, socioeconomic factors, academic success, reading habits, television viewing time*

Introduction

As a form of management, democracy is the system that relies on the legislative, executive, and judicial independence, has freedom of press and expression, and operates balance and control mechanisms. In terms of the relationship between individuals, democracy is a concept that includes a set of values and a way of life that regulates how individuals should behave each other (Doğan, 2017; Tezcan, 1994; Gözütok, 1995). In this sense, democracy is more than a form of government (Dewey, 1996). As for democratic attitude, it means that individuals have a positive world view of both dimensions of democracy. In order for democracy to be settled in society as a form of management and lifestyle, it is important to bring values such as justice, equality, freedom, respect to human beings, life and nature, tolerance, and responsibility to citizens.

The main functions of educational institutions include bringing democratic life and democratic values to society (Davis, 2010). The Turkish education system can be said to prioritize this issue. Indeed, the Ministry of National Education (MoNE) clearly put the emphasis on democracy in the Basic Objectives of the Law of National Education No. 1739 as well as in the Basic Principles of National Education (MoNE, 1973).

Citizens' being equipped with democratic attitudes and behaviors starts with basic education. All kinds of knowledge and experience about democracy that could be provided to the children through their schools and environments will shape their social lifestyles in the future. At this point, primary school is of great importance in the formation and internalization of these values (Yeşil, 2003). To help students gain democratic attitudes and behaviors at the primary school level, lessons such as life sciences and social studies are essential. The

---

[*] This study was produced from the thesis titled "Investigation of the relationship between the critical thinking dispositions and democratic attitudes of primary school 4th grade students" by Mehmet Kara under the supervision of Cüneyit Akar.

[†] Corresponding Author: *Cüneyit Akar, akarcuneyt@gmail.com*

constructivist approach, which is in line with the democratic approach, has been taken as the basis of the programs renewed in 2005 (MEB, 2005; 2015). When considered from this point of view, it is seen that more than half of the basic skills, values and learning objectives identified in the social studies program are related to democracy (Tonga, Keçe & Kılıçoğlu, 2013; Polat & Güler, 2020). Besides, it has been found that more than half of the concepts in the 4th-grade Social Studies Teacher Manual are related to democracy (Akar & Doğan, 2016). Some studies conducted in Turkey found that primary school students (Aslan, 2018; Tiryaki, 2018) and secondary school students (Doganay & Yellow, 2004; Ural & Scott, 2011) have a high level of democratic attitudes. These findings point out the importance of gaining democratic attitude and behavior. Undoubtedly, the programs are not sufficient alone in acquiring attitudes and values. Family and teachers' attitude is especially effective in gaining values (Subba, 2014; Bradshaw, 2014).

One of the essential characteristics of democracy is that it allows different thoughts, attitudes and information sources to coexist. Democracy also offers an atmosphere where many ideas and ideologies compete with each other. With the mass media becoming widespread, this competitive environment can bring enormous intellectual wealth to individuals. However, it should be taken into consideration that this situation also comes with some disadvantages that must be handled. The widespread use of mass media also poses serious risks and threats, especially for children and young people. It is seen that this platform is also used as a propaganda center for malicious ideas and thoughts. The purpose of educational institutions is to equip their citizens with essential instruments to know how to benefit from these means and to protect themselves from risks and dangers. At this point, critical thinking, one of the basic intellectual skills that will support a democratic attitude, is needed.

At this stage, the relationship between critical thinking, one of the independent variables of the research, and the democratic attitude, the dependent variable, was tried to be based on theoretical explanations and field studies in the light of the information obtained as a result of the literature review.

Relationship between Democratic Attitude and Critical Thinking

Critical thinking has sub-dimensions of skill, tendency, and habit (Gündoğdu, 2009). Critical thinking as a skill can be defined as "the individual's thinking rationally and thoroughly when deciding what s/he is doing or what s/he believes" (Ennis, 1996). Skill is mostly seen as the mental and analytical thinking activity of individuals, whereas tendency/attitude is defined as a consistent internal motivation in relation to problems and decision making by using their thoughts (Good, 2011). Tendency turns into behaviors and attitudes and it is a feature that affects and directs individuals' lives (McCarthy, 1992). Ennis (1996) defined critical thinking disposition as a tendency to critically thinking. People who are inclined to think critically have certain characteristics such as looking for the clear statement of the thesis or problem, trying to be well informed, using reliable sources, giving references to the sources used, considering the situation completely, trying to stick to the main point, keeping the main problem in mind, searching for options, and being open-minded (Ennis, 1996). Although having a critical thinking disposition/attitude does not mean that individuals use critical thinking skills effectively in their daily lives, many studies have found that there is a significant relationship between critical thinking disposition/attitude and critical thinking skill (Facione, 2000; Sulaiman, Rahman & Dzulkifli, 2009).

It is seen that the development of critical thinking skills in students is given importance with renovated programs in Turkey (MoNE, 2005). It is possible to say that the studies on this subject vary. Some studies in Turkey found students' critical thinking skills higher (Demir, 2006; Kalkan, 2008; Karabacak, 2011; Yıldız, 2011; Aslan, 2018), while some others found critical thinking levels low or moderate (Özdemir, 2005; Gülveren, 2007; Akar, 2007; Görücü, 2014; Akar & Kara, 2016). Although the number of studies on this subject has increased in recent years, it is difficult to say that these are sufficient. It is expected that such an important skill should be included in academic studies more because critical thinking is one of the most important requirements for success in a democratic society (Siegel, 1988; Yamaguchi & Maguth, 2005; Doğanay, Akbulut, Taş & Erden, 2007; Cırık, 2008; Boone, 2008; Edwards, 2010).

According to Smith (2009), the main purpose of democratic education is not only to improve students' cognitive abilities but also to help them have democratic and critical thinking attitudes. This is because democratic attitude requires individuals to have characteristics such as searching for information, critical thinking, questioning, and taking responsibility. In democratic environments, it is possible to encounter various situations such as participation in decisions, making individual choices with free will, having or developing a unique world view, and making political, social, educational, or commercial choices among different options. Under these circumstances, individuals need to have a critical and questioning perspective in order to make the most appropriate decision by themselves. According to Tung and Chang (2009), individuals can only have competitive skills in a global society with good critical thinking skills. Likewise, Tsui (1999) underlined that developing critical thinking in students is of importance in preserving the desired democratic society and increasing the workforce in an increasingly complex world. This enables individuals to evaluate their mindset, personal decisions, and actions (Hawkins, 2006). On the other hand, to be a good critical thinker, it is necessary

to think freely, express their feelings and discuss what they think, which requires a democratic environment (Celep, 1995; Yeşil, 2003). In this sense, the development of critical thinking without democracy is out of the question and vice versa (Akar, 2017).

Apart from the studies that theoretically state that there is a relationship between critical thinking and democratic attitude, there are also studies that test the relationship between these variables. In a study by Acun, Demir, and Göz (2010), a positive correlation (p = 0.31) was found between pre-service teachers' critical thinking skills and citizenship behaviors. Uluçınar (2012) found that pre-service teachers' critical thinking disposition predicted their democratic attitudes at the level of 18%. Likewise, in a study conducted on prospective teachers, it was pointed out that critical thinking tendencies of prospective teachers predicted their multiculturalism tendencies at the level of 25% (Akar, 2017). Tubarik and Gün (2016), who discussed the democratic classroom environment as an independent variable, concluded that perceiving the classroom environment democratically affected students' critical thinking disposition at the level of 7%. In the study conducted by Aslan (2018), he concluded that the critical thinking skills of elementary school 4th-grade students account for 11% of their democratic attitudes.

These findings point out that there is a relationship between critical thinking and democratic attitude and similar features. However, it is seen that the findings of the studies carried out especially in primary school are limited. Based on the aforementioned reasons, this study aims to determine to what extent students' critical thinking attitudes predict their democratic attitudes on a different sample group. It is of great importance to determine the relationships between such variables with a wide range of studies in Turkey, a country having different geographical and socio-economic characteristics, and having a population of 83 million citizens consisting of young people in particular.

The Relationship between Democratic Attitude and Socio-Economic Factors and Some Personal Variables

A large number of variables that affect the democratic attitude can be listed. However, it is not possible to handle all variables in a single study at once. For this reason, along with the critical thinking attitude, the variables such as gender, socio-economic characteristics (parents' education levels, household income level, place of residence), academic success, reading habits, and television watching time were discussed as independent variables. The reasons for choosing these variables as predictive independent variables of the study were tried to be grounded on in the light of the findings obtained from the field studies.

Gender

It has been observed that the gender variable was included in all studies on democratic attitude. Some studies have found that democratic attitudes vary by gender. As a result of the literature review, the democratic attitudes of female students were found higher than that of male students in the studies conducted in primary school (Aslan, 2018), secondary school (Sağlam, 2000; Kaldırım, 2003; Gürbüz, 2006; Ural & Sağlam, 2011; Şimşek, 2011; Yüksel, Bağcı & Vatansever, 2013; Kuş & Çetin, 2014; Öztaşkın & İçen, 2015) and high school (Doğanay & Sarı, 2004; Ulusoy, 2007; Diker, 2012). Similar results were also reached in the studies carried out with the participation of higher age groups. Likewise, females were found to have higher democratic attitudes compared to males in the studies on prostective teachers studying at various departments (Saracaloğlu, Evin & Varol, 2004; Turan & Taşpınar 2004; Genç & Kalafat, 2007; Gömleksiz & Kan, 2008; Akın & Özdemir, 2009; Aydemir & Aksoy, 2010; Merey, Kaymakcı & Kılıçoğlu, 2011; Ekici, 2014), teachers (Büyükkaragöz & Kesici, 1996; Taçman, 2006; Üstün, 2011; Kurnaz, 2011; Karatekin, Merey & Kuş, 2013; Özdaş, Ekinci & Bindak, 2014; Arslan & Çalmaşur, 2017) and school administrators (Ozan, Türkoğlu & Şener, 2010; Arslan & Çalmaşur, 2018). In contrast to these findings, some studies did not find any difference by gender. There was no difference in democratic attitudes by gender in the studies on the students in primary (Tiryaki, 2018) and secondary schools (Çırak, 2019), prospective teachers (Dilekmen, 2000; Ömerustaoğlu, 2004; Karahan, Sardoğan, Özkamalı, Dicle & Dicle, 2006; Ercoşkun & Nalçacı, 2008; Ektem & Sünbül, 2011; Nazıroğlu & Çetin, 2014; Elkatmış & Toptaş, 2015; Akar, İnel & Yalçıntaş, 2017) and teachers (Yalçın, 2007; Şahin, 2008; Yılmaz, 2011; Karatekin, Merey & Kuş, 2013; Kontaş, Selçuk & Polat, 2016). In another study conducted with the participation of prospective teachers, Çermik (2013), however, found that the difference favored the men.

Findings obtained from the research results show that gender factor can be a predictive variable on the democratic attitude in all age groups. Based on these reasons, gender variable was included in the research.

Socioeconomic Factors

Parents' educational level, income level and the place of residence were included in the study as socioeconomic factors. In a research within the scope of the Project on Monitoring and Evaluation of Academic Skills (ABIDE) conducted by MoNE (2019) in 2018, it was determined that the mother's education level is an important

variable affecting the fourth-grade students' overall success. There are also studies suggesting that the democratic attitudes of students have increased with the rise in the mother's education level. In the study conducted by Tiryaki (2018) on the 3rd-grade elementary school students, it was concluded that the democratic attitudes of the students differ according to the mother's education level. Likewise, some other studies in the context of elementary school level (Ural & Sağlam, 2011; Kuş & Çetin, 2014), it was found that the mother's education level affects the democratic attitudes of elementary school students. Similar results were obtained in the studies conducted on prospective teachers studying in various departments. Nazıroğlu and Çetin (2014), Gömleksiz and Çetintaş (2011) and Evcimik (2009) concluded that the mother's education level affects the democratic attitudes of the students. In contrast to these findings, some studies reported that no significant difference was foundaccording to the mother's education level. In some studies conducted at elementary school level (Kardaş, 2013; Sağlam, 2000; Aycan & Çalık, 2003; Şimşek, 2011), it was concluded that the democratic attitudes of elementary school students do not differ significantly according to the mother's education level. Some other studies conducted on university students (Kılıç, Ercoşkun & Nalçacı, 2004; Bulut, 2006; Genç & Kalafat, 2008; Aydemir & Aksoy, 2010; Merey, Kaymakcı & Kılıçoğlu, 2011; Yazıcı, 2011; Ekici, 2014; Elkatmış & Toptaş, 2015) also found that the democratic attitudes of prospective teachers do not differ significantly according to the mother's education level.

There are studies indicating the democratic attitudes differ positively according to the father's education level. In a study conducted by Kuş and Çetin (2014) on elementary school students, it was concluded that the father's education level influences the students' perceptions of democracy. Also in some studies (Özbek, 2004; Saracaloğlu, Evin & Varol, 2004; Bulut, 2006; Gömleksiz & Kan, 2008; Evcimik, 2009; Yazıcı, 2011) conducted on prospective teachersstudying in various departments, it was concluded that the father's education level influences the democratic attitudes of students. In contrast to these findings, no significant difference was found according to the father's education level. In the context of elementary school level, there are studies suggesting that democratic attitudes do not differ according to the father's education level (Ural & Sağlam, 2011; Kardaş, 2013).

In the literature, some studies pointed out that the household income level affects the democratic attitudes of the students. In some studies conducted on elementary school students (Kuş & Çetin, 2014; Şimşek, 2011), it was proved that the household income level positively affects the students' perceptions of democracy. Likewise, the studies carrid out by Güven (2005) on secondary school students, and by Evcimik (2009) on education faculty students concluded that the household monthly income has a positive impact on the students' democratic attitudes. On the other hand, in a study conducted in elementary schools(Konak, 2012), no significant difference was found in the democratic attitudes of students according to their household income levels. Furthermore, some other studies conducted on prospective teachers (Kılıç, Ercoşkun & Nalçacı, 2004; Ömerustaoğlu, 2004; Gömleksiz & Kan, 2008; Akın & Özdemir, 2009; Merey, Kaymakcı & Kılıçoğlu, 2011; Elkatmış & Toptaş 2015) found that household monthly income do not have a significant effect on students' democratic attitudes. Apart from these studies, there are various studies suggesting that income level and democratic attitude are inversely related; democratic attitudes decrease as the income level increases (Karatekin, Merey & Kuş, 2013; Gömleksiz & Çetintaş, 2011).

Some studies suggest that the democratic attitudes differ according to the place of residence. Güven (2005) conducted a study on secondary school studentsand found that in the history lesson, the democracy understanding of the students who live in urban areas was found to be higher. In a study conducted by Arslan and Çalmaşur (2017) on teachers, they concluded that the democratic attitude scores of elementary and secondary school teachers differ significantly according to the variable of the place of the residence. In the same study, the democratic attitude scores of the teachers working in districts were found to be higher than those who work in the province. On the other hand, in a study conducted by Ercoşkun ve Nalçacı (2008) on prospective teachers, it was concluded that the place of residence has no significant effect on their democratic attitudes.

The research findings indicate that variables such as parents' education level, income level and the place of residence may have a predictive effect in the democratic attitudes.

Academic Success, Reading and Television Viewing Habits

No studies were found to analyze the relationship between students' democratic attitudes and their academic success. However, there are studies that found democratic family (Alpoğuz & Şahin, 2014; Karadağ, 2007; Kaya, Bozaslan & Genç, 2012; Karakaş, 2008; Baltacı, 2010) and democratic classroom environment (Tubarik & Gün, 2016) have a positive effect on students' overall success. Families and teachers' having democratic attitudes is extremely crucial for the children to adopt democracy and develop a positive attitude. The students

having this attitude also acquire the awareness of the the fact that the freedoms provided by democracy also impose certain responsibilities on them. Therefore, it is expected from students who have adopted this attitude to be successful in academic tasks, which is one of their prior responsibilities.

During the literature review, no studies were found on the relationship between the reading habits and the democratic attitude. However, a study suggesting that there is a relationship between empathic sensitivity, which is one of the characteristics of democratic attitude, and reading habits (Zorlu, Öneren, & Çiftçi, 2019). In another study on teachers, it was concluded that reading habits affect the students' happiness (Altuntaş & Genç, 2018). In a study on aggressive behaviors that can be considered as the opposite of the democratic attitude, it was determined that the students with the highest aggression scores are the ones who read books the least (Sili, 2012). According to the ABİDE report (2018), it was found that the students who display the least bullying behavior read more books than the ones displaying bullying behavior the most. Since reading books improves the ability of thinking and empathizing in the individual, students who stay away from reading books have relatively uncontrolled attitudes (Sili, 2012).

Although there were no studies on the relationship between the television viewing time and democratic attitude, there are studies suggesting that spending a lot of time watching TV and using mobile phones or tablets have negative effects on students' academic success (Akar & Kara, 2017). According to a study on elementary students by MoNE (2008), the biggest harm of television watching was determined to be eliminating the opportunity of having a conversation in the family. It can be suggested that the decreased communication in the family will create a negative impact on the emergence and development of democratic values in family members.

These findings suggest that academic success, reading and television viewing habits may be related to democratic attitudes. In the light of this information, these variables have been included in the scope of the research.

As can be seen from the literature findings, sampling of studies discussing the democratic attitude and critical thinking in Turkey generally consist of prospective teachers, secondary and high school students. It is seen that very few studies were done in the context of elementary school level. It should be noted that the studies are mostly based on the difference between averages, and there is a limited number of studies that aims to explain or predict the variance in democratic attitudes. Therefore, it can be said that such studies are extremely insufficient and further studies are required. From this point of view, this study firstly aims to determine the students' democratic attitude levels and critical thinking attitude levels, and then to determine to what extent the students' critical thinking attitudes and some other variables can predict their democratic attitudes. Specifically, this study sought to answers the following research questions:

1. What is the level of democratic and critical thinking attitudes of the fourth-grade elementary school students?
2. To what extent the following variables predict the fourth-grade students' democratic attitudes;
   a) Critical thinking attitudes,
   b) Gender,
   c) Mother's education level,
   d) Father's education level,
   e) Household monthly income,
   f) Place of residence,
   g) Academic success,
   h) Number of the books read monthly and
   i) Daily television viewing time.

Method

Research Model

This study, which aims to examine the fourth-grade students' democratic attitudes in terms of critical thinking attitudes and other (demographic) variables, was carried out using the relational screening model, which is one of the quantitative research methods. Relational screening model "aims to determine the existence and degree of co-variation between two or more variables" (Karasar, 2005: 81). The dependent variable of the research is the students' democratic attitudes. The independent variables of the research are a) critical thinking attitudes, b)

gender, c) socioeconomic factors (parental educational background, income, place of residence), d) academic success, monthly reading habits and television viewing time.

Population and Sampling

The population of this research consists of 22.478 fourth-grade students in the primary schools in the central districts of Diyarbakır in the 2017-2018 academic year. The sampling of the research was selected using the stratified sampling method based on the socio-economic levels of schools. Prior to the sampling, the schools were stratified into sub-population groups as lower, middle, and high socioeconomic groups with the help of the provincial directorate of national education. Afterwards, 25 schools, which can be reached by the researchers easily, were selected as the sampling. These 25 schools consisted of 8 low socioeconomic, 9 middle socioeconomic, and 8 high socioeconomic groups for the sake of accurate distribution. As a result, the sampling of the research consisted of 1066 students studying in these 25 schools. After the schools were determined, no specific student selection was made in the schools; all fourth-graders within the scope of the research were included in the sampling. Information about the participants is given in Table 1.

Table 1. Descriptive Information about the Participants

	N	%		N	%
Mother's Education Level			***Father's Education Level***		
Illiterate	179	16.8	Illiterate	62	5.8
Primary school	322	30.2	Primary school	187	17.5
Secondary school	188	17.6	Secondary school	234	22.0
High school	227	21.3	High school	315	29.5
Undergraduate/Graduate	150	14.1	Undergraduate/Graduate	268	25.1
Household Income			***Place of Residence***		
Minimum Wage	277	26.0	Village	32	3.0
1601-2000 TL	343	32.2	Town/District	126	11.8
2001-3000 TL	211	19.8	Province	908	85.2
3000 TL and above	235	22.0	***Number of Book Read Monthly***		
Academic Success			0-5 books	44	4.1
0-40	11	1,0	6-10 books	168	15.8
0-60	43	4,0	11-15 books	246	23.1
60-80	194	18,2	16-20 books	248	23.3
80-100	818	76,7	21 books and more	360	33.8
Daily TV Viewing Time			***Gender***		
Never watch	45	4.2	Girl	531	48,8
1-2 hours	457	71.0	Boy	535	50,02
3-4 hours	206	19.3			
5 hours and above	58	5.4			
Total	**1066**	**100**		**1066**	**100**

Data Collection Tools

A personal information questionnaire, democratic attitude scale and critical thinking scale were used in the study.

Personal Information Questionnaire

A Personal Information Questionnaire developed by the researchers was used to collect personal information of the students participating in the research. In the personal information questionnaire, the gender of the students, the educational background of the parents, the place where the family lives (village, town/county, province), academic achievements (the average of grades of Turkish, Mathematics, Science and Social Studies lessons), the number of books read per month, daily television viewing times (in hours).

Democratic Attitude Scale

The "Democratic Attitude Scale (DAS)" developed by Erbil and Kocabaş (2017) was used to determine the democratic attitude levels of the students constituting the study group. Composed of a total of 10 items, the DAS is scored in three-point Likert scale as 'no' (1), 'sometimes' (2), 'yes' (3). The total scores that can be obtained

from the scale vary between 10 and 30. 7 of the 10 items are positive items and 3 of them are negative items. As a result of the exploratory factor analysis, it was determined that 10 items grouped under a single factor and the factor loads of the scale items ranged between 0.41 and 0.66. As a result of confirmatory factor analysis, the fit indices were determined to fit well ($x2$ = 42.52, $x2$ / sd = 1.214, RMSEA = 0.03, SRMR = 0.035, GFI = 0.98, AGFI 0.96, CFI = 0.99 and NNFI = 0.99). The total score correlations of the DAS were found between 0.24 and 0.50. Cronbach Alpha reliability value of the scale was found to be .76. In the light of the data obtained in this study, the Cronbach Alpha reliability coefficient of the scale was found .76, suggesting that the scale is reliable.

Critical Thinking Attitude Scale

"Critical Thinking Attitude Scale" was used to determine the critical thinking attitudes of the students. The scale was firstly developed by Akar Vural (2005) to measure the effect of two different programs on the 9th-grade students' attitudes towards critical thinking. The scale, which has 20 items, consists of seven dimensions (open-mindedness, problem-solving, fatalism, paranormal beliefs, dogmatism, conservatism, and discussion). 12 of 20 items are negative and 8 of them are positive. The scale is scored in the five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The correlations between the sub-dimensions of the scale and the total score were found to range between .43 and .67.

In this study, 3 of the 20 items in the scale (items 2, 13 and 14) were not found appropriate by the Provincial Directorate of National Education. Therefore, the scale was implemented based on 17 items. The lowest score that can be obtained from the scale is 17, and the highest score is 85. Cronbach Alpha reliability value of the scale was found .76 after it was revised. The scale was evaluated on a single dimension, and analyzes and interpretation were made accordingly.

Implementation, Analysis and Interpretation of Data

The scales were implemented by the researcher himself. Prior to the implementation, necessary permissions were obtained from the Directorate of National Education. In addition, school administrators and classroom teachers supported the researcher during the implementation process and the students were asked to fill in the form on a volunteer basis.

The collected data were transformed into computerized data and analyzed using a statistical software package. The data obtained before the data analysis were examined and revised in terms of missing or incorrect values, outliers, normality, and multivariate outliers; 45 pieces of data were not suitable and they were eliminated. After these procedures, the remaining data of 1066 students were analyzed. Prior to the analyzes, the negative items in the scales were reversed.

During the analysis, the average and standard deviation values of the students' democratic and critical thinking attitudes were calculated. The formula "number of choices - 1 / number of choices" was used to calculate the score ranges, and the result was 0.80 for the critical thinking attitude scale (5-1/5) and 0.66 for the democratic attitude scale (3 - 1/3). Besides, the standard values obtained were converted into percentile scores.

Since gender and the place of residence are categorical variables, they have been converted into artificial variables before the correlation analysis. For this purpose, gender was numbered as 1 (girl) and 2 (boy); and place of residence was numbered as 1 (village), 2 (district), and 3 (province). Interpretations were made based on this classification. Pearson moment correlation coefficient technique was used to determine the relationship between democratic attitude and critical thinking attitude, and Spearman Rho coefficients were calculated to explain the relationship between democratic attitudes and other rank-order variables such as parents' educational background. The point-biserial correlation technique was used for the gender variable. After the correlation analysis, multiple regression analyzes were performed.

Regression analysis can determine how much of the variance in the dependent variable can be explained or predicted by variables, as well as to what extent these variables are effective (Büyüköztürk, 2004). In this study, the available data were subjected to multiple hierarchical regression analysis. The analyzes were carried out in 4 stages. In the first stage, the gender variable was included in the regression model. In the second stage, mother's and father's education level, place of residence, and household monthly income level were included in the model. In the third stage, books read monthly, academic success and television watching time were included in the model. In the fourth and last stage, the scores related to the students' critical thinking attitude scale were included in the model.

Findings

The findings related to the average, standard deviation and percentile values of the fourth-graders' democratic and critical thinking attitudes are shown in Table 2.

Table 2. Average, standard deviation and percentile values of the scores that students get from the scales

Scales	N	$\overline{X}$	Ss	%
Democratic Attitude	1066	2.67	0.38	89
Critical Thinking Attitude	1066	3.67	0.61	73,4

The average score of the students on the Democratic Attitude Scale is 2.67 (89%) on the 3-point Likert type scale, and the average score they got on the Critical Thinking Scale is 3.67 (73.4%) on the 5-point Likert type scale. According to this, it can be said that the democratic attitude levels of the fourth-grade students who participated in the study are "very good" and their critical thinking levels are "good". The 'very good' level of democratic attitude scores can be interpreted in a way that the level reached is "sufficient". Critical thinking scores of the students were determined as "good". Although this statement may seem positive at first glance, there is an important difference between students' democratic attitudes and their critical thinking attitudes considering the percentages. The scores related to critical thinking attitude are very low compared to the democratic attitudes. In addition, given the importance and necessity of critical thinking, these results do not make it possible to suggest students' critical thinking attitude levels are "sufficient".

Correlations between the variables predicting students' democratic attitude levels

The correlations between the predictive variables of the research and the predicted variable were calculated before proceeding to the regression analysis, and the correlation coefficients are listed in Table 3.

Table 3. Correlations between the variables

	1	2	3	4	5	6	7	8	9	10
1- Democratic attitude	1									
2- Critical thinking attitude	,22**	1,00								
3- Mother's education level	,13**	,08**	1,00							
4- Father's education level	,20**	,11**	,52**	1,00						
5- Household monthly income	,01	,00	,31**	,29**	1,00					
6- Place of residence	,18**	,11**	,10**	,15**	,08**	1,00				
7- Gender	-,04	-,04	,00	,02	,03	-,06	1,00			
8- Academic success	,24**	,07*	,27**	,31**	,21**	,17**	-,01	1,00		
9- Number of books read monthly	,15**	,08**	,21**	,22**	,18**	,07*	-,03	,28**	1,00	
10- Daily TV viewing time	-,24**	-,08**	-,10**	-,14**	-,06	-,02	,11**	-,11**	-,12**	1,00

**. 0.01 level of significance; *. 0.05 level of significance

Having analyzed Table 3, it is seen that there are low-level significant relationships between students' democratic attitude scores and critical thinking attitude scores and some other variables. The level of the relationship between democratic attitude and critical thinking attitude was calculated as r=,22. The critical thinking attitude variable was the third variable that had the highest relationship with democratic attitude following television viewing time and academic success.

It can be seen that the daily television viewing time (r=-24) and academic success (r=,24) have the highest relationship with the students' democratic attitude scores. As the students' television viewing time increases, their democratic attitude scores decrease; however, as their academic success increases, their democratic attitude scores increase. It is seen that there is a positive relationship between the number of books read monthly and democratic attitude (r=,15). As for the gender variable, no relationship was found between gender and democratic attitudes.

It is also noteworthy that there are significant relationships between democratic attitudes and the variables that show the socioeconomic status of students. The highest relationship among these is with the father's education level (r =, 20). Then comes the place of residence (r =, 18). The residential area where the students' families live is numbered from the village to the province. For this reason, it can be interpreted that the democratic attitudes increase as the place of residence shifts from village to the province. To put it in simple terms, as the people live in cities more, their democratic attitudes increase. Another socioeconomic variable is the mother's education level. As the mother's educational status increases, an increase is observed in the democratic attitudes of the students (r =, 13). In terms of the household income, no significant relationship was found between income and democratic attitudes.

In brief, the students' democratic and critical thinking attitudes are positively related to the parents' educational status, the students' living in urban areas rather than rural areas, their academic success and reading habits. On the other hand, as the students' television viewing time increases, their democratic attitudes decrease. No relationship was found between the students' democratic attitudes and the variables of gender and household income.

Regression Analysis regarding Students' Democratic Attitude Scores

The data collected in this study were subjected to hierarchical regression analysis. At this step, independent variables were included in the regression model in turn and each variable set's level of predicting the democratic attitude scores was calculated. The results are shown in Table 4. Afterwards, the regression coefficients for the data were calculated in order to understand to what extent the variables subjected to regression analysis were effective in predicting the scores. The results are shown in Table 5.

Table 4. Hierarchical regression model for the total scores of students' democratic attitude scale

| R | R Square | Adjusted R Square | Std. Error of theEstimate | ChangeStatistics | | | | | p |
				R SquareChange	F Change	df1	df2	Sig. F Change	
1	,037[a]	,001	,000	3,82	,001	1,427	1	1064	,233
2	,257[b]	,066	,062	3,70	,065	18,405	4	1060	,000
3	,374[c]	,140	,133	3,55	,074	30,237	3	1057	,000
4	,408[d]	,167	,160	3,50	,027	33,844	1	1056	,000

1st Predictor: (Constant), Gender
2nd Predictor: (Constant), Gender, Mother's Education Level, Place of Residence, Household Monthly Income, Father's Education Level
3rd Predictor: (Constant), Gender, Mother's Education Level, Place of Residence, Household Monthly Income, Father's Education Level, Books Read Monthly, Academic Success Scores, Daily TV Viewing Time
4th Predictor: (Constant), Gender, Mother's Education Level, Place of Residence, Household Monthly Income, Father's Education Level, Books Read Monthly, Academic Success Scores, Critical Thinking Attitude Scores

In regression analysis, gender was the first variable included in the model. The gender variable alone explains 0.1% of the students' democratic attitudes ($R^2 =, 001$). Based on this result, p value is not significant (p = 0.23) considering the change in F value. It is seen that gender does not have a significant effect on predicting students' democratic attitudes.

Socioeconomic factors (mother's education level, father's education level, the residence of place and household income) were included in the model as the second set of variables along with gender. It is seen that socioeconomic factors predicted students' democratic attitudes at 6.6% ($R^2 =, 066$). When the gender effect excluded, socioeconomic variables alone predict the democratic attitude variance at the level of 6.5%. P-value seems to be significant (p = 0.00) considering the change in F value, suggesting that socioeconomic variables significantly predict students' democratic attitudes.

The third set of variables included in the model was the number of books read monthly, academic success, and television viewing times. It is seen that the number of books read monthly and academic success and TV viewing time, along with the previous variables, predicted students' democratic attitude scores at the level of 14% ($R^2 =, 14$). The power of the set of variables to predict the students' democratic attitude scores alone was determined as 7.4% ($R^2 =, 074$). Considering the change in F value, it is seen that p value is significant (p = 0.00), suggesting that the number of books read monthly, academic success and TV viewing time variables predict students' democratic attitudes significantly.

Students's scores regarding the critical thinking attitude scale were included in the model the last. It was found that the critical thinking scores along with the previous variables predicted approximately 16.7% of the variance in the students' democratic attitude scores. Critical thinking attitude scores alone predict the variance at about

2.7% ($R^2 =, 027$). It is seen that p value is significant considering the change in F value (p = 0.00), suggesting that critical thinking attitude scores significantly predict students' democratic attitude scores.

Considering all variables together, as a result, it can be suggested that all variables predicted 17% of the variance in students' democratic attitude scores. The most predictive variable of students' democratic attitude levels is the academic success, reading habit and TV viewing time with 7.4%. Then comes mother and father's education level, income status and place of residence (6.5%). Students' critical thinking attitudes predict their democratic attitude level at the level of 2.7%. Gender appears to have no predictive effect.

Regression coefficients relating to the data are listed in Table 5 to understand to what extent the variables subjected to regression analysis are effective.

Table 5. Regression coefficients of students'total scores on democratic attitude scale

Model	Unstandardized Coefficients		Standardized Coefficients	T	P (Sig.)
	B	Std.Error	Beta		
(Constant)	18,451	1,172		15,745	,000
Gender	,005	,217	,001	,024	,981
Mother's Education Level	,038	,098	,013	,386	,699
Father's Education Level	,291	,109	,092	2,670	,008
Household Monthly Income	-,297	,105	-,086	-2,824	,005
Place of Residence	,960	,242	,115	3,972	,000
Academic Success	,236	,045	,162	5,252	,000
Books Read Monthly	,161	,095	,051	1,705	,089
Daily TV Viewing Time	-1,150	,177	-,187	-6,503	,000
Critical Thinking Attitude	,061	,010	,166	5,818	,000

Considering the regression coefficients, the variable that affects the students' democratic attitudes the most is the daily TV viewing time (β = -, 187; p = 0.00). The second highest predictive variable appears to be the critical thinking attitude (β =, 166; p = 0.00). Then comes the students' academic achievements (β =, 162; p = 0,00), the place of residence (β =, 115; p = 0,00), father's education level (β =, 092; p = 0,00), and household monthly income (β=-,086; p=0,05); these variables were found to predict democratic attitudes significantly.

Increasing television viewing time affects students' democratic attitudes negatively. On the other hand, it is seen that the variable that affects the democratic attitude most positively is the critical thinking attitude. It is seen that the variable that affects democratic attitudes most positively after the critical thinking attitude is academic success.

Regression coefficients show that the father rather than the mother is effective on students' democratic attitudes. It is seen that the household income affects democratic attitudes negatively. Increased financial status reduces students' democratic attitudes. As the place where the family lives shifts from rural areas to urban areas, the democratic attitudes of the students are affected more positively, suggesting that the democratic attitudes of the children whose families reside in urban areas rise.

It is seen that gender, the number of books read monthly, and the mother's education level that is significantly correlated with the students' democratic attitudes have no significant effects on the students' democratic attitudes.

Conclusion and Discussion

Conclusion and Discussion on Primary School Students' Democratic and Critical Thinking Attitude Levels

Research results suggest that the primary school fourth-grade students' democratic attitude levels are "very good". The results of the current study are the same as the results of the study conducted by Aslan (2018) using the same scale on the fourth-graders. Both studies found the average of the students' democratic attitudes as 2.67. Tiryaki (2018) concluded that primary school third graders' democratic attitude levels were high. It is seen that some studies on primary and secondary school students found similar results (Doğanay and Sarı, 2004; Ural and Sağlam, 2011). Similar results were reported in the studies on teachers and prospective teachers (Yurtseven, 2003; Özdaş, Ekinci & Bindak, 2014; Elkatmış & Toptaş, 2015; Kesici, Pesen & Oral, 2017). Yuksel, Bağcılar and Vatansever (2013) conducted a study on the last graders in grade schools in seven different regions of

Turkey, and Diyarbakır was found to be the province with the highest "tolerance" value. Our findings coincide with the results of these studies.

As a result of the research, students' critical thinking attitudes were found "good" (73%). This result is quite low compared to students' democratic attitude scores. Besides, given the importance and necessity of critical thinking, it is not possible to say that this result reflects the desired level.

Having examined the studies conducted in primary, secondary and high schools in Turkey, it is seen that some studies found students' critical thinking skills high (Demir, 2006; Kalkan, 2008; Karabacak, 2011; Yıldız, 2011; Aslan, 2018), while some others found students' critical thinking skills low or medium-level (Özdemir, 2005; Gülveren, 2007; Akar, 2007; Görücü, 2014; Akar & Kara, 2016). A similar conclusion was reached in the studies that examine "tendency/attitude". Some studies found students' tendency to think critically at a high or medium-high level (Kartal, 2012; Evin Gencel & Güzel Candan, 2014; Akar, 2017), while other studies found students' tendencies at medium or low levels (Tümkaya, 2011; Demir & Aybek, 2014; Kuvaç & Koç, 2014). The differing results may be due to the structure of the measurement tools and the characteristics of the sampling groups.

Conclusion and Discussion on the Predictive Variables' Level of Predicting Democratic Attitudes

As a result of the research, all of the independent variables predicted about 17% of the variance in students' democratic attitudes. Of the data included in the regression model as 4 blocks, the variables that have the highest predictive power in terms of predicting democratic attitudes were the students' academic success, the number of books read monthly, and TV viewing time. It was concluded that these variables predicted students' democratic attitude scores at the level of 7.4%. The second variable group with the highest predictive power was socioeconomic variables. It was found that socioeconomic variables predicted students' democratic attitude scores at the level of 6.5%. The scores relating to the critical thinking attitude included in the model alone predicted students' democratic attitude scores at the level of 2.7%. When the regression coefficients were analyzed, it was observed that gender, mother's education level and the number of books read monthly had no significant effect on predicting students' democratic attitude scores.

When the regression coefficients were analyzed, the scores relating to the critical thinking attitude included in the model alone significantly predicted students' democratic attitude scores at the level of 2.7%. Similar results were obtained by different studies in the literature. Akar (2017) concluded that prospective teachers' critical thinking attitudes predicted 25% of their multiculturalism perceptions reflecting their democratic values and in a similar study by Uluçınar (2012), it was concluded that prospective teachers' critical thinking tendencies predicted their democratic values by 18%. Acun et al. (2010) found the level of significance of the positive relationship between prospective teachers' critical thinking skills and good citizenship behavior at the level of 0.31. Turabik and Gün (2016) concluded that democratic perception of the classroom environment affected students' critical thinking disposition at the level of 7%. In the study conducted by Aslan (2018) on the fourth graders in primary school, students' critical thinking attitudes accounted for 11% of their democratic attitudes. In the current study, students' critical thinking attitude scores account for 2.7% of the variance in their democratic attitude scores; the level of explanation was found lower than the results of the studies mentioned above. However, it can be said that the conclusion that critical thinking attitudes affect democratic attitudes was supported by this study as well and students' critical thinking attitudes and behaviors are a factor that affects their democratic attitudes. Students' democratic attitude scores increase as their critical thinking attitude scores increase. Critical thinking is closely related to active citizenship skills and is a prerequisite for raising individuals with democratic skills. Students who can think critically can use these skills more effectively when required (Doğanay, Akbulut, Taş & Erden, 2007). Therefore, in many democratic societies, individuals are expected to think critically and employ this skill in decision-making process to come up with a solution to social problems (Seferoğlu & Akbıyık, 2006). In light of this information, it can be said that the results of this study coincide with the results of various studies in the literature.

Research results show that gender alone does not have a significant effect on predicting democratic attitudes. It is seen that field studies, in this regard, have resulted in two ways. Approximately half of the field studies conducted from primary school to university level found that girls/women have higher democratic attitudes (Sağlam, 2000; Bingöl, 2000; Saracaloğlu, 2001; Kaldırım, 2003; Saracaloğlu, Evin & Varol, 2004; Turan & Taşpınar, 2004; Gürbüz, 2006; Ulusoy, 2007; Genç & Kalafat, 2007; Akın & Özdemir, 2009; Aydemir & Aksoy, 2010; Ural & Sağlam, 2011; Şimşek, 2011; Merey, Kaymakcı & Kılıçoğlu, 2011; Diker, 2012; Yüksel, Bağcı & Vatansever, 2013; Ekici, 2014; Kuş & Çetin, 2014; Öztaşkın & İçen, 2015; Aslan, 2018;). However, other half of the studies found no significant difference by gender (Dilekmen, 2000; Kılıç, Erçoşkun, & Nalçacı, 2004; Ömerustaoğlu, 2004; Karahan, Sardoğan, Özkamalı, Dicle & Dicle, 2006; Bulut, 2006; Ercoşkun & Nalçacı, 2008; Ektem & Sünbül, 2011; 2011; Nazıroğlu & Çetin, 2014; Çetin, 2014; Elkatmış & Toptaş, 2015; Akar, İnel & Yalçıntaş, 2017; Tiryaki, 2018; Çırak, 2018). Only one study (Çermik, 2013) found a significant

different by gender in favor of men. According to the findings of these studies, girls can be said to be more prone to have democratic attitudes; however, this suggestion was not confirmed by the current study.

The findings suggest that the predictive power of the parents' educational background is due to the father's educational status, and the mother's educational status has no significant effect. This finding is similar to the findings of the study conducted by Kuş and Çetin (2014). In addition, there are also other studies suggesting that father's education status affects democratic attitudes (Saracaloğlu, Evin & Varol, 2004; Güven, 2005; Bulut, 2006; Evcimik, 2009; Yazıcı, 2011; Kuş & Çetin, 2014). On the other hand, in many studies carried out in different settings ranging from primary school to university level, mother's education status seems to be more effective (Güven, 2005; Gömleksiz & Kan, 2008; Gömleksiz & Çetintaş, 2011; Ural & Sağlam, 2011; Kardaş, 2013; Nazıroğlu & Çetin, 2014; Aslan, 2018; Tiryaki, 2018). In this study, it was concluded that the democratic attitudes of the students in the sampling group are affected by the father's education status. This situation may be due to the characteristics of the sampling group. Considering the demographic data, it is seen that the educational status of the mothers in the sample group of the research is quite low. Approximately 17% of the mothers in the sampling group are illiterate and about 30% of the mothers are primary school graduates. In the Southeastern Anatolia region of Turkey, where the sampling group is located, girls' schooling rates are the lowest of the overall rates. In addition, the native language of the region is mostly not Turkish and they learn Turkish in schools. Therefore, it can be said that illiterate mothers do not know Turkish either or do not have enough knowledge of Turkish. Father's education level is higher than the mother's. The rate of illiterate fathers is about 6%. The rate of fathers having a high-school degree or above is approximately 55%. The high rate of illiterate mothers, in particular, may lead fathers to be more prominent in children's education.

There are studies in the literature suggesting that income status positively affects students' democratic attitudes. These studies were carried out in primary schools (Şimşek, 2011; Kuş & Çetin, 2014), secondary schools (Güven, 2005) and education faculties of universities (Evcimik, 2009) and they concluded that monthly income status had a positive effect on students' democratic attitudes. In some other studies, the income status was found to have no effect on democratic attitudes. The studies carried out on grading school students (Konak, 2012) and prospective teachers (Kılıç, Ercoşkun & Nalçacı, 2004; Ömerustaoğlu, 2004; Gömleksiz & Kan, 2008; Akın & Özdemir, 2009; Merey, Kaymakcı & Kılıçoğlu, 2011; Elkatmış & Toptaş 2015) concluded that the monthly income status did not cause any significant difference in democratic attitudes.

In this study, it was found that monthly income status had a negative effect on democratic attitudes. Albeit few, there are studies in the literature that have similar results. Gömleksiz and Çetintaş (2011) and Karatekin, Merey and Kuş (2013) concluded that the preservice teachers who have high income have low democratic attitude scores. One of the reasons behind this situation may be that the increase in financial status negatively affects the democratic attitude behaviors of individuals depending on their personality traits. Democracy embodies certain traits such as tolerance, respect, and patience towards others. The individuals standing on their own legs financially may feel the need for others less. In parallel, it may result in behaviors such as decreasing tolerance and respect towards others, and even behaviors such as looking down on others and mocking them. As a matter of fact, there are studies suggesting that peer-bullying observed among children can be demonstrated by students with high financial possibilities (Can, 2007; Pişkin, 2010; Ayık, 2014; Sarı & Demirbağ, 2019). Şimşek (2011) concluded that students with the least democratic attitudes were among the highest income group. In the same study, however, the democratic attitude scores of the group, which is located in a lower step of the highest income group, were found to be significantly higher than the lower groups. These findings suggest that the income status may have a positive effect on democratic attitudes to an extent and after a point, it may have negative effects on democratic attitudes. However, it cannot be said that income status is a factor that directly affects students' democratic attitudes. Financial possibilities, in this regard, should be interpreted by considering the general personality characteristics of the students, their periodical development characteristics, family attitudes and environmental conditions. Therefore, considering these factors, it can be argued that income status can have different effects on different students.

The current study has found that there is an increase in students' democratic attitude scores as the place of residence shifts from village to city. These findings coincide with the results of the study on high school students conducted by Güven (2005). Democratic values are more city-oriented by nature. Therefore, the students whose families live in urban areas are expected to have more democratic values.

As a result of the research, it was concluded that students' democratic attitude scores were affected by academic success. It can be said that one reason for this result is that students who are more successful in their lessons may also be more harmonious in terms of following the rules in school and having good human relations. In the literature, there was no study directly related to the relationship between these two variables. In addition, some

studies that consider the democratic family and school environment as independent variables found that the democratic environment affects student achievement positively (Karadağ, 2007; Kaya, 2012; Alpoğuz & Şahin, 2014; Tubarik & Gün, 2016). Democracy provides students with opportunities to think freely, express what they think, and make decisions and choices freely. However, democracy does not provide unlimited freedom to individuals. It also imposes responsibilities of the consequences of each decision made on individuals. Therefore, democracy provides individuals with freedoms on the one hand and brings a strong sense of responsibility on the other hand. Considering the findings from this point of view, it can be interpreted that students with high democratic attitudes can also be successful in academic tasks based on their sense of responsibility.

Although the reading habit variable shows a significant relationship in the correlation table, the power of the relationship was not found significant considering the regression coefficients. Although this study did not report the expected results, the studies in the literature show that reading habits can also have a positive effect on affective characteristics. A study reported that there is a positive correlation between empathic sensitivity, which is one of the characteristics of democratic attitude, and reading habits (Zorlu, Öneren & Çiftçi, 2019). In a study on aggressive behaviors that can be considered the opposite of democratic attitude, it was found that the students with the highest aggression scores were the ones who read books the least (Sili, 2012). The ABİDE report (2018) also found that the students who displayed the least bullying behavior read more books and those who displayed this behavior read few books. Since reading books improves the ability of thinking and empathy in the individual, students staying away from books have relatively uncontrolled attitudes (Sili, 2012). The fact that a similar result was not reported by this study may have resulted from the characteristics of the sampling group or the characteristics of the measurement instruments used in the study.

According to the results of the research, the variable that predicted the democratic attitudes of the students the most was the daily television viewing time. Democratic attitude scores decrease as students' television viewing time increases. No study investigating the relationship between television viewing time and democratic attitude has been found in the literature. However, there are studies that found spending too much time watching TV and using mobile phones or tablets have a negative effect on students' overall success (Akar & Kara, 2017). A study conducted by MoNE (2008) on elementary school students found that the most harmful side of watching television was that it eliminates the opportunity to have a conversation in the family. It can be suggested that decreasing communication within the family harms the emergence and development of the democratic values in the family members. It can be suggested that the findings of both studies coincide with each other.

Suggestions

Based on the results of the research, the following recommendations are made:

➢ Democratic attitudes of students are at a "very good" level. On the other hand, although students' critical thinking attitudes are at a "good" level, they were found to be quite low compared to democratic attitude scores. Critical thinking attitude affects students' democratic attitudes positively. In this respect, efforts should be made to raise students' critical thinking attitudes to a higher level. For this reason, critical thinking activities should be paid more attention and given more room. However, it should be taken into consideration that critical thinking attitude, by its very nature, is influenced by informal activities and role-model behaviors exhibited by teachers. To illustrate, rumor-based expressions, complaints, events, and etc. are frequently experienced in school and classroom settings, and students often talk to their teachers about such events. The teacher can consider such situations as a case study and evaluate it as an informal activity. In such cases, the critical and democratic perspective that teachers will exhibit will motivate students to have such perspectives in similar settings. For this reason, in-service seminars, conferences, workshops, and similar activities aimed at raising teachers' awareness on these issues should be held for displaying general role model behaviors both in general and during the implementation of the activities in particular.

➢ It has been reported by many studies that family education is a variable that affects students. This result was also observed in this research. The effect of the family is great especially in terms of the emergence and retention of attitudes. It is seen that the mothers' educational background is low in the context of the region. The rate of illiterate people is quite high in particular. For this reason, various projects should be developed for families, especially for mothers. In addition, communication and cooperation between the school, the teacher, and the family are extremely important in the emergence and retention of attitudes. Students should see similar role model behaviors by their teachers and

parents in similar social settings such as schools and their families. For this reason, the required importance should be given to school and family cooperation not only in achieving academic success but also in bringing positive attitudes and values.

> Too much screen time on smartphones, tablets, and TV causes a negative effect on children. In this regard, teachers and families should act in cooperation and preventive measures should be taken to raise awareness among families. Children should be offered alternatives that would benefit them to spend their time and meet their needs for entertainment. It is seen that reading habit, in particular, is effective not only in academic success but also in gaining positive attitudes. Attention should be paid to the development of children's reading habits. In addition, it can be suggested that parents should get support from counseling services, and education specialists and psychologists who are experts in their fields should be invited to schools to give conferences and seminars to families and teachers.

> Research results suggest that the democratic attitude level of those who live in urban areas is high. Provincial national education administrators, school administrators, and teachers need to be more sensitive about this issue and make more efforts to help students who live in villages and towns to gain democratic attitudes.

> It is seen that the studies on these research subjects are insufficient, especially in the context of elementary education. For this reason, it may be suggested that further studies should be made with similar or different variables in elementary school settings. Apart from quantitative research, qualitative research should also be made. Therefore, it can be suggested that future studies could fruitfully explore this issue further by both quantitative and qualitative research designs.

Referensces

Acun, İ., Demir, M. & Göz, N. L. (2010). Öğretmen adaylarının vatandaşlık yeterlilikleri ile eleştirel düşünme becerileri arasındaki ilişki [Relationship between teacher candidates' citizenship competencies and critical thinking]. *Journal of Social Studies Education Research*, 1(1), 107-123.

Akar, C. (2007). *İlköğretim öğrencilerinde eleştirel düşünme becerileri[Critical thinking skills in primary school students]* (Unpublished doctoral dissertation). Gazi University, Institute of Education Sciences, Ankara.

Akar, C., & Yalçın, D. (2017). İlkokul 4.Sınıf Sosyal Bilgiler Öğretmen Kılavuz Kitabında yer alan kavramlar listesinin demokrasi ve çokkültürlülük kavramları açısından incelenmesi [Examining the list of concepts in Primary School 4th Grade Social Studies Teacher Guide Book in terms of democracy and multicultural concepts]. *Turkey Social Studies Journal,* 21(2), 435-450. Retrieved from https://dergipark.org.tr/tr/pub/tsadergisi/issue/31710/347503.

Akar, C. (2017). Sınıf öğretmeni adaylarının eleştirel düşünme eğilimlerinin çok kültürlülük değerlerini yordama düzeyi [Level of pre-service teachers' critical thinking disposition to predict multicultural values]. *Journal of Ahi Evran University Kırşehir Faculty of Education (KEFAD)*, 18(1), 741-762.

Akar, C., & Kara, M. (2016). İlkokul 4. sınıf öğrencilerinin eleştirel düşünme becerilerinin bazı değişkenlere göre değerlendirilmesi [Evaluation of critical thinking skills of primary school 4th grade students according to some variables]. *International Turkish Literature Culture Education Journal*, 5(3), 1339-1355.

Akar, C., İnel, Y., & Yalçıntaş, E. (2017). Research of levels of democratic values of classroom teacher candidates on various variables. *Journal of Current Researches on Educational Studies*, 7(2), 41-57. doi: 10.26579/jocures-7.2.4

Akar-Vural, R. (2005). *Bertolt Brecht'in öğretici oyunlarının "eğitimde drama" ve "sahneleme" yöntemleri temelinde hazırlanan iki farklı programın ortaöğretim hazırlık sınıfı öğrencilerinin eleştirel düşünmeye yönelik tutumlarına etkisi [The effects of two different programs prepared on the basis of "drama in education" and "staging" methods of Bertolt Brecht's tutorial games on secondary school prep students' attitudes towards critical thinking]* (Unpublished doctoral dissertation). Çukurova University, Adana.

Akın, U. & Özdemir, M. (2009). Öğretmen adaylarının demokratik değerlerinin çeşitli değişkenler açısından incelenmesi: Eğitim bilimleri fakültesi örneği [The effects of two different programs prepared on the basis of "drama in education" and "staging" methods of Bertolt Brecht's tutorial games on secondary school prep students' attitudes towards critical thinking]. *Journal of Ankara University Faculty of Educational Sciences*, 42(2), 183-198.

Alpoğuz, D., & Şahin, A. (2014). Demokratik veya otoriter olarak algılanan ana-baba tutumlarının ilköğretim öğrencilerinin Türkçe dersi akademik başarılarına etkisi [The effects of parental attitudes perceived as

democratic or authoritarian on the academic success of primary school students in Turkish language]. *Journal of Mother Tongue Education*, 2(3), 53-67. doi: 10.16916/aded.92531

Altuntaş, S., & Genç, H. (2018). Mutluluğun yordayıcısı olarak psikolojik sağlamlık: öğretmen örnekleminin incelenmesi [Resilience as Predictor of Happiness: Investigation of Teacher Sample]. *Hacettepe University Journal of Education Faculty,*1-15. Advance online publication. doi: 10.16986/HUJE.2018046021

Arslan, A., & Çalmaşur, H. (2018). Okul yöneticilerinin çokkültürlü eğitim tutumlarına ve demokratik tutumlarına ilişkin görüşleri: bir karma yöntem çalışması [School administrators' views on multicultural education attitudes and democratic attitudes: a mixed method study]. Journal of Computer and Education Research, 6 (12) , 163-194. doi: 10.18009/jcer.425761

Arslan, A., & Çalmaşur, H. (2017) İlkokul ve ortaokul öğretmenlerinin çokkültürlü eğitim tutumları ve demokratik tutumlarının çeşitli değişkenler açısından incelenmesi [Multicultural education attitudes and democratic attitudes of primary and secondary school teachers in terms of various variables]. *Journal of Atatürk University Kazım Karabekir Education Faculty,* 3(4), 90-109.

Aslan, S. (2018). The relationship between critical thinking skills and democratic attitudes of 4th class primary school students. *International Journal of Progressive Education,* 14(6), 61-69. doi: 10.29329/ijpe.2018.179.5

Aycan, N., & Çalık D. (2003). İlköğretim okullarında demokrasi eğitimi: Manisa örneği [Democracy education in primary schools: An example of Manisa.]. Manisa Celal Bayar University Journal of Social Sciences, 1(2). Retrieved from https://dergipark.org.tr/tr/pub/cbayarsos/issue/4075/53822

Aydemir, H., & Aksoy, D. A. (2010). Eğitim fakültesi öğrencilerinin demokratik tutumlarının bazı değişkenlerle ilişkisi: Malatya örneği [The relationship between democratic attitudes of education faculty students with some variables: Example of Malatya]. *Journal of Erzincan Education Faculty,* 12(1).

Ayık, İ. (2014). *Ortaokullarda şiddet ve okul iklimi arasındaki ilişki: Konya ili örneği [Relationship between violence and school climate in secondary schools: Konya province example]* (Unpublished master's thesis). Mevlana University Institute of Social Sciences, Konya.

Baltacı, Ö. (2010). *Üniversite öğrencilerinin sosyal kaygı, sosyal destek ve problem çözme yaklaşımları arasındaki ilişkinin incelenmesi [Examining the relationship between university students' social anxiety, social support and problem solving approaches]* (Unpublished master's thesis). Selcuk University Institute of Educational Sciences, Konya.

Bay, E., Gündoğdu, K., & Kaya, H. İ. (2010). The perceptions of prospective teachers on the democratic aspects of the constructivist learning environment. *Electronic Journal of Research in Educational Psychology*, *8*(2), 617-642.

Boone, M. (2008). A democratic education: Three interpretations. *Journal Of Philosophy And History Of Education,* Vol. 58, 16-22.

Bradshaw, R. (2014). Democratic teaching: an incomplete job description. *Democracy and Education*, 22 (2), 1-5.

Bulut, N. (2006). Bir grup üniversite öğrencisinin demokratik tutumları ile çeşitli değişkenler arasındaki ilişkiler [Democratic attitudes of a group of university students and relationships between various variables]. *Education Management in Theory and Practice*, Sayı 45, 37-59.

Büyükkaragöz, P., Kesici, Ş. (1996). Öğretmenlerin hoşgörü ve demokratik tutumları [Teachers' tolerance and democratic attitudes]. *Education Management in Theory and Practice*, 7 (7) , 353-365. Retrieved from https://dergipark.org.tr/tr/pub/kuey/issue/10388/127091

Can, Y. (2007). *İlköğretim okullarında şiddet öğrenci, öğretmen ve yöneticilerin yaklaşımları [Approaches of violence students, teachers and administrators in primary schools]* (Unpublished master's thesis). Beykent University Institute of Social Sciences, İstanbul.

Celep, C. (1995). Demokratik bir eğitim sistemi için [For a democratic education system]. *Journal of Education for Life*. 38, 20-24.

Cırık, İ. (2008). Çok kültürlü eğitim ve yansımaları [Multicultural education and its reflections]. *Hacettepe University Journal of Education Faculty, 34* (34), 27-40. Retrieved from https://dergipark.org.tr/tr/pub/hunefd/issue/7802/102246

Çırak, A. (2019). Ortaokul altıncı sınıf öğrencilerinin temel demokratik değerlerinin çeşitli değişkenler açısından incelenmesi [Examining the basic democratic values of sixth grade students in secondary school in terms of various variables]. *International Turkish Journal of Educational Sciences,* 7(12).

Davis, M. (2010). *Practicing in democracy in the NCLB Elementary classroom*, submitted in Partial Fulfilment of the Requirements for The Degree Master of Science in Education, Domanician University of California, San Rafael, CA.

Demir, M. (2006). İlköğretim dördüncü ve beşinci sınıf öğrencilerinin sosyal bilgiler derslerinde eleştirel düşünme düzeylerinin çeşitli değişkenler açısından incelenmesi [Investigation of critical thinking levels

of primary school fourth and fifth grade students in social studies lessons in terms of various variables]. *Gazi University Journal of Gazi Education Faculty*, 26 (3), 155-170. Retrieved from http://www.gefad.gazi.edu.tr/en/issue/6752/90797

Demir, R., & Aybek, B. (2014). Lise öğrencilerinin eleştirel düşünme eğilimlerinin çeşitli değişkenler açısından incelenmesi [Investigation of critical thinking tendencies of high school students in terms of various variables]. *Journal of Muğla University Social Sciences Institute, (32), 122-140.*

Diker, Z. (2012). Demokrasi eğitimi ve lise öğrencilerinin demokratik davranışları üzerine bir araştırma[A research on democracy education and democratic behavior of high school students]. *Journal of Education and Teaching Research,* 1(3), 229-238.

Doğanay A., Akbulut-Taş M. A., & Erden Ş. (2007). Üniversite öğrencilerinin bir güncel tartışmalı konu bağlamında eleştirel düşünme becerilerinin değerlendirilmesi [Evaluation of critical thinking skills of university students in the context of a current controversial subject]. *Education Management in Theory and Practice*, 52(52), 511-546.

Doğanay, A., & Sarı, M. (2004). İlköğretim ikinci kademe öğrencilerine temel demokratik değerlerin kazandırılma düzeyi ve bu değerlerin kazandırılması sürecinde açık ve örtük programın etkilerinin karşılaştırılması [Comparison of the implications of the open and implicit program in the process of providing basic democratic values to the second grade students of primary education and the process of bringing these values]. *Education Management, 10*(39), 356-383.

Edwards, D. B. (2010). Critical pedagogy and democratic education: Possibilities for cross-pollination. *Urban Rev,* 42, 221–242.

Ekici, Y. F. (2014). Öğretmen adaylarının demokratik tutumları ve demografik özellikleri arasındaki ilişkinin incelenmesi [Examining the relationship between pre-service teachers' democratic attitudes and demographic characteristics]. *International Periodical For The Languages, Literature and History of Turkish or Turkic,* 9(11), 593-602.

Ektem, I. S., & Sünbül, A. M. (2011). Öğretmen adaylarının demokratik tutumları üzerine bir araştırma [A research on the prospective teachers' democratic attitudes]. *Journal of Selçuk University Ahmet Keleşoğlu Education Faculty*, 31, 159- 168.

Elkatmış, M., & Toptaş, V. (2015). Sınıf öğretmeni adaylarının demokratik tutumlarının incelenmesi [Examining the democratic attitudes of elementary teacher candidates]. *Yüzüncü Yıl University Journal of Education*, 12 (1), 128-144. Retrieved from https://dergipark.org.tr/tr/pub/yyuefd/issue/25854/272575

Ennis, R. H. (1996). *Critical thinking*. Upper Saddle River, NJ: Prentice-Hall.

Erbil, D. G., & Kocabaş, A. (2017). İlkokul 3. sınıf öğrencileri için bir ölçek geliştirme çalışması: demokratik tutum ölçeği [A scale development study for primary school third grade students: the democratic attitude scale]. *Kastamonu Education Magazine, 1*(25), 221-232.

Ercoşkun, M. H., & Nalçacı, A. (2008). Sınıf öğretmeni adaylarının empatik beceri ve demokratik tutumlarının incelenmesi [Examination of elementary teacher candidates' empathic skills and democratic attitudes]. *Journal of National Education,* Sayı: 180, 204-215.

Evcimik, S. Ç. (2009). *Eğitim Fakültesi Öğrencilerinin Demokratik Tutumları [Democratic Attitudes of Faculty of Education Students]* (Unpublished Master Thesis). Fırat University, Institute of Social Sciences, Elâzığ.

Evin-Gencel, İ., & Güzel-Candan, D. (2014). Öğretmen adaylarının eleştirel düşünme eğilimleri ve yansıtıcı düşünme düzeylerinin incelenmesi [Investigation of pre-service teachers' critical thinking tendencies and reflective thinking levels]. *International Journal of Educational Programs and Instructional Studies, 4*(8), 55-68.

Facione, P. A. (2000). The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill. *Informal Logic*, 20(1), 61-84.

Genç, S. Z., & Kalafat T. (2007). Öğretmen adaylarının demokratik tutumları ile empatik becerilerinin değerlendirilmesi üzerine bir araştırma [A research on the evaluation of the teachers' democratic attitudes and empathic skills]. *Journal of Social Sciences*, 19, 119- 222.

Gömleksiz, M. N., & Çetintaş, S. (2011). Öğretmen adaylarının demokratik tutumları (Fırat, Dicle,7 Aralık, Cumhuriyet ve Erzincan Üniversiteleri Örneği) [Democratic attitudes of teacher candidates (Fırat, Dicle, 7 December, Cumhuriyet and Erzincan Universities Example)]. *Journal of Dicle University Ziya Gökalp Faculty of Education, 17*(2011)*,* 1-14.

Gömleksiz, M. N., & Kan, Ü. A. (2008). Eğitim fakültesi ve tezsiz yüksek lisans programlarına kayıtlı öğretmen adaylarının demokratik tutumlarının değerlendirilmesi (Fırat Üniversitesi örneği) [Evaluation of democratic attitudes of prospective teachers enrolled in education faculty and master's programs without thesis (Fırat University example)]. *Journal of National Education*, 178. 44-64.

Görücü, E. (2014). *Altıncı, yedinci ve sekizinci sınıf öğrencilerinin okuma alışkanlıkları ile eleştirel düşünme becerileri arasındaki ilişkinin incelenmesi [Examining the relationship between reading habits and*

critical thinking skills of sixth, seventh and eighth grade students] (Unpublished master thesis). Yeditepe University, Istanbul.

Gülveren, H. (2007). *Eğitim fakültesi öğrencilerinin eleştirel düşünme becerileri ve bu becerileri etkileyen eleştirel düşünme faktörleri [Critical thinking skills of education faculty students and critical thinking factors affecting these skills]* (Unpublished Doctoral Thesis). Dokuz Eylul University Institute of Educational Sciences, Izmir.

Gündoğdu, H. (2009). Eleştirel düşünme ve eleştirel düşünme öğretimine dair bazı yanılgılar [Some misconceptions about teaching critical thinking and critical thinking]. *Celal Bayar University Institute of Social Sciences,* 7(1).

Gürbüz, G. (2006). *İlköğretim 7. ve 8. sınıflarda vatandaşlık bilgisi dersinde demokrasi eğitimi [Democracy in the 7th and 8th grades primary education]* (Unpublished master thesis). Abant İzzet Baysal University, Institute of Social Sciences, Bolu.

Güven, A. (2005). *Tarih öğretiminde öğrencilerin demokratik tutumlarının çeşitli değişkenler açısından incelenmesi (Erzurum örneği) [Examining the democratic attitudes of students in history teaching in terms of various variables (Erzurum example)]* (Unpublished Doctoral Thesis). Atatürk University, Institute of Social Sciences, Erzurum.

Hawkins, J. (2006). Accessing multicultural issues through critical thinking, critical inquiry, and the student research process. *Urban Education Journal,* 41(169).

İbrahimoğlu, Z., & Öztürk C. (2013). Sosyal bilgiler dersinden örnek olay yöntemi kullanımının öğrencilerin akademik başarı derse karşı tutum ve eleştirel düşünme becerileri üzerine etkileri [The effects of using case study method from social studies lesson on students' academic achievement lesson attitude and critical thinking skills]. *Journal of Uludağ University Faculty of Education*, 26(2), 523-548.

Kaldırım, E. (2003). *İlköğretim 8.sınıf öğrencilerinin demokrasi algıları [Democracy perceptions of elementary school 8th grade students]* (Unpublished master thesis). Gazi University Institute of Educational Sciences, Ankara.

Kalkan, G. (2008). *Yedinci ve sekizinci sınıf öğrencilerinin eleştirel düşünme düzeyleri [Critical thinking levels of seventh and eighth grade students]* (Unpublished master thesis). Osmangazi Üniversitesi Eğitim Bilimleri Enstitüsü, Eskişehir.

Karabacak, H. (2011). *İlköğretim öğrencilerinin eleştirel düşünme becerileri ve beşinci sınıf öğrencilerinin eleştirel düşünme düzeyi (Erzurum ili örneği) [Critical thinking skills of primary school students and the critical thinking level of fifth grade students (Erzurum province example)]* (Unpublished master thesis). Atatürk University, Erzurum.

Karadağ, İ. (2007). *İlköğretim beşinci sınıf öğrencilerinin akademik başarılarının sosyal destek kaynakları açısından incelenmesi [Examining the academic success of primary school fifth grade students in terms of social support resources]* (Unpublished master thesis). Çukurova University, Adana.

Karahan, T., Sardoğan, E., Özkamalı, E., Dicle, N., & Dicle, A. (2006). Öğretmen adaylarında demokratik tutum, nevrotik eğilimler ve kendini gerçekleştirme [Democratic attitude, neurotic tendencies and self-actualization among teacher candidates]. *Hacettepe University Journal of Education Faculty*, 30 (30) , 149-158 . Retrieved from https://dergipark.org.tr/tr/pub/hunefd/issue/7806/102375

Karakaş, Y. (2008). *Lise öğrencilerinin mükemmeliyetçilik düzeyleri ile sosyal kaygı düzeyleri arasındaki ilişki [The relationship between high school students' perfectionism levels and social anxiety levels]* (Unpublished master thesis). Mugla University Institute of Social Sciences, Muğla.

Karasar, N. (2005). Scientific Research Method. Ankara: Nobel Publishing.

Karatekin, K., Merey, Z., & Kuş, Z. (2013). Öğretmen adayları ve öğretmenlerin demokratik tutumlarının çeşitli değişkenler açısından incelenmesi [Examining the democratic attitudes of teacher candidates and teachers in terms of various variables]. *Kastamonu Education Magazine*, 21(2).

Kardaş, S. (2013). *6. Sınıf öğrencilerinin demokratik tutumları ile ahlaki olgunluk düzeyi arasındaki ilişkinin incelenmesi (Kadıköy ilçesi örneği) [Examining the relationship between the democratic attitudes of 6th grade students and the level of moral maturity (Example of Kadıköy district)]* (Unpublished master thesis). Marmara University Institute of Educational Sciences, Istanbul.

Kartal, T. (2012). İlköğretim fen bilgisi öğretmen adaylarının eleştirel düşünme eğilimlerinin incelenmesi [Investigation of critical thinking tendencies of elementary science teacher candidates]. *Journal of Ahi Evran University Kırşehir Faculty of Education,* 13(2), 279-297.

Kaya, A., Bozaslan, H., & Genç, G. (2012). Üniversite öğrencilerinin anne-baba tutumlarının problem çözme becerilerine, sosyal kaygı düzeylerine ve akademik başarılarına etkisi [The effect of university students 'parents' attitudes on problem solving skills, social anxiety levels and academic success]. *Journal of Dicle University Ziya Gökalp Faculty of Education*, 18, 208-225.

Kesici, A., Pesen, A., & Oral, B. (2017). Öğretmenlerin sınıf içi demokratik davranışlarının çeşitli değişkenlere göre incelenmesi [Examination of teachers' classroom democratic behavior according to various variables]. *Electronic Journal of Social Sciences*, 16(60), 34-45.

Kılıç, D., Erçoşkun, H., & Nalçacı, A. (2004). Sınıf öğretmenliği öğretmen adaylarının demokratik tutumları [Democratic attitudes of prospective classroom teachers]. *Turkey Social Studies Journal,* 8(2-3), 9-19.

Koç, I., & Kuvaç, M. (2016). Reservice science teachers' metacognitive awareness levels. *Europen Journal of Education Studies*, 2(3), 44-63.

Kontaş, H., Selçuk, E., & Polat, M. (2016). Ortaokul öğretmenlerinin demokratik tutumları ile öğrencilerin demokratik değerlere sahip olma düzeylerinin incelenmesi [Democratic attitudes of secondary school teachers and the level of students' having democratic values]. *The Journal of Academic Social Science Studies*, 44, 141-151.

Kurnaz, A. (2011). *İlköğretim öğretmenlerinin yaratıcılık düzeyleri ve demokratik tutumları arasındaki ilişkinin değerlendirilmesi [Evaluation of the relationship between primary school teachers' creativity levels and democratic attitudes]* (Unpublished master thesis). Kahramanmaraş Sütçü İmam University, Kahramanmaraş.

Kuş, Z., & Çetin, T. (2014). Perceptions of democracy of primary school students. *Educational Sciences: Theory & Practice*, 14(2), 786-790. Doi: 10.12738/estp.2014 .2.1493.

MEB, Education Research and Development Department, (2008). Television watching habits of students. Ankara: MEB publications.

MEB (1973). Millî Eğitim Temel Kanunu (1973). *Resmî Gazete*. Yayım Tarihi: 24.06.1973. Sayısı: 14574. Numarası: 1739. [National Education Basic Law (1973). *Official newspaper*. Publication Date: 24.06.1973. Issue: 14574. Number: 1739].

Merey, Z., Kaymakçı, S., & Kılıçoğlu, G. (2011). Öğrenci konseyi seçimlerinde sosyal bilgiler öğretmen adaylarının demokratik tutumları (Gazi Üniversitesi örneği) [Democratic attitudes of social studies teacher candidates in student council elections (Gazi University example)]. *Journal of National Education*, 41 (191), 191-209. Retrieved from https://dergipark.org.tr/tr/pub/milliegitim/issue/36191/406900

Nazıroğlu, B., & Çetin, M. (2014). Yüksek din eğitimi öğrencilerinin demokratik tutumları üzerine bir inceleme [A review on the democratic attitudes of higher education students]. *Journal of Hitit University Faculty of Theology*, 13(25), 57-91. doi: 10.14395/jdiv109

Ozan, M. B., Türkoğlu, A. Z., & Şener, G. (2010). Okul yöneticilerinin sergiledikleri demokratik tutum ve davranışlarının öğretmenlerin motivasyonuna etkisi [The effect of democratic attitudes and behaviors of school administrators on the motivation of teachers]. *Fırat University Journal of Social Sciences*, 20(1), 275-294.

Ömerustaoğlu, A. (2004). Felsefe grubu öğretmenliği ve diğer anabilim dallarında öğrenim gören öğrencilerin demokratik tutumlarının çeşitli değişkenler açısından incelenmesi [Examining the democratic attitudes of students studying in philosophy group teaching and other departments in terms of various variables]. *Sakarya University Faculty of Education Journal,* 7, 193-204.

Özdaş, F., Ekinci, A., & Bindak, R. (2014). Öğretmenlerin demokrasiye ilişkin tutumlarının bazı değişkenler açısından incelenmesi [Examining teachers' attitudes towards democracy in terms of some variables]. *Journal of İnönü University Education Faculty,* 15(3).

Özdemir, S. M. (2005). Üniversite öğrencilerinin eleştirel düşünme becerilerinin çeşitli değişkenler açısından değerlendirilmesi [Evaluation of critical thinking skills of university students in terms of various variables]. *Turkish Journal of Educational Sciences*, 3(3), 297-316.

Öztaşkın, Ö. B., & İçen, M. (2015). Ortaokul öğrencilerinin hoşgörü eğilimleri ile demokrasi algıları arasındaki ilişkinin incelenmesi [Examining the relationship between middle school students' tolerance tendencies and perceptions of democracy]. *Journal of Black Sea Social Sciences*, 7(02).

Pişkin, M. (2010). Examination of Peer bullying among primary and middle school children in Ankara. *Education and Science*, 35(156), 176-189.

Polat, S., & Güler, H. (2020). Democracy in social studies curricula (2005-2018) from the perspective of teachers: a qualitative study. *Journal of Computer and Education Research*, 8(15), 156-180. doi: 10.18009/jcer.661105

Sağlam, H. İ. (2000). Sosyal Bilgiler Dersi'nin demokratik tutum geliştirmedeki rolü [The role of Social Studies Course in developing democratic attitude]. *Journal of National Education*, 146, 67- 71.

Saracaloğlu, A. S., Evin, İ., & Varol, S. R. (2004). İzmir ilinde çeşitli kurumlarda görev yapan öğretmenler ile öğretmen adaylarının demokratik tutumları üzerine karşılaştırmalı bir araştırma [A comparative research on the democratic attitudes of teachers working at various institutions in İzmir province and prospective teachers.]. *Educational Sciences in Theory and Practice,* 4 (2), 335-364.

Sarı, C., & Demirbağ, B. (2019). İlk ve ortaokul öğrencilerinde akran zorbalığının sosyodemografik değişkenler açısından değerlendirilmesi: ilçe örneği [Evaluation of peer bullying in terms of sociodemographic variables in primary and secondary school students: a sample of the district]. *Journal of Public Health Nursing*, 1(3), 119-131. Retrieved from https://dergipark.org.tr/tr/pub/jphn/issue/48387/642804

Seferoğlu, S., & Akbıyık, C. (2006). Eleştirel düşünme ve öğretimi [Critical thinking and teaching]. *Hacettepe University Journal of Education Faculty,* (30), 193-200.

Siegel, H. (1988). *Educating Reason: Rationality, Critical Thinking, and Education* (New York: Routledge).

Sili, A. (2012). Lise öğrencilerinin saldırganlık eğilimleri üzerine sosyolojik bir değerlendirme: Erzurum örneği [A sociological assessment of high school students' aggression tendencies: an example of Erzurum]. *Ekev Academy Journal*, (51), 261-271.

Smith, K. M. (2009). Dualism with in progressive pedagogies: The dynamic nature of democratic education. *Social Studies Review, 48*(1), 38-44.

Subba, D. (2014). Democratic values and democratic approach in teaching: a perspective. *American Journal of Educational Research*, 2(12), 37-40.

Sulaiman, W. S. W., Rahman, W. R. A., & Dzulkifli, M. A. (2009). *Relationship between critical thinking dispositions, perceptions towards teachers, learning approaches and critical thinking skills among university students,* The 4th International Postgraduate Research Colloquium.

Büyüköztürk, Ş. (2004). Veri analizi el kitabı. *[Data analysis handbook].* Ankara: PegemA

Şahin, G. (2008). *İlköğretim Okullarında Görev Yapan Öğretmenlerin Demokratik Tutumları ile Dogmatik Düşünce Biçimleri Arasındaki İlişkinin İncelenmesi [Investigation of the Relationship Between Democratic Attitudes and Dogmatic Thought Forms of Teachers Working in Primary Schools]* (Unpublished Master Thesis). Yeditepe University Institute of Social Sciences, İstanbul.

Şimşek, N. (2011). İlköğretim sosyal bilgiler dersinde öğrencilerin demokratik kazanımlarının çeşitli değişkenler açısından incelenmesi [Examining the democratic achievements of students in terms of various variables in elementary social studies course]. *Journal of National Education*, 41(190), 188-201.

Taçman, M. (2006). İlköğretim sınıf öğretmenlerinin demokratik tutumları [Democratic attitudes of primary school teachers]. *Cypriot Journal of Educational Sciences*, 1(1), 60–46.

Tiryaki, B. (2018). *İlkokul 3. Sınıf öğrencilerinin hayat bilgisi dersine yönelik tutumlarıyla demokratik tutumları arasındaki ilişki [The relationship between elementary school 3rd grade students' attitudes towards life studies and their democratic attitudes]* (Unpublished master thesis). Fırat University, Elazığ.

Tonga, D., & Kılıçoğlu, M. (2013). İlköğretim Sosyal Bilgiler Dersinde yer alan kazanımların vatandaşlık ve insan hakları konuları açısından değerlendirilmesi [Evaluation of the acquisitions in Primary Education Social Studies Course in terms of citizenship and human rights issues]. *Turkey Social Studies Journal*, 171 (171), 289-308. Retrieved from https://dergipark.org.tr/tr/pub/tsadergisi/issue/21497/230501

Tsui, L. (1999). Courses and Instruction Affecting Critical Thinking. *Research in Higher Education,* 40(2).

Tung, C. A., & Chang, S. Y. (2009). Developing critical thinking through literature reading, *Feng Chia Journal of Humanities and Social Sciences*, 19, 287-317.

Turabik, T., & Gün, F. (2016). The relationship between teachers' democratic classroom management attitudes and students' critical thinking dispositions. *Journal of Educationand Training Studies*, 4(12), 45-57.

Turan, M., & Taşpınar M. (2004). The effect of citizenship and human rights course on students' views on democracy and human rights. International Democracy Education Symposium (20-21 May), Çanakkale.

Tümkaya, S. (2011). Fen bilimleri öğrencilerinin eleştirel düşünme eğilimleri ve öğrenme stillerinin incelenmesi [The study of science students' critical thinking disposition and learning styles]. *Journal of Ahi Evran University Faculty of Education*, 12(3), 215-234.

Uluçınar, U. (2012). *Öğretmen adaylarının eleştirel düşünme eğilimlerinin demokratik değerlerini yordama düzeyi [The level of predicting teacher candidates' democratic values of critical thinking disposition]* (Unpublished master thesis). Eskişehir Osmangazi University Educational Sciences Institute, Eskişehir.

Ulusoy, K. (2007). *Lise tarih programında yer alan geleneksel ve demokratik değerlere yönelik öğrenci tutumlarının ve görüşlerinin çeşitli değişkenler açısından değerlendirilmesi [Evaluation of student attitudes and views towards traditional and democratic values in high school history program in terms of various variables]* (Unpublished Doctoral Thesis) Gazi University Institute of Educational Sciences, Ankara.

Ural, S. N., & Sağlam, H. İ. (2011). İlköğretim öğrencilerinin demokratik tutum düzeylerinin bazı değişkenler açısından incelenmesi [Investigation of elementary school students' democratic attitude levels in terms of some variables]. *Journal of Values Education,* 9(22), 161- 180.

Üstün, G. (2011). *Sınıf öğretmenlerinin yansıtıcı düşünme eğilimleri ile demokratik tutumları arasındaki ilişkinin incelenmesi [Examining the relationship between reflective thinking tendencies and democratic attitudes of classroom teachers]* (Unpublished master thesis). Marmara University Institute of Educational Sciences, İstanbul.

Yalçın, G. (2007). *Ortaöğretim öğretmenlerin sınıf yönetiminde gösterdikleri davranışların demokratikliğine ilişkin öğretmen ve öğrenci görüşleri [Teachers and students' views on the democracy of the behaviors of secondary education teachers in classroom management]* (Unpublished master thesis). İnönü University, Malatya.

Yamaguchi, M., & Maguth, B. (2005). Students as critical thinkers: Multidimensional citizenship in a global era, *Ohio Council for the Social Studies*, 31-35.

Yazıcı, K. (2011). Sosyal bilgiler öğretmen adaylarının demokratik değerlerinin çeşitli değişkenler açısından incelenmesi [Examining the democratic values of social studies teacher candidates in terms of various variables]. *Education and Science Journal*, 26(159), 165-178.

Yeşil, R. (2003). Demokratik eğitim ortamının insan hakları temeli [Human rights basis of democratic education environment]. *Journal of Kırşehir Faculty of Education*, 4(2), 45-54.

Yıldız, N. (2011). *İlköğretim altıncı sınıf öğrencilerinin fen ve teknoloji dersinde eleştirel düşünme düzeylerinin çeşitli değişkenler açısından incelenmesi [Critical thinking levels of sixth grade elementary school students in science and technology lesson in terms of various variables]* (Unpublished master thesis). Selcuk University, Konya.

Yılmaz, K. (2011). İlköğretim okulu öğretmenlerinin sınıf yönetimi tarzları ile demokratik değerlere ilişkin görüşleri arasındaki ilişki [The relationship between primary school teachers' classroom management styles and their views on democratic values]. *Journal of Values Education*, 9 (21), 147-170.

Yurtseven, R. (2003). *Ortaöğretim kurumlarında çalışan öğretmenlerin demokratik tutumları [Democratic attitudes of teachers working in secondary education institutions]* (Unpublished master thesis). Çukurova University Institute of Social Sciences, Adana.

Yüksel, İ., Bağcı, Ş., & Vatansever, E. (2013). İlköğretim son sınıf öğrencilerinin demokratik değerlere sahip olma düzeylerinin belirlenmesi [Determining the level of elementary school students having democratic values]. *Adıyaman University Journal of Social Sciences Institute.* (12) , 309-334. doi: 10.14520/adyusbd.427

Zorlu, G., Öneren, M., & Çiftçi, G. E. (2019) Çalışanların kitap okuma alışkanlıkları ile duygusal zekâ arasındaki ilişkiye yönelik bir araştırma [A research on the relationship between employees' reading habits and emotional intelligence]. 4th International EMI Entrepreneurship & Social Sciences Congress, 29-30 November 2019, İstanbul.

Identifying the Challenges of Involvement in Entrepreneurship Activities among a Group of Undergraduates

Sofeana Mohd Farid[1], Shukran Abdul Rahman[1]
[1]International Islamic University Malaysia

To cite this article:

Identifying the Challenges of Involvement in Entrepreneurship Activities among a Group of Undergraduates

Sofeana Mohd Farid[1*], Shukran Abdul Rahman[1]
[1]International Islamic University Malaysia

Abstract

Entrepreneurship activities have been promoted in Malaysian universities through entrepreneurship education, mainly with the aim to minimize youth unemployment rate in Malaysia. The effort may have been hindered due to many reasons that may explain the lack of effectiveness of entrepreneurship education found in previous studies. The present study was conducted to identify the challenges of involvement in entrepreneurship activities among a group of undergraduates from a public university in Malaysia. Qualitative data was collected from three males and seven female undergraduates who were selected to participate in semi-structured interviews. They were asked to provide insights on their lives as students who participated in entrepreneurship activities under the entrepreneurship center in the university they attended. Information gathered in the interviews was then analyzed using thematic analysis. There were two major themes found in this study which are external challenges and personal challenges. The findings are expected to provide insight on the key challenges faced by students as a result of their involvement in entrepreneurship activities. It could also contribute to the efforts to improve entrepreneurship education in Malaysia, and in turn inform the development of programs to develop interventions that improve entrepreneurship activities in higher learning institutions.

Key words: Entrepreneurship education, Entrepreneurship education in Malaysia, Young entrepreneurs, Higher learning institution

Introduction

As a response to the impending slowdown of the Malaysian economy, the Government of Malaysia has developed several plans to prepare the people, especially youths, in facing the bad economic condition: lack of job opportunities and even increasingly high standards of living. In this regard, an entrepreneurship ecosystem has been developed in Malaysian public universities, aimed at helping university students to equip themselves with competencies that enable them to successfully deal with the scenario (Rahim et al., 2015). It helps students to face the adversities which lie ahead in life by being productive rather than only staying put without doing anything (Din et al., 2015). The initiative does not only help the students to improve their financial condition but boost the nation's economic development It also provide opportunities for people to create jobs instead of being employed by others. In fact, the introduction to the world of entrepreneurship in universities is seen to be a suitable intervention as it exposes youths to many entrepreneurial experiences, hence preparing them for the real world outside of their universities (Rahim et al., 2015).

Entrepreneurship education is also being practiced in other higher learning institutions all around the world. In the western countries, particularly in the United States of America, many universities have employed various techniques of teaching entrepreneurship such as the provision of real-world entrepreneurship experiences, and allocation of entrepreneurship-related resources for universities or colleges (Morris et al., 2017; Barnard et al., 2018). According to Zhou and Xu (2012), colleges in China developed the awareness of entrepreneurship among their students by organizing systematic programs and competitions. Finland, in its efforts to encourage entrepreneurship activities, introduced multiple types of entrepreneurship education among its youths. Finland has adopted various strategies to avail entrepreneurship education programs, such as through learning, business competitions and co-ops education (Laurikainen et al., 2018).

---
[*] Corresponding Author: *Shukran Abdul Rahman, shukran@iium.edu.my*

Despite the availability of various initiatives, it seems that the programs have yet to produce fruitful impacts (Khazanah Research Institute, 2019), reflected in the increase of unemployment rate among graduates despite the existence of entrepreneurship programs in universities. There were many studies conducted to assess undergraduates' lack of interest and successes in entrepreneurship programs during or post-university life. Some have attributed it to the readiness of students and universities in undergoing or executing the program, respectively (e.g., Othman et al., 2012; Iqbal et al., 2012). Apart from studies centered in Malaysia, there were indeed international studies that have discussed the subject of entrepreneurship education and challenges behind said program. Written below are studies amassing from researchers in America, Brazil, China, United Kingdom, Indonesia, Iran, Lebanon, Nigeria, Spain that have all but generally agreed on a few reasons as to why entrepreneurship education has not yet flourished as planned.

Challenges of Involvement in Entrepreneurship Activities: Literature Review

Khazanah Research Institute (2019) reports that the number of unemployed youths has increased 0.33% in the year 2018, showing 10.9% of the total youths unemployed, the highest ever recorded data in Malaysian history of youth unemployment. Beginning of the year 2010, the Government of Malaysia introduced various initiatives to reduce unemployment rates by introducing and promoting entrepreneurship program in universities (Din et al., 2016). This plan aimed to help students becoming entrepreneurs, or at least acquiring entrepreneurial skills to be utilized later in life so that they may resort to self-employment, and in turn gradually overcome unemployment issues among graduates. Despite the efforts, unemployment issue remains a worrying concern to many, warranting many researchers to conduct studies which attempted to find answers for the issues, especially on why entrepreneurship education is not able to produce entrepreneurs among graduates. There were many underlying problems over the implementation of entrepreneurship program in Malaysia (Yusoff et al., 2014). There are studies from literature all over the world which has reported several aspects that contribute to the drawbacks: including the following:

1. Entrepreneurship education has been found to have low effectiveness and weak efficiency, which are attributed to:
   a. the lack of skills among staff who promote, teach and train entrepreneurship programs in universities (Karimi et al., 2010; Olorundare & Kayode, 2014; Welsh & Drăguşin, 2011; Yusoff et al., 2014);
   b. the low commitment of staff and management towards the government's entrepreneurship agenda (Rahim et al., 2015);
   c. the low effectiveness of entrepreneurship education in motivating or equipping youths to survive outside their university life (Chiekezie et al., 2016);
   d. the limited entrepreneurial skills among entrepreneurship educators (Othman et al., 2012);
   e. create meaningful experiences in in the program with such limited time (McGuigan, 2016);
   f. non-blending of the subject matter in entrepreneurship education with different majors at the university, while also being based on the realities of life (Karimi et al., 2010; Welsh & Drăguşin, 2011);
   g. traditional methods of teaching and learning that makes entrepreneurship stimulation difficult (Sanchez et al., 2017);
   h. lack of training provision for instructors (Ozdemir, 2018; Wang, 2013).

2. Entrepreneurship development ecosystem is shaped by universities' internal environment despite the availability of Government's supports and supplies for universities to conduct the program. The unconducive or poor ecosystem are attributable to:
   a. universities inefficiency to promote or greaten entrepreneurship culture. The rigid university rules and administrator's negative attitude towards entrepreneurship programs have caused difficulties for entrepreneurship education program to run smoothly (Othman et al., 2012);
   b. students' family role on students' entrepreneurial intentions. Some parents have unfavourable attitude towards entrepreneurship program, hence influenced their children's attitude and motivation towards entrepreneurship program (Rengiah & Sentosa, 2016);
   c. roles of the society, the absence of positive perception on entrepreneurship program might hinder the development of entrepreneurial activities. On the contrary, it was found that the support provided by the society would enable students to develop positive attitudes towards entrepreneurship; and motivate them to acquire relevant skills to survive as entrepreneur (Ferreira et al., 2016);

      d.   the difficulties of teaching a large number of students with different learning rates, motives, backgrounds and capabilities (Vanevenhoven, 2013);

      e.   poor entrepreneurship culture within the community (Karimi et al., 2010; Olorundare & Kayode, 2014; Welsh & Drăguşin, 2011).

3. There are personal psychological issues confronting students who participated in entrepreneurship education programs, some of them are:

      a.   low readiness to undergo entrepreneurship education (Othman et al., 2012);

      b.   lack of skills to become participants of entrepreneurship programs (Ustyuzhina et al., 2019; Yusoff et al., 2014);

      c.   lack of commitment towards the Government's entrepreneurship agenda (Yusoff et al., 2014);

      d.   inability to deal with self and psychological barriers, making it challenging for them to convert their entrepreneurial intentions into a reality (Sandhu et al., 2011);

      e.   lack of skills to deal with stress, leading them to avoid risk that are present in entrepreneurship activities (Sandhu et al., 2011);

      f.   lack of knowledge and skills in financial matters. They need financial literacy and other entrepreneurial skills in order to develop entrepreneur mindset (Bagheri & Pihie, 2012);

      g.   lack of favourable attitude towards, and weak intention in, entrepreneurship program. Both attitude and strong intention are important predictors for the program to become successful and effective (Rengiah & Sentosa, 2016);

      h.   lack of interest on entrepreneurship activities despite the many initiatives to nurture entrepreneurship skills. They are reluctant to invest their time and energy in entrepreneurship programs (Bagheri et al., 2013);

      i.   absence of prior knowledge and experience on entrepreneurship, dampening the entrepreneurship education process even if they have high desires or intentions towards entrepreneurship (Bagheri et al., 2013);

      j.   lack of motivation to acquire entrepreneurial skills, potentially due to their low needs for achievement (Chiekezie et al., 2016).

Given the difference of time and ecosystem between the time when the above researches were conducted and the current context of undergraduates' entrepreneurship involvement, the scenario encountered by the latter might be different from that of the former. Besides the potential obsolescence of explanation from the previous research findings (e.g psychological capital; or the mediating role of attitude towards entrepreneurship) there has been less research that was conducted to identify issues and challenges faced by undergraduates who participated in entrepreneurship activities. There were only few researches that dove into the subjective realm of this topic such as the study to identify real life challenges faced by the students during their involvement in university entrepreneurship activities. Furthermore, research on entrepreneurship programs at universities is still at an infancy stage as researchers have yet to understand the topic well (Frese & Gielnik, 2014), underscoring the need to investigate entrepreneurship programs in a deeper sense. Additionally, Welsh and Drăguşin (2011) has noted that there is a lack of feedback coming from the students who go through said program in order to improve the curriculum. This was also echoed by Lorcu and Erduran (2016) in their writing of the need to further analyse the effect of entrepreneurship education on the students as well as institutions. The same was written in a report by the Department for Business, Innovation and Skills of the UK government in 2013, which was to record students' individual experiences of enrolling in enterprise and entrepreneurship education programs. That said, it is imperative that an in-depth study to be conducted to identify the consequences of undergraduates' involvement in entrepreneurship activities, giving more focus to assess the challenges faced by undergraduates when joining entrepreneurship activities.

This study is conducted by anchoring on the study by Oparaocha and Pokidko (2013) which focused on assessing the impact of entrepreneurship education on Finnish students in the MSc Entrepreneurship program (see Table 1). This study explores information beyond the quantitative data and identifies issues that have hindered the process of cultivating entrepreneurship competencies amongst undergraduates through entrepreneurship education. It is expected that the findings from this study could expand the body of knowledge on entrepreneurship education in Malaysia; and provide necessary information for universities and the Government, on issues which concern the development of entrepreneurial culture among undergraduates in the country. The insight could be used to inform the design of better entrepreneurship intervention programs in the future.

Method

Participants

10 participants from a Malaysian public university consisting three males and seven females were interviewed for this study. All of them are student entrepreneurs who also involved themselves in activities under the University's Entrepreneurship Development Centre, namely Entrepreneurship Club, Social Entrepreneurship Club, Agrosis and Kosiswa. This study employed a purposive sampling technique where participants were selected from the Entrepreneurship Development Centre of the University. The reason for these participant criteria is that they directly receive entrepreneurship education from the university, which is the objective of the establishment of each entrepreneurship center at universities in Malaysia. This would enable them to relay rich information about the subject of this study better than other students as they have under gone the entire learning process first hand (Creswell, 2011).

Design

According to the authors, Oparaocha and Pokidko (2013), qualitative research is suitable in order to capture information that transcends numbers such as experiences of participants and more. As this research attempts to explore the challenges faced by entrepreneurship education students, this research adopted semi-structured interview which facilitates a flexible conversation between the researchers and participants, leading to open ended answers and giving deeper analysis on issues and challenges face by students who participated in entrepreneurship program (Oparaocha & Pokidko, 2013). This interview enabled the researchers to gather verbal information and capture non-verbal elements of face-to-face communication from the selected participants (Saunders et al., 2012). This technique allows different participants to comment on the subject-in-question based on their point of views and emphasize on details that they consider important (Charmaz, 2006). The researchers have developed and used the following interview protocol in the semi-structured interview.

Interview Protocol

Based on Oparaocha and Pokidko (2013), the researchers have structured their interview as follows:

Table 1. Interview protocol

Phase		Purpose
Initial contact by phone or e-mail	a)	Making short introduction of the researchers
	b)	Introducing the identified participants to the research by informing them about research on entrepreneurship education program.
	c)	Placing request to interview the identified participants
Confirmation of interview (phone/e-mail)	a)	Making gentle reminder about the time for interview
	b)	Informing the participants about the content of the interview: best and worst times
First phase of the interview	a)	Initiating the interview session by doing warm-up with the participants. "*Tell us a bit about yourself*" question.
	b)	Gathering information on memorable events and incidents during their involvement in
	c)	Placing supportive questions to the participants - "Why and How"- for deeper reflection
Break	a)	Serving coffee or tea with snacks
Second phase of the interview	b)	Soliciting meaning on the information provided over the subject-in-question from the participant's current life- stage
Wrap-up	c)	Gathering additional comments from the participants. Allowing each to elaborate any information they provided earlier.

Procedure

The researchers have secured approval from the Ethics Committee to conduct the study. Participants were approached face-to-face and also through digital communication medium (WhatsApp messages). The

researchers briefed the participants on the topic of this study, the objectives of the study and also informed them of their rights in having their information not disclosed. This is done by assigning each participant to a participant code, for example, Participant 001 or participant one, so that the participants remain anonymous (Fleming & Zegwaard, 2018). They were informed that each of them is protected at all times and that they have the right to withdraw from the interview without being given any penalty. Participants were also informed that the interview is to be done in places or times that they deem comfortable; and it would be conducted without exposing them to any harm or danger. The participants were allowed to contact the researchers if they want to know the result of the study or for any kind of enquiries. They were also invited to ask the researchers if they wanted to know more about the study.

Before the interview session began, participants were given an informed consent form which required their signature if they agreed with the term of reference stated in the form. The informed consent entails the information of the study, the right to withdraw and the confidentiality of the information given by the participants (Roshaidai, 2018). The interview took 15 to 30 minutes per participant and the interview was guided by the researchers with open-ended questions. This type of interview technique is considered to be the most effective in gaining open-ended answers and at the same time providing a guidance for participants to answer (Oparaocha & Pokidko, 2013). The data of this interview is recorded after securing permission from the participants. Each of them was given a token of appreciation for participating in the study.

Data Analysis

The data of this study has been analyzed using thematic analysis (TA). TA, as mentioned by Braun and Clarke (2014), is a suitable approach to reflect qualitative data that are more application-based, such as practices or policies. TA is a useful toolkit to analyze large sets of data for its well-structured guidelines; and to reflect the core information from the data collected in a study (Nowell at al., 2017). The data of this study was analyzed in six stages, as proposed by Braun and Clarke (as cited in Howitt & Cramer, 2017). The data is first familiarized by researchers through reading and re-reading of the information available in the transcription of the 10 interviews for several times. This is to see whether any information occurs regularly over several reviews. Next, chunks of data from the 10 transcriptions were initially coded to summarize the key parts of the texts. Then, coding that share similar meaning were joined together to form themes. This were done several times by the researchers, to ensure that the similarities of the coding are in congruent and clear. The themes were then further refined and inspected repetitively, so that each data fits with the themes well. After that, themes and sub-themes were identified as well as labelled, which were then scrutinized again for the final report. The finalized themes and sub-themes were reported in the findings section.

Findings

The findings of this research have yielded several themes which serve as a basis to understand the challenges faced by undergraduates on their involvement in entrepreneurship. Two categories of main themes have been derived from the data, external challenges and personal challenges.

External Challenges

The challenges have been extrapolated from the interview with the participants, presented below.

Student Participation for Entrepreneurship Programs

Many participants stated that entrepreneurship program is such an important program and most of them are organized by entrepreneurship clubs at the University. Nevertheless, some of the participants highlighted a number of issues pertaining to managing the participants of entrepreneurship activities, especially on the difficulties to manage students who were in entrepreneurship clubs. The participant said that "it was quite segregated in two groups. We communicated but we did not bond well. We don't always meet each other. Also, there are some who do not feel a sense of responsibility".

Among the issues gathered in this study is the recruitment of students to participate in entrepreneurship programs. There were difficulties in attracting participation from students. Participant two and three said that it was a difficult effort to make them feel interested to attend entrepreneurship activities. They said that it was "difficult to get an audience for these programs".

The other issue is on the development of good entrepreneurship activities. Equipping students with entrepreneurship activities is a huge task, especially in grabbing students' attention towards the program. Participant two, three, four, five and eight have given similar factors as to why it is difficult to attract students in joining entrepreneurship programs. It is found that this is attributable to the mismatch between the entrepreneurship activities and students' program of studies. This causes very little chance for students to enhance their interest in entrepreneurship activities, especially when entrepreneurship activities are only suitable for undergraduates in certain field of knowledge. One participant has told that he "realized that in this university nobody knows about it unlike in other places. This is because no one is a student of agriculture here, no project venues too".

Entrepreneurship Education Program

The participants expected they would gain the much-needed knowledge by joining entrepreneurship development program. Nevertheless, participant four, five and nine have also reported that they felt that the office which is in charge of entrepreneurship education did not help them much. In effect, the participants have to search for entrepreneurship knowledge on their own, outside the university. A participant said that she "find things on my own. I go for outside programs myself".

Participant ten has said that the students need to themselves find a way to gain extra knowledge, thus, they are involved in any entrepreneurial activities mainly as their accomplishment of work, instead of the effort to acquire knowledge of entrepreneurship. The participant highlighted that "the students need to find a way themselves to gain extra knowledge". They have also attained information on their respective clubs or even entrepreneurial knowledge on their own; most of them were not given by the management. Participant ten said that the Office "does not play that much of a part in shaping a student to become an entrepreneur".

This is also echoed by participant one who has said that "they did not explain anything to us, we had to dig the information ourselves". Participant four felt disappointed with the roles of entrepreneurship education program. Some students enrolled into entrepreneurship niche class but were quite disappointed with it because they already know most of the theories taught in the class. The participant said that the "niche area class is all theory which I already known before I enroll in the class. I expect more from that class that is why I enrolled". One of the participants have said that enrolling in entrepreneurship niche class "can be a bit boring".

Incompetent administrative or management staff

One of the main challenges that the students faced during their involvement in entrepreneurship program at their university is difficulties in dealing with the staff at the university. This may stem from the incompetence of the management in terms of dealing with students and entrepreneurial activities. The management may have not been able to connect with the students well and may not have managed the programs properly. These can be exemplified in many instances in the data where students have complained about the difficulty to communicate with the university staff who are in charge of managing their programs or businesses. Participant one talked at length about how tough it is to be on the same page with the management, with regard to the club that the participant had joined. One participant has noted that "they don't explain anything about it, how the business was used and how much capital, how much is left. I have asked personally, to get this clear. They don't give us the information".

The negative vibes shown by trainers to participants who attend entrepreneurship programs is also a challenge. Some participants in this study highlighted that they could not even understand what the trainers wanted to do with them. Participant four said: "Those negative vibes, mostly negative thought against us, don't understand what we are trying to do and do not want to listen to what we want to do".

The other issue is on the unpleasant experiences when dealing with the University management. Participant five have experienced confusion with the university over the permission to utilize facilities (e.g venue) in the University. Participant five and eight have also stated that the explanation they received from the management is irrelevant and should not have happened considering all the factors they have taken into consideration to avoid the problem. The participant told that "people came and questioned us on the approval of using this venue. We did not understand why that became an issue where we were banned from doing this project when we have clearly inquired them on it".

Some reported that they have not been politely or respectfully attended to for the mistakes they did when handling application processes, like proposal making or meetings, which they felt should not have happened (Participant three, four, seven and eight). One said that "there were mistakes in the proposal but no need to yell at us". Management are also not equipped with the real knowledge of entrepreneurship. Participant one explained that the management did not understand certain information with regards to their entrepreneurship clubs. The participant said that for them "the challenge is the administration, I don't know. Business cooperatives have resources, network hence there is no problem in facilities. It's just the administration".

Participant four and ten stated that the management does not really help student entrepreneurs to flourish in the university by saying that "the office is just for managerial stuff, like when you want to open up a business here. You need them for the place to set up and operate your business. The rest is on the students".

The office is said to not being objective when processing the application for approval of programs (participant one, four and five). One of them has said that "they don't really take these four clubs seriously; they don't manage them well".

Personal Challenges

With regard to the personal challenges, the following are the findings which present the experiences of participants' personal competencies and attitudes towards entrepreneurship programs.

Lack of interest in university entrepreneurship program

There is a lack of interest in entrepreneurship programs organized by the University. It has been reported that some students show low favorable attitude towards being tied with the university's entrepreneurship agenda. Participant one has said that the office "makes simple things complicated which in my opinion may make people feel that they don't want to become an entrepreneur here. They just do it alone, not wanting to affiliate as university student entrepreneurs".

Low understanding of the concept of entrepreneurship

Lack of understanding on important concepts in entrepreneurship activities, such as the concept of business cooperatives or social entrepreneurship have been reported as unpleasant issues which affect participants' motivation to involve in entrepreneurship activities. Participant four explained the difficulties to make students understand certain concepts in entrepreneurship, such as executing business in a cooperative. The participant said, "to re-correct the foundations of the cooperative here, which stems since its opening in 2011. That's the struggle, to correct the systems".

Participant two, three and five have given a set of similar factors to this issue. Firstly, there is little manpower to manage entrepreneurship program. This has made the program a burden and a struggle to them. One of them said that they "were still having a program that day that left us with very little manpower". All of them have also recalled of times when club members do not engage with no matter what promotions they have put out for the program. "The club members we have is around 70. We spread the message to ask them to join, but they hardly engage with us", said participant two. Secondly, low level of cooperation among members of the entrepreneurship society. Participant six said that many of them left the WhatsApp Group' of the entrepreneurship society so that they will not be bothered by any messages disseminated by the club". The participant said that "we already have their data, but they do not participate. Some of them left the group".

Group Conformity

This is about conformity to friends with less favorable attitude towards entrepreneurship program. This has caused some participants who joined entrepreneurship program to follow their friends who left the entrepreneurship activities. This is also admitted by participant three and ten. One of them highlighted that their "friend asked me to join. It was an induction, so I think that was what is not good. We just go there for an induction so we did not think of the responsibilities that come with it". Another said that the "main reason for joining this is to add more experience, but really, I got into this because of a friend. They asked me to tag along so I came".

Physical exhaustion

Almost all participants complained about physical exhaustion, in that they feel extremely tired from working in the entrepreneurship club or handling entrepreneurship programs. They said that it is tiring to go from their residential college to the program venue, classes and back to the program venue again after their class. They said that "because of the things we had to do, the schedule, the venue. It was all time consuming. The distance from place to place is far". One also said that it "kind of puts pressure on me. It is really tiring and it also affects my schedule".

Emotional exhaustion

Participants have also talked about how emotionally drained they were due to their involvement in entrepreneurship activities. One of the participants said that she and her team members were affected by all the negativities that surround them. Participants two and three felt anxious from all the work due to lack of knowledge and experience for being new in this area. They have a lot more to learn. So, when the work piles, they feel the pressure to complete them well. The participant noted that she "also have not joined anything before, so I felt pressured. I don't understand anything". They also felt pressured to keep up with all the work and responsibilities so much so that they thought of quitting from the position they held in the entrepreneurship club. One of them said "it makes me sad". They also said that they "cried a lot".

Academic struggles

Almost all participants have told that they struggled with academic workload during their involvement in entrepreneurship activities. Some participants had to skip a lot of classes whenever they have entrepreneurship programs. The participant on this matter said that "when the program comes, we always skipped class". They had to handle huge responsibilities in entrepreneurship club alone because there were limited number of committee members to help them. They recalled that they "had to resort to outside help, we interviewed students other than our club to participate". They have also had to sacrifice studies and priorities work in entrepreneurship club as said by participant two and three. They said that "it affects our study. When things go wrong with the proposal, we have to spend time on that".

Participant five reported that their "examination result dropped after being involved in entrepreneurship activities". Participant six have said that they continued thinking about work in entrepreneurship club even if they have already gone back to hostel room. The participant said that "when we go back to the dorm that is all we think about. That's the biggest concern, I also need to focus on my studies". Participant seven added that they "had to provide too much time for entrepreneurship activities, and less time on work and personal management".

Discussion

Students who took the decision to involve themselves in entrepreneurship activities in their university years may risk facing challenges in their journey. This research was based from studies by Rahim et al. (2015), Bagheri et al. (2013), Rengiah & Sentosa (2016) and Ridzwan et al. (2017) which have discussed the challenges of youth entrepreneurs, especially those who participated in entrepreneurship activities at their universities. Findings from the research show no conclusive reason for students' lack of interest to become entrepreneurs after they graduate from universities. Building from there, this research follows the recommendation of Frese and Gielnik (2014) as well as Rengiah and Sentosa (2016) who believe that a qualitative study should be done to further investigate this matter.

The findings show that most participants went through similar experience despite being affiliated to different clubs under the university's entrepreneurship development programs. Most of them have encountered problems when communicating with the management, and that each of them has talked about how such situation affected them mentally throughout their involvement in entrepreneurship activities. They have also reported that their difficulties in engaging their fellow students or convincing them to join entrepreneurship activities.

On the findings that concern the administrative or management staff, there were previous research which found that there were lecturers who taught entrepreneurship in universities who lack readiness or credibility in the area (Karimi et al., 2010; Olorundare & Kayode, 2014; Rahim et al., 2015; Welsh & Drăguşin, 2011; Yusoff et al.,

2014). This is in line with the data collected from the participants in this study, who said that the management is not really ready to take on this job and that some classes were too easy for them that it bores them. The fact that most students had to find information themselves shows that the management is still yet to master the knowledge of this area. This has also caused the students to not be exposed to the much-needed knowledge. This could also be one of the causes of miscommunication between the management and the students. This has been identified as one of the factors that may hinder the delivery of a quality entrepreneurship education in Nigeria by Chukwudi and Nwosu (2018) and in India by Joshi (2014) as well as in Finland (Seikkula-Leino et al., 2019). In comparison, universities in the United States of America have frequently employed real entrepreneurs to become educators of the subject, giving the students an opportunity to learn directly from the people in the industry (Moses & Akinbode, 2014). Students who participated in entrepreneurship clubs were also facing problems that relate to the management of students. The researchers found that it was difficult for students to convince other students to seek entrepreneurial knowledge. This has caused student participants to avoid purely joining because they wanted to learn, to just working in achieving all the club's main goals. This was also a similar concern that was brought up in previous researches involving the culture of the community surrounding the students in terms of entrepreneurship (Ferreira et al., 2016; Karimi et al., 2010; Olorundare & Kayode, 2014; Welsh & Drăguşin, 2011). Similarly, this research also found that participants struggle to find loyal and committed members who will not flee off whenever there is a problem. This might be because entrepreneurship may not be something that university students are interested in because of their different majors, or personal background and more. This was stated by one of the participants in the findings of this study, which can also be seen in the writings of previous research (Karimi et al., 2010; Ustyuzhina et al., 2019; Vanevenhoven, 2013; Welsh & Drăguşin, 2011; Yusoff et al., 2014).

Most of the participants also agreed that their experience can be regarded as participating in a life lesson. They felt that their hardships have made them better people, although it does not necessarily make them to become entrepreneurs in the future with one of them referring to entrepreneurship as only a potential 'side job'. Only a couple of them really stated that they have taken entrepreneurship activities seriously and as a job, referring to how they have already decided to become even before involving in the university's entrepreneurship activities as a couple of participants said that they may 'continue with entrepreneurship' and that they 'have always wanted to become an entrepreneur'. Participants also commented on their struggles in their studies. Some of them missed classes, some of them have seen the deterioration of their academic performance, as reflected in their decreased cumulative grade point average score (CGPA). However, they think that it is part of student life which does not need to be all academic-related activities, and a number of them also think that the life outside the university is harsher than it is inside. Pittaway et al. (2015) and Wang (2013) have also said that despite the challenges of managing other students as club leaders, these students walked away feeling that they achieved more in terms of experience and skills rather than when they are not involved. The researchers have also reported that students felt like they are valued more in employment than their counterparts in university due to their experiences in these student entrepreneurship clubs.

There are limitations to this study, one of it being the type students who became the participants to this research. There are many other students who do not affiliate themselves to the university's entrepreneurship development centre but still got involved in entrepreneurship. They may present different types of data than the ones collected from the students who become members of entrepreneurial clubs of the university. This may lead to the inappropriateness to generalise the findings to other settings, as this may also not be the case in other universities. Other limitations to this study are that it is interviewing participants who has time constraints. This has caused the interview to be conducted in a limited time.

Conclusion

Students who have involved themselves in entrepreneurship at their universities have indeed gone through challenges along the way. There were some key findings gathered in this research, divided into two categories; external and personal challenges. Students have found that it is a difficult task to deal with the management aspects of entrepreneurship activities, which have resulted to other problems such as the feeling of stress and the inability to continue as members of their club. One finding in particular talks about how the environment in the university has hindered them from expanding the club to a larger scale. Most of them have also told about the way they searched for certain information, which was not given to them by the management. They took it to themselves to add more knowledge and experience in entrepreneurship. They have also explained how they were drained physically and mentally from all the work that they have done. Most of them have also struggled academically but felt that it was understandable and acceptable considering how students should actually be spending their student life by not only dedicating themselves only to academic work. These findings are

consistent with. Rahim et al. (2015) and Rengiah and Sentosa (2016) have discussed the issues of entrepreneurship educator and the readiness of universities in entrepreneurship education.

The findings are essential to inform the development, change or improvement of entrepreneurship policies. Entrepreneurship education programs must continue to be reviewed so as to ensure that students are equipped to meet the challenges they might face when participating in the program. It is, hence, recommended that this type of research to be conducted at other universities and perhaps with wider set of criteria for participant selection. This is because this finding is generalizable to every setting. Other universities or students who do not affiliate themselves with the entrepreneurship development center may show different data due to the different ways they operate. These diverse selections of participants may bring researchers to new findings.

In conclusion, the research has shed some light into identifying the challenges these students face in their entrepreneurial journey. The researchers believe that there are still more to be explored in order to understand why graduates do not opt to become an entrepreneur later after graduation as suggested in many researches. More work in this area of research will surely bring researchers closer to the understanding of the situation the youth around the world have at hand, especially young Malaysians, in terms of entrepreneurship involvement. Future researchers are recommended to collect data from student entrepreneurs who are not affiliated to university entrepreneurship clubs to explore the bigger picture of the challenges faced by university students in entrepreneurship. It is also recommended that researchers consider better planning in interviewing participants. This is so in order to enable researchers to collect a more detailed and in-depth data from the participants, avoiding matters such as participants' time restraint, environmental noises and more. Researchers may also interview educators of entrepreneurship education or relevant university personnel who are responsible for entrepreneurship in universities. This would really help relevant authorities to work about the ways to solve existing issues faced in the efforts in nurturing entrepreneurial competencies among university students and graduates.

References

Bagheri, A., & Pihie, Z. A. L. (2012). An exploratory study of entrepreneurial attributes among Malaysian university students. *Life Science Journal 2012*, *9*(3), 2358-2365.

Bagheri, A., Pihie, Z. A. L., & Krauss, S. E. (2013). Entrepreneurial leadership competencies among Malaysian university student entrepreneurial leaders. *Asia Pacific Journal of Education*, *33*(4), 493–508. https://doi.org/10.1080/02188791.2013.822789

Barnard, A., Pittz, T., & Vanevenhoven, J. (2018). Entrepreneurship education in U.S. community colleges: A review and analysis. *Journal of Small Business and Enterprise Development*. https://doi.org/10.1108/JSBED-06-2018-0178

Braun, V., & Clarke, V. (2014). What can "thematic analysis" offer health and wellbeing researchers? *International Journal of Qualitative Studies on Health and Well-being, 9*(1). https://doi.org/10.3402/qhw.v9.26152

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage Publications.

Chiekezie, O., Nzewi, H., & Erhinmwionose, I. (2016). Entrepreneurial skill acquisition and job creation in Benin City, Nigeria. *EPRA International Journal of Economic and Business Review*, *4*(6).

Chukwudi, J., & Nwosu, J. (2018). Entrepreneurship education and the challenges of graduate employability in Nigeria. *International Journal of Innovation Management*, *9*, 189-193. https://doi.org/10.18178/ijimt.2018.9.5.812

Creswell, J. W. (2011). *Educational research: Planning, conducting, and evaluating quantitative*. Prentice Hall.

Din, B. H., Anuar, A. R., & Usman, M. (2016). The effectiveness of the entrepreneurship education program in upgrading entrepreneurial skills among public university students. *Procedia - Social and Behavioral Sciences*, *224*, 117–123. https://doi.org/10.1016/j.sbspro.2016.05.413

Department for Business, Innovation and Skills. (2013). *Enterprise education impact in HE and FE - final report*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/2087 15/bis-13-904-enterprise-education-impact-in-higher-education-and-further-education.pdf

Fleming, J., & Zegwaard, K. (2018). Methodologies, methods and ethical considerations for conducting research in work-integrated learning. *International Journal of Work-Integrated Learning*, *19*, 205-213.

Frese, M., & Gielnik, M. (2014). The Psychology of entrepreneurship. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*, 413-438. https://doi.org/10.1146/annurev-orgpsych-031413-091326

Ferreira, A., Loiola, E., & Guedes-Gondim, S. (2017). Motivations, business planning, and risk management: Entrepreneurship among university students. *RAI Revista de Administração e Inovação*.

Howitt, D., & Cramer, D. (2017). *Research methods in psychology* (5th ed.). Pearson Education Ltd.

Iqbal, A., Melhem, Y., & Kokash, H. (2012). Readiness of the students towards entrepreneurship: A case of Saudi private university. *European Scientific Journal, 8*(5).

Joshi, R. (2014). Entrepreneurship education: Core, context and challenges. *Journal of Entrepreneurship & Management, 3*(2).

Karimi, S., Chizari, M., Biemans, H., & Mulder, M. (2010). Entrepreneurship education in Iranian higher education: The current state and challenges. *European Journal of Scientific Research, 48*.

Khazanah Research Institute. (2019). *A statistical snapshot of youth unemployment, 2011 to 2018*. http://http://www.krinstitute.org/Views-@-A_Statistical_Snapshot_of_Youth_Unemployment,_2011_to_2018.aspx

Laurikainen, M., Silva, F., Schlember, P., Soares, J., & Melo, L. (2018). Entrepreneurship education: What can we learn from the Brazilian and the Finnish examples? *Revista Ibero-Americana de Estudos em Educação, 13,* 337-358. https://doi.org/10.21723/riaee.nesp1.v13.2018.11414

Lorcu, F., & Erduran, G. Y. (2016). Soul of entrepreneurship, entrepreneurship education? *International Journal of Social Sciences and Education Research, 2*(3).

McGuigan, P. (2016). Practicing what we preach: Entrepreneurship in entrepreneurship education. *Journal of Entrepreneurship Education, 19*.

Morris, M., Shirokova, G., & Tsukanova, T. (2017). Student entrepreneurship and the university ecosystem: A multi-country empirical exploration. *European Journal of International Management, 11,* 65-85. https://doi.org/10.1504/EJIM.2017.081251

Moses, C., & Akinbode, M. (2014). Entrepreneurship curriculum and pedagogical challenges in captivating students' interest towards entrepreneurship. *The International Journal Research Publication's Research Journal of Economics & Business Studies, 4*(1).

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*. https://doi.org/10.1177/1609406917733847

Olorundare, S., & Kayode, D. (2014). Entrepreneurship education in Nigerian universities: A tool for national transformation. *Asia Pacific Journal of Educators and Education, 29*.

Othman, N., Hashim, N., & Wahid, H. A. (2012). Readiness towards entrepreneurship education. *Education Training, 54,* 697–708. https://doi.org/10.1108/00400911211274837

Oparaocha, G.O., & Pokidko, D. (2013). Exploring the impact of higher entrepreneurship education: Critical experiences as triggers or deterrents for career perceptions of graduates.

Ozdemir, P. (2018). Internal and external interactions of entrepreneurial education ecosystems of universities: A study at Turkish universities. *The Online Journal of Quality in Higher Education, 5*(3).

Pittaway, L., Gazzard, J., Shore, A., & Williamson, T. (2015). Student clubs: Experiences in entrepreneurial learning, *27,* 1-27. https://doi.org/10.1080/08985626.2015.1014865

Rahim, H. L., Abdul Kadir, M. A. B., Zainal Abidin, Z., Junid, J., Mohd Kamaruddin, L., Mohd Lajin, N. F., & Ahmad Bakri, A. (2015). Entrepreneurship education in Malaysia: A critical review. *Journal of Technology Management and Business, 2*(02).

Rengiah, P., & Sentosa, I. (2016). An empirical study on the effectiveness of entrepreneurship education in developing entrepreneurial intention among Malaysian university students. *Education Journal, 5*(4), 53-61. https://doi.org/10.11648/j.edu.20160504.13

Ridzwan, R., Muhammad, N. M. N., & Rahman, A. A. A. (2017). Issues and challenges among young entrepreneurs in Malaysia. *IOSR Journal of Business and Management, 19*(03), 80–84. https://doi.org/10.9790/487x-1903028084

Roshaidai, S. (2018). Ethical considerations in qualitative study. *International Journal of Care Scholars, 1*(2).

Sanchez, J.C., Ward, A., Hernández, B., & Florez, J. (2017). Entrepreneurship education: State of the art. *Propósitos y Representaciones, 5*(2), 401 - 473. http://dx.doi.org/10.20511/pyr2017.v5n2.190

Sandhu, M. S., Sidique, S. F., & Riaz, S. (2011). Entrepreneurship barriers and entrepreneurial inclination among Malaysian postgraduate students. *International Journal of Entrepreneurial Behavior & Research, 17*(4). 428–449. https://doi.org/10.1108/13552551111139656

Saunders, M., Lewis, P., & Thronhill, A. (2012). *Research methods for business students* (4th ed.). Pearson Education Ltd.

Seikkula-Leino, J., Ruskovaara, E., Pihkala, T., Rodríguez, I., & Delfino, J. (2019). Developing entrepreneurship education in Europe: Teachers' commitment to entrepreneurship education in the UK, Finland and Spain. *In the role and impact of entrepreneurship education*. Edward Elgar Publishing. https://doi.org/10.4337/9781786438232.00014

Ustyuzhina, O., Mikhaylova, A., & Abdimomynova, A. (2019). Entrepreneurial competencies in higher education. *Journal of Entrepreneurship Education, 22*(1).

Vanevenhoven, J. (2013). Advances and challenges in entrepreneurship education. *Journal of Small Business Management*. http://dx.doi.org/10.1111/jsbm.12043

Wang, X. (2013). Reflection on challenges and countermeasure of entrepreneurship education in China. *Cross-Cultural Communication*, *9*(2), 18-22.

Welsh, D. H. B., & Drăguşin, M. (2011). Entrepreneurship education in higher education institutions as a requirement in building excellence in business: The case of the University of North Carolina at Greensboro. *Forum Ware International Journal*, *1*, 266-272.

Yusoff, M. N. H. B., Zainol, F. A., & Ibrahim, M. D. B. (2014). Entrepreneurship education in Malaysia's public institutions of higher learning—A review of the current practices. *International Education Studies*, *8*(1). https://doi.org/10.5539/ies.v8n1p17

Zhou, M., & Xu, H. (2012). A review of entrepreneurship education for college students in China. *Administrative Sciences, 2*, 82-98. https://doi.org/10.3390/admsci2010082

The Development and Validation of the Emotional Literacy Skills Scale

Melek Alemdar[1], Hüseyin Anılan[2]
[1]National Ministry of Education, Turkey
[2]Eskişehir Osmangazi University

To cite this article:

The Development and Validation of the Emotional Literacy Skills Scale*

Melek Alemdar[1], Hüseyin Anılan[2]
[1]National Ministry of Education, Turkey
[2]Eskişehir Osmangazi University

Abstract

Emotional literacy is the ability to understand, express and regulate emotions in social contexts. It emphasizes the ability to communicate with certain feeling words in interpersonal relationships. It stands as a bridge between the thoughts and emotions of school stakeholders, contributing to more effective learning, safer schools, and a democratic climate (O'Hara, 2011). It is possible to teach emotions academically, even to make them a part of the curriculum (Antidote, 2003); however, it is essential to also see emotional literacy as a vital skill through the values of the school and the behaviors of teachers. Therefore, it is valuable to evaluate the emotions of teachers in the school environment. To do this, it was the aim of this study to develop a reliable and valid measurement tool -- the Emotional Literacy Skills Scale (ELSS). The validity of the model was confirmed via exploratory factor analysis and confirmatory factor analysis. We found a Cronbach Alpha reliability coefficient of .85. The findings showed that the Emotional Literacy Skills Scale is valid and reliable and formed by five factors: Motivation, empathy, self-regulation, emotional awareness, and social skills.

Key words: Emotional literacy skills, Scale development, Affective domain

Introduction

A high-quality education in a country facilitates the economic, political, social, and cultural development of the society. Education does not offer a mass production system, but it is the most powerful and effective instrument that provides a "human" element for the needs of every institution and organization (Kayadibi, 2001, p. 74). "Teachers, the new generation will be your devotion," said Mustafa Kemal Ataturk, the founder of the Turkish Republic, to emphasize the teachers' qualifications in educating people. As seen in the Great Leader's words, the quality of education is directly related to the quality of teachers. Due to economic, social, and technological developments, new demands in education have also led to different expectations from the teachers. This has caused drastic changes and transformations in the teachers' identity. In the light of scientific and technological developments, it is of course a priority for teachers to adopt novel pedagogies and to support their social and professional development. However, teaching and learning are not just processes of cognition, knowledge, and skill -- they are emotional practices as well (Hargreaves, 2001). Teachers, the mediators of achieving affective goals in society, need to have competencies to recognize, regulate and express their feelings. Emotionally competent teachers set the tone of the classroom by establishing supportive and encouraging relationships with their students, coaching in conflict resolutions, encouraging collaborative work among students and acting as a role model for healthy communication and prosocial behavior (Gong et al., 2013).

In Turkey, teacher competencies are determined by the Ministry of National Education (MONE). The report, which sets a framework for teacher education policies, published by MONE (2017) announced general competencies for teaching profession as:

> Every teacher should have strong communication skills, be able to plan the teaching effectively, possess the required professional skills, and have substantial intellectual knowledge regarding his/her subject field. These qualifications, which are expected from a teacher to perform his/her profession properly, form the basis of teacher competencies (p. 7).

When we consider the competencies explained above, we can infer that the social processes, in which teachers use the affective skills, are ignored. In the past Şişman (2009, p. 76) criticized this situation as "Values corresponding to the affective domain were not included among the teacher competencies; perhaps because it was not possible to measure". Concerned by a lack of focus on the affective domain in teachers we decided to develop a valid and reliable scale to measure, interpret, or define teachers' emotional states, based on the concept of emotional literacy.

In the spirit of this research, emotional literacy refers to the sum of social, emotional, and behavioral skills that are necessary in all areas of the individual's school, family, and social life (Alemdar, 2018). This article reviews critically what we currently know (and do not know) about the concept and creates a measurement tool to determine how it is schooled in teachers' behaviors.

Theoretical Framework

Emotional Literacy

> *Fish swim, birds fly, and people feel.* Sometimes we are happy, sometimes we are not; but sometimes in our lives, we are sure to *feel anger and fear, sadness and joy, greed and guilt, lust and scorn, delight and disgust.* While we are not free to choose the emotions that arise in us, we are free to choose how and when to express them, provided we know what they are (Ginott, Ginott & Goddard, 2003, p. 27).

Ginott et al. (2003) originates emotional literacy implicitly by emphasizing that emotions are a genetic heritage for humans. Emotions, if used as data sources, may serve to add meaning to real-life situations and to regulate relationships. So, it is important to be aware of emotions and to make the best use of their power. To this end, a significant growing body of research has recently been directed toward resolving conceptual and measurement issues related to emotional literacy skills (Antidote, 2003; Killick, 2006; Park, 1999; Weare, 2004).

Emotional literacy was first used by Steiner (1979) in his book 'Healing Alcoholism', referring to a skill which may solve affective complexities in and around people and requires behavioral results in social contexts. After its foundations laid by American Humanist Psychology studies (Park, 1999; Weare and Gray, 2003), the term gained popularity in Britain with psychotherapist Susie Orbach's article in the 1970s. Steiner's (2003, p. 11) famous definition "understanding your personal strength, quality of life, and your emotions to improve the quality of life of the people around you" highlighted the term and introduced it into the international research agenda. Various studies stress the ability to recognize, understand, handle and express emotions appropriately (Sharp, 2001; Weare, 2004). Antidote (2003), which is known for its emotional literacy studies, refers to emotional literacy as the practice of thinking about how emotions shape behaviors at an individual and collective level. This strong rationalist assumption positions the term at the core of the action, by stressing that thoughts enriched with emotions are the key to increasing emotional competence. Emotions prompt thinking skills rather than interfering with them, and creative, critical, and positive thinking styles, in turn, serve to shape emotional situations. Joseph, Strain, and Ostrosky (2005) emphasized the capability to describe, understand and respond to feelings in oneself and others. As well as depicting an active role on an individual basis, this expression underlines an effective use of the skills mentioned above to react to the emotional states of others. Emotional literacy requires competency to communicate with certain emotion words in interpersonal relationships. Based on that, Orbach (1998) simply summarizes it as the ability to ask "how are you?" and listen to the answer. In this research, the concept of emotional literacy has been used in a context that emphasizes a person's ability to acknowledge and express feelings, and to communicate through specific emotion words in interpersonal relationships. This definition differentiates between personal and social talents, underlining a whole set of skills.

Knowing how to express feelings tactfully is vital to explaining how we communicate and behave, what kind of a person we want to be, and why we should behave in a certain way on particular occasions (Alemdar, 2014). Understanding the universe within the context of emotions, here, involves enjoying our own emotions, listening to and responding to the needs of others, and correcting our emotional damage (Matthews, 2006). To meet all these emotional needs, it is necessary to carefully consider "how we feel", that is, to read emotions like a book. The definitions above indicate a similarity between the theories of emotional literacy and emotional intelligence. In the literature, although some studies suppose that these two concepts can be used in parallel or interchangeably (Bocchino, 1999; Claxton, 2005; Killick, 2006; Perry, Lenny, & Humphrey, 2008), especially in the field of education, a clear distinction has been made by researchers (Dickson & Burton, 2011; Matthews, 2006; Park, 1999; Weare, 2004).

The Differences between Emotional Literacy and Emotional Intelligence

Emotional literacy is a term that was first developed and mentioned in American Humanist Psychology studies in the 1970s. So, it is claimed that it has a longer history compared to emotional intelligence (Park, 1999). Also, these two concepts are used to indicate similar meanings in different geographies. Emotional literacy is frequently used in various projects, researches, publications, and conferences in the UK, while emotional intelligence appears more dominant in the United States (Carnwell & Baker, 2007; Holmes, 2016; Weare, 2004).

Emotional intelligence touches on the ability to deal thoroughly with emotions, while emotional literacy refers to communicating through emotional vocabulary (Alemdar, 2014). Steiner (2003) stated that emotional literacy is centered in the heart and this is the most important thing distinguishing it from emotional intelligence (Holmes, 2016). While emotional intelligence expresses an innate personality dynamic to be nurtured, emotional literacy is the unity of understanding, strategy, and skills that a person can develop throughout life (Mader, 2005). We can say that emotional intelligence is a characteristic or potential, on the other hand, emotional literacy is an understandable, learnable, and improvable skill. Emotional literacy may contribute to processing emotional intelligence capacity, e.g., empathy as a personality dynamic represents the emotional intelligence and using specific strategies and methods to increase this potential empathy and also employing it in socio-cultural processes represent emotional literacy (Kandemir & Dündar, 2008).

Emotional literacy requires taking responsibility both to understand our feelings and to organize our social relationships. Some studies have differentiated between the two concepts based on the semantic differences between literacy and intelligence terms (Matthews, 2006; Southampton Psychology Service, 2003). The term "intelligence" has accrued a negative connotation that would undermine the positive message (Sharp, 2001) and tends to suggest a capacity that is innate and fixed, not teachable (Ripley & Simpson, 2007; Weare & Gray, 2003). Matthew's (2006) semantic differentiation has asserted that the term "literacy" is more related to language and the culture that can be improved by the use of language. Weare (2004, p. 2), also, has stated that those who are familiar with the term "literacy" (especially primary school teachers and language teachers) can get ideas from verbal literacy on how to define and teach social and emotional skills in the context of emotional literacy. The Southampton Psychology Service (2001) has supported this idea as the emotionally literate person should be able to name and read symbols and signs (psychological signs, facial expressions, other forms of non-verbal communication). Flynn (2010) also thinks that the term "literacy" is malleable for emotional development while emotional intelligence evokes a traditional fixed manner.

Emotional literacy is more relevant to education than the more commonly used emotional intelligence (Meekums, 2008; Stone, 2005); it focuses on the emotional health of the learner in emotionally related situations and qualifies the social environment in educational institutions (Coşkun, 2015). When we focus on the literature derived to differentiate between these two terms, emotional literacy seems more appropriate to the idea we want to impart through this study, as it emphasizes the ability to understand, manage and develop these skills over time in educational contexts (Rae, 2012).

Key Elements of Emotional Literacy

Various models of emotional literacy offer a certain level of welfare for schools and individuals' social life in many different contexts (Alemdar, 2014; Coşkun, 2015; Matthews, 2006). The notion of emotional literacy and its prominent components was first introduced by Steiner (1979, p. 19) as:
- Knowing your feelings,

- Having a sense of empathy,
- Learning to manage emotions,
- Eliminating emotional damage,
- The ability to integrate these four traits: emotional interactivity.

One of the major models to emerge was that of Faupel (2003). The two main domains in his model are personal competencies and social competencies. Personal competencies include self-awareness, self-regulation, and motivation, while social competencies cover empathy and social skills, and thus the model consists of five components. (Cited in, Killick, 2006, p. 12).

Weare (2004, p. 23) suggests that an individual's emotional literacy skill consists of some overlapping social and emotional competencies that can be divided into three basic groups:
- Self-understanding,
- Understanding, expressing and managing our emotions,
- Understanding and making relationships.

These three dimensions contain sub-skills that can be developed throughout childhood and adulthood in school and learning life. They include, positive and realistic self-concept, a sense of optimism, expressing emotions, social bonding (loving and trusting others), empathy, effective communication, etc. (Weare, 2004). This model focuses on practices and social environments aiming to improve the individual's emotional abilities.

Uzan (2018) summed up skills included in the emotional literacy models as: the individual capacity to name and characterize emotions felt, understanding why those emotions are felt, expressing them to the other people plainly, and taking control of emotions. Similarly, in their study, Kandemir and Dündar (2008) gathered emotional literacy components as: empathy, self-regulation, self-motivation, social skills, emotional awareness, managing emotions, removing emotional damage (emotional regulation), and problem-solving.

When the literature is examined, we can see two scale development studies (Akbağ, Küçüktepe, & Özmercan, 2016; Palanci, Kandemir, Dundar, & Özpolat, 2014). Both of those scales were developed focusing on similar traits in college students' samples. As we are more interested in the affective domain in teachers' context, in which we want to determine the current condition of teachers in Turkey and the effects of other variables, this study aims to develop a scale of emotional literacy with the validity and reliability studies.

Method

This study is a scale development study. This part includes stages of development of the Emotional Literacy Skills Scale and the features of the participants.

Study Group

The study involves teachers working in central districts of Eskişehir in the 2017-2018 academic year; 373 participants for exploratory factor analysis, 399 participants for confirmatory factor analysis. The teachers were chosen from different types of schools, e.g., Science High School, Social Sciences High School, Anatolian High School, Vocational and Technical Anatolian High School, Tourism Vocational High School, Trade Vocational High School, and Imam Hatip High School. The demographics of the research sample are presented in Table 1.

Table 1. Demographics of Participants

		Exploratory Factor Analysis		Confirmatory Factor Analysis	
	Category	N	%	N	%
Gender	Female	144	38,6	223	55,9
	Male	215	57,6	176	44,1
	Unstated	14	3,8	-	-
	Total	373	100	399	100
Graduation	Education Faculty	254	68,2	226	56,6
	Faculty of Arts and Sciences	99	26,5	149	37,4
	Others (Vocational Ed. Fac., etc.)	12	3,2	24	6,0
	Unstated	8	2,1	-	-
	Total	373	100	399	100

	1-5 Years	22	5,9	52	13,0
	6-10 Years	51	13,6	76	19,0
Length of	11-15 Years	61	16,3	72	18,0
Service	16-20 Years	78	20,9	107	26,8
	21 Years +	68	18,2	82	20,6
	Unstated	93	25,1	10	2,5
	Total	373	100	399	100

The data regarding the gender of the teachers participating in the study shows that more males attended in exploratory factor analysis, while more females took part in confirmatory factor analysis. In both studies, the majority of the teachers are graduates of the Education Faculty and their length of service seems more or less the same.

Process

The Emotional Literacy Skills Scale (ELSS) was developed in the context of teachers, and it is hoped to be used in adult samples by conducting validity and reliability studies. The scale development process should comprise several stages (DeVellis, 2012). The first step is to determine the purpose of the scale and to define the target audience (Şencan, 2005). Since the target audience of this study is teachers and the aim is to explore their emotional literacy skills, a detailed literature review has been done on the issue. The characteristics found in the most prominent studies were identified as "motivation, empathy, self-regulation, emotional awareness, and social skills" (Antidote, 2003; Killick, 2006; Steiner, 2003; Weare, 2004). In the following step, the items based on expressing the extent and content of sub-skills were written. Based on five dimensions, a total of 63- item pool was created. As stated in DeVellis (2014), having replacement items that represent the same item in the testing process provides strong items in the final version of the scale in case of omitting other items. The items were examined through a focus group interview with three faculty members working in separate fields and eight teachers who have a Master's or Ph.D. in Curriculum and Instruction. Five teachers in the focus group were also language experts, so they examined the items in terms of semantics. One of the faculty members works as an Assessment and Evaluation Specialist in education, the second one is a subject area specialist who had theses in emotional intelligence, and the last one is experienced in scale development in curriculum and instruction studies.

The items were evaluated in terms of content, meaning, and clarity of expression. Through 63 items, three items were removed due to the overlapping and contradictory expressions. Doctoral thesis monitoring committee members analyzed the items to affirm the content validity. The final 60-item-draft was revised in line with experts' opinions and feedbacks and was arranged using a five-point Likert scale ("1- Never True", "2- Rarely True", "3-Often True", "4- Usually True" and "5- Almost Always True ").

IBM-SPSS 21 program for exploratory factor analysis, Lisrel 8.7 program for confirmatory factor analysis were used. The mean series were assigned for the missing data and then the normality of the data was tested. The data showed a normal distribution for each item since the values of kurtosis (ranging from .00 to -.61) and skewness (ranging from -.25 to -.43) were within acceptable limits (Kline, 2005; Hair et al., 2010). The conducted validity and reliability studies of the scale have been reported.

Results

Validity

Factor analysis is used to provide clues about the structure of the relationships between variables that are thought to be related (DeVellis, 2012). When the scale items match up with the theoretical knowledge in the literature, the factor analysis method should be applied before the reliability analysis if the aim is to develop a multi-dimensional scale measuring complex conceptual structures (Şencan, 2005). There are two types of factor analyses: exploratory and confirmatory. Exploratory factor analysis, which was often used to develop psychological tests at the very beginning, is an effective technique used to reveal longitudinal relationships in complex, multiple factored structures like intelligence, skills, etc. (Rasch, Kubinger, & Yanagida, 2011). Considering that the emotional literacy skill is made up of multiple dimensions, firstly exploratory factor

analysis was applied to determine the number and the nature of the factor structure. Principal Component Analysis method, which is used in social sciences as a factoring technique, has been used (Büyüköztürk, 2011).

Exploratory factor analysis (EFA)

For Exploratory Factor Analysis, the suitability of the sample for factor analysis was evaluated first. Suggested minimums for sample size include views like there should be at least five participants for each item (Şencan, 2005) or 300 people in total are sufficient for factor analysis (Field, 2013). So, a sample group of 373 people was considered as suitable for factor analysis, supported by a sufficient .83 Kaiser Meyer Oklin value. Also, Bartlett Test of Sphericity results ($\chi 2$ (373)= 2817.718 df:435 p<.01) indicated that the data set ensures multivariate normality. In the Principal Components Analysis, direct oblimin, was used as the factor rotation. The analysis was repeated several times; two items were removed due to the factor load below .30 and loading to more than one factor. In the EFA, the slope graph and the eigenvalues of the factors are also used for determining the number of factors. Figure 1 suggests a five-factor structure.
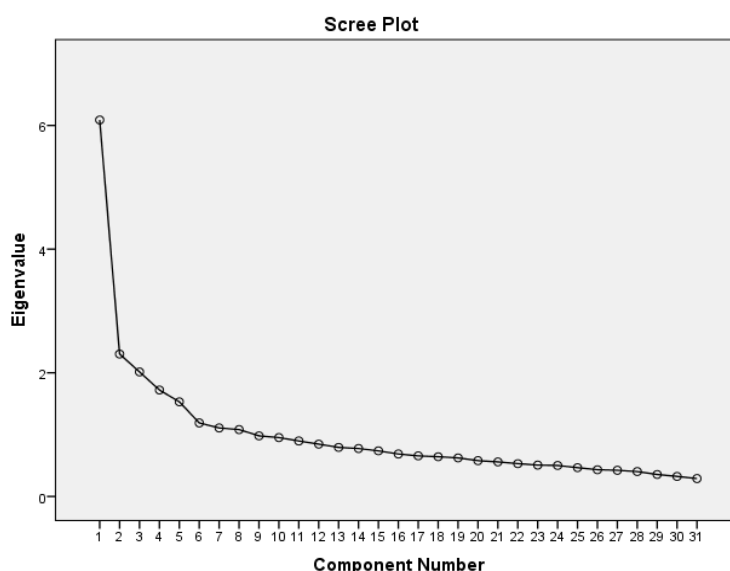
Figure 1. Scree Plot Chart of ELSS

Table 2. Exploratory Factor Analysis Results of ELSS.

Items	Factor Loadings				
	Motivation	Empathy	Self-regulation	Emotional awareness	Social skills
1. I believe I have to work hard to achieve a quality of life.	.689				
2. I consider most of the things I do as a waste of time.	.590				
3. I fulfill my duties or responsibilities properly.	.580				
4. I easily get distracted while completing a task.	.539				
5. I do everything with the thought that it contributes to me.	.527				
6. I believe each new day is a new opportunity to improve myself.	.519				
7. I leave things to the last minute.	.426				
8. I have the ability to articulate		.611			

the feelings of the people around me.					
9. I easily realize when someone is hiding his/her true feelings.	.577				
10. In an argument, I take into consideration the opinions of other people.	.480				
11. When I hurt someone, I can express my sadness.	.478				
12. I behave with an awareness of how impacts my attitudes have on other people.		.670			
13. I discern the difference between my feelings and behaviors.		.600			
14. When I lose a game or race, I question the underlying reasons.		.553			
15. If I feel inadequate about something, I immediately try to make up this deficiency.		.533			
16. Even when I'm angry, I weigh up all the options before doing something.		.531			
17. If the method I adopt while doing something does not prove effective, I try other ways out.		.521			
18. It is hard for me to identify my feelings.			.690		
19. I have difficulty in expressing my worries.			.632		
20. Sometimes I say, "I wish I were someone else."			.589		
21. I can express my feelings clearly.			.560		
22. I consider people around me have difficulty in understanding me.			.549		
23. I can plainly say why I am sorry.			.475		
24. I enjoy spending time with a circle of friends.				.716	
25. I care about the way I communicate with my friends.				.707	
26. I feel happy when I share things with my friends.				.662	
27. I often feel left out by my friends.				.623	
28. I spend most of my time alone.				.528	
29. I fail in social relations with my friends.				.524	
30. I care about having eye-contact during a communication.				.443	
31. I'm regarded as "social" by those around me.				.436	
Eigenvalue	2,01	1,53	1,72	2,30	6,08
Variance Explained	8,52	7,67	8,31	9,69	9,85

The results of the analysis are presented in Table 2. The factor loadings related to ELSS show that "Motivation" dimension has seven items (1, 2, 3, 4, 5, 6, 7), "Empathy" dimension has four items (8, 9, 10, 11), "Self-regulation" dimension has six items (12, 13, 14, 15, 16, 17) "Emotional awareness" dimension has six items (18, 19, 20, 21, 22, 23) and "Social skills" dimension has eight items (24, 25, 26, 27, 28, 29, 30, 31). The factor loading value, which is the coefficient that explains the relationship between items and factors, should be above the limit value of 0.30 (Field, 2013; Şencan, 2005). Factor structures in ELSS seem to be appropriate since the factor loads for each item are above .40 and there are no contradictory items.

The eigenvalues of the factors range between 6.08 and 1.53, and the total variance explained by five factors is 44%. In the studies of social sciences, a total variance explanation above 40% in factor analysis is regarded as acceptable (Büyüköztürk, 2011; Kline, 1994). As a result of factor analysis, it was seen that emotional literacy indicated a five-factor structure. These factors are named as "Motivation", "Empathy", "Self-regulation", "Emotional Awareness" and "Social Skills", in line with the characteristics of items and statements in the literature (Antidote, 2003; Killick, 2006; Steiner, 2003; Weare, 2004).

Confirmatory Factor Analysis (CFA)

In order to strengthen the construct validity of the model which was obtained through EFA, a confirmatory factor analysis was applied on a second data set taken from 399 teachers (N=399). This analysis was carried out by using the scale form consisting of 31 items. The results achieved by analysis of the constituted model via CFA were [$\chi2$/sd (1840.36/424) = 4.34, p = .001, IFI = .87; CFI = .87; RMSEA = .092] and they were found out of acceptable limits (Çokluk, Şekercioğlu, & Büyüköztürk, 2014). The modification recommendations showed that the fit indices could be improved by performing six modifications on the model The modifications were made between the items belonging the same dimensions (2 and 4; 18 and 19; 8 and 9; 5 and 6; 27 and 29; 30 and 31). When the items to be modified were checked out, it was seen that they reflect similar situations. The model was retested by adding error covariance among the items.

Table 3. The Fit Parameters of ELSS related to the CFA Model

Fit Parameters	Coefficient
IFI	0.90
NFI	0.89
GFI	0.89
CFI	0.90
RMSEA	0.07
Sd	418
SRMR	0.07
$\chi2$	1423.34
$\chi2$/sd	3.4

The fit indices obtained were [$\chi2$ / sd (1423.34/418) = 3.41, p = .001, CFI = .90; GFI = .89, IFI = .90; RMSEA = .07 (Confidence interval for RMSEA = .073 – .082)] and the model indicated a good fit. Standard values for the indices were: GFI and IFI values should be between 0 and 1. Although there is no agreement in the literature concerning these values, if the value is close to 1, it indicates excellent fit (West, Taylor & Wu, 2012) and values between 0.80 and 0.89 indicate a good fit (Doll, Xia & Tarkzadeh, 1994; Frias & Dixon, 2005). RMSEA value also varies between 0 and 1 (Cole, 1987). If the value is closer to 0, it indicates a good fit and the value between 0.08 and 0.10 gives a moderate fit (Byrne, 2016). $\chi2$/df ratio indicates an excellent fit if it is lower than 2 and the value between 2-5 is a good fit (Jöreskog & Sörbom, 2001). Thus, all standardized fit indices indicated that the model factor structure was confirmed. The range of fit index for CFA is displayed in Table 3.
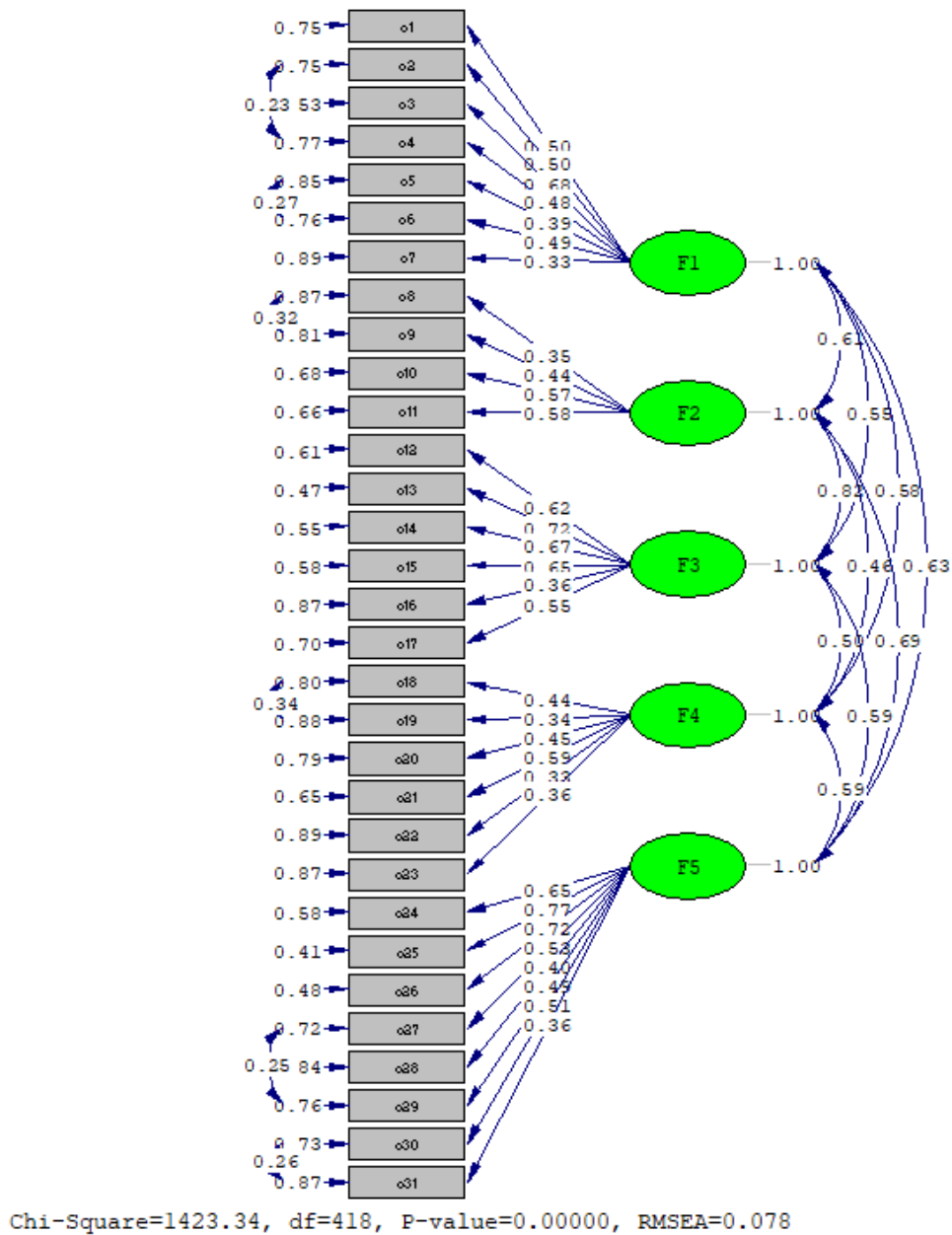
Figure 2. Path Diagram for CFA

The model obtained as a result of the analysis is presented in Figure 2. The item factor loads of ELSS varied between .33 and .77 and all factor loads were significant at the level of .001.

Sub-Upper scores at 27%, Anti-image correlations and t-test results between item-total correlations

Another method used to test the validity of the scale is the item-total correlation (item discrimination). Item total correlation scores are used in terms of interpreting how each item contributes to the phenomenon to be explained and how much it distinguishes in terms of the measured property. For item discrimination on Likert scales, generally, techniques based on the mean differences between 27% sub-upper groups and correlation values are used (Şahin & Gülleroğlu, 2013). In this study, both methods were used to determine item discrimination. The scores obtained from the scale were sorted in ascending, and two groups, sub-upper 27%, were formed. As a result of the independent group t-test, it was found that the difference between the lower and upper group means was significant (p <.01). Thus, the scale is distinctive in measuring the intended feature.

Table 4. Sub-Upper scores at 27%, Anti-image correlations and t-test results between item-total correlations

Items	Item correlation	Anti-image correlation	t	Items	Item correlation	Anti-image correlation	t
I-1	.300	.756	6,60	I-17	.511	.858	10,4
I-2	.402	.840	8,03	I-18	.348	.841	7,42
I-3	.407	.849	8,47	I-19	.373	.832	7,64
I-4	.375	.848	7,80	I-20	.352	.876	7,56
I-5	.281	.826	6,12	I-21	.477	.840	11,1
I-6	.458	.906	10,3	I-22	.342	.823	7,33
I-7	.330	.824	7,24	I-23	.399	.818	8,94
I-8	.316	.858	7,09	I-24	.347	.818	6,86
I-9	.254	.788	6,34	I-25	.415	.806	8,29
I-10	.310	.790	7,71	I-26	.422	.792	9,60
I-11	.341	.856	7,02	I-27	.462	.822	10,0
I-12	.321	.792	7,64	I-28	.383	.786	7,97
I-13	.452	.872	10,8	I-29	.442	.849	9,16
I-14	.359	.858	7,83	I-30	.352	.833	8,38
I-15	.395	.836	9,49	I-31	.409	.873	8,74
I-16	.300	.770	5,82				

In Table 4, the total correlation values of the items vary between .300 and .511 except for two items, and thus the items are well distinguished. Items with a total correlation of 0.30 and higher are said to distinguish well (Büyüköztürk, 2011). It was observed that the anti-image correlation values of the two items with the item-total correlation of .281 and .254 were above .70 and it was not necessary to subtract them because of their contribution to the scale. The values in the anti-image correlation matrix must be at least 0.50 (Sipahi, Yurtkoru, & Çinko, 2008). As seen in Table 4, the anti-image correlation values of the items range between 0.75 and 0.90, showing that items contribute to the factor structure of the scale at a high rate.

Reliability

To determine the scale's internal consistency, Cronbach's alpha (α) reliability coefficient was calculated. The Cronbach Alpha value takes the variance values of the research items into account and it is used to see if the responses of the participant are consistent within the scale (Bryman, 2012). Values related to reliability analysis are presented in Table 5.

Table 5. Number of Items, Correlations between Factors, Internal Consistency Coefficients for TABS Factors

Factors	Item Number	Cronbach's alpha
1- Motivation	7	.70
2- Empathy	4	.60
3- Self-regulation	6	.70
4- Emotional awareness	6	.71
5- Social skills	8	.77
Total	31	.85

N=373, **p<.01

The Cronbach's alpha (α) reliability coefficient for the scale applied to 373 teachers was found as .85. This value shows that the scale is highly reliable (Bryman, 2012, p. 170). The Cronbach Alpha internal consistency coefficients of the scale changed between .60 and .77 while the correlations between the factors were found to change between .20 and .50. Empathy sub scale's Cronbach's coefficient value is lower than the others, this can be interpreted as the number of items in that subscale are insufficient, yet it still appears to be within the lower limit of .60 in the literature (DeVellis, 2012).

Conclusion and Discussion

This study aimed to determine the validity and reliability of ELSS with 373 teachers for exploratory factor analysis and 399 teachers for confirmatory factor analysis. The study was carried out in seven stages: (i) detailed analysis of the theoretical structure of emotional literacy skills, (ii) creating the item pool, (iii) clarifying the format of the instrument, (iv) the experts' reviews of the items, (v) the validity and reliability analyses of the items (exploratory factor analysis, confirmatory factor analysis, internal consistency, and correlations between the factors.) (vi) applying the scale and (vii) evaluating the items to finalize the scale.

According to the studies on the theoretical structure, emotional literacy requires knowing the taste of feelings, and understanding the feelings in a way that can respond to the wishes and needs of others. Emotional literacy satisfies the need to clearly define and convey our emotions. The skills reflecting the characteristics of an emotionally literate person are frequently repeated in the literature as "motivation, empathy, emotional awareness, self-regulation, and social skills" (Antidote, 2003; Killick, 2006; Steiner, 2003; Weare, 2004). To operationalize these skills, an item pool was constructed and the items, with five-point Likert scale options, were evaluated by experts.

In the next stage of the study, the factor structure of the ELSS was examined and a 5-factor structure was revealed by EFA. Factors named in line with the literature are motivation (7 items), empathy (4 items), self-regulation (6 items), emotional awareness (6 items), and social skills (8 items). When the scale development studies on emotional literacy skills in national literature are examined, it was seen that three factors (self-regulation, emotional awareness, and social skills) identified in this study were also obtained in other studies (Akbağ, Küçüktepe, & Özmercan, 2016; Palancı et al., 2014). This implies that our findings are consistent with the scales developed in the context of Turkish culture samples. In addition to these three factors, it is thought that the motivation and empathy dimensions included in the study will have valuable contributions to emotional literacy skill studies in the sample of teachers. Because, in the context of emotional literacy skills, competent teachers are guides who develop encouraging relationships with their students, support students' strengths and abilities, solve problems in conflict situations, and act in a way to increase intrinsic motivation (Alemdar, 2019). In their study, Perry, Lenny, & Humphrey (2008, p. 35) defines the characteristics of an emotional literate teacher as a person who knows and understands his own feelings, listens to children, can use emotional language, understands that children have emotions and considers these, empathizes and provides a safe and comfortable environment for children to learn better.

Through CFA, statistically significant results and appropriate the Chi-square ($\chi2$) value relevant to the constructed model were determined [$\chi2$ = 1423.34, sd = 418, p <.01]. Depending on the degree of freedom, the low Chi-square ($\chi2$) value indicates that the scale items are suitable for the data collected (Jöreskog & Sörbom, 2001). In addition, the other fit indices of the model [CFI = .90, GFI = .89, IFI = .90, RMSEA = .07,] indicate a good fit with the proposed model. The coefficient obtained from CFI, IFI, GFI varies between 0 and 1. It is stated in the literature that these values are very sensitive to the sample size and reflect the perfect fit as they approach 1 (Raykov & Marcoulides, 2006, p. 46). The RMSEA value expressing the margin of error among the observed and produced matrices should be less than .10 (Cole, 1987, p. 586). For the RMSEA; the value that equals to or less than 0.05 is a good fit, between 0.05 and 0.08 is a sufficient fit and between 0.08 and 0.10 is a moderate fit (Byrne, 2016, p. 98).
The results of the reliability coefficients of each factor in the ELSS (.70, .60, .70, .71, .77) were statistically acceptable. These values show that ELSS can be used to evaluate teachers' emotional literacy skills (Bryman, 2012; DeVellis, 2012).

Consequently, the constructed theoretical model can be used for teachers, since the factors formed based on the data obtained from teachers for ELSS were confirmed by validity and reliability findings. ELSS serves to measure the desired feature and can be used in adult groups especially in teacher samples.

Recommendations

➢ The data to develop emotional literacy skills scale for this study was gathered from teachers as professionals. For further studies, different professional groups of samples can be used to have a valid and reliable tool.

➢ The study group of this research consists of teachers working in public high schools. In future researches, teachers from different school levels and also from private schools can be sampled.
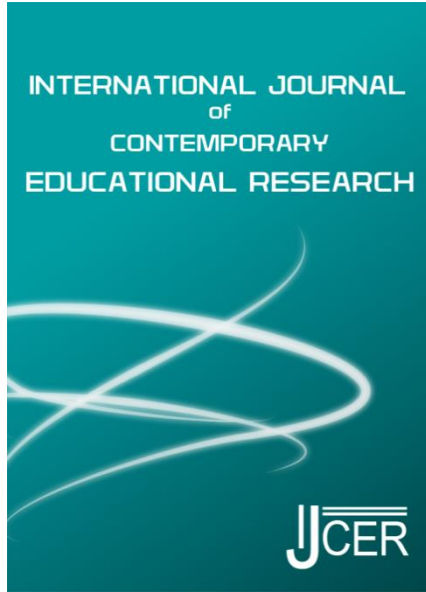
Acknowledgements

References

Alemdar, M. (2014). *Examining the emotional literacy skill levels of high school students* (Unpublished master's dissertation). Eskişehir Osmangazi University, Eskişehir.

Alemdar, M. (2018). Examining the emotional literacy skill levels of high school students. *Educational Policy Analysis and Strategic Research*, *13*(2), 69-86. DOI: 10.29329/epasr.2018.143.4

Alemdar, M. (2019). *Reflection of social capital in educational processes: Emotional literacy and emotional labor context* (Unpublished doctoral dissertation). Eskişehir Osmangazi University, Eskişehir.

Akbağ, M., Eminoğlu Küçüktepe, S., & Eminoğlu Özmercan, E. (2016). A study on Emotional Literacy Scale Development. *Journal of Education and Training Studies*, *4*(5), 85-91. https://doi.org/10.11114/jets.v4i5.1419

Antidote (2003). *The emotional literacy handbook: Promoting whole-school strategies*. London: David Fulton in association with Antidote.

Bocchino, R. (1999). *Emotional literacy: To be different kind of smart*. London: Sage.

Bryman, A. (2012). *Social research methods*. New York: Oxford University.

Byrne, B. M. (2016). *Structural equation modeling with AMOS: basic concepts, applications, and programming*. New York: Routledge

Büyüköztürk, Ş. (2011). *Handbook of data analysis for social sciences*. Ankara: Pegem Akademi.

Carnwell, R., & Baker, S. A. (2007). A qualitative evaluation of a project to enhance pupils' emotional literacy through a student assistance program. *Pastoral Care in Education: An International Journal of Personal, Social and Emotional Development*, *25*(1), 33-41.

Cole, D. A. (1987). Utility of confirmatory factor analysis in test validation research. *Journal of Consulting and Clinical Psychology*, *55*, 584-594.

Claxton, G. (2005). *An intelligent look at emotional intelligence*. London: ATL. Retrieved from http://www.atl.org.uk/Images/Emotional%20intelligence.pdf

Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2014). *Multivariate statistics-Spss and Lisrel applications for social sciences*. Ankara: Pegem Akademi.

Coşkun, K. (2015). *Impact of emotional literacy training on students' emotional intelligence level in primary schools* (Unpublished doctoral dissertation). Ondokuz Mayıs University, Samsun.

DeVellis, R. F. (2012). *Scale development: Theory and application*. Newbury Park, CA: Sage.

Dickson, E., & Burton, N. (2011). 'You looking at me?': Investigating 9 and 13 year-olds' ability to encode and decode nonverbal communication and demonstrate 'emotional literacy'. *Education 3-13*, *39*(3), 265-276. DOI: 10.1080/03004270903505062

Doll, W. J., Xia, W., & Torkzadeh, G. (1994). A confirmatory factor analysis of the end-user computing satisfaction instrument. *MIS Quarterly*, *18*(4), 43–461.

Faupel, A. (2003). *Emotional literacy assessment and intervention ages 11-16*. Southampton: Nfer-Nelson.

Field, A. (2013). *Discovering statistics using SPSS*. London: Sage.

Flynn, S. J. (2010). *A comparative and exploratory study of the nfer-nelson emotional literacy scale in an Irish context* (Unpublished doctoral dissertation). University of Exeter, The UK.

Ginott, H. (1965). *Between parent and child*. New York: Macmillan Company.

Ginott, H., Ginott, A., & Goddard, H. W. (2003). *Between parent and child* (Revised Edition). New York: Three Rivers.

Gong, S., Chai, X., Duan, T., Zhong, L., & Jiao, Y. (2013). Chinese teachers' emotion regulation goals and strategies. *Psychology*, *4*(11), 870–877.

Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis*. London: Prentice-Hall.

Hargreaves, A. (2001). The emotional geographies of teaching. *Teachers College Record*, *103*(6), 1056-1080.

Holmes, G. (2016). *Emotional literacy: an investigatıon into changes in outcomes and mechanisms following a small group targeted intervention. A case study with primary aged children* (Unpublished doctoral dissertation). University of Birmingham, The UK.

Joseph, G., Strain, P., & Ostrosky, M. M. (2005). *Fostering emotional literacy in young children: Labeling emotions*. What Works Brief, Center on the Social and Emotional Foundations for Early Learning. Retrieved from https://challengingbehavior.cbcs.usf.edu/docs/whatworks/WhatWorksBrief_21.pdf

Jöreskog, K., & Sörbom, D. (2001). *LISREL 8.51*. Mooresvile: Scientific Software.

Kandemir, M. ve Dündar, H. (2008). Emotional literacy and emotional literacy learning environments. *Sakarya University Journal of Education Faculty*, (16), 83-90.

Kayadibi, F. (2001). Factors affecting the quality of education and the effect of quality education on production. *Journal of Istanbul University Faculty of Theology*, *3*(1), 71-94.

Killick, S. (2006). *Emotional literacy: at the heart of the school ethos*. London: Paul Chapman.

Kline, P. (1994). *An easy guide to factor analysis*. New York: Routledge.

Kline, R. B. (2005). *Principle and practice of structural equation modeling*. New York: Guilford.

Mader, L. (2005). *Men who abuse: Examining emotional literacy in men's story* (Master's thesis). Acadia University, Scotland.

Matthews, B. (2006). *Engaging education: Developing emotional literacy, equity, and co-education*. England: Open University

Meekums, B. (2008). Developing emotional literacy through ındividual dance movement therapy: A pilot study. *Emotional and Behavioural Difficulties*, *13*(2), 95-110. DOI: 10.1080/13632750802027614

Ministry of National Education (2017). General competencies for teaching profession. Retrieved from https://oygm.meb.gov.tr/meb_iys_dosyalar/2018_06/29111119_TeachersGeneralCompetencies.pdf

O'Hara, D. (2011). The impact of peer mentoring on pupils' emotional literacy competencies. *Educational Psychology in Practice*, *27*(3), 271–291. DOI: 10.1080/02667363.2011.603533

Orbach, S. (1998). Emotional literacy. *Young Minds Magazine*, (33), 12-13.

Park, J. (1999). Emotional literacy: Education for meaning. *International Journal of Children's Spirituality*, *4*(1), 19-28. DOI:10.1080/1364436990040103

Palancı, M., Kandemir, M., Dündar, H., & Özpolat, A. R. (2014). Validity and reliability study of emotional literacy scale. *International Journal of Human Sciences*, *11*(1), ss. 481-494. doi: 10.14687/ijhs.v11i1.2806

Perry, L., Lennie, C., & Humphrey, N. (2008). Emotional literacy in the primary classroom: teacher perceptions and practices. *Education 3-13*, *36*, 27-37. DOI: 10.1080/03004270701576851

Rae, T. (2012). Developing emotional literacy – approaches for staff and students developing an approach in the SEBD School. *International Perspectives on Inclusive Education*, 2, 1-18. DOI:10.1108/S1479-3636(2012)0000002004

Rasch, D., Kubinger, K. D., & Yanagida, T. (2011). *Statistics in psychology using R and SPSS*. Chichester: J. Wiley.

Raykov T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling*. NJ: L. Erlbaum Associates

Ripley, K., & Simpson, E. (2007). *First steps to emotional literacy: A program for children in the Foundation Stage and Key Stage 1 and for older children who have language and/or social communication difficulties*. London: A David Fulton.

Sipahi B.,Yurtkoru, E. S., & Çinko M. (2008). *Data analysis with SPSS in social sciences*. Istanbul: Beta.

Sharp, P. (2001). *Nurturing emotional literacy*. London: David Fulton.

Southampton Psychology Service. (2001). *Emotional literacy: Assessment and intervention: Ages 11–16*. London: nferNelson.

Steiner, C. (2003). *Emotional literacy: Intelligence with a heart*. CA: Personhood.

Steiner, C. M. (1979). *Healing alcoholism*. New York: Grove.

Stone, V. M. (2005). *An exploration of the validity of a measure of emotional literacy*. Unpublished Master's Dissertation, Cardiff University, Wales.

Şahin, D. B., & Gülleroğlu, H. D. (2013). Investigation of psychometric properties of scales created with different item analysis techniques used to select items for Likert type scales. *Asian Journal of Instruction*, *1*(2), 18-28.

Şencan, H. (2005). *Reliability and validity in social and behavioral measures*. Ankara: Seçkin.

Şişman, M. (2009). Öğretmen yeterlilikleri: Modern bir söylem ve retorik. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, *10*(3), 63–82

Uzan, M. F. (2018). *Evaluation of emotional perception skill levels of learners of Turkish as a foreign language* (Unpublished master's thesis). Çanakkale Onsekiz Mart University, Çanakkale.

Weare, K. (2004). *Developing the emotionally literate school*. London: Paul Chapman.

Weare, K., & Gray, G. (2003). *What works in developing children's emotional and social competence and wellbeing?* London: Department for Educational and Skills.

West, S. G., Taylor, A. B., & Wu, W. (2012). Model Fit and Model Selection in Structural Equation Modeling. Hoyle, R H. (Ed), *Handbook of Structural Equation Modeling* içinde (209-231). New York: Guilford.

Do Career Decidedness and Career Distress Influence Psychological and Subjective Wellbeing? The Mediating Role of Mindfulness

Ümüt Arslan[1], Burcu Bayraktar-Uyar[1]
[1]İzmir Democracy University

To cite this article:

Do Career Decidedness and Career Distress Influence Psychological and Subjective Wellbeing? The Mediating Role of Mindfulness

Ümüt Arslan[1*], Burcu Bayraktar-Uyar[1]
[1]İzmir Democracy University

Abstract

In this study, the mediating role of mindfulness in the relationship of career decidedness and career distress with psychological wellbeing (PWB) and subjective wellbeing (SWB) was investigated. 338 university students (245 females and 93 males, M = 22.6 years, SD = 5.25) from Izmir, Turkey participated in the study. Career Decidedness Scale, Career Distress Scale, The Flourishing Scale, Positive and Negative Affect Scale, The Satisfaction with Life Scale and The Mindful Attention Awareness Scale were used in the study. The Pearson Correlation Test was applied for correlations and PROCESS Model 4 analysis tools were used to test the mediating models. The results showed that (a) career decidedness predicted mindfulness, PWB, and SWB, and (b) mindfulness played a mediating role in the link between career decidedness and PWB and SWB. While career distress predicted PWB and SWB, it did not predict mindfulness. In addition, mindfulness did not have a mediating role in the relationships of career distress with PWB and SWB.

Key words: Career Distress, Career Decidedness, Subjective Wellbeing, Psychological Wellbeing, Mindfulness

Introduction

One's career is a constantly changing and developing structure in relation to their education, the expectations of the individual and their environment from life, their interpersonal relationships and future plans (Eryılmaz & Mutlu, 2017). Career choice is not only limited to work life; it is also a combination of roles one plays throughout their life (Super, 1980). Individuals experience many problems due to their personal characteristics, social values and career choices during their career development (Vondracek, 1998). Individuals who confront with these problems constantly strive to cope with many stressful and challenging factors and to adapt to the changes they encounter (Gati et al.,1996). It is seen that people who are successful in this process have high levels of wellbeing (Uthayakumar et al., 2010). People who have a high level of wellbeing and easily adapt to the changes in their career decision can also develop healthy ways of coping with problems encountered in different areas of life (Demirci et al., 2017). Mindfulness, which is effective in individuals' developing healthy coping skills, is defined as the individuals' focusing on everything that is happening at the moment with toleration, without judging themselves and the outside world, and it is one of the important factors that support wellbeing (Brown & Ryan, 2003). It is known that mindfulness-based exercises are used to assist career work (Galles et al., 2019; Zhang, 2011), and that wellbeing interacts with developments in the career field (Joo & Lee, 2017). Despite this, studies that examine the relationship between career processes affecting individuals' lives to this extent, and their wellbeing and mindfulness levels as a whole are limited.

Career Decidedness and Career Distress

Career choice and decision are individuals' determination of current job opportunities and professions, the evaluation of whether these professions will meet their own expectations, and the decision on the most suitable one among the options (Doğan & Bacanlı, 2012; Sampson et al., 1992; Yeşilyaprak, 2011). During the process of career decision, individuals may have both cognitive and psychological difficulties, and so this situation causes them to feel tense and stressed and affects their psychological and physical wellbeing (Mann et al., 1989). Career decidedness, on the other hand, is a concept that can vary depending on individual and environmental factors such as individuals' personal characteristics, the culture they live in, their abilities and interests, and that indicates the commitment to the career decision they made and satisfaction with this decision

---

(Miller, 2011; Gati et al., 1996). In other words, career decidedness is the level of confidence and certainty regarding a particular career decision (Gordon, 1998). However, individuals may feel uncomfortable in making a career decision, and avoid thinking about career and setting a career goal (Creed & Gagliardi, 2015). Career distress is defined as feeling some uncomfortable affects during the career decision-making process and avoidance of determining a career; and this situation leads individuals to feel a series of negative affects (Creed & Gagliardi, 2015; Şensoy & Siyez, 2018). Due to career distress, individuals feel anxiety, depression, and various negative effects that may influence daily life (Larson et al., 1994). Furthermore, career distress is a concept that can be experienced in this process and includes feelings such as helplessness, depression, aimlessness, anxiety, guilt, and hopelessness that negatively affect university students (Şensoy & Siyez, 2018; Larson et al., 1994). Career distress is one of the most common sources of stress among university students, and 25% of the students who apply to psychological counseling departments in universities seek psychological help for this reason (Benton et al., 2003; Creed et al., 2016).

Psychological Wellbeing (PWB) and Subjective Wellbeing (SWB)

One of the concepts associated with career choices in individuals' lives is wellbeing (Uthayakumar et al., 2010). The concept of wellbeing, which is frequently used together with such concepts as happiness, welfare, and virtue, has two separate and measurable sub-dimensions that are psychological wellbeing (PWB) and subjective wellbeing (SWB; Graham & Crown, 2014; Grant et al., 2009). PWB is described as the ability of the individual to fulfil their goals, to maintain their own development and to establish healthy relationships with people (Keyes et al., 2002). SWB, on the other hand, can simply be defined as a way of evaluating one's own life (Diener et al., 2009). Research has shown that SWB is associated with career indecision, difficulties in career decision-making, and career decision self-efficacy (Bacanlı, 2016; Uthayakumar et al., 2010). Career distress (Günay & Çelik, 2019) and career indecision (Viola et al., 2016) have been found to negatively affect PWB. In addition, conscious mindfulness exercises (Schutte & Malouff, 2011), which are known to be highly associated with SWB, are known to have beneficial effects for individuals including reducing anxiety during their career choices, improving problem solving skills, and enabling them to have less negative career thoughts (Galles et al., 2019).

Mindfulness as a Mediator

Kabat-Zinn (2003) described one of most commented definitions of the mindfulness as 'the awareness that emerges through paying attention on purpose, in the present moment, and nonjudgmentally to the unfolding of experience moment by moment' (p. 145). The concept of mindfulness is defined as recognizing and accepting various affects that individuals feel without suppressing or trying to change them (Neff, 2003). Increasing mindfulness helps individuals be happier, more comfortable and conscious (Hollis-Walker & Colosimo, 2011), reduce depression symptoms (McCracken & Thompson, 2009), and develop and use healthier techniques to cope with stress (Weinstein et al., 2009). This study was conducted because it was thought that SWB and PWB of individuals would be positively affected by ensuring these developments in the process of making a career decision and continuing the career. It is known that mindfulness is associated with positive career thoughts, and the use of holistic career counseling interventions including mindfulness practices is recommended (Galles et al., 2019). In addition, mindfulness exercises can be taught to help individuals to cope with negative thoughts they feel due to their concerns for the future and career instability (Jacobs & Blustein, 2008). Also, it is known that individuals with high level of mindfulness experience fewer difficulties in career decision-making (Zhang, 2011). Thus, in addition to positive correlation with wellbeing, mindfulness also has positive effects on career developments of individuals (Galles et al., 2019; Zhang, 2011)

Purpose of the Study

In order to fill the gap in the literature, it is important to study university students' career development and wellbeing. Also, investigating possible variables which may play a role in this relationship between career and wellbeing of university students is of great importance. Although studies reveal that awareness is related to both career-related thoughts (Galles et al., 2019; Zhang, 2011) and wellbeing (Bacanlı, 2016; Schutte & Maalouff, 2011; Uthayakumar et al., 2010), there is no study examining the mediating role of mindfulness in the relationship with wellbeing. However, studies provide clues that mindfulness may have a significant mediating role in the relationship between career decidedness and career distress, and wellbeing. Therefore, this study was conducted to reveal in detail the relationships among career decidedness, career distress, wellbeing and mindfulness, which are stated to affect each other in previous studies, and also to contribute to the literature.
The four hypotheses in this study are as follows:

1. Mindfulness plays a mediating role in the relationship between career decidedness and SWB.
2. Mindfulness plays a mediating role in the relationship between career decidedness and PWB.
3. Mindfulness plays a mediating role in the relationship between career distress and SWB.
4. Mindfulness plays a mediating role in the relationship between career distress and PWB.

Method

Participants

In this study, data were collected from 338 university students (72.5% are female, 27.5% are male) studying in Izmir, whose ages ranged between 18 and26 years (*M* = 22.6 years, *SD* = 5.25*)*. The sample covers general background of university students in Turkey. The sample of the study was determined through convenience sampling. Convenience sampling is one of the non-random sampling methods and defined as collecting data from a sample that the researcher can easily access (Büyüköztürk et al., 2019).

Research Design

In order to use the data collection tools in the study, an ethical committee approval was obtained from Izmir Democracy University Social and Human Sciences Scientific Research and Publication Ethics Committee (Decision No: 2020/06.05). The scales and the Informed Consent Form (information on the voluntary principle, the purpose of the study, how the study will be conducted, how the participants' information will be kept confidential, and that the results will be used for educational purposes) were uploaded on Google Forms, where data can be collected online. The researchers emailed faculty members at the university twice with a three-week interval in between, with an invitation containing the purpose of the study and an invitation link to direct their students to this study.

Assessment Tools

Mindful Attention Awareness Scale

The Mindful Attention Awareness Scale developed by Brown and Ryan (2003) and adapted into Turkish by Özyeşil et al. (2011) is a unidimensional, 6-point Likert scale (1 = almost always and 6 = almost never) consisting of 15 items. As the scores obtained from the scale increase, the level of mindfulness increases as well. The lowest score that can be obtained from the scale is 15, while the highest score is 90. The Cronbach's Alpha internal consistency coefficient of the scale is .82 in its original language (Brown & Ryan, 2003), and .80 in its Turkish adaptation (Özyeşil et al., 2011). The Cronbach's Alpha internal consistency coefficient calculated in this study is .83.

Career Distress Scale

The Career Distress Scale is a 6-point Likert scale (1 = I totally disagree and 6 = I strongly agree) consisting of 9 items (Creed et al., 2016). The higher scores obtained from the scale indicate higher career distress. The scale consists of two sub-dimensions as "negative affects" and "perceived barriers". The Cronbach's Alpha internal consistency coefficient of the scale is .90 in the original study, and in the Turkish adaptation, it is .81 (Şensoy & Siyez, 2018). The Cronbach's Alpha internal consistency coefficient calculated in this study is .84.

Career Decidedness Scale

The Career Decidedness Scale is a five-point Likert scale (1 = I strongly disagree and 5 = I strongly agree), consisting of 6 items and one dimension, developed by Lounsbury et al. (1999) and adapted into Turkish by Akçakanat and Uzunbacak (2019). As the scores obtained from the scale increase, the level of individuals' career decidedness also increases. The Cronbach's Alpha internal consistency coefficient of the scale is .95 in its original language (Lounsbury et al., 1999), and .80 in its Turkish adaptation (Akçakanat & Uzunbacak, 2019). The Cronbach's Alpha internal consistency coefficient calculated in this study is .80.

The Flourishing Scale

The Flourishing Scale (TFS) was developed by Diener et al. (2009) to measure PWB. Telef (2013) adapted TFS into Turkish. TFS consists of 8 items and one dimension. The scale is a 7-point Likert scale (1 = I strongly disagree and 7 = I strongly agree), and it can be used to interpret that the individual has many psychological

resources and strengths with the increase in the score obtained from the scale. The lowest score to be obtained from the scale is 8, while the highest score is 56. The Cronbach's Alpha internal consistency coefficient of the scale is .87 in its original language (Diener et al., 2009), and .80 in its Turkish adaptation (Telef, 2013). The Cronbach's Alpha internal consistency coefficient calculated in this study is .83.

The Positive and Negative Affect Schedule (PANAS)

Positive-Negative Affect Scale (PANAS) developed by Watson et al. (1988) and adapted into Turkish by Gençöz (2000) is a 5-point Likert scale consisting of 20 affective items, 10 of which are positive and 10 negative. Separate calculations are made for positive and negative affects in the scale. The Cronbach's Alpha internal consistency coefficient of the scale is .88 for positive affects and .85 for negative affects in its original language (Watson et al., 1988). In the adaptation study of the scale, Cronbach's Alpha coefficients were .86 for positive affects and .83 for negative affects (Gençöz, 2000). In this study, the Cronbach's Alpha values are .87 for positive affects and .83 for negative affects.

The Satisfaction with Life Scale

The Satisfaction with Life Scale developed by Diener et al. (1985) and adapted into Turkish by Köker (1991) is a 7-point Likert-type scale (1 = Not at all suitable and 7 = Very suitable) and consists of 5 items in total. The lowest score that can be obtained in the scale is 5 while the highest score is 35. The higher scores obtained from the scale indicate higher life satisfaction. The Cronbach's Alpha internal consistency coefficient of the scale is .84 in the Turkish adaptation (Köker, 1991). The Cronbach's Alpha internal consistency coefficient calculated in this study is .82.

Data Analysis

All variables to be measured in the study, except for SWB, are measured with one scale. The Satisfaction with Life Scale and PANAS were used together to measure SWB. In this study, SWB was calculated by taking the total score over the model proposed by Busseri and Sadava (2011) with the concept of "Subjective Wellbeing (SWB) as a Composite'' (p. 296), which positive emotions affect life satisfaction positively (+) and negative emotions negatively (-) . The Pearson Product Moment Correlation Coefficient was used to examine the relationships between career distress, career decision-making, mindfulness, PWB, and SWB. The scores of skewness and kurtosis coefficients of all scales were measured between +1 and -1, which shows measure normality assumptions. Lastly, PROCESS Macro Model 4 with 5,000 bootstraps (Hayes, 2017) was used for testing the mediating role to examine four hypothesis models. PROCESS is a tool that can calculate the effect of observed variable path analysis based on mediator and moderator variables (Hayes, 2017). At the same time, it can create direct and indirect effects with single or multiple intermediaries in models using mediator variables (Hayes, 2017).

Findings

The validity of the scales, their correlations with each other, and the mediating role indicated in the hypotheses were analysed respectively. First, as stated in the method section, the Cronbach's Alpha values of the scales used in the study were at a valid level. Second, the correlation results showed that career decidedness had (a) a significant, low-level positive relationship with mindfulness (r =.190, p <.001), (b) a significant, moderate-level positive relationship with PWB (r =.416, p <.001), and (c) a significant, moderate-level positive relationship with SWB (r =.437, p <.001). Another correlation result/other correlation results showed that career distress had (a) a significant and low-level negative relationship with PWB (r = -.240, p <.001), (b) a significant and low level negative relationship with SWB (r = -291, p <.001), and non-significant relationship with mindfulness (r =.66). Third, mindfulness and PWB (r =.319, p <.001) and mindfulness and SWB (r =.369, p <.001) showed a significant and moderate-level positive relationship. Lastly, the relationship between PWB and SWB (r =.517, p <.001) was statistically significant and moderately positive.

Table 1. Bivariate correlations

	CDE	CDS	PWB	SWB	MIND
CDE	-	-	-	-	-
CDS	-.518[*]	-	-	-	-

PWB	.416*	-.240*	-	-	-
SWB	.437*	-.291*	.517*	-	-
MIND	.190*	.066	.319*	.362*	-
M	21.65	34.19	42.65	27.17	57.04
SD	4.18	5.92	6.51	14.54	10.39
Skewness	-.05	.08	-1.01	-.39	.04
Kurtosis	.91	.63	1.12	.57	.07

Note: CDE: Career Decidedness, CDS: Career Distress, SWB: Subjective Wellbeing, PWB: Psychological Wellbeing, MIND: Mindfulness, M: Mean, SD: Standard Deviation, * $p < 0.001$

This study has four hypothesis models. In the first and second hypotheses, the mediating role of mindfulness in the relationship of career decidedness with PWB and SWB were tested. In the third and fourth hypotheses, the mediating role of mindfulness in the relationship of career distress with PWB and SWB were tested. Fig. 1 and Fig. 2 show the outcomes of the hypothesis models.

The results of Hypothesis 1 and 2 showed that career decidedness predicted mindfulness ($a_i = .4769$, p <.001), PWB ($c' = .5789$, p <.001) and SWB ($c'' = 1.3382$, p <.001) significantly and positively (Fig. 1). Also, mindfulness predicted PWB ($b_i = .1561$, p <.001) and SWB ($b_{ii} = .4046$, p <.001) positively and significantly. In PWB, the mediating role of mindfulness was .0745 (LLCI = .0247 - ULCI =.1358), and the total mediating role of career decidedness and mindfulness was .6533 (LLCI =.5002 - ULCI = .8065). In SWB, the mediating role of mindfulness was .1929 (LLCI = .0639 - ULCI = .3586), and the total mediating role of career decidedness and mindfulness was calculated as 1.5312 (LLCI = 1.1928 - ULCI = 1.8696). As the p-value of both predictive effects were lower than .05 and there was no zero value between the LLCI and ULCI values of the mediating roles, these values were significant (Hayes, 2017).
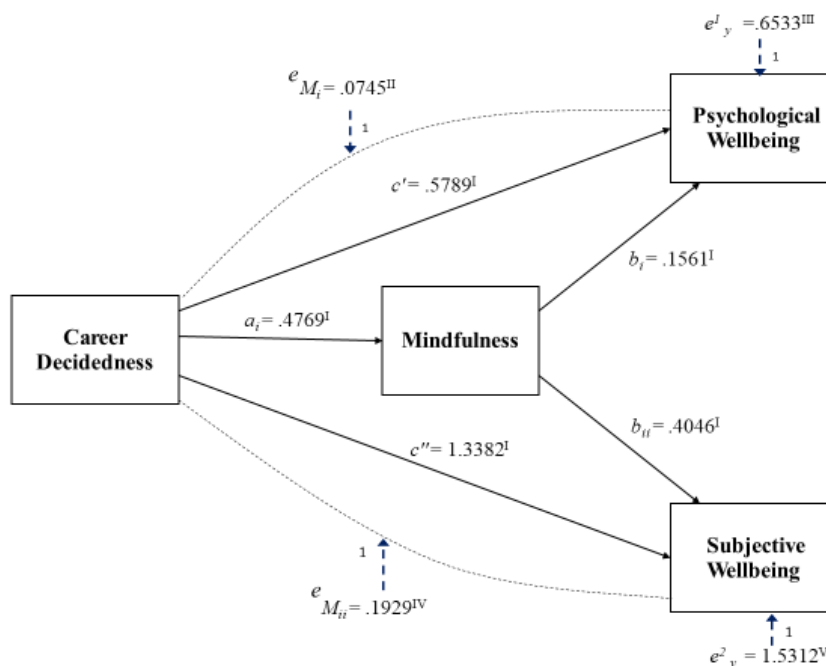

Figure 1. [I]p < .001; [II]LLCI= ,0247-ULCI= .1358; [III]LLCI= .5002-ULCI= .8065; [IV]LLCI= ,0247-ULCI= .1358; [V]LLCI= .5002-ULCI= .8065

The results of Hypothesis 3 and 4 showed that career distress predicted PWB ($c' = -.2888$, p <.001) and SWB ($c'' = -.7760$, p <.001) negatively and significantly (Fig. 2). Mindfulness predicted PWB ($b_i = .2111$, p <.001)

and SWB (b_{ii}= .5357, p <.001) positively and significantly. However, career distress did not significantly predict mindfulness (a_i= 1164). In the relationship between career distress and PWB (LLCI = -, 0152 - ULCI = .0675) and SWB (LLCI = -. 0379 - ULCI = .1658), the mediating role of mindfulness was not significant since it had a zero value between LLCI and ULCI values (Hayes, 2017).
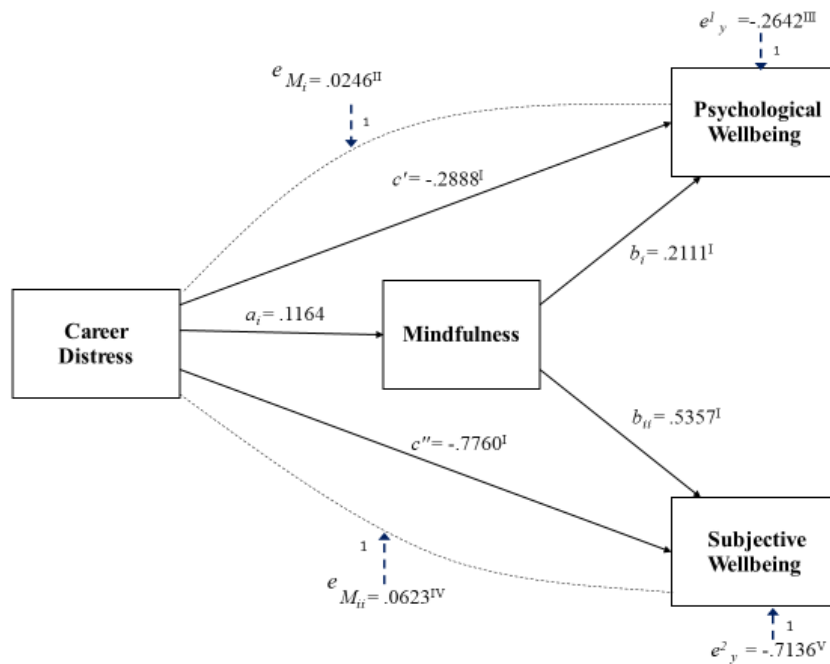

Figure 2. [I]p < .001; [II]LLCI= -,0152-ULCI= .0675; [III]LLCI= -.3787-ULCI= -1497.; [IV]LLCI= -,0379-ULCI= .1658; [V]LLCI= .9657-ULCI= .4616

Discussion

The aim of this study was to investigate the relationship between career thoughts and wellbeing, and the mediating role of mindfulness in these relationships. The results of the correlation analysis showed that there was a negative relationship between career distress, PWB and SWB; a positive relationship between career decidedness, PWB and SWB; and a positive relationship between career decidedness and mindfulness. In the mediating role analyses, it was seen that (a) career decidedness positively predicted mindfulness, PWB and SWB, (b) mindfulness played a significant mediating role between career decidedness and PWB, and (c) mindfulness played a mediating role between career decidedness and SWB. In addition, the mediating role analyses performed with career distress showed that (a) career distress negatively predicted PWB and SWB, (b) career distress did not predict mindfulness significantly, and (c) mindfulness did not play a mediating role between career distress, and PWB and SWB.

The findings of negative prediction of career distress on wellbeing support the literature. It was found that individuals who did not have career difficulties or who could cope well with this situation made their career choices comfortably and competently (Yavuz, 2006). The negative relationship between career distress and career decidedness overlaps with Yavuz's (2006) study. At the same time, individuals who have career decidedness feel more secure, and individuals' career decidedness predicts their SWB (Uthayakumar et al., 2010). The findings of this study showed that career decidedness positively predicts not only SWB but also PWB.

This study showed that career decidedness predicted mindfulness positively and significantly. Career decidedness is a process that requires individuals to know themselves very well and to have high awareness of their interests, abilities, skills and values (Doğan & Bacanlı, 2012; Yeşilyaprak, 2011); and mindfulness is defined as the fact that individuals have the skills to focus on the moment they are in, thus making them happier and more comfortable and develop healthier strategies to cope with stressful situations (Hollis-Walker & Colosimo, 2011; Özyeşil et al., 2011; Ülev, 2014; Weinstein et al., 2009). It is also known that individuals with high levels of mindfulness experience less difficulty when making a career decision than the individuals with low levels of mindfulness (Zhang, 2011). One of the findings of this study, i.e., 'career decidedness positively predicts mindfulness', is supported by other studies in the literature (Dobkin & Hutchinson, 2013; Jacobs &

Blustein, 2008). Individuals with high level of career decidedness are able to be mindful of the current moment and to give attention to possible options.

On the other hand, the prediction of mindfulness by career distress was not statistically significant. Career distress is a concept that includes negative affects experienced by individuals while making a decision about their future (Creed & Gagliardi, 2015). While mindfulness focuses on the present moment (Özyeşil et al., 2011), career distress points at the future and the plans for the future (Şensoy & Siyez, 2018). At the same time, mindfulness requires a full focus on both positive and negative experiences and affects by developing acceptance (Cash & Whittingham, 2010). Career distress is a process that ends with individual's making a career decision and developing career decidedness (Larson et al., 1994). For these reasons, the predictive relationship between these two variables may have turned out to be insignificant due to their focus on different time periods. From another point of view, mindfulness includes accepting and embracing all positive and negative experiences, while career distress is a part of a process that the individual aims to escape from all negative affects and puts an end to these negative affects.

The Mediating Role of Mindfulness

Although the relationship between career and mindfulness has been studied in the literature, there are no studies revealing the mediating role of mindfulness in the relationship between career decidedness and PWB and SWB. With this research, it has been found out that (a) career decidedness predicts mindfulness, PWB and SWB, and (b) mindfulness plays a mediating role in the relationship between (a) career decidedness and PWB, and (b) career decidedness and SWB. Mindfulness affects career decidedness positively and significantly, and when the career decidedness and mindfulness of university students increase, both their PWB and SWB increase as well. Considering the findings of this study, it can be said that mindfulness has a positive effect on the wellbeing of university students. In a period when career decisions and choices are questioned, focusing on the moment increases students' PWB and SWB. This is because university students experience many negative affects and situations such as depression, unhappiness, hopelessness and anxiety due to being concerned for the future, job opportunities and trying to make the right decision, and whether the major they choose will respond to their interests and abilities (Doğan & Bacanlı, 2012). However, mindfulness does not involve focusing on the past and the future, but gently bringing the mind to the present moment and accepting some negative affects that may be experienced without judging them (Özyeşil, 2011). For this reason, it is seen that the wellbeing of the university students who move away from their future anxiety, fulfil the needs of the moment and think about the past less is high, as this study shows. Previous research confirms that individuals with high levels of mindfulness lead a more comfortable and conscious life, which is known to contribute to the individual's wellbeing (Brown & Ryan, 2009). The mediating role of mindfulness revealed by this research has also been identified as a factor that increases the wellbeing of university students in the career decision-making process, where they may experience negative affects such as anxiety and hopelessness. In sum, students with high career decidedness have higher PWB and SWB if they are mindful.

Conclusion

This initiate study in the field of career counseling investigates both PWB and SWB at the same time and contains results regarding the individuals' career situations. Career distress has a negative effect and career decidedness has a positive effect on both PWB and SWB. Importantly, it also found that mindfulness can play an important role in practices that will increase wellbeing. Mindfulness is a factor that positively affects PWB and SWB of individuals. Lastly, this study provides evidences that mindfulness can mediate effect positively on the relationship between career decidedness and wellbeing.

Contribution of the Study to the Field, Limitations and Recommendations

The positive relationship between mindfulness and career decidedness, which is one of the results of this study, proves that the addition of mindfulness exercises to intervention programs (Talib et al., 2015) of the psychological counseling service received by students in the field of career counseling will be beneficial. As this study was carried out with university students in a city in Turkey, its effect value is limited to university students. Conducting similar qualitative or quantitative studies with participants who are already working, seeking a job, or having problems with their career may contribute to studies on the effect of career distress and career decidedness on PWB and SWB.

References

Akçakanat, T., & Uzunbacak, H. H. (2019). Kariyer kararlılığı ölçeği: Türkçe'ye uyarlama, geçerlik ve güvenirlik çalışması [Career decidedness scale: adaptation to Turkish, validity and reliability study]. *İktisadi İdari ve Siyasal Araştırmalar Dergisi [Journal of Economic, Administrative and Political Studies]*, 4(9), 159-170.

Bacanlı, F. (2012). Kariyer karar verme güçlükleri ve meslek seçimine ilişkin akılcı olmayan inançların ilişkisi [Relationship between career decision making difficulties and irrational beliefs about career choice]. *Türk Psikolojik Danışma ve Rehberlik Dergisi [Turkish Journal of Psychological Counseling and Guidance]*, 4(37), 86-95.

Benton, S. A., Benton, S. L., Newton, F. B., Benton, K. L., & Robertson, J. M. (2004). Changes in Client Problems: Contributions and Limitations From a 13-Year Study. *Professional Psychology: Research and Practice, 35*(3), 317–319. https://doi.org/10.1037/0735-7028.35.3.317

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84*(4), 822–848. https://doi.org/10.1037/0022-3514.84.4.822

Busseri, M. A., & Sadava, S. W. (2011). A Review of the Tripartite Structure of Subjective Well-Being: Implications for Conceptualization, Operationalization, Analysis, and Synthesis. *Personality and Social Psychology Review, 15*(3), 290–314. https://doi.org/10.1177/1088868310391271

Büyüköztürk, Ş., Kılıç, E.Ç., Akgün, Ö. E., Karadeniz, Ş. ve Demirel, F. (2012). *Bilimsel Araştırma Yöntemleri [Scientific Research Methods]*. Pegem Academy.

Cash, M., & Whittingham, K. (2010). What facets of mindfulness contribute to psychological well-being and depressive, anxious, and stress-related symptomatology? *Mindfulness, 1*(3), 177–182. https://doi.org/10.1007/s12671-010-0023-4

Creed, P. A., & Gagliardi, R.-E. (2015). Career compromise, career distress, and perceptions of employability: The moderating roles of social capital and core self-evaluations. *Journal of Career Assessment, 23*(1), 20–34. https://doi.org/10.1177/1069072714523082

Creed, P. A., & Hood, M. (2014). Disengaging From Unattainable Career Goals and Reengaging in More Achievable Ones. *Journal of Career Development, 41*(1), 24–42. https://doi.org/10.1177/0894845312471195

Creed, P. A., Hood, M., Praskova, A. ve Makransky, G. (2016). The career distress scale: Using rasch measurement theory to evaluate a brief measure of career distress. *Journal of Career Assessment, 24*(4), 732–746. https://doi.org/10.1177/1069072715616126.

Demirci, İ., Ekşi, H., Dinçer, D. ve Kardaş, S. (2017). Beş boyutlu iyi oluş modeli: PERMA Ölçeği Türkçe Formunun geçerlik ve güvenirliği [Five-dimensional wellbeing model: validity and reliability of the Turkish form of the PERMA Scale]. *The Journal of Happiness & Wellbeing*, 5(1), 60-77.

Deniz, M. E., Erus, S. M. ve Büyükcebeci, A. (2017). Bilinçli farkındalık ile psikolojik iyi oluş ilişkisinde duygusal zekanın aracılık rolü [Bilinçli farkındalık ile psikolojik iyi oluş ilişkisinde duygusal zekanın aracılık rolü]. *Türk Psikolojik Danışma ve Rehberlik Dergisi [Turkish Journal of Psychological Counseling and Guidance]*, 7(47), 17-31.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment, 49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Diener, E., Lucas, R. E. ve Scollon, C. N. (2009). Beyond the hedonic treadmill: Revising the adaptation theory of wellbeing. In The science of wellbeing (pp. 103-118). Springer, Dordrecht.

Dobkin, P. L., & Hutchinson, T. A. (2013). Teaching mindfulness in medical school: where are we now and where are we going?. *Medical education*, 47(8), 768-779. https://doi.org/10.1111/medu.12200

Doğan, H. ve Bacanlı, F. (2012). The effect of a career decision-making group guidance program on career decision-making difficulties [Special Issue]. *Energy Education Science and Technology Part B: Social and Educational Studies*, 1, 912-916.

Eryılmaz, A. ve Ercan, L. (2016). Öznel iyi oluşun cinsiyet, yaş grupları ve kişilik özellikleri açısından incelenmesi [Examining subjective wellbeing in terms of gender, age groups and personality traits]. *Türk Psikolojik Danışma ve Rehberlik Dergisi [Turkish Journal of Psychological Counseling and Guidance]*, 4(36), 139-151.

Eryılmaz, A. ve Mutlu, T. (2017). Yaşam boyu gelişim yaklaşımı perspektifinden kariyer gelişimi ve ruh sağlığı [Career development and mental health from a lifelong development perspective]. *Psikiyatride Güncel Yaklaşımlar [Current Approaches in Psychiatry]*, 9(2), 227-249.

Fuqua, D. R. ve Hartman, B. W. (1983). Differential diagnosis and treatment of career indecision. *The Personnel and Guidance Journal*, 62(1), 27-29. https://doi.org/10.1111/j.2164-4918.1983.tb00112.x

Galles, J., Lenz, J., Peterson, G. W. ve Sampson Jr, J. P. (2019). Mindfulness and Decision-Making Style: Predicting Career Thoughts and Vocational Identity. *The Career Development Quarterly*, *67*(1), 77-91. https://doi.org/10.1002/cdq.12164

Gati, I., Krausz, M. ve Osipow, S. H. (1996). A taxonomy of difficulties in career decision making. *Journal of Counseling Psychology, 43*(4), 510–526. https://doi.org/10.1037/0022-0167.43.4.510.

Gençöz, T. (2000). Pozitif ve negatif duygu ölçeği: Geçerlik ve güvenirlik çalışması [Positive and negative affect scale: validity and reliability study]. *Türk Psikoloji Dergisi [Turkish Journal of Psychology]*, *15*(46), 19-26.

Gordon, V. N. (1998). Career decidedness types: A literature review. *The Career Development Quarterly, 46*(4), 386-403. https://doi.org/10.1002/j.2161-0045.1998.tb00715.x

Graham, C. ve Crown, S. (2014). Religion and wellbeing around the world: social purpose, social time, or social insurance? *International Journal of Wellbeing. 4*(1), 1-27.

Grant, S., Langan-Fox, J., & Anglim, J. (2009). The Big Five traits as predictors of subjective and psychological well-being. *Psychological Reports, 105*(1), 205–231. https://doi.org/10.2466/pr0.105.1.205-231

Günay, A., & Çelik, R. (2019). Kariyer stresinin psikolojik iyi oluş ve iyimserlik değişkenleriyle yordanması [Prediction of career stress with psychological wellbeing and optimism variables]. *Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü Dergisi [Süleyman Demirel University Journal of Social Sciences Institute], 2*(33), 205-218.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford.

Hinkelman, J. M. ve Luzzo, D. A. (2007). Mental health and career development of college students. *Journal of Counseling & Development*, *85*(2), 143-147. https://doi.org/10.1002/j.1556-6678.2007.tb00456.x

Hollis-Walker, L. ve Colosimo, K. (2011). Mindfulness, self-compassion, and happiness in non-meditators: A theoretical and empirical examination. *Personality and Individual Differences*, *50*(2), 222-227. https://doi.org/10.1016/j.paid.2010.09.033

Jacobs, S. J., & Blustein, D. L. (2008). Mindfulness as a coping mechanism for employment uncertainty. *The Career Development Quarterly*, *57*(2), 174-180. https://doi.org/10.1002/j.2161-0045.2008.tb00045.x

Joo, B. K. ve Lee, I. (2017). Workplace happiness: work engagement, career satisfaction, and subjective wellbeing. In *Evidence-based HRM: A Global Forum for Empirical Scholarship*. Emerald Publishing Limited. Vol. 5 Issue:2, pp. 206-221.

Kabat-Zinn, J. (2003), Mindfulness-Based Interventions in Context: Past, Present, and Future. Clinical Psychology: *Science and Practice, 10*: 144-156. https://doi.org/10.1093/clipsy.bpg016

Keyes, C. L. M., Shmotkin, D., & Ryff, C. D. (2002). Optimizing well-being: The empirical encounter of two traditions. Journal of Personality and Social Psychology, 82(6), 1007–1022. https://doi.org/10.1037/0022-3514.82.6.1007

Köker, S. (1991). *Normal ve sorunlu ergenlerin yaşam doyumu düzeyinin karşılaştırılması [Comparison of life satisfaction levels of normal and problematic adolescents].* (Publication No. 16802) [Master thesis, Ankara University]. Council of Higher Education National Thesis Center.

Larson, L. M., Toulouse, A. L., Ngumba, W. E., Fitzpatrick, L. A. ve Heppner, P. P. (1994). The development and validation of coping with career indecision. *Journal of Career Assessment, 2*, 91–110. https://doi.org/10.1177/106907279400200201.

Lounsbury, J. W., Tatum, H. E., Chambers, W., Owens, K. S., & Gibson, L. W. (1999). An investigation of career decidedness in relation to "Big Five" personality constructs and life satisfaction. *College Student Journal, 33*(4), 646–652.

Lucas, M. S., Skokowski, C. T., & Ancis, J. R. (2000). Contextual themes in career decision making of female clients who indicate depression. *Journal of Counseling & Development, 78*(3), 316–325. https://doi.org/10.1002/j.1556-6676.2000.tb01913.x

Mann, L., Harmoni, R., & Power, C. (1989). Adolescent decision-making: The development of competence. *Journal of Adolescence, 12*(3), 265–278. https://doi.org/10.1016/0140-1971(89)90077-8

McCracken, L. M., & Thompson, M. (2009). Components of mindfulness in patients with chronic pain. *Journal of Psychopathology and Behavioral Assessment, 31*(2), 75–82. https://doi.org/10.1007/s10862-008-9099-8

Miller, A. D. (2011). *Career Decidedness, Meaning in Life, and Anxiety: A Mediation/Moderation Model.* (Publication No. 1514295) [Master thesis, Southern Illinois University]. ProQuest Dissertations and Theses publishing.

Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity, 2*(3), 223–250. https://doi.org/10.1080/15298860309027

Özyeşil, Z. (2011). *Öz- Anlayış ve Bilinçli Farkındalık* [Self-Understanding and Conscious Awareness]. Maya Akademy.

Özyeşil, Z., Arslan, C., Kesici, Ş. ve Deniz, M. E. (2011). Bilinçli farkındalık ölçeği'ni Türkçeye uyarlama çalışması [The study of adapting the conscious awareness scale to Turkish]. *Eğitim ve Bilim [*Education and Science], *36*(160).

Sampson, J. P., Peterson, G. W., Lenz, J. G. ve Reardon, R. C. (1992). A cognitive approach to career services: translating concepts ınto practice. *The Career Development Quarterly, 41(1), 67–74.* https://doi.org/10.1002/j.2161-0045.1992.tb00360.x

Schutte, N. S., & Malouff, J. M. (2011). Emotional intelligence mediates the relationship between mindfulness and subjective well-being. *Personality and Individual Differences, 50*(7), 1116–1119. https://doi.org/10.1016/j.paid.2011.01.037

Super, D. E. (1980). A life-span, life-space approach to career development. *Journal of Vocational Behavior, 16*(3), 282-298.

Şensoy, G. ve Siyez, D, M. (2018). The career distress scale: structure, concurrent and discriminant validity, and internal reliability in a Turkish sample. *International Journal for Educational and Vocational Guidance, 19, 203–216.* https://doi.org/10.1007/s10775-018-9376-3.

Talib, J. A., Salleh, A., Amat, S., Ghavifekr, S., & Ariff, A. M. (2015). Effect of career education module on career development of community college students. *International Journal for Educational and Vocational Guidance, 15*(1), 37–55. https://doi.org/10.1007/s10775-014-9279-x

Telef, B. B. (2013). Psikolojik iyi oluş ölçeği: Türkçeye uyarlama, geçerlik ve güvenirlik çalışması [Psychological wellbeing scale: adaptation to Turkish, validity and reliability study]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi [Hacettepe University Journal of Education Faculty], 28*(28-3), 374-384.

Uthayakumar, R., Schimmack, U., Hartung, P. J., & Rogers, J. R. (2010). Career decidedness as a predictor of subjective wellbeing. *Journal of Vocational Behavior, 77*(2), 196-204. https://doi.org/10.1016/j.jvb.2010.07.002

Ülev, E. (2014). *Üniversite öğrencilerinde bilinçli farkındalık düzeyi ile stresle başa çıkma tarzının depresyon, kaygı ve stres belirtileriyle ilişkisi [The relationship between the level of conscious awareness and the style of coping with stress with depression, anxiety and stress symptoms in university students].* (Publication No. 365039) [Master thesis, Hacettepe University]. Council of Higher Education National Thesis Center.

Viola, M. M., Musso, P., Inguglia, C., & Lo Coco, A. (2016). Psychological well-being and career indecision in emerging adulthood: The moderating role of hardiness. *The Career Development Quarterly, 64*(4), 387-396. https://doi.org/10.1002/cdq.12073

Vondracek, F. W. (1998). Career development: A lifespan perspective (Introduction to the special section). *International Journal of Behavioral Development*, *22*(1), 1-6. https://doi.org/10.1080/016502598384487

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Weinstein, N., Brown, K. W., & Ryan, R. M. (2009). A multi-method examination of the effects of mindfulness on stress attribution, coping, and emotional well-being. *Journal of Research in Personality, 43*(3), 374–385. https://doi.org/10.1016/j.jrp.2008.12.008

Yavuz, S. (2006). *İşletmelerde kariyer yönetimi ve uygulamadan bir örnek [An example from career management and implementation in businesses].* (Publication No. 210014) [Master thesis, Ankara University]. Council of Higher Education National Thesis Center.

Yeşilyaprak, B. (2011). *Mesleki rehberlik ve kariyer danışmanlığı: Kuramdan uygulamaya [Vocational guidance and career counseling: From theory to practice.].* Pegem Publishing.

Zhang, Q. (2011). *The role of dispositional mindfulness on occupational engagement and emotional career indecision among college students.* (Publication No. 879417209) [Master thesis, Southern Illinois University]. ProQuest Dissertations and Theses publishing.

The Adaptation of Professional Learning Activities Scale to Turkish: The Validity and Reliability Study

Mahmut Polatcan[1]
[1]Karabuk University

To cite this article:

The Adaptation of Professional Learning Activities Scale to Turkish: The Validity and Reliability Study[*][†]

Mahmut Polatcan[‡],
[1]Karabuk University

Abstract

The purpose of this study is to adapt the Professional Learning Activities Scale (PLAS) developed by Geijsel, Sleegers, Stoel, and Krüger (2009) into Turkish through conducting the relevant validity and reliability analyses. This study followed the pathway recommended by Hambleton and Patsula (1999) for the adaptation process. The data we used came from a total of 256 teachers working in 16 pre-schools, primary and secondary schools located in the Karabuk and Istanbul provinces of Turkey. We performed Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) to test the construct validity of the Turkish-adapted form of the scale. In addition, Cronbach's Alpha coefficient was calculated for the reliability of the scale. The results of the study demonstrated that The Turkish Form of PLA is a reliable and valid instrument to evaluate the extent to which teachers are engaged in professional learning activities.

Key words: Professional learning activities scale, adaptation, validity-reliability, teacher professional learning.

Introduction

A large body of research has accumulated evidence which advocates the notion that teacher and thereby teaching is a crucial factor in terms of elevating student learning and achievement (Barth, 1990; Darling-Hammond, 2000; Darling-Hammond & Richardson, 2009; Kwakman, 2003; Leithwood, Patten, & Jantzi, 2010). Educational scholars have long agreed that teachers play a significant role in transforming large-scale reforms into effective instructional practices conductive to augmented school success (Supovitz, Sirinides, & May, 2010; Thoonen et al., 2011). Given these findings, a substantial body of research has been devoted to unveiling the factors that impact the quality of teachers' instructional practices (Geijsel et al., 2009; Liu, Hallinger, & Feng, 2016; Sebastian & Allensworth, 2012). Extant research reveals that principal leadership (Sebastian & Allensworth, 2012; Thoonen et al., 2011), teacher collaboration (Li, Hallinger, Kennedy, & Walker, 2016), teacher agency (Frost, 2006) and peer influence (Supovitz et al., 2010) were important determinants of the extent to which teachers change and improve their instructional practices. Among varied factors, the construct of teacher professional learning has garnered specific attention from educational scholars in the last decade (e.g. Qian & Walker, 2013; Hallinger & Kulophas, 2019; Shengnan & Hallinger, 2020). This vein of research has concluded that when teachers are more engaged in professional learning activities, they are more likely to improve their skills and content knowledge to better meet the diverse needs of students.

Albeit the theoretical foundations of teacher professional learning were grounded in English-speaking western countries (e.g. Barth, 1990; Darling-Hammond, 2000; Little, 1990; DuFour, 2004; Hargreaves, 1994), research on the construct has found "a welcome reception" in some East Asian counties such as China (Hallinger, Liu, & Piyaman, 2017, p. 2) and Thailand (Hallinger, Piyaman, & Viseshsiri, 2017). However, only few Turkish scholars have conducted empirical investigations on teacher professional learning and related constructs (Bektaş, Kılınç, & Gümüş, 2020; Bellibaş, Bulut & Gedik, 2017; Karacabey, Bellibaş, & Adams, 2020; Özdemir, 2020). One possible explanation for the scarcity of this genre of research could be the lack of a valid and reliable scale to measure the engagement of teachers into professional learning activities.

---

[*] The adapted scale form (Öğretmen Mesleki Öğrenme Etkinlikleri Ölçeği – ÖMÖEÖ) is found in the Appendix and can be used in future research without any permission from the authors.
[†] The original scale is available at the Appendix A section of the paper entitled "The effect of teacher psychological, school organizational and leadership factors on teachers' professional learning in Dutch schools."
[‡] Corresponding Author: *Mahmut Polatcan, mahmutpolatcan78@gmail.com*

Upon examining the exiting local research base on teacher professional learning, we came across only one piece of research that aimed to adapt a relevant scale into Turkish language and culture (Gümüş, Apaydın, & Bellibaş, 2018). In this particular study, the authors attempted to adapt the Teacher Professional Learning Scale originally developed by Liu et al. (2016), which measured the extent to which teachers participate in professional learning activities under four dimensions entitled collaboration, reflection, experimentation, and reaching out to the knowledge base. We consider this scholarly effort valuable in terms of accelerating the local research on teacher professional learning and also agree that some of the items and dimensions of the scale that the current study tried to adapt into Turkish language seem similar to those of Gümüş et al.'s adapted scale (for instance both scales include reflection/experimentation dimensions). It is however here to note that the scale we adapted is different from the previous one in terms of its ability to capture "the extent to which teachers change their practice during the last years toward promoting process-oriented student learning, focusing on strategic, meaningful, and social learning as well as on pupils motivation for learning" (Geijsel et al., 2009, p. 424). This may enable local researchers to investigate the factors influential on teacher change in instructional practices as well as uncovering the factors related to teacher professional learning. In a recent study, for instance, Özdemir (2020) used this scale in a somewhat unorthodox way, taking the first two dimensions – 'keeping up to date' and 'experimentation and reflective practice' – as independent factors that may affect the third dimension entitled teachers' changed practice and concluded that an increase in the first two resulted an increase in teacher efforts to modify instructional practices. Therefore, the purpose of this study is to adapt the "Professional Learning Activities Scale" developed by Geijsel et al. (2009) into Turkish. We believe this study, along with relevant past research (Gümüş et al., 2018), and would help increase the size and quality of local research efforts on teacher professional learning with its nuanced approach toward measuring teacher engagement in professional learning activities.

Teacher Professional Learning

The construct of teacher development or learning has long been debated in EDLM (Educational Leadership and Management) field in the context of school improvement (Barth, 1990; Darling-Hammond, 2000; Flores, 2004). The traditional wisdom has regarded teacher learning as a series of mostly externally-oriented courses, workshops, or training sessions in which teachers are sent randomly in order to increase their subject matter knowledge and skills (Kwakman, 2003). Under the conceptualization of teacher professional development, external experts or field specialists are seen as core actors who transport the given knowledge or information to teachers to help them change and improve teaching practices (Easton, 2008). However, the relevant research provided contradicting results regarding the effectiveness of teacher professional development on teacher and student learning, which provoked scholarly attention to generate fresh ways of conceptualizing teacher learning (Desimone, 2009; Thoonen et al., 2011).

As a newer form of teacher development, therefore, teacher professional learning represents a broader perspective in terms of leveraging teacher quality to influence school success (Liu et al., 2016; Parise & Spillane, 2010). Grounded in adult learning and situated learning theories (Thoonen et al., 2012), this extended perspective on teacher professional learning suggests that teachers learn in a variety of ways from formally-organized teacher development programs to school-embedded practices such as classroom visits, action research teams or mentoring sessions (Kwakman, 2003; Qian & Walker, 2013). This perspective considers school as a social learning environment in which teachers demonstrate a greater amount of agency through interacting and collaborating with colleagues to change teaching practices and reflecting on them (Borko, 2004; Vescio, Ross, & Adams, 2008).

The PLAS

Geijsel et al. (2009) developed the PLAS depending on some relevant literature pieces (Geijsel et al., 2001; Kwakman, 2003; Leithwood, Dart, Jantzi, & Steinbach, 1993). The scale included 17 items under three dimensions entitled Keeping up to Date (4 items), Experimentation/Reflective Practice (5 items), and Changed Practice (8 items). The response range of the 4-point Likert-type subscale ranges from 1 (almost) never to 4 (almost) always for Keeping up to Date and Experimentation/Reflective Practice sub-dimensions and from 1 (disagree) to 4 (agree) for Changed Practice sub-dimension. The first sub-dimension, keeping up to date, measures the extent to which teachers follow up the developments in their field through various activities such as participating in in-service training activities even when they are not obligatory and reading the professional publications. This sub-dimension addresses the importance of being informed about the recent developments in the field. Experimentation and reflective practice, as the second, includes five items to investigate the extent to which teachers shoulder the responsibility for sustaining professional learning through using student feedbacks and classroom visits. This component suggests that teachers take action and reflect on their practice to provide

their students with higher-quality instructional practices (Geijsel et al., 2001, 2009). The last sub-dimension, changed practice, focuses on whether and to what extent teachers modify their teaching practices over a period of time to meet the diverse learning needs of students. This component, therefore, includes items related to motivating students, extending the instructional strategies used in the classroom, and showing care and consideration to emotional states of students along with their cultural backgrounds (Geijsel et al., 2009; Kwakman, 2003).

Method

Participants

A total of 256 teachers working in primary and secondary schools located in Karabuk and Istanbul provinces participated in the study. 59 (23%) of the participants were male and 197 (77%) were female. Educational attainment of the teachers was distributed with 11.7% (n =30) bachelor's degrees and 88.3% (n=226) graduate degrees. The majority of the respondents are working in primary schools (n = 134; 48.5%), and pre-schools (n = 85; 33.2%) while the number of teacher working in secondary schools is the smallest of all (n = 37; 14.5%). The average professional experience of the participants was 12.63 years, with a SD = 8.18.

The Adaptation

This process includes the adaptation of the scale items with its scoring directions and answering options into Turkish. Following the principles of Hambleton and Patsula (1999), first, we reached a consensus that the construct that is aimed to be measured (professional learning activities) makes sense in Turkish educational setting and that the adapted scale holds promise for contributing well to the relevant knowledge base on EDLM in Turkey. Given that recent educational reforms place a specific emphasis on teacher professional activities (Ministry of National Education [MoNE], 2018), we consider local EDLM researchers would benefit from administrating this scale to uncover whether and to what extent teacher professional learning activities occur and even to explore the variables which the construct is associated with. The authors followed a line of steps during the adaptation and the following section elaborates on each step.

The Process

Translation

The first step constitutes the translation of the original scale into Turkish. Thus the original scale was administered to three experts working in the EDLM field and fluent in both languages. Alongside, each author translated the scale on separate sessions. After the experts translated the items together with scoring directions and answering options into Turkish, we obtained six different translation forms and combined them into a single form. We then gathered online twice to discuss the translated scale items and reached an agreement on a single form.

Linguistic Equivalence - First Round

The second step refers to scholarly efforts to ensure the linguistic equivalence of the scale. For this, we prepared a form for expert review and delivered the scale together with a review form through e-mail to five academics. It is here to note that throughout the process, we paid attention to incorporating the experts into the study based on a line of explicit principles such as working with different experts across different steps, choosing the experts from EDLM or Educational Assessment and Evaluation fields who are fluent in both languages (Turkish and English). In this step, these academics were asked to evaluate the scale in terms of its language equivalence by filling out the review form or taking separate notes on the scale sheet. All the experts sent back the forms including suggestions for revision. We then conducted the necessary amendments on four items (4, 6, 11, 16) based on expert opinions.

Linguistic Equivalence - Second Round

In the third step, we implemented a back-translation strategy for ensuring the language validity of the scale. For back-translation, we sent the final Turkish items to a scholar who received his bachelor's degree from an English Language Education Department and his Ph.D. in the EDLM field. We then consulted three field experts to check for the language accuracy between the original and the back-translated form. We saw a high

level of congruence among experts' views on the language accuracy of the scale. After revising the suggestions, we moved to another step for furthering language accuracy.

Linguistic Equivalence - Third Round

In this fourth step, we delivered the scale to a group of target respondents who did not participate in the original study and asked them to point misunderstood items if any. Depending on respondents' opinions, we revised two items in terms of language accuracy.

Content Validity

The fifth step refers to testing the content validity of the scale. Content validity addresses "the extent to which an empirical measurement reflects a specific domain of content" (Carmines & Zeller, 1979, p. 20). In other words, we aimed to test whether the items of PLAS are representative enough to include the integral parts of teacher professional learning (Chen, 2020). In this step, we organized an expert panel including three EDLM scholars, who conducted a line of research on school improvement – more specifically on teacher professional development/learning – to evaluate the extent to which the scale items could measure the specific components of the related domain. Each participating expert received their Ph.D.'s in EDLM field and had more than ten years of research experience. We sent the items to the experts via e-mail and asked them to assess the ability of items' representativeness of teachers' professional learning activities and to identify if any amendments are required. Although the experts seemed almost satisfied with the form by suggesting no discarding or adding items, we still benefited from the panel, modifying three items located in two separate sub-dimensions.

Structural Validity

In the sixth stage, we aimed to check the structural validity of the scale using a same sample of a total of 256 teachers. Structural validity denotes "the degree to which the scores of an instrument are an adequate reflection of the dimensionality of the construct to be measured" (Mokkink et al., 2010, p. 742). This study, therefore, tested the structural validity of the PLAS by employing Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). By doing EFA, we aimed to observe the factor structure of the scale and to decide if any reduction is necessary. Then by performing CFA, we checked whether the data confirms the existing factor structure of PLAS. Furthermore, we calculated Cronbach's Alpha internal consistency values and item-total correlations to evaluate the reliability of the scale.

Findings

In this section, we reported the construct validity and reliability analyzes of the measurement tool.

Construct Validity

Explanatory Factor Analysis

In order to perform EFA, the distribution of the data set in the population must be normal. This assumption is for the multivariate normal distribution of the linear components of all variables. The Kaiser-Meyer-Olkin (KMO) and Bartlett's Test of Sphericity was used to examine the multivariate normal distribution of the data. The Bartlett's Sphericity test result indicates significance (p<.001), and a minimum KMO test result of .70 is another prerequisite for conducting EFA. Results from KMO and Bartlett's Test of Sphericity provided evidence of the factorability of the scale items (Kline, 2011; Thompson, 2004). Therefore, our preliminary findings from Bartlett's Sphericity test ($\chi2$= 2294.547, df=136; p<.001) and KMO value (.880) complied with the expected criteria, showing that EFA can be performed (Table 1).

Table 1. PLAS factors, items, factor loadings, eigenvalues, and total variance explained

Factor Items			
Changed Practice	Factor loading	Eigenvalues	Total variance explained (%)
i13	.864		
i14	.841		
i17	.804		
i16	.793	5.999	29.840
i12	.784		
i15	.773		

i11	.728		
i10	.706		
Keeping up to Date			
i3	.805		
i4	.765	3.232	16.128
i1	.747		
i2	.651		
Experimentation /Reflective Practice			
i9	.709		
i6	.689		
i8	.673	1.269	15.790
i5	.642		
i7	.582		
Professional Learning Activities			61.758

Varimax Rotation was applied due to the multi-factor structure of the PLAS. As a result of the analysis, it was seen that the scale constitutes a three-factor structure, as in the original scale. The three factors in the measurement tool together explained about 61.75% of the total variance. The factor loads of the items in Changed Practice dimension ranged between .70 and .86 and constituted approximately 29.84% of the total variance. The factor loads of items in Keeping up to Date dimension varied between .65 and .80 and contributed 16.12% of the total variance. Experimentation/Reflective Practice dimension explained approximately 15.79% of the total variance and the factor loads of the items are ranked between .58 and .70.

Confirmatory Factor Analysis

We conducted CFA to validate the factor structure of PLAS. A multi-criteria approach has been adopted in order to evaluate the model fit. As a reference in this approach, model fit indexes such as the ratio of chi-square value to degrees of freedom ($\chi 2 / df$) <3; comparative fit index (CFI), Tucker – Lewis index (TLI), normed fit index (NFI), incremental fit index (IFI), goodness of fit index (GFI)> .90; root mean square error of approximation (RMSEA) <.08; standardized root mean residual (SRMR) <.08 are examined (Brown, 2015; Hu & Bentler, 1999). The standardized factor loads of the items related to the factors are shown in Figure 1.
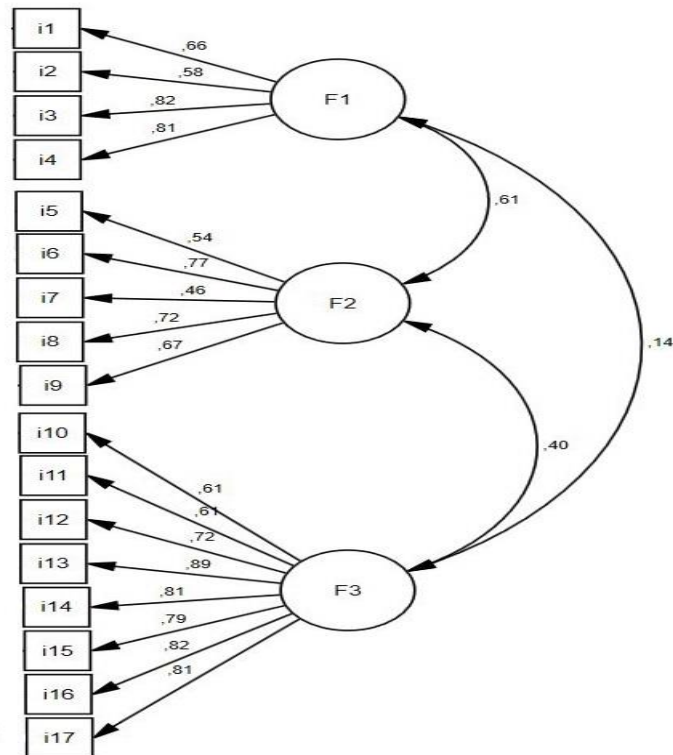

Figure 1. CFA analysis results

The results from CFA showed that GFI and NFI indexes were at the acceptable level and that other indices yielded perfect fit to the data ($\chi2 = 191.057$; df = 111; p <.001 ($\chi2$/df =1.72); CFI = .96; GFI=.92; IFI=.96; NFI= .92; TLI =.96; RMSEA = .053; SRMR = .052). This result points to the utility of PLAS to measure the extent to which teachers are engaged in professional learning activities. Furthermore, the three-factor structure was confirmed in the adapted scale, as in the original scale. The standard coefficients of the four items in Keeping up to Date dimension ranged from .58 to .81, whilst from .46 to .77 for the five items in the Experimentation/Reflective Practice dimension and from .61 to .89 for the eight items in the Changed Practice dimension.

Reliability Analysis

First, we tested the internal reliability, mean and standard deviation, and the correlations between scale factors. The fact that the internal consistency coefficients of the factors are higher than .70 indicates that the factors of the scale are reliable. We also found Cronbach's Alpha coefficient score higher than .88, revealing that scale items are related to teacher professional learning activities (Table 2).

Table 2. Descriptive statistics, Cronbach's coefficient, and factor inter-correlations.

Factor	M^a (SD)	α	1	2	3	PLAS
1: Keeping up to date	3.02(.62)	.82	-	.520**	.144*	.743**
2: Experimentation/reflective practice	3.09(.52)	.75		-	.339**	.801**
3: Changed practice	3.48(.65)	.92			-	.690**
PLAS	3.20(.44)	.88				-

a: Maximum frequency is 4, SD: Standard deviation, **: Correlation is significant at the 0.01 level (2-tailed), *: Correlation is significant at the 0.05 level (2-tailed), α: Cronbach's Alpha coefficient.

When the descriptive statistics of the PLAS are examined, it is observed that the factor average scores are at a high level (Keeping up to Date: M= 3.02, SD= .62; Experimentation/Reflective Practice: M= 3.09, SD= .52; Changed Practice: M= 3.48, SD= .65; PLAS: M= 3.20, SD= .44). As a result of the internal consistency analysis, it was seen that Cronbach's Alpha coefficients varied between .75 and .92. Third, we observed that the correlations between the factors changed from .14 and .52. Considering the correlations between the three factors, it was found that only the correlation between Keeping up to Date and Changed Practice is weak. The other correlations among factors are at moderate level. Finally, we examined the item-total correlations and 27% lower (n = 69) and 27% upper (n = 69) group item discrimination values for the reliability of the PLAS (Byrne, 2010). The item-total correlations of the items in all dimensions and t-test values related to the difference between 27% lower and 27% upper group scores are shown in Table 3.

Table 3. Item total corrected correlations and 27% lower-upper group t-test scores

	Item	Group	n	Mean	SD	t	Item total corrected correlation	p
Keeping up to date	i1	Lower	69	2.52	.55	21.98	.670	.00
		Upper	69	4.00	.00			
	i2	Lower	69	1.78	.41	27.74	.587	.00
		Upper	69	3.76	.42			
	i3	Lower	69	1.92	.26	65.92	.681	.00
		Upper	69	4.00	.00			
	i4	Lower	69	2.34	.63	21.52	.654	.00
		Upper	69	4.00	.00			
Experimentation/ Reflection Practice	i5	Lower	69	1.60	.49	23.41	.472	.00
		Upper	69	3.57	.49			
	i6	Lower	69	2.82	.41	23.29	.635	.00
		Upper	69	4.00	.00			
	i7	Lower	69	1.95	.20	62.90	.404	.00
		Upper	69	3.97	.16			
	i8	Lower	69	2.63	.51	12.45	.589	.00
		Upper	69	3.68	.46			
	i9	Lower	69	2.34	.56	24.32	.538	.00
		Upper	69	4.00	.00			

Changed practice	i10	Lower	69	2.55	.69	17.26	.627	.00
		Upper	69	4.00	.00			
	i11	Lower	69	2.49	.69	17.90	.628	.00
		Upper	69	4.00	.00			
	i12	Lower	69	2.30	.79	17.79	.711	.00
		Upper	69	4.00	.00			
	i13	Lower	69	2.47	.73	17.08	.835	.00
		Upper	69	4.00	.00			
	i14	Lower	69	2.52	.69	17.56	.794	.00
		Upper	69	4.00	.00			
	i15	Lower	69	2.33	.81	16.95	.731	.00
		Upper	69	4.00	.00			
	i16	Lower	69	2.66	.91	9.16	.755	.00
		Upper	69	4.00	.00			
	i17	Lower	69	2.17	.89	17.03	.765	.00
		Upper	69	4.00	.00			

It was determined that the corrected item total correlations of the Keeping up to Data dimension were between .58 and .68 and all correlations were significant at the $p < .01$ level. Item total correlations ranged from .40 to .58 for Experimentation/Reflective Practice dimension and from .62 to .83 for Changed Practice dimension, which were found to be significant at the $p < .01$ level. These findings show that each item in the scale has a moderate to strong positive relationship with the scale, which proves that the items are consistent with the scale. It is also seen that the t-test values of all items in the lower-upper 27% groups varied between 9.16 and 65.92 and that all t values are significant at the $p < .01$ level. These results show that all items are discriminative.

Discussion and Conclusion

This study attempted to adapt the "Professional Learning Activities Scale" developed by Geijsel et al. (2009) into Turkish. We, therefore, produced validity and reliability estimates of the scale through conducting EFA and CFA as well as calculating Cronbach's Alpha coefficient. The results of EFA revealed that the scale yielded a three-factor structure entitled keeping up to date, experimentation/reflective practice, and changed practice. CFA results also indicated that the model provided a good fit to the data, with four-factor and 17 items. We found Cronbach's Alpha reliability coefficient over the standard of .70, referring to a good level of reliability for the scale (George & Mallery, 2003). In addition, low or moderate associations among factors addressed that each scale factor measured different properties. The reliability and validity results are all indicative of the notion that the scale could be regarded as a valid and reliable tool for measuring the extent to which teachers are engaged in professional learning activities.

As an emerging avenue of research pertaining to school improvement, the construct of teacher professional learning has consumed increasing scholarly attention for the last decade (e.g. Thoonen et al., 2011). Educational scholars put a specific emphasis on teachers' sustaining professional learning in various ways to improve the quality of teaching, thereby leveraging student learning (Desimone, 2009; Easton, 2008; Hallinger et al., 2017; Kwakman, 2003). Albeit a substantial body of research provided evidence of a robust link that supports the salience of teacher professional learning in terms of augmenting school success over the last decade (e.g. Hallinger & Kulophas, 2019), the local research evidence on the construct is scarce (Karacabey et al., 2020). Therefore, the current scale holds promise to contribute to the local research, providing a valid and reliable tool for measuring the extent to which teachers keep abreast of the recent development in their field, to embark upon new instructional initiatives and reflect on existing practices to provide students with a higher-quality instruction and to change teaching practices to promote student learning. The local researchers may also benefit from the scale to uncover the link between teacher learning and other constructs pertaining to affective states of teachers such as teacher trust and teacher self-efficacy or some other school-related factors as collaborative school culture or structure. Finally, Turkish EDLM researchers may use the scale to discover the factors that influence teacher change in practice.

References

Barth, R. S. (1990). *Improving schools from within. Teachers, parents, and principals can make the difference*. San Francisco, CA: Jossey-Bass.

Bektaş, F., Kılınç, A. Ç., & Gümüş, S. (2020). The effects of distributed leadership on teacher professional learning: mediating roles of teacher trust in principal and teacher motivation, *Educational Studies*, doi: 10.1080/03055698.2020.1793301

Bellibas, M. S., Bulut, O., & Gedik, S. (2017). Investigating professional learning communities in Turkish schools: The effects of contextual factors. *Professional Development in Education*, *43*(3), 353-374. doi:10.1080/19415257.2016.1182937

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, *33*(8), 3-15. doi: 10.3102/0013189X033008003

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. New York: Guilford publications.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications and programming (2nd Edition)*. New York: Routledge.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment* (Vol. 17). Thousand Oaks, CA: Sage. doi: 10.4135/9781412985642

Chen, J. (2020). Understanding teacher leaders' behaviours: Development and validation of the Teacher Leadership Inventory. *Educational Management Administration & Leadership*, 1-19. doi: 10.1177/1741143220945704

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives*, *8*(1), 1-44.

Darling-Hammond, L., & Richardson, N. (2009). Teacher learning: What matters? *Journal of the Department of Supervision and Curriculum Development*, *66*(5), 46-53.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181-199. doi: 10.3102/0013189X08331140

DuFour, R. (2004). What is a professional learning community? *Educational Leadership*, *61*(8), 6-11.

Easton, L. B. (2008). From professional development to professional learning. *Phi Delta Kappan*, *89*(10), 755-761.

Flores, M.A. (2004). The impact of school culture and leadership on new teachers' learning in the workplace. *International Journal of Leadership in Education*, *7*(4), 297-318. doi:10.1080/1360312042000226918

Frost, D. (2006). The concept of 'agency' in leadership for learning. *Leading & Managing*, *12*(2), 19-28.

Geijsel, F. P., Sleegers, P. J. C., Stoel, R. D., & Krüger, M. L. (2009). The effect of teacher psychological and school organizational and leadership factors on teachers' professional learning in Dutch schools. *Elementary School Journal*, *109*, 406-427. doi: 10.1086/593940

Geijsel, F. P., Sleegers, P. J. C., van den Berg, R. M., & Kelchtermans, G. (2001). Conditions fostering the implementation of large-scale innovations in schools: Teachers' perspectives. *Educational Administration Quarterly*, 37, 130-166. doi: 10.1177/00131610121969262

George, D. & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston: Allyn & Bacon.

Gümüş, S., Apaydın, Ç., & Bellibaş, M. (2018). Adaptation of teacher professional learning scale to Turkish: The validity and reliability study. *Journal of Education and Humanities*, *9*(17), 108-124.

Hallinger, P., & Kulophas, D. (2019). The evolving knowledge base on leadership and teacher professional learning: a bibliometric analysis of the literature, 1960-2018. *Professional Development in Education*, doi: 10.1080/19415257.2019.1623287

Hallinger, P., Liu, S., & Piyaman, P. (2017). Does principal leadership make a difference in teacher professional learning? A comparative study China and Thailand. *Compare: A Journal of Comparative and International Education*, *49*(3), 341-257. doi: 10.1080/03057925.2017.1407237

Hallinger, P., Piyaman, P., & Viseshsiri, P. (2017). Assessing the effects of learning-centered leadership on teacher professional learning in Thailand. *Teaching and Teacher Education*, 67, 464-476. doi:10.1016/j.tate.2017.07.008

Hambleton, R. K., & Patsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology*, *1*(1), 1-30.

Hargreaves, A. (1994). *Changing teachers, changing times: Teachers' work and culture in the postmodern age*. London: Cassell.

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

Karacabey, M. F., Bellibaş, M. Ş., & Adams, D. (2020). Principal leadership and teacher professional learning in Turkish schools: Examining the mediating effects of collective teacher efficacy and teacher trust. *Educational Studies*, doi:10.1080/03055698.2020.1749835

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York: Guilford publications.

Kwakman, K. (2003). Factors affecting teachers' participation in professional learning activities. *Teaching and Teacher Education*, *19*(2), 149-170. doi: 10.1016/S0742-051X(02)00101-4

Leithwood, K., Dart, B., Jantzi, D., & Steinbach, R. (1993). *Building commitment for change and fostering organizational learning* (Final report for phase four of the research project, Implementing British Columbia's Education Policy). Victoria: British Columbia Ministry of Education.

Leithwood, K., Patten, S., & Jantzi, D. (2010). Testing a conception of how school leadership influences student learning. *Educational Administration Quarterly*, *46*(5), 671-706. doi.org/10.1177/0013161X10377347

Li, L., Hallinger, P., Kennedy, K. J., & Walker, A. (2016). Mediating effects of trust, communication, and collaboration on teacher professional learning in Hong Kong primary schools. *International Journal of Leadership in Education*, *20*(6), 697-716. doi: 10.1080/13603124.2016.1139188

Little, J. W. (1990). The persistence of privacy: Autonomy and initiative in teachers' professional relations. *Teachers College Record*, *91*(4), 509-536.

Liu, S., Hallinger, P., & Feng, D. (2016). Supporting the professional learning of teachers in China: Does principal leadership make a difference? *Teaching and Teacher Education*, *59*, 79-91. doi: 10.1016/j.tate.2016.05.023

Ministry of Education [MoNE] (2018). *Turkey's 2023 Education Vision*. Retrieved from http://2023vizyonu.meb.gov. tr/doc/2023_EGITIM_VIZYONU.pdf

Mokkink, L., Terwee, C., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63(7), 737-745. doi: 10.1016/j.jclinepi.2010.02.006

Özdemir, N. (2020). How to improve teachers' instructional practices: the role of professional learning activities, classroom observation and leadership content knowledge in Turkey. *Journal of Educational Administration*, doi 10.1108/JEA-10-2019-0189

Qian, H., & Walker, A. D. (2013). How principals promote and understand teacher development under curriculum reform in China. *Asia-Pacific Journal of Teacher Education*, *41*(3), 304-315. doi: 10.1080/1359866X.2013.809050

Parise, L. M., & Spillane, J. P. (2010). Teacher learning and instructional change: How formal and on-the-job learning opportunities predict change in elementary school teachers' practice. *The Elementary School Journal*, *110*(3), 323-346. doi:10.1086/648981

Sebastian, J., & Allensworth, E. (2012). The influence of principal leadership on classroom instruction and student learning: A study of mediated pathways to learning. *Educational Administration Quarterly*, *48*(4), 626-663. doi: 10.1177/0013161X11436273

Shengnan, L., & Hallinger, P. (2020). Unpacking the effects of culture on school leadership and teacher learning in China. *Educational Management, Administration & Leadership*, doi: 10.1177/1741143219896042

Supovitz, J., Sirinides, P., & May, H. (2010). How principals and peers influence teaching and learning. *Educational Administration Quarterly*, *46*(1), 31-56. doi/10.1177/1094670509353043

Thompson, B. (2004). *Exploratory and confirmatory factor analysis.* Washington, DC: American Psychological Association.

Thoonen, E. E., Sleegers, P. J., Oort, F. J., & Peetsma, T. T. (2012). Building school-wide capacity for improvement: The role of leadership, school organizational conditions, and teacher factors. *School Effectiveness and School Improvement*, *23*(4), 441-460. doi: 10.1080/09243453.2012.678867

Thoonen, E. E., Sleegers, P. J., Oort, F. J., & Peetsma, T. T., & Geijsel, F. P. (2011). How to improve teaching practices: The role of teacher motivation, organizational factors, and leadership practices. *Educational Administration Quarterly*, *47*(3), 496-536. doi: 10.1177/0013161X11400185

Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education*, *24*(1), 80-91. doi: 10.1016/j.tate.2007.01.004

APPENDIX

ÖĞRETMEN MESLEKİ ÖĞRENME ETKİNLİKLERİ ÖLÇEĞİ (ÖMÖEÖ)	Hiçbir zaman	Bazen	Sıklıkla	Her zaman

GÜNCEL KALMA

1 Mesleki açıdan kendimi geliştirmek için sorumluluk alırım.

2 Zorunlu olmasa bile hizmet içi eğitim etkinliklerine katılırım.

3 Mesleğime ilişkin yayınları okurum.

4 Mesleğime ilişkin öğretim materyallerini (ders kitabı, harita, elektronik cihaz vb.) düzenli olarak incelerim.

DENEYİM VE YANSITMA

5 Meslektaşlarımdan öğrenebilmek için onların derslerini gözlemlerim.

6 Edindiğim yeni bilgi ve becerileri, derslerime aktarırım.

7 Derste kullanacağım öğretim materyallerini kendim hazırlarım.

8 Öğretimin niteliğini artırmak için öğrenci dönütlerinden yararlanırım.

9 Meslektaşlarımdan öğrenebilmek için öğretim sürecinde yaşadığım sorunları onlarla tartışırım.

ÖĞRETİMİ DEĞİŞTİRME

SON 3-5 YILDIR;

Not: 3 yıldan daha az kıdeme bulunan öğretmenler "Son 3-5 yıldır" ifadesini göreve başladığımdan bu yana şeklinde değerlendirebilirler.

	Katılmıyorum	Az katılıyorum	Biraz katılıyorum	Katılıyorum

10 Öğrencilerin motivasyonunu artırmaya daha fazla odaklanmaya başladım.

11 Sınıf içinde daha fazla öğretim stratejisinden (sunuş, buluş, araştırma-inceleme) yararlanmaya başladım.

12 Öğrencilerimle etkileşimim arttı.

13 Ders işleme hızımı, farklı düzeydeki öğrencilerin öğrenme ihtiyaçlarına göre ayarlıyorum.

14 Derslerimde farklı öğretim yönteminden (anlatım, problem çözme, gösterip yaptırma vb.) yararlanıyorum.

15 Öğrencilerimin duygusal durumlarına daha fazla önem veriyorum.

16 Öğrencilerime birlikte çalışma yapmaları için daha fazla süre veriyorum.

17 Öğrencilerin kültürel farklılıklarına (köy-kent, bölgesel vb.) daha fazla hassasiyet gösteriyorum.

Computer-Aided Argument Mapping for Improving Critical Thinking: Think Better! Discuss Better! Write Better!

Elif Sönmez[1], Büşra Nur Çakan Akkaş[1], Esra Kabataş Memiş[1]
[1]Kastamonu University

Computer-Aided Argument Mapping for Improving Critical Thinking: Think Better! Discuss Better! Write Better!

Elif Sönmez[1*], Büşra Nur Çakan Akkaş[1,] Esra Kabataş Memiş[1]
[1]Kastamonu University

Abstract

The aim of this study is to investigate the impacts of the use of computer-aided argument maps as a tool to promote prospective teachers' critical thinking skills and dispositions. In this regard, qualitative research method was used in the study. The data of the research were collected through semi-structured interviews. Study group consists of 30 senior prospective teachers from three different classes studying science teaching at a university in Turkey in 2017-2018 academic year. At these three classes mentioned, individual and collaborative argument maps were created in addition to the ABI (Argumentation-based inquiry) activities. The study group was formed on a volunteer basis with 10 students from each class who were selected from, their course performances into consideration. A computer software was used to create the argument maps. Data obtained from interviews were analyzed through NVIVO program. The results obtained indicated that critical thinking skill sub-dimensions such as explanation, analysis, interpretation, evaluation, Self-correction and critical thinking disposition sub-dimensions such as questioning the reliability of sources, being open and fair-minded, being respectful of differences were emphasized more in the group in which the prospective teachers performed computer-aided, individual and collaborative mapping rather than the group in which only ABI activities were performed.

Key words: Argumentation, Argument map, Argumentation-based inquiry (ABI), Critical thinking skills and dispositions.

Introduction

In this age, most of human life is spent by solving the problems that change every day and making decisions regarding the sudden situations that develop. With the ease of access the Internet, there has been an excessive increase in information produced and shared. This information may mislead us about the decisions we make. Hence, when individuals cannot distinguish the accurate information they obtain from many sources from the inaccurate one, it causes more harm than benefit for them.

It becomes even more important for individuals to solve their increasingly complex personal and social problems. This is because individuals are needed to think free of prejudices and intuitive ideas, rationally, in a word, critically, to make healthier decisions with regard to society. In the same direction, Partnership for 21st century Skills (2017) emphasizes the need for the US education system to raise every student with critical thinking skills to ensure their success in their daily lives. In this regard, in line with the trends in the world, Turkish Education System aims to raise generations that have acquired skills such as critical thinking, problem solving, innovativeness, creativity, communication and collaboration within the scope of learning and innovation skills stated by Fadel (2008) as the skills of this century (Ministry of National Education [MoNE], 2017). When considered from this point of view, it is possible to say that students equipped with 21st century skills are prepared for the uncertainty of the future beyond today.

Especially, many educators stated the importance of critical thinking and its encouragement should be one of the most important goals in higher education (Davies, 2011; Harrell, 2011; McMillan, 1987). Besides, it has been emphasized in many national and international studies that critical thinking needs to be improved in higher education (Çınar, 2009; Davies, 2011; Doğan, 2006; Harrell, 2011; McMillan, 1987; Higher Education Council [YOK], 2011; MoNE, 2017, P21, 2017; Common Core State Standards Initiative [CCSS], 2017). It would not be the right approach to care for students attending higher education to graduate as individuals only with the

---

content knowledge of their field. Individuals' ability to think critically enables them to judge their knowledge rationally and accurately (Mc Millan, 1987), and thus it makes individuals to be more successful in performing various tasks in business life (Davies, 2013). Accordingly, it can be said that critical thinking is a preferred quality in individuals graduating from higher education.

Despite the constant emphasis on gaining critical thinking, some researchers have argued that higher education programs do not offer the experience to support critical thinking which is necessary for students to solve complex problems (Reimold, Slifstein, Heinz, Mueller-Schauenburg, & Bares, 2006). Considering the studies examining the critical thinking of prospective teachers; it is seen that the results of the mean scores from the critical thinking scales are mostly low (Akgün & Duruk, 2016; Can & Kaymakçı, 2015; Grosser & Lombard, 2008; Halpern, 1998; Hayırsever & Oğuz, 2017; Kuhn, 1999; Tümkaya, 2011) and medium (Çakırlar-Altuntaş, Yılmaz & Turan, 2017; Çevik, 2013; Deniz & Kaptan, 2011; Korkmaz, 2009; Kürüm, 2002; Tufan, 2008; Uluçınar, 2012; Yıldırım & Şensoy, 2017). Nevertheless, it is known that students' learning to think critically depends on the competence of teachers on this subject (Demirci, 2000). The ability of teachers to express themselves clearly in a free and democratic environment, their perspectives on situations or events, and to discuss them within the frame of causality can also be reflected in education and training activities. In this regard, the idea of training teachers thinking critically becomes more of an issue to deal with in teacher education programs.

When teacher training programs in Turkey are examined, it draws attention that activities improving critical thinking of prospective teachers are emphasized to be performed in pre-service, teaching practices and content of academic courses. Along the same line, when examining the qualifications that form the basis of teacher competencies, it is seen that the following statements are specified (MoNE, 2017):

- "They obtain information from a questioning point of view in their field (p, 13)"
- "They make self-evaluation by benefiting from the opinions and suggestions received (p, 16)",
- "They respect individual and cultural differences (p, 16)",
- "They cooperate with the relevant institution, person and colleagues in education and training activities (p, 14)",
- "They create democratic learning environments where students can communicate effectively (p, 14)".

This perception requires teachers to assume the role of a teacher who can think critically and carry out activities that support critical thinking in the educational environment.

Critical Thinking

Critical thinking is the evaluation of how accurate a decision is or the results achieved in solving a problem (Halpern, 1996). It is mentioned in the literature that critical thinking has skill (cognitive) and disposition (affective) dimensions. For instance, Facione (1990) stated that critical thinking includes cognitive skills such as interpretation, analysis, evaluation, inference, and self-correction. Halpern (1998) emphasized that in order to acquire these skills it is important for students to have these competencies as well as to want to use them. It is believed that critical thinking is not only limited to a proper use of a skill in a given case. Ennis (1991) stated that besides the skills of critical thinking, it also includes dispositional sub-dimensions such as questioning the reliability of sources, being open-minded, being sensitive and respectful to others' emotional, informational and cultural situations. To explain the relationship between the skill and dispositional dimensions of critical thinking, Sears and Parsons (1991) pointed out that those with critical thinking skills might not be inclined to use any of them. However, it should be known that individuals cannot develop expertise in any field without the willingness to make the necessary mental effort to use a skill (Wagner, Leana, Locke, & Schweiger, 1997). This information reveals that both the skill and disposition dimensions of critical thinking are important for students to grow up as good critical thinkers. Therefore, both skill and disposition dimensions of critical thinking are approached together in this study.

The fact that teachers have knowledge about what critical thinking is and its scope, why it is needed and how it should be improved is important in raising critical thinkers. In this sense, the first question that comes to mind should be "Do teachers have the competence to improve critical thinking?". In their study, Paul, Elder and Bartell (1997) examined to which extent prospective teachers were ready to teach critical thinking skills. Within this scope, the researchers interviewed academicians at university. The results showed that faculty members teaching prospective teachers were not able to make a clear explanation of critical thinking mostly and that they did not have information about the skills that should be developed in students. When recent studies are examined, the study of Janssen, Mainhard, Buisman, Verkoeijen, Heijltjes, van Peppen and van Gog (2019) draws attention. This study reveals conclusions that teachers know very little about how to improve their critical

thinking skills and attitudes towards critical thinking teaching. This result shows that prospective teachers are not equipped with critical thinking competencies.

Critical thinking is the ability to evaluate the evidence and rationale for a claim (van Gelder, 2001). For this reason, the ability to mount arguments, analyzing and evaluating arguments are considered as essential components of critical thinking (Ennis, 1987). In this case, it is important to look at the role played by argumentation in developing critical thinking.

In the development of an argument, stages such as making research on the subject, examining the subject from different perspectives, making a claim and identifying concrete evidences that support or refute this claim, mounting an argument accordingly and examining the factors that can improve this argument are followed (Freely & Steinberg, 2000; Toulmin, 2003). In order to realize this process, the individual should use critical thinking skills. Therefore, what enables the development of students' critical thinking is the effective use of these skills in the argumentation process (Allen, Berkowitz, Hunt & Louden, 1999; van Gelder, 2001; Twardy, 2004). Given this need, the Argumentation-Based Inquiry (ABI) approach provides an environment that supports the use of high-level cognitive skills. ABI is an important tool for creating an effective learning environment in which students create arguments, support their claims and actively use their speaking and writing skills in this process (Hand & Keys, 1999). Students' performing activities accompanied by inquiry conducting interactive group work, exchanging ideas and arguments through collective negotiation and creating meaning and reflective writing are series of activities required by the ABI approach (Keys, Hand, Prain, & Collins, 1999). Two templates have been developed to be used by students and teachers in the ABI approach (Keys et al., 1999). In this context, there are some activities including meaningful thinking, writing, reading and discussion skills of students in the teacher template (Table 1.). As another component of ABI, the student template is used individually or as a group during the negotiation phases.

Table 1. Teacher and student template

A template for teacher-designed activities to promote laboratory understanding.	A template for student.
1. Exploration of pre-instruction understanding through individual or group concept mapping.	1. Beginning ideas- What are my questions?
2. Pre-laboratory activities, including informal writing, making observations, brainstorming, and posing questions.	2. Tests- What did I do?
3. Participation in laboratory activity.	3. Observations-What did I see?
4. Negotiation phase I-writing personal meanings for laboratory activity. (For example, writing journals.)	4. Claims-What can I claim?
5. Negotiation phase II-sharing and comparing data interpretations in small groups. (For example, making group charts.)	5. Evidence- How do I know? Why am I making these claims?
6. Negotiation phase III-comparing science ideas to textbooks for other printed resources. (For example, writing group notes in response to focus questions.)	6. Reading- How do my ideas compare with other ideas?
7. Negotiation phase IV-individual reflection and writing. (For example, creating a presentation such as a poster or report for a larger audience.)	7. Reflection-How have my ideas changed?
8. Exploration of post-instruction understanding through concept mapping.	

In simplest terms, an argument is a structure consisting of justified claims after considering different perspectives and data. Argument map is defined as a clear presentation of the reasoning elements regarding this structure and the relations between them using graphics or other non-verbal techniques (van Gelder, 2003). That is to say, in classroom practices, students see their own reasoning manners when they create an argument map using graphical representations. From this point of view, it seems that the educational value of creating an argument map comes from allowing students to explore different views in the process to support the validity and logic of their reasoning.

Argument maps have some advantages over traditionally generated arguments. Hoffman (2005) was the first to state that the use of graphical techniques in argument mapping facilitates the analysis of the argument structure. Additionally, the structure of argument maps, which reveals the hierarchical and clear relations between the argument elements, helps to reduce the complexity of the problems. van Gelder (2001) stated that those who

created argument maps were provided with a considerable ease in evaluating and organizing their thoughts. For these reasons, it can be said that argument mapping can be used to support the development of critical thinking. In some studies, effects of the ABI approach on thinking have been examined. For instance, Keys, et al. (1999) and Opstal and Daubenmire (2015) indicated that laboratory practices based on the ABI approach had a positive effect on students' use of metacognitive thinking skills. Roviati, Widodo, Purwianingsih and Riandi (2019) shared conclusions in their study that ABI-based laboratory activities significantly improved university students' critical thinking skills. In another study, it was revealed that ABI, which is an argument-based approach to science teaching of primary school students, improved critical thinking skills (Hand, Shelley, Laugerman, Fostvedt, & Therrien, 2018). Similarly, in a study conducted with 8th grade students, the ABI approach was reported to have a positive effect on the acquirements in critical thinking skills (Jang & Nam, 2013). These studies indicate that the practices based on the ABI approach improve students' critical thinking at different grade levels.

It is crucial for students to see and evaluate their thoughts and reasoning in a concrete way to think critically. van Gelder (2002) emphasized the necessity of creating a mental picture of an entire argument using tools such as an argument map in the application of these activities. This is because during the argumentation process, a student's ability to distinguish between weak and strong arguments and reporting the effectiveness of his own arguments affect his critical thinking (Sanders, Wiseman, & Gass, 1994).

In line with this information, argument mapping, which is a tool used in creating an argument, should be examined.

Argument Mapping

Argument mapping is defined as a completely explicit presentation of the elements of reasoning about the structure of an argument and the relationships between them by using graphics or other non-verbal techniques (van Gelder, 2003). Compared to developing traditional arguments, the visual structure of argument maps enables students to evaluate themselves by revealing the reflection of their thinking. By seeing the visual representation of the thought, a student can better evaluate the logic used and make adjustments (van Gelder, 2001). The repetition of this reasoning process is called deliberate practice according to the Quality Practice Hypothesis (van Gelder, 2001; van Gelder, Bissett & Cumming, 2004). According to the hypothesis: deliberate practices are often required to improve critical thinking. Practices have been associated with gaining expertise in physical and cognitive skills (Charness, Tuffiash, Krampe, Reingold, & Vasyukova, 2005). To this respect, Plant, Ericsson, Hill and Asberg (2005) expressed deliberate practice as students' performing activities intensely from easy to difficult by guidance based on the skills that were aimed to be improved. In order for these practices to be effective, it has been reported that it is important to provide sufficient feedback and redo them until the skill is acquired (van Gelder, Bissett & Cumming, 2004).

Computer-aided argument maps provide more opportunities for deliberate practices than argument maps created using paper and pencil. This is because in computer software developed for the argument map, guidance is provided by offering scaffolding that will ensure the necessary support in individual studies. For instance; when students select a component while creating argument maps with the help of scaffolds, the system can give them some advice on what to do next (van Gelder, 2001). Accordingly, computer-aided argument mapping allows intensive practice within a limited time, thanks to its visual and easily editable structure that reduces the complexity of the arguments. In this respect, computer-aided argument map is a tool to support comprehensive deliberate practices in the improvement of critical thinking.

When the literature is examined, it is seen that computer programs are developed in which argument mapping can be done to help develop quality arguments. For instance, van Gelder (2001) stated that by using the "Reason! Able" argument mapping software, university students made a significant progress in their critical thinking skills as a result of their 12-week practice. In addition, the results of the study indicated that the improvement of critical thinking was at a high level compared to the results of the studies performing teaching practices to improve critical thinking. Similarly, in the study of Donohue, van Gelder, Cumming, and Bissett (2002) university students created argument maps using "Reason!Able" software. California Critical Thinking Skills Test was used to evaluate students' critical thinking skills after they engaged in argument maps activities over a term and the results indicated that argument mapping significantly contributed to students' critical thinking skills. In parallel with other studies, Twardy (2004) also found that the argument maps created through software significantly improved the critical thinking skills of university students. In addition to these results, the fact that some of the students had difficulty in creating an argument map and more time was required to give a

wide feedback to these maps were stated as the limitations of using a software. Nonetheless, the researcher reported that the software was a good experience for most students in terms of improving critical thinking skills.

Some software offers students some features to create argument maps collaboratively and individually. Considering the cognitive processes required to create an argument, it is seen that they include skills such as students' creating a conceptual understanding of the subject, collecting evidence and analyzing it; justifying, sharing and defending discussions; trying to persuade their peers and reaching an agreement at the end of the discussion (Voss & Means, 1991). In this regard, it is possible to say that creating an argument involves both social and cognitive activities. The participation of students in collaborative discussion enables them to increase their social and cognitive interaction. In this way, students can develop their higher-order cognitive skills by having to reason, reflect and synthesize (Akyol, Garrison & Özden, 2009; Darabi et al., 2013; Ioannou, Demetriou & Mama, 2014). In addition to these skills, it is crucial for teachers to provide environments for their students which will encourage them to assume responsibility and support many affective features such as being respectful for different views, seeking for reason, seeking for accuracy.

In the light of this information, this study was carried out to examine the impacts of the use of computer-aided argument maps as a tool to promote prospective teachers' critical thinking skills and dispositions.

Research questions are as follows:
How do ABI activities influence prospective teachers' critical thinking skills and dispositions?
How do computer-aided individual argument mapping in addition to ABI activities influence prospective teachers' critical thinking skills and dispositions?
How do computer-aided collaborative argument mapping in addition to ABI activities influence prospective teachers' critical thinking skills and dispositions?

Method

Qualitative research method was used in the study. In this context, semi-structured interviews were held at the end of the processes in order to collect data on critical thinking skills and dispositions of prospective teachers during the activities.

Study Group

Study group consists of 30 senior prospective teachers from three different classes studying science teaching at a medium-scaled state university in Turkey in 2017-2018 academic year. In one of these three classes mentioned, only activities based on the ABI approach were conducted; in the second one, individual argument mapping were conducted in addition to the ABI activities; and in the third one individual and collaborative argument mapping were conducted in addition to the ABI activities. Study group consists of 30 prospective teachers in total, with 10 prospective teachers from each class. While determining prospective teachers, the fact that their performances were at all levels (good, intermediate and weak) considering experimental reports and voluntary basis were taken into account.

Semi-Structured Interviews

Semi-structured interviews were carried out by the researchers to evaluate to which extent prospective teachers emphasized critical thinking during activities based on ABI approach, individual and collaborative argument mapping. For example, questions were asked to the students whether they made preparations before participating in the activities and about the situations they paid attention to when choosing the resources for those who did, thus existence of the dimension of questioning the reliability of the sources under the disposition dimension of critical thinking was tried to be examined. Each interview was recorded using a voice recorder with the permission of the prospective teacher. Prospective teachers in groups, in which individual and collaborative argument maps were created, were asked questions to determine the skill and disposition dimensions of critical thinking, which they emphasized in the process of creating an argument map (individual, collaborative) and ABI. In the other group, prospective teachers were asked questions about their critical thinking during the ABI experiment activities. For instance, the existence of questioning the reliability of the sources sub-dimension under the disposition dimension of critical thinking was tried to be examined by asking

the prospective teachers in all three groups about whether they made preparations before participating in the activities, and the situations taken into consideration by those who did when choosing the sources.

Research Settings

In all three classes, lessons were carried out for eight weeks based on the ABI approach. The treatment process of the study consists of two parts: ABI and argument mapping activities. The research process was summarized in Figure 1.
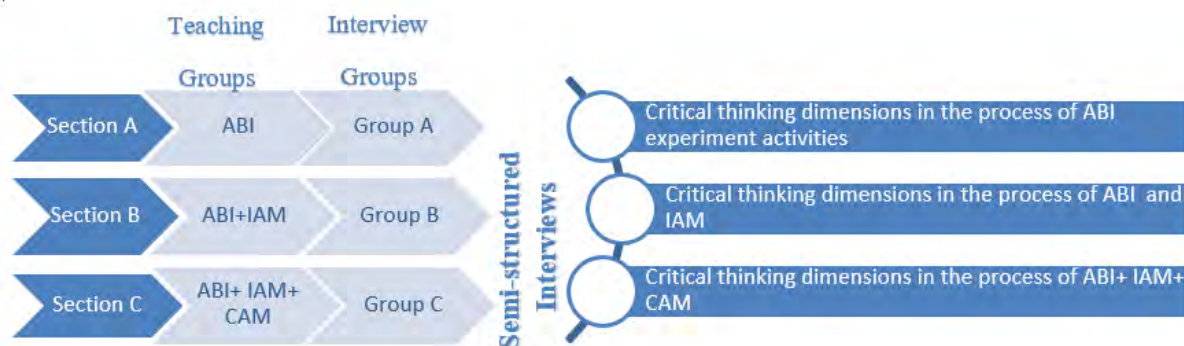

Figure 1. Research Process

Procedures of the ABI Approach

As part of ABI activities, prospective teachers carried out a total of eight activities within the scope of a preparatory activity in small groups and seven "Optics" subjects (light and shadow, mirrors (plane and spherical mirrors), mirror systems, lenses, lens systems and refraction). The preparatory activity was implemented to introduce ABI process and to make the process of creating arguments more efficient by understanding the structure between the claim, the evidence and the claim-evidence. Subsequently, students were asked to make preparations in the scope of the specified topics and come to the class by forming initial questions. In the class, whether the questions created by the prospective teachers were proper to search was evaluated. In order to investigate the questions, they determined, they designed an experiment by making small group discussions consisting of – people. Afterwards, prospective teachers made their claims in line with the experiment they carried out by the nature of the ABI approach. After small group discussions, the students made a large group discussion by presenting the questions they researched, their claims they made as a result of the data they obtained and observations, and their evidence supporting these claims to the class. In the meantime, the researcher asked questions to start and continue the negotiation phase. At the end of the large group discussion, while collecting information on the subject, the researcher asked questions to enable students to make preparations for the next lesson and to draw their attention to the subject. Then, she asked the students to come to class with the questions they wanted to investigate. Throughout this process, the researcher asked questions that sometimes started or continued the discussion and that triggered the prospective teachers to think higher; and acted as a guide.

Procedures of Individual and Collaborative Argument Mapping

An introduction course on argument mapping was given before the experiment activities started in groups (Group B and Group C) that were asked to create an argument map. In this course, students were explained what the argument map was, for what purpose it was created, how it would be used, and the computer program (Rationale-Argument Mapping) where they would create an argument map was introduced. In order to log into the program, each student was provided with pre-created account information and asked to create weekly argument maps through this account. The prospective teachers on both groups created an individual argument map pre-treatment on the subject of "Mysterious Death" activity first so that they could get used to the process. These maps were evaluated by the researchers immediately and deficiencies were eliminated by providing individual feedback. In this way, students made claims, supported these claims with evidence, explained their thoughts with reasonable grounds, and practiced to create their arguments by establishing connections of claim-reason-evidence. Afterwards, the prospective teachers were given the task of creating a total of seven individual argument maps (see Appendix A) on the subject that reflected the main idea of that activity after each

experiment activity. They were asked to send the argument maps they created individually to the e-mail address stated by the researchers within the specified timeframe. The researchers gave the students feedback about the process by evaluating their individual argument maps every week. One of these two groups (Group C) created an collaborative argument map (see Appendix B) in addition to the individual argument map. So as to make the interaction more efficient in a short time while creating a collaborative argument map, the students worked in small groups of two people. Two students within a group shared one computer to generate their argument maps in a computer laboratory. The prospective teachers started to create an argument map within the framework of a claim made by the researchers before the experiment activity (For example, reflection occurs in the lenses). There was another small group where each small group together would create an argument map. The groups had the opportunity to discuss and create collaborative argument maps simultaneously. The students attempted to use the data in line with their claims or to persuade the other group by referencing different sources (own information, book or internet) and sharing their visuals. The researchers, on the other hand, were involved in the simultaneous discussions of them with the opportunity to access the argument maps of all groups. By taking the role of a guide, they asked questions triggering the prospective teachers to make inquiries regarding their claims, reasons and evidence; and added supportive or disproving statements to their maps. Throughout this process, the prospective teachers produced 4 collaborative argument maps in accordance with the nature of the subjects (shadow, mirror, refraction and lens). Each collaborative argument mapping was carried out in about two-hour time period. At the end of each collaborative and simultaneous argument mapping, the students were given feedback after the evaluation of their argument maps by the researchers in terms of the accuracy and explanation of the statements, whether the claims were hierarchical, the validity of the evidence, and the interaction levels of the groups.

Data Analyses

In order to evaluate the data obtained through semi-structured interviews, the interviews were transcribed into written documents. The data obtained from the interviews were analyzed based on definitions of critical thinking skills and dispositions published by Facione (1990) as the Statement of Expert Consensus on Critical Thinking Educational Assessment and Instructional Purposes within the scope of Delphi research project. Content analysis was conducted using NVivo 11 program.

Trustworthiness of The Study

In this study, the procedures stated by Creswell and Miller (2000) that should be used to ensure validity in qualitative research such as long-term observation in the research environment, detailed description of the research environment, participants and the themes created, and supervision of the research process by both someone involving in the research and someone outside the research were taken into consideration. Since the treatments were carried out by the researchers, sufficient time was spent with the participants in the research environment. The researchers met prospective teachers before starting the practice and interacted with them during process. It is one of the measures to be taken in order to increase the quality of the research by asking the experts having general knowledge of the study and specialized on qualitative research methods to examine the research from various perspectives. In this regard, the researchers aimed at ensuring internal audit with two field experts involving in all stages of the research process. In the method of the research, the research environment and the characteristics of the prospective teachers participating in the interviews were tried to be defined. Interview questions were created with two of the field experts on the subject, and the points to be considered in semi-structured interviews were emphasized. Examinations were made on the significance and integrity of the findings obtained from the interviews. The consistency of the themes between sub-themes and codes and with other themes was evaluated and whether they constitute a meaningful whole was examined. The results were presented by giving direct quotations from the interview text.

Results and Discussion

As a result of the analyses of the interviews with the prospective teachers participating in the ABI experimental activities, collaborative and individual argument mapping, Critical Thinking theme and, within this scope, two sub-themes as Critical Thinking Skill and Critical Thinking Disposition were determined. The codes of explanation, analysis, inference, interpretation, evaluation and Self-correction as part of Critical Thinking Skill sub-theme; the codes of open and fair mindedness, seeking for alternatives, seeking for accuracy, seeking for causes, curiosity, evaluating a complex subject regularly, seeking for a clear expression of the problem statement, questioning the reliability of the sources, seeking for the certainty of the subject, evaluating the

subject holistically, being respectful and sensitive to the knowledge, emotion and cultural status of someone else and attitude towards decision- making in cases where the reason and evidence are insufficient as part of Critical Thinking Disposition sub-theme were formed. Themes, sub-themes, frequently mentioned codes and sample expressions regarding these were presented in Table 2.

Table 2. Results Related to Critical Thinking Theme

Theme	Sub Theme	Code	Sub Code	Frequency (*f*)[*]		
				Section A	Section B	Section C
Critical Thinking	Critical Thinking Skills	Explanation	Presenting Arguments	12	14	17
			Justifying Procedures	15	17	23
			Stating Results	2	4	6
		Analysis	Identifying Arguments	1	5	7
			Analyzing Arguments	1	10	13
			Examining Ideas	9	11	15
		Inference	Conjecturing Alternatives	-	1	4
			Querying Evidence	3	5	7
			Drawing Conclusions	7	7	8
		Interpretation	Clarifying Meaning	3	5	7
		Evaluation	Assesing Claims	12	14	17
			Assesing Arguments	3	5	9
		Self-Discipline	Self-Examination	5	15	16
			Self-Correction	-	2	3
		Total		**73**	**116**	**152**
	Critical Thinking Dispositions	Questioning the reliability of the sources		5	6	8
		Seeking for causes		1	-	-
		Curiosity		1	-	-
		Attitude towards decision- making in cases where the reason and evidence are insufficient		1	1	2
		Being respectful and sensitive to the knowledge, emotion and cultural status of someone else		3	3	15
		Evaluating a complex subject regularly		-	1	1
		Evaluating the subject holistically		-	-	3
		Seeking for a clear expression of the problem statement		3	4	8
		Seeking for alternatives		5	4	5
		Seeking for the certainty of the subject		3	6	5
		Open and fair mindedness		3	3	15
		Seeking for accuracy		5	8	6
		Total		**28**	**36**	**70**

* Each prospective teacher can highlight the sub-codes more than once during the interviews.

Results Related to the Critical Thinking of Prospective Teachers (Group C) Participating in ABI, IAM and CAM in the Process

Sub-Theme of Critical Thinking Skills

When the data in Table 2 are examined, it is noteworthy that under the sub-theme of Critical Thinking Skill, the prospective teachers who expressed their opinions about ABI, IAM and CAM processes made statements mostly on the situation "*Justifying Procedures*" under the code of "*Explanation*". They stated that while evaluating the claims presented by other groups, the experiment process should be examined in order to be convinced of the suitability of the experimental environment, the accuracy of the methods used and the results they obtained. A prospective teacher stated the following regarding this situation: "*After all, we have information. We conducted the experiments, too. It's the same thing after all. We state what is missing in their claims. It's okay if they can show the things that are not clear in our minds on the experiment.*" Another prospective teacher explains as in the following that he or she uses the "Justifying Procedures " situation to

persuade other students who evaluate their claims: "*When only our observations were sufficient to explain, we tried to convince them with our observations; when that was not enough, we tried to convince them with the experiment we carried out. In other words, when they were not convinced of the statement we made we were showing it through experiment*." A prospective teacher emphasized that the process should be proved in order to ensure the accuracy of the results they obtained during the phase of explaining the arguments with this statement: "*When someone says something, I think they can see that situation inaccurately. Or, they may misunderstand and misinterpret. But the reason why we are five people there is to correct when it is wrong or to prove it if it is accurate. We were always testing this among us. So, it's nice to have different views*."

One of the codes under Critical Thinking Skill sub-theme that the prospective teachers frequently emphasized regarding the practices was "*Analysis*". For example, a prospective teacher used the following expressions related to the sub-code of "*Examining Ideas*": "*Creating an argument map helped me a lot to understand and establish connections between concepts. Apart from questioning on my own, when making the claim, I was always questioning and discussing in the group on the reason, accuracy and provability*." Another prospective teacher mentioned that creating an argument during the process has a positive effect on himself regarding the sub-code "*Identifying Arguments*" as follows: "*The first argument map I created and the present one is very different. I always asked questions based on my claims on the most recent argument map. I can understand whether I have created a valid argument based on my answers. Unnecessary information creates confusion. Thus, just adding supportive information is important for a good argument. Now I'm paying attention to these*."

It was seen that the prospective teachers expressing their opinions on the practice highlighted the explanations about the "*Assessing Claims*" situation under the code of "*Evaluation*" within the scope of Critical Thinking Skill sub-theme. The prospective teachers mostly made evaluations about the quality of the claims they made in the ABI and CAM processes. They were asked questions about how they made the evaluation of the experimental stages of the claims that put through in small and large group discussions and what their evaluation criteria were. It was stated that, while evaluating the claims made by their friends, the prospective teachers paid attention to the accuracy of the results they obtained by conducting experiments and the use of those results as evidence. In this regard, they stated that the claims which could be proved with sufficient evidence, comprehensible and clearly formed were found to be of higher quality. For example, a prospective teacher explained this situation as follows: "*The claims made should be clear and understandable. They should refute the opposite view completely. So, I think there should be more supporters of the claims. In order to be a qualified claim, they must have supporters explaining every situation*." Another prospective teacher mentioned the criteria he or she took into consideration for the claim to be quality as follows: "*You need to have a good evidence for a good claim. There must be a lot of evidences*". When the prospective teachers were asked about the effect of argument mapping on the process, those in Group C stated that they frequently made "*Self-Examination*" and, in this regard, some went through "*Self-correction*". For instance, a prospective teacher's opinion on Self-correction is as follows: "I think it's good to see that my idea is wrong. Because you attend the class to learn and to correct your mistakes. I mean, you may assume something is correct. However, when you try it there you may say what I saw was wrong. I think it is not a bad thing." The statement of another prospective teacher on making self-examination is as follows: "In this process, I was constantly sounding my own knowledge. You cover a topic this week and then move on to other topics. You don't need to study it over again. But, it is not like that in these activities. The information is interconnected. I cannot find support for it or make a claim without knowing that subject. You learn the subject completely by finding the right way to correct that mistake and support it." Another prospective teacher pointed out the "Self-Regulation" code while evaluating ABI, IAM and CAM process: "At the beginning of the treatment, I did not know how to express myself and create a map to refute or support my claim. Especially, collaborative mapping is very different from the individual one in this respect. Now, I know better how to provide an evidence to support my claim. I had an effort to show what I know on the individual argument map, but I think there is more in the collaborative one. Because you are proceeding mutually. It proceeds in an interrogatory way. People may question themselves, but may not be able to look from a different angle. When it is collaborative, we can respond differently." Additionally, the prospective teachers were asked to compare the report preparation process, which was a writing activity in ABI process, and the argument mapping process, which was another writing activity used in addition to that process. For example, a prospective teacher made comparison with the following statements: "What we said in one of the CAM activities was refuted. On one occasion we refuted ourselves, too. Because what we wrote about the concepts of reflection and refraction were confused at first. When we wrote something inaccurate at that point, we realized our mistake about those concepts and refuted ourselves. Since the information was always in front of our eyes on the argument map, I could see the difference more clearly. As there is only one experiment in the report, we just wrote it immediately. But, on the argument map, we considered whether it was okay or not. There was the time another idea could come to our mind during the process. Therefore, I think the argument map was a bit more effective than the report" The prospective teachers

were asked questions about at which stage and how they expressed their ideas in small group discussions. In this regard, while the prospective teachers in Group C mentioned that they had small group discussions at the stage of obtaining data in order to create a quality claim in the ABI activities, they frequently pointed out "*Drawing Conclusions*" situation under the code of "*Making Inferences*" within the scope of Critical Thinking Skill sub-theme. For example, a prospective teacher mentioned that they paid attention to determining the correct results by organizing their experiments repeatedly: "*For example, when two thin-edged lenses were used and the object was placed between the two lenses, the reverse image was reflected on the screen. No, erect image was reflected. But, I was examining whether it was erect or reverse when we reflected on the wall. We questioned ourselves about where exactly the image was. "No," one of us said. "First we need to bring the lens closer to the object". Then another friend said, "No, we need to bring the lenses closer to each other or keep them a little further away from each other. We observed what was correct by experimenting* "

Sub-Theme of Critical Thinking Dispositions

During the semi-structured interviews, prospective teachers were asked whether they made preparations before coming to class; and if they did, what resources they scanned and according to which criteria they selected these resources. It was seen that prospective teachers frequently emphasized the code of "*Questioning the Reliability of the Sources*" by stating that they paid attention to the reliability of the resources while choosing them. Additionally, they pointed out the situation of "*Open and Fair-Mindedness*" by stating that they did not hesitate to explain their ideas in small and large group discussions and that they reached a compromise by discussing. They stated that preparing a collaborative argument map provided more opportunities to be open and fair-minded. For example, a prospective teacher stated the following sentences about this situation: "*If somebody is shy, they will not speak out in the class. With collaborative argument map, they can express their thoughts more easily in an electronic environment.*" Another prospective teacher expressed this situation as follows: "*When on the board, not everyone can talk to the class because they cannot be comfortable. They may be shy. We can object to each other very comfortably in groups where we create collaborative argument map; we present our knowledge although it is accurate or not at that moment. If it is wrong, they can correct me. In this way, they reinforce their knowledge, as well.*"

In the interviews, when the prospective teachers were asked about the contribution of small group discussions to the process, they stated that situation enabled different ideas to arise. They emphasized that having different ideas enabled them to make the inquiry more comfortably and to reach more accurate results. They pointed out the code of "*Being Respectful and Sensitive to the Knowledge, Emotion and Cultural Status of Someone Else*" within the scope of Critical Thinking Disposition dimension by stating that situation was important in terms of creating stronger arguments, and therefore they respected different ideas. A prospective teacher who stated that his or her ideas are cared by the group members and that he or she also cared about that situation presented the following sentences: "*Differentiation of ideas prompted us to think. Because when you say something, group ideas are consulted. When I say anything in this group, it can be put into perspective. Actually, it makes me feel good because I can express myself.*" The prospective teachers were asked to compare the processes of creating collaborative and individual argument maps. In this regard, they stated that creating an argument map contributed more to the "*Evaluating the Subject Holistically*" situation. For example, a prospective teacher explained this situation as follows: "*If I look at it from different aspects, there is something that they all contributed to me. In the individual argument map, I create a map which can show me what I know and have learnt, and which can contribute to me in the future. I convey all I know to there. I convey what I know in the collaborative one, as well. But when a contradictory thought is presented, my map gets wider because it leads me to the different things that I do not know. Since my knowledge is expanding, I think collaborative one is more useful. But we cannot say the other one is useless.*" Another prospective teacher made the following explanations about this situation: "*I am planning to use it when I become a teacher, as well. Because you look at things from a general aspect. That is to say, all the information is included.*"

Results Related to the Critical Thinking of Prospective Teachers (Group B) Participating in ABI and CAM in the Process

Sub-Theme of Critical Thinking Skills

As a result of the analyses of the interviews regarding ABI, IAM and CAM treatment, the prospective teachers mostly emphasized the importance of "*Justifying Procedures*" situation within the scope of Critical Thinking Skill dimension by mentioning having knowledge of the experimental process in order for them to be able to evaluate the claims. A prospective teacher explained this situation with the following sentences: "*For example,*

if it is necessary to conduct an experiment in a large group discussion, it is very important for that person to do the experiment and to tell the calculation steps. Have those steps handled correctly? In the meantime, the result has been found as 10, but was it 8+2 or 6+4?" With this statement, the prospective teacher emphasized that the control of the valid method and obtaining reliable results, that is to say "*Justifying Procedures*", was effective in evaluating the claims. In addition to this situation, the prospective teachers in group B mentioned that comparing ideas and concepts within the scope of "*Analysis*" code was important in identifying problems and investigating the effect of the part on the whole. For example, a prospective teacher stated the following expressions for the sub-code of "*Examining Ideas*": "*Sometimes the groups put forward the same claims. In that case, we compared our own process with theirs. We compared our results with theirs. In this process, we questioned and discussed theirs as well as ours in a large group. We tried to find what changed the result we reached.*"

The prospective teachers in group B frequently emphasized the sub-code of "*Assesing Claims*" under the "*Evaluation*" code by mentioning the quality of the claims and the characteristics that a good claim should have. A prospective teacher mentioned that he or she evaluated his or her own claims according to these criteria in order to make a strong claim while preparing an argument map as follows: "*After finding a claim, we consider it more comprehensively. We evaluate it with different interpretations and approaches. In this process, we need to have an opposing idea to refute this claim.*"

It was observed that the prospective teachers emphasized the "*Self-Examination*" situation under the code of "*Self-Regulation*" during the practice of argument mapping. For instance, a prospective teacher stated the following expressions about the situations he or she noticed individually when evaluating the argument map processes that he or she created regarding the same subject with the claims made in the ABI activities: "*During the large group discussion, the least criticized claim was ours, which could not be refuted, in the mysterious death activity we conducted before the experiments. While creating argument map, I was almost going to refute my own claim. I realized what I did not notice neither in the large group discussion nor in the report while creating argument map. I thought I needed more supporters.*" Another prospective teacher emphasized having difficulties in associating the ideas he or she thought was correct and noticing this situation while preparing an argument map individually: "*The argument map was good for reviewing the subject and it made us realize our mistakes. It is related to where I put the information I give on the map and with which one I establish a connection. Actually, the information is not wrong but the connections are wrong. I noticed that.*"

Sub-Theme of Critical Thinking Dispositions

The prospective teachers, who were asked about their thoughts on having different ideas in small group discussions, emphasized the code of "*Seeking for the Certainty of the Subject*" by stating that they dealt with a situation in different aspects and investigated the certainty of the situation. A prospective teacher stated the following about this situation: "*At the end of the process, I realized that we needed to make evaluations through these discussions. Because you have to base things upon something. Things have to be certain for the result to be solid.*" Another prospective teacher explained with the following sentences that the use of reasoning and objections in the process of individual argument mapping affected the certainty of the claim: "*In the first argument map, I did not form a sentence starting with "But", "However". I went on using the conjunction "Because", that was how I formed the sentences. It is like a ladder. However, in the last one, I knew that I should use all of them to increase the strength and certainty of the information. I will use them accordingly.*"

Results Related to the Critical Thinking of Prospective Teachers (Group A) Participating in the ABI Activities in the Process

Sub-Theme of Critical Thinking Skills

Within the scope of Critical Thinking Skill dimension, the prospective teachers in Group A emphasized the sub-codes of "Justifying Procedures", "*Presenting Arguments*" under the code of "Explanation" and the sub-code of "*Assessing Claims*" under the code of "*Evaluation*". The prospective teachers in Group A emphasized the situation of "*Drawing Conclusions*" under the code of "*Explanation*" more than those in Group B and C. When the prospective teachers were asked about the characteristics of a good claim, they stated that making it with valid methods and reliable results was crucial. They stated that this situation could only be evaluated by indicating the results. For instance, a prospective teacher expressed the following statements about this situation: "*The results have to be explained. For example, there should not be different results. Everyone should reach the*

same result." Another prospective teacher expressed with the following sentences that it caused the claim to be weakened when the results obtained were not recorded or stated: "*When our friends asked questions and we could not answer them, we would be refuted. Instructor sometimes asked questions but we could not give an answer. Actually, we conducted that experiment, but sometimes we could not support our claim because we could not find what she wanted from us. We needed to answer the questions of our friends. If we were not able to give an answer, we would be automatically refuted. So, we did the experiment correctly. However, since we could not answer the questions, we were refuted.*" When the prospective teachers were asked about their opinions regarding the reports they created during the ABI process, those in group A stated that stage provided the opportunity of "*Self-Examination*". For example, a prospective teacher explained this situation as follows: "*It enabled me to see the change in myself. In my first report, there was nothing in my mind about what to ask or what to do. I was not aware of anything till now as to what I knew or what I learnt. The report writing process made me be aware of myself. I learnt that I needed to ask questions and to think about how I reached a result. During this process, I realized what I knew before and what I gained. That's why I cared about the report a lot and I filled it carefully. As I said before, my first and last report is different from each other. I noticed the improvement in myself.*"

Sub-Theme of Critical Thinking Dispositions

As a result of the analyses of the semi-structured interviews, it was observed that prospective teachers in group A mostly emphasized the code of "*Conjecturing Alternatives*" within the sub-theme of Critical Thinking Disposition dimension. For example, when asked about the effect of small group discussions on the experimental process, a prospective teacher explained as in the following that they experienced the situation of seeking for alternatives during the design and implementation phase of the experiment due to different ideas: "*Everyone in the group expressed their ideas, and then we conducted the experiment again accordingly. When someone stated that something was wrong or when we could not reach a result, they thought of trying again in another way. So, we tried in another way, as well. We proceeded with a common idea of everyone. We did not experience anything that would affect the experiment negatively.*" Another prospective teacher similarly emphasized the code of "*Conjecturing Alternatives*" with the following expressions while experiencing the same process: "*For example, in experiments that we could not carry out, we tried to reach a result through changing the way or experiment setup by looking at our drawings, reviewing our thoughts again or exchanging ideas with our friends again.*"

Conclusion and Recommendations

In this study, which sub-dimensions were indicated by the prospective teachers, who participated in ABI, IAM and CAM practices, within the scope of skill and disposition components of critical thinking during the practices was addressed in a holistic approach. Although the tools used to measure critical thinking in the literature mostly examine only one dimension of critical thinking, they do not provide a deeper knowledge about its sub-dimensions (E.g. The California Critical Thinking Disposition Inventory, Cornell Critical Thinking Test, Watson-Glaser Reasoning Test). Also, this study presents the idea that interviews can be used as an alternative in the analysis of critical thinking, since studies of adapting scales to every language and level are either not conducted or are not up-to-date. In this respect, it is possible to say that this study will contribute to the relevant literature.

Additionally, it is important to conduct studies in which different approaches are applied in learning environments and the effects of argument mapping are examined. This study was limited to interviews as the data sources used. For future research, the evaluation of in-class observations on argument mapping practices and the argument maps produced at the end of the process is recommended now that it will provide evidence not only for the quality of the arguments, but also for the process itself. Consequently, it is emphasized that students and teachers' knowing the structure of an argument and creating quality arguments are crucial both for their academic and business lives. Therefore, more integration of the argument maps in learning environments will be an important step to accomplish this aim.

References

Akgun, A., & Duruk, U. (2016). The investigation of preservice science teachers' critical thinking dispositions in the context of personal and social factors. *Science Education International*, 27(1), 3-15.

Akyol, Z., Garrison, D. R., & Ozden, M. Y. (2009). Online and blended communities of inquiry: Exploring the developmental and perceptional differences. *The International Review of Research in Open and Distributed Learning*, 10(6), 65-83.

Allen, M., Berkowitz, S., Hunt, S., & Louden, A. (1999). A meta-analysis of the impact of forensics and communication education on critical thinking. *Communication Education*, 48(1), 18-30.

Can, Ş., & Kaymakçı, G. (2015). Pre-service teacher's critical thinking tendencies. *Education Sciences*, 10(2), 66-83.

Charness, N., Tuffiash, M., Krampe, R., Reingold, E., & Vasyukova, E. (2005). The role of deliberate practice in chess expertise. *Applied Cognitive Psychology*, 19(2), 151-165.

Common Core State Standards Initiative. (2017). *Common core state standards for English language arts & literacy in history/social studies, science, and technical subjects*. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers.

Çakırlar-Altuntaş, E., Yılmaz, M., & Turan, S. L. (2017). An investigation on critical thinking tendencies of pre-service biology teachers. *Ege Journal of Education*, 19(1), 34-45

Çevik, S. (2013). An investigation of the critical thinking dispositions of pre-service teachers at a private non-profit university. Doctoral dissertation, Bilkent University, Ankara.

Çınar, İ. (2009). Küreselleşme, eğitim ve gelecek[Globalization, education and the future]. *Kuramsal Eğitimbilim Dergisi*, 2(1).

Darabi, A., Liang, X., Suryavanshi, R., & Yurekli, H. (2013). Effectiveness of online discussion strategies: A meta-analysis. *American Journal of Distance Education*, 27(4), 228-241.

Davies, M. (2011). Introduction to the special issue on critical thinking in higher education. *Higher Education Research & Development*, 30(3), 255-260.

Davies, M. (2013). Critical thinking and the disciplines reconsidered. *Higher Education Research & Development*, 32(4), 529-544.

Demirci, C. (2000). Critical Thinking. *Ankara: Eğitim ve Bilim, Türk Eğitim Derneği*, 115, 25.

Deniz, E., & Kaptan, F. (2011). Öğretmen adaylarinin eleştirel düşünme beceri düzeyleri üzerine bir inceleme. *Cagdas Egitim Dergisi*, (389).

Donohue, A., van Gelder, T., Cumming, G., & Bissett, M. (2002). *Reason! project studies 1999–2002*. Melbourne: The University of Melbourne.

Ennis, R. H. (1987). *A taxonomy of critical thinking dispositions and abilities*. In J. B. Baron & R. J. Sternberg (Eds.), Teaching thinking skills: Theory and practice (pp. 9–26). New York: Freeman.

Ennis, R. (1991). Critical thinking: A streamlined conception. *Teaching philosophy*, 14(1), 5-24.

Facione, P. (1990). *Critical thinking: a statement of expert consensus for purposes of educational assessment and instruction (The delphi report)*. Millbrae, CA: The California Academic Press; 1990.

Fadel, C. (2008). 21st Century Skills: *How can you prepare students for the new global economy*. Paris: OECD. Retrieved February, 20, 2019. http://www.aacc.nche.edu/Resources/aaccprograms/ate/conf2010/Documents/NSF%20ATE%20-%2021stCS%20-%20STEM%20-%20Charles%20Fadel.pdf

Freeley, A.J. & Steinberg, D.L (2000). *Argumentation and debate: Critical thinking for reasoned decision making*. Belmont, CA: Wadsworth.

Grosser, M.M., & Lombard, B.J.J. (2008). The relationship between culture and the development of critical thinking abilities of prospective teachers. *Teaching and Teacher Education*, 24(5), 1364-1375.

Hand, B., & Keys, C.W. (1999). Inquiry investigation. *The Science Teacher*, 66(4), 27.

Hand, B., Shelley, M.C., Laugerman, M., Fostvedt, L., & Therrien, W. (2018). Improving critical thinking growth for disadvantaged groups within elementary school science: A randomized controlled trial using the Science Writing Heuristic approach. *Science Education*, 102(4), 693-710.

Halpern, D.F. (1998). Teaching critical thinking for transfer across domains: disposition, skills, structure training, and metacognitive monitoring. *American psychologist*, 53(4), 449.

Harrell, M. (2011). Argument diagramming and critical thinking in introductory philosophy. *Higher Education Research & Development*, 30(3), 371-385.

Hayırsever, F., & Oğuz, E. (2017). Effect of teacher candidates' educational beliefs on their critical thinking tendencies. *Bolu Abant İzzet Baysal University Journal of Faculty of Education*, 17 (2), 757-778.

Higher Education Council (2011). *Turkey higher education qualifications framework*. Last accessed November 3, 2019, http://tyyc.yok.gov.tr.

Ioannou, A., Demetriou, S., & Mama, M. (2014). Exploring factors influencing collaborative knowledge construction in online discussions: Student facilitation and quality of initial postings. *American Journal of Distance Education*, 28(3), 183-195.

Jang, J.Y., & Nam, J. (2013). Examining the relationship between a structured reading framework and students' critical thinking ability within an argument-based inquiry approach. *Journal of The Korean Association For Science Education*, 33(3), 569-580.

Janssen, E.M., Mainhard, T., Buisman, R.S., Verkoeijen, P.P., Heijltjes, A.E., van Peppen, L.M., & van Gog, T. (2019). Training higher education teachers' critical thinking and attitudes towards teaching it. *Contemporary Educational Psychology*, 58, 310-322.

Keys, C.W., Hand, B., Prain, V., & Collins, S. (1999). Using the science writing heuristic as a tool for learning from laboratory investigations in secondary science. *Journal of Research in Science Teaching*, 36(10), 1065-1084.

Korkmaz, Ö. (2009). Teachers' critical thinking level and dispositions. *Journal of Kirsehir Education Faculty*, 10(1), 1-13.

Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher*, 28(2), 16-46.

Kürüm, D. (2002). *Critical thinking abilities of teacher trainees*. (Master's thesis, Anadolu University).

Lee, M., Kim, H., & Kim, M. (2014). The effects of Socratic questioning on critical thinking in web-based collaborative learning. *Education as Change*, 18(2), 285-302.

McMillan, J.H. (1987). Enhancing college students' critical thinking: a review of studies. *Research in Higher Education*, 26(1), 3-29.

Ministry of National Education (MoNE) (2017). *General Competencies for Teaching Profession*. Retrieved December 3, 2019 from https://oygm.meb.gov.tr/www/ogretmenlik-meslegi-genel-yeterlikleri/icerik/39

Oh, E., & Kim, H. (2016). Understanding cognitive engagement in online discussion: Use of a scaffolded, audio-based argumentation activity. *International Review of Research in Open and Distributed Learning: IRRODL*, 17(5), 28-48.

Partnership for 21st Century Skills. (2017). *Partnership for 21st century skills*. Retrieved December 3, 2019 from http://www.p21.org/

Paul, R.W., Elder, L., & Bartell, T. (1997). *California teacher preparation for instruction in critical thinking: Research findings and policy recommendations*. California Commission of Teacher Credentialing, Sacramento, CA (1997)

Plant, E.A., Ericsson, K.A., Hill, L., & Asberg, K. (2005). Why study time does not predict grade point average across college students: Implications of deliberate practice for academic performance. *Contemporary Educational Psychology*, 30(1), 96-116.

Reimold, M., Slifstein, M., Heinz, A., Mueller-Schauenburg, W., & Bares, R. (2006). Effect of spatial smoothing on t-maps: arguments for going back from t-maps to masked contrast images. *Journal of Cerebral Blood Flow & Metabolism*, 26(6), 751-759.

Roviati, E., Widodo, A., Purwianingsih, W., & Riandi, R. (2019, February). Prospective biology teachers' critical thinking skills in microbiology argument-based inquiry laboratory activities. *In Journal of Physics: Conference Series*, 1157 (2), 022108. IOP Publishing.

Sears, A., & Parsons, J. (1991). Towards critical thinking as an ethic. *Theory & Research in Social Education*, 19(1), 45-68.

ter Berg, T., van der Brugge, E., & Cullen, S. (2013). Teaching critical thinking with rationale. *In 7th International Technology, Education and Development Conference*.

Toulmin, S.E. (2003). *The uses of argument*. Cambridge, England: Cambridge University Press.

Tufan, D. (2008). *Critical thinking skills of prospective teachers: Foreign language education case at the Middle East Technical University*. Master's thesis, Middle East Technical University, Ankara.

Tümkaya, S. (2011). Comparison of college science major students' learning styles and critical thinking disposition. *Journal of Kirsehir Education Faculty*, 12(3), 215-234.

Twardy, C. (2004). Argument maps improve critical thinking. *Teaching Philosophy*, 27(2), 95-116.

Uluçınar, U. (2012). *Prediction levels of preservice teachers' critical thinking dispositions on democratic values*. Master's thesis, Osman Gazi University, Eskişehir.

van Gelder, T. (2001). *A Reason! Able approach to critical thinking*. Principal Matters: The Journal for Australasian Secondary School Leaders, May 2002: 34–36, 2001.

van Gelder, T., Bissett, M., & Cumming, G. (2004). Cultivating expertise in informal reasoning. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 58(2), 142.

van Opstal, M.T., & Daubenmire, P.L. (2015). Extending students' practice of metacognitive regulation skills with the science writing heuristic. *International Journal of Science Education*, 37(7), 1089-1112.

Voss, J.F., & Means, M.L. (1991). Learning to reason via instruction in argumentation. *Learning and Instruction*, 1(4), 337-350.

Wagner III, J.A., Leana, C.R., Locke, E.A., & Schweiger, D. M. (1997). Cognitive and motivational frameworks in US research on participation: a meta-analysis of primary effects. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 18(1), 49-65.

Yıldırım, H.İ., & Şensoy, Ö. (2017). Effects of inquiry based learning approach on creative thinking and scientific process skills. *Cumhuriyet International Journal of Education*, 6 (1), 34 – 46

Appendix A. An Example of Indiviual Argument Map
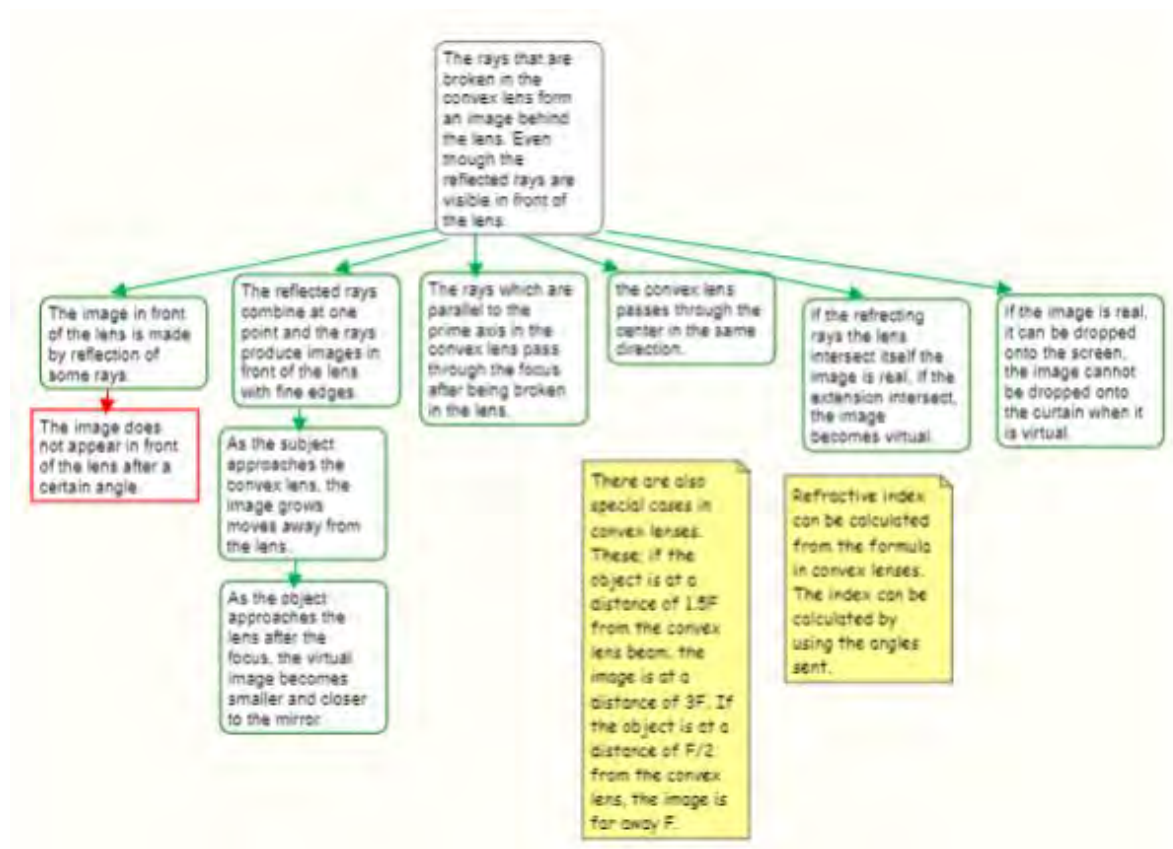
Appendix B. An Example of Collaborative Argument Map

Foreign Language Writing as a Developmental Process (Foundation, Expansion, Development, and Completion): The FEDCom Model

Ali Eryılmaz[1], Yusuf Emre Yeşilyurt[2]
[1] Yildiz Technical University
[2] Gazi University

To cite this article:

Foreign Language Writing as a Developmental Process (Foundation, Expansion, Development, and Completion): The FEDCom Model

Ali Eryılmaz[1*], Yusuf Emre Yeşilyurt[2]
[1] Yildiz Technical University
[2] Gazi University

Abstract

This study investigates the problems that are encountered in learning writing in a foreign language from primary school to university in the Turkish context. It focuses particularly on the problems that teachers/instructors face in teaching writing in English to students concerning students' challenges. Forty teachers were interviewed in all levels, ten participants for each. Here, the challenges that the teachers observed their students faced in the learning process of writing in English were identified. Then, ten students at a tertiary level institution were interviewed about their experiences in learning writing in English in a retrospective manner. The devised FEDCom model illustrated the development of EFL writing of learners in each level of education. The model may be considered as the big picture of learning EFL writing in all education level of Turkey. It may help researchers, teacher trainers, and teachers who want to put forth solutions to the difficulties that learners may experience in learning writing in English.

Key words: EFL writing, Developmental writing, FEDCom, Holistic writing model, L2 writing

Introduction

Writing is an important part of communication and writing skill is considered as the most difficult skill (Phuket & Othman, 2015; Richards & Renandya, 2002) among the four skills. Nunan (1999) stated that producing a coherent, fluent, extended piece of writing is the most difficult task for language learners. Although writing seems arduous for both native and non-native learners, it becomes almost the most challenging task (Pajares, 2003) to the learners in contexts where English is taught as a foreign language (EFL henceforth).

English has been taught as a compulsory subject from primary school to tertiary education in Turkey for about two decades with the 1997 education reform (Kirkgoz, 2007). However, despite the innovations in the curriculum and the efforts spent, it is difficult to claim that English is taught well across the country. Many studies have shown that there are still serious problems in teaching EFL in Turkey (Incecay, 2012; Oktay, 2015; Solak & Bayar, 2015). Since benefitting from emerging approaches, such as CLT (Ozsevik, 2010), which are more effective in teaching a foreign language is considered challenging, traditional methods are highly implemented in language teaching in this context. Accordingly, some components of language (grammar, reading, and vocabulary) regarding receptive skills are more focused on than the other components regarding productive ones (i.e. speaking and writing) (Kırkgöz, 2007). Thus, some skills are learned better than the others. Writing in a foreign language is among the skills which have not been focused as it deserves, and consequently the desired level of success in this skill cannot be achieved by learners of English nationwide Turkey (Dogançay-Aktuna, 1998).

EFL writing, the issue under investigation of this study, is different from writing in English as a second language (ESL writing) in some aspects although they take place under the same roof in the relevant literature (Cumming, 2001a; Hasan & Akhand, 2010; Lee, 2007; Nation, 2009; Susser, 1994). English is taught as a foreign language in Turkey, which means learners learn the language in classroom-like environments in general, and they have very limited opportunity to practice it outside the classroom. This has several critical negative impacts on learning both receptive and productive skills of language. Unlike in ESL contexts where English is the dominant language, in EFL contexts people outside class environments barely use the language, and the

---

language only rarely plays a role in the mass media and in advertisements (Sivell, 2013). Thus, learners cannot be exposed to adequate amount of target language input and cannot become fully familiar with that language. Therefore, they cannot improve their listening and reading skills, which are essential for improving writing and speaking skills considering that listening and speaking, and reading and writing are interdependent (Moon, 2008).

Although there have been many studies focusing on writing in a second language, the relevant literature has paid insufficient attention to "the particular situation of writing instruction in non-English dominant countries" (Leki, 2001, p.197). The existing studies investigated such issues as peer and teacher feedback, instructor practices, improving students' composing skills, improving students' writing through study of semantic concepts, improving writing through micro-blogs, and using strategies to improve writing (Akyel & Kamisli, 1996; Alagozlu, 2007; Arslan & Sahin-Kizil, 2010; Bozkurt et. al., 2016; Erkan & Saban, 2011; Kirmizi & Kirmizi, 2015; Ozturk & Cecen, 2007; Trotman, 2010; Uysal, 2008). Since most studies concentrated on the ESL writing problems of students in center countries (Leki, 2001), little has been done about writing in EFL contexts.

Moreover, the majority of the research studies on EFL writing have focused merely on higher education. Additionally, the literature lacks a comprehensive picture of the issue in a holistic manner by taking all education levels into consideration. Therefore, this study is an attempt to explore the problems that learners encounter in EFL writing, from primary school to higher education. It intends to illustrate learners' developmental stages in improving EFL writing in each education level, and to suggest ways (a model) to achieve success in learning EFL writing.

Literature Review

Writing requires considerable number of skills and conventions such as writing readiness and grammatical rules, and only by achieving these prerequisites well, may learners get ready to be proficient and effective writers (Emmons, 2003). Considering that one's writing in his/her mother tongue is a complicated process, it can be claimed that writing in a second language (henceforth L2) should be more difficult and challenging. Cumming (2001b) states that "writing in a second language forms a focus for individuals to learn ways of cooperating with and seeking assistance from diverse people and resources; to adapt to and reflect on new situations, knowledge and abilities; to negotiate relations of work and power; and to gain and modify new senses of self" (p. 7). Consequently, it can be asserted that writing in a foreign language is a much more sophisticated issue.

Scholars like Zamel (1983), Flower and Hayes (1981), Choi (2016), and Li (2018) view writing as a complicated recursive process rather than a simple linear one. According to Zamel (1983) the process of writing is "non-linear, exploratory, and generative whereby writers discover and generate ideas as they attempt to approximate meaning". He states that this process involves such sub-processes as planning, collecting data, drafting, revising, rewriting and editing. It is understood that in order for a learner to write well in an L2, it is necessary to control over all this process along with the sub-processes, which require them to possess specialized skills (Hyland, 2019).

In addition to the complicated nature of writing, another factor that makes EFL writing more challenging is social and cultural factors which affect the quality of EFL writings of learners. These entail both the distance between the languages, the first language (L1 henceforth) of the learner and the target language, and the cultural differences. A good example of this is the study by Steffenson, Joag-Dev and Anderson (1979) and Yu and Lee (2016) in which the researchers illustrate the impact that culture has on schemata of people, which affect people's process of composing their writing.

Moreover, since different languages may have different rhetorical organizations, there may be negative transfer from L1 writing to L2. For instance, Ostler (1987) declared that "various cultures organize the development of ideas differently when writing expository prose" (p. 169). Bennui (2016), Kaplan (1966) and Derakhshan and Karimian (2020) express that failure in the second-language discourse rules is because of interference from the writer's L1 rhetorical organization.

Each foreign language context may have different complication in learning writing in the target language. One of them is the structural and cultural differences between the mother tongue and the target language regarding some aspects in writing. For the current context, for example, while there is almost always one-to-one sound-symbol correspondence in Turkish, the case in English is different, which complicates the work of learners especially in dictation. Another problematic issue the results of which reflect negatively on learning EFL writing

is that profession choice in Turkey generally depends on the placement scores in the university entrance exam where foreign language competence is either not measured or only partially measured excluding writing. The professional self is not the structure that emerges suddenly. It progresses through certain stages from childhood to adulthood (Super, 1990; Ginzberg, 1972). In other words, it has a developmental nature. Although it is located among the objectives of the Ministry of National Education, applications for students to develop their interests and skills from a very early age is not sufficient in Turkey (Yesilyaprak, 2007, 2012). The lack of clarity of interests and abilities causes students to have a low level of commitment to academic goals and choices. This reality emerges as an important problem for teaching English with four skills as it is for other courses. In this sense, instead of developing skills that are not measured in high-stake exams such as writing and speaking in the target language, students generally study just to pass the course or get high marks from these high-stake exams. As a result, the goal of learning to write in English generally fails.

Public school students in Turkey start engaging in EFL learning at primary schools. They have 2 hours of English classes a week at primary school (starting from the second grade), an average of 3,5 hours and 3 hours a week at secondary school, and high school respectively for each grade. The students must answer English language questions which measures just reading, vocabulary, and grammar knowledge of the students in High School Entrance Exam (LGS henceforth) before high school. After high school, students except those who want to study at English language teaching departments or similar ones do not have to take the language test component of the university entrance exam, which again measures only reading skill and grammar knowledge. The students who attend to the universities that have English medium instruction must join preparation classes if they fail in the language proficiency exam; others just take 3 hours of English courses a week in the first grade of their departments. Considering the above information, it is clear that just reading component of four skills is assessed at high stake exams at all levels, other components including writing are not.

To conclude, what make writing in a foreign language a challenging skill on the part of its learners is that it is much more demanding than writing in someone's mother tongue. This is because social and cultural differences impact writing, and different languages have different rhetorical organization. All these factors cause EFL writing to be considered as a challenging skill to be learned and mastered.

Problems in Learning Foreign Language Writing

Since writing in a foreign language is a challenging skill, problems have always been observed in various foreign language contexts concerning writing in English. In her study, Leki (2001) put the challenges in writing, which have been frequently faced in any EFL context, into two main categories: daily experiences of writing teachers and ideological challenges. While the first includes crowded classes, time constraints, and accommodating local needs, the second group comprises justifying the large investment to teach English, resisting center-imposed and top-down methods and materials, and dealing with students about the role that writing plays in their lives.

Ezza (2010) tried to display that "educational policies can have their role in the learner's writing problems" in a foreign language (p. 33). According to him, learners alone should not be blamed for their failure in writing. He stated that such factors as "teacher/student ratio, the number of students in the classroom, the number of writing courses, course materials, and teaching methodology, (p. 33)" have influence on Arab EFL learners' failure in writing.

Moreover, Younes and Albalawi (2015) explored the most common types of writing problems of English language learners at higher education level. Grammar, punctuation, and spelling problems were found among the most salient problematic areas in their study. Chen (2002) investigated the problems of university EFL writing in Taiwan. The study uncovered students' inability to use words properly and precisely. It also revealed problems in vocabulary, grammar, generating ideas, and writing organization due to the differences between their own language and English. The study of Ahmed (2010) discovered that students encounter problems concerning cohesion and coherence in their EFL essay writings. All these studies tried to demonstrate the problems in learning EFL writing.

There have been many studies focusing on writing in a second language, however, the relevant literature has paid insufficient attention to "the particular situation of writing instruction in non-English dominant countries" (Leki, 2001, p.197). The existing studies on EFL writing investigated such issues as peer and teacher feedback, instructor practices for writing assessment, improving students' composing skills, improving students' writing through study of semantic concepts, improving writing through micro-blogs, and using strategies to improve

writing (Cumming, 2001a; McMullen, 2009; Mermelstein, 2015; Muslim, 2014; Schenck & Choi, 2015; Yang, Badger & Yu, 2006).

As for the research on EFL writing, there have been insufficient studies informing about the reality of EFL writing learning across all education levels in the Turkish context despite a clear need for finding out the problems in this issue. Additionally, the majority of studies in the relevant local literature focus on foreign language writing at higher education level. These studies tried to explore writing anxiety (Kirmizi & Kirmizi, 2015; Ozturk & Cecen, 2007) and writing apprehension (Erkan & Saban, 2011), use of blogs to improve writing (Arslan & Sahin-Kizil, 2010; Bozkurt et. al., 2016), teacher oral feedback (Trotman, 2010) and peer feedback (Mıstık, 1994) on student writing, the comparison of L1 and L2 writing (Akyel & Kamisli, 1996; Uysal, 2008), motivating factors behind writing (Buyukyavuz & Cakir, 2014), and critical thinking in EFL writing (Alagozlu, 2007). L2 writing education in the levels before higher education, namely, in primary, secondary, and high schools, has been under-researched in the Turkish context.

Human development is an important process. Developmental psychology has explored this process within some perspectives (Miller, 2002). Many theorists have investigated the developmental changes in people's emotions, thoughts, and behaviors. For example, while Piaget (1964) examined cognitive development of human beings, Erikson (1956) addressed the development of human ego. As an important field of behavior in foreign language learning, EFL writing should also be examined from a developmental point of view. In this regard, various scientists have come up with theoretical explanations at macro and micro level in this subject. For example, in macro level, Cumming (2001b) pointed out that writing in a second language is a sophisticated issue, and he hinted that it is a developmental process. To him, in this developmental process, each education level is a new situation requiring particular knowledge and abilities. On the other hand, it may be deduced from what Zamel (1983) stated on the issue that writing process has a developmental nature in micro level; and planning, collecting data, drafting, revising, rewriting, and editing are developmental stages of writing in the micro level.

Many researchers have studied the problems of EFL writing and solutions to these problems (Ahmed, 2010; Alharbi, 2019; Al Seyabi and Tuzlukova, 2014; Barkaoui, 2019; Chen, 2002; Ezza, 1990; Leki, 2001; Younes & Albalawi, 2015;). However, these studies are not developmental in nature, and also, they have not investigated the issue in a holistic manner. The purpose of this study is not to focus merely on one education level like the existing studies in the relevant literature, but to depict the whole big picture of learning EFL writing in each education level within a developmental perspective. It aims to display the problematic areas and recurring challenges in learning EFL writing of students from primary school to higher education from the viewpoints of teachers and students. It then intends to present a clear model of the learning of EFL writing in the Turkish context within a developmental perspective. The research questions that guide the study are given below:

1. What challenges do Turkish learners of EFL writing come across in their education life from primary school to university?
2. What are the reasons for the challenges that are encountered by the learners of EFL writing in each education level?
3. What can be the solutions to the challenges that are encountered by the learners of EFL writing in each education level?

Method

This study has two groups of participants. The first group is teachers of English at different levels of education. Then, the second one comprises a group of first-year university students.

The teachers: This research study was carried out interviewing teachers/instructors who teach EFL at any education level. The researchers interviewed with 40 teachers in total; ten teachers at primary school, ten at secondary school, ten at high school, and ten at university. Half of the teachers from each stage of education were male and the other half were female. The teachers' ages ranged from 33 to 55. All of the participants were EFL teachers or instructors in a major city of Turkey.

The students: After completing the interviews with the teachers, the researchers also interviewed with five female and five male students at prep class in a state university. The students' ages ranged from 17 to 20. All participants were public primary-secondary and public high school graduates. The participants started learning English from the beginning level at university although they had taken English courses since primary school.

Criterion sampling was employed in the current study. According to Creswell (2007) this type of sampling "works well when all individuals studied represent people who have experienced the phenomenon" under investigation (p. 128). Teachers and students were randomly selected as participants among the individuals who meet certain criteria. Here, teachers should have at least ten years of experience. As for students, they should be university students who have low-level of English proficiency in spite of their previous language education. Since the teachers who have at least ten years of experience and who have taught English at different levels of education (primary, secondary, etc.), and the students themselves can be accepted as the most important source of information about students' learning experiences of EFL writing, they were determined as the participants of the study.

Interviews

Before the interviews with the teachers and the students, their consents were taken, and they were assured that they can withdraw their consent at any phase of the study. Each teacher and student participated in an interview with the same researcher. A semi-structured interview protocol was designed so that some developing topics during the interviews could also be included in the questions. After informed consent was taken from each participant, interviews were tape-recorded and transcribed subsequently.

Firstly, the teachers and instructors were interviewed one-on-one starting from the primary school EFL teachers to university EFL instructors. While the interviews with the teachers lasted about 20 minutes for each, the ones with university students took about 15 minutes on average. Researchers tried to receive as much information as possible about teachers' experiences on the difficulties and challenges that their students encounter in learning EFL writing. After these interviews, the students from university were also interviewed. In these interviews, students were asked about the current experiences as well as the practices they did in the past school years on EFL writing. The purpose of this second series of retrospective interviews with the students was to identify the students' own experiences concerning the challenges in learning EFL writing, and to triangulate the data that were obtained from the interviews with the teachers.

Trustworthiness

The researchers applied multiple procedures to ensure the trustworthiness and the credibility of the study (Guba, 1981; Strauss & Corbin, 1998). In the process of member checking, ten per cent of both groups of participants reviewed and responded to their transcripts of the interviews. Then, some colleagues of the researchers carried out peer reviews. The intercoder reliability was found as 88.88 percent based on the calculations stated in Miles and Huberman (2015). Additionally, feedback was received from many colleagues on the evolving theory and the interpretation of the data. The researchers identified the concepts in the data and examined them across the stages of the model. Grounded theory is not generalizable, but it may be transferable to other contexts. The degree to which this grounded theory is transferable is based on the thick descriptions provided in the study.

Data Analysis

Constant comparison method within a systematic design (Glaser & Strauss, 1967) was applied in this research study. The data were started to be analyzed as soon as the data collection started. The data were analyzed by using open, axial, and selective coding (Strauss & Corbin, 1998). The content was analyzed in sentences or groups of sentences converging in a single idea in the open coding. These units were coded to indicate that idea. All items were identified, and these were formed into abstract concepts in the axial coding. During selective coding, these abstract concepts were combined into one central category. Any new data were analyzed and constantly compared to the emerging model through the existing data. It was either integrated into the emerging model, or shaped the existing model through new emerging themes within itself. This process continued till the data saturated. Then, the FEDCom Model was formed along with four categories: a- foundation b-expansion c-development d-completion. Attributes of each of these categories were also identified. Developing codes, categories and themes were formed in an emerging manner rather than applying predetermined categories on the data (Creswell, 2007).

Findings and Emerging Theory

The data from each education level were analyzed first, and then the FEDCom model was produced by integrating the findings. For this reason, firstly, the findings for each level of education are provided. Following this, the integrated findings are presented within the devised model. The findings are provided in a way that each part of the all quadripartite (i.e. relating to primary-secondary-high school and university) research questions are addressed in the consecutive sections in terms of the level in question. The titles in the findings are presented from the most stated to the least one, the frequency numbers are not provided since they are thought not to illustrate much, though.

The Difficulties of Learning EFL Writing at Primary School

After the whole analysis, it was discovered that primary school is the *foundation* level for writing in English. The participant teachers stated that they help students establish foundation at this level for EFL writing. For example, teacher T_6 said this:

"Students have just begun writing in English at primary school, and this is the basis of the problems. It would be unfair to expect too much from new learners. They have been establishing the basis yet." (T_6)

As an answer to the first parts (i.e. relating to primary school) of the research questions, this section provides the causal and intervening conditions of the challenges that Turkish EFL writing learners come across in primary school level and the strategies that they apply to overcome these challenges.

Causal Conditions

Five causal conditions were determined based on the data about why students at primary school cannot improve their EFL writing. These causal conditions and some representative examples pointing to them were listed below:

Negative transfer: While writing in English, students try to write using Turkish language components. Most primary teachers stated that students transfer some of graphemic, syntactic, and semantic components of their L1 in their EFL writing. The teachers observed that primary school students mostly transfer graphemic structures. The effect of negative transfer of some graphemic forms can be illustrated in the below expression of T_6 who is an English teacher at primary school:

"For instance, they do this: You know the letter 'i'. In English, it is written as 'i' if it is not capitalized, and as 'I' (without a dot over it) if capitalized. We cannot get the students accept this. "Why is not there a dot in the capitalized one?" they say. Another example is that they use the letters found in Turkish but not in English like "ç" and "ğ". They do not want to use the letters 'w' and 'q' because they were not in the Turkish alphabet until quite recently. They match up English with Turkish." (T_6)

(Notes: 1- The letter "i" is written as "İ" with a dot over it. If you do not put the dot it represents another letter. 2- These letters and numbers in the parentheses point to the owner of the quotes.)

Lack of transfer of writing skills: Since the students do not have enough writing skill in their mother tongue, they cannot transfer those skills to writing in an L2. The teachers claimed that the students do not write enough in their L1 to gain such basic skills of writing as making sentences and putting them together to have a meaningful unit; thus, they cannot transfer such basic elements of writing skills to their EFL writing. One of the primary school teachers states that:

"We mostly do not write in English in primary school. We most particularly never write in the second grade since the students get difficulty writing even in Turkish…" (T_1)

Lack of writing experience: As students do not have enough writing experience in English, they cannot achieve the desired level of success in writing. It was observed that students generally do not try to write in English outside school. Teachers relate this to the fact that they cannot allocate enough time to practice different writing genres in classes so the students cannot do any practice on their own. The below expression of an English teacher at primary school displays this case clearly;

"We have so little writing activity in primary school… that we do not have them write at all. In fact, writing can be improved with homework and by writing more and more, but we do not assign homework. The more they write, the more they improve their writing skill. But do they write?" (T_5)

Lack of phonological awareness: The fact that students cannot comprehend enough the sound-symbol relationship in English language affects their writing negatively. The below answer to a question by a primary school teacher demonstrates students' lack of phonological awareness:

"They want to write whatever sounds they hear. They do not recognize sound differences. For example, when they hear "how are you", they write it as /hav ar yu/ since they got used to writing as they hear." (T_2)

(Note: In Turkish, generally there is one-to-one sound-symbol correspondence.)

Lack of imitative writing: Primary school students sometimes cannot even copy the writings on the board. Considering that this is a very basic sub-skill for writing, it can be claimed that such students are not ready to write even in their L1. T_8 says:

"English writing of our students at second and third grades are very problematic. They certainly cannot copy the writings on the board." (T_8)

The above findings indicate that primary school students get difficulty in writing in English due to negative language transfer, less writing experience, and their being unready for writing in an L2. Now, the following section will present the strategies of students that they use to cope with these difficulty areas.

Strategies to cope with the problems

In the analysis, it was found out that there were four strategies to cope with the challenges in learning EFL writing at primary school. These strategies can be seen as the positive counterparts of the abovementioned problems. These are positive transfer, enhancing writing experience, increasing phonological awareness, and achieving imitative writing.

Positive transfer: It is a fact that students' application of L1 writing knowledge and the strategies that they make use of while writing in their mother tongue sometimes facilitate their writing in English. Transfer of sub-skills of writing helped them to write well in the L2. The L1 composition classes seemed to promote their EFL writing:

"There are writing activities in Turkish that can facilitate writing. For example, they write compositions in Turkish. Since there are writing practices in other subjects, they get accustomed to writing, and this eases their writing in English." (T_9)

Writing Experience: The teachers observed that the more they do writing practice with their students and the more their students engage in writing activities, the better the students learn and practice EFL writing. Some teachers also asserted that even writing in L1 enhance students' L2 writing. To some teachers, the key word to improve EFL writing is 'writing more and more'.

Phonological awareness: As mentioned above the phonological differences between Turkish and English make writing challenging in some respect. Becoming aware of the phonological differences between these languages make EFL writing more convenient for the students.

"At first, they write whatever and however they hear. This happens when they first get acquainted with English. Then, they recognize phonological differences, especially at 4th grade. However, this problem keeps going in the least." (T_2)

Imitative writing: In some activities of junior primary school writing classes, students are given example sentences including some target chunks and they are supposed to write some other sentences by imitating the given examples. This imitative writing gets them accustomed to writing in full sentences and in a coherent way. One of the primary school teachers explains the reason why achieving imitative writing is important in learning EFL writing is as follows:

"When we give a chunk and tell them 'you will write like this.', they can write then at 4th grade. For example; introducing family. When I ask them to introduce all members of their families, if I do not give the model in chunks, they fill just the names of their fathers, or age in the blanks like 'Ahmet, 38' etc. rather than in full sentences." (T_6)

So far, the causal conditions behind the problems of EFL writing in the primary school and the strategies to cope with these problems have been presented. In addition to these conditions some other factors that influence students' EFL writing were discovered. These factors are provided in the following section under the name of intervening conditions.

Intervening Conditions

The analyses of the data displayed that some factors can either complicate or facilitate the process of learning EFL writing. Nevertheless, the majority of these intervening conditions emerged from the data were evaluated as complicating factors in that they aggravate students' writing in English. For the first developmental stage, the data revealed nine intervening conditions:

Curriculum: Curriculum has frequently been criticized by the teachers and instructors as being irrelevant and intense. Although defined as intense, it is evaluated as lacking in guiding the writing process. Primary school teachers' statements show the important role that the curriculum plays in learning EFL writing:

"As for our curriculum… Our books, the topics in them and topic distribution are definitely not suitable for primary school level. There are unnecessary topics. For example, the kid does not know what the simple present tense is in Turkish, but we try to teach it in English. But the kid does not know! Again, the subject 'adjectives' is given in the 3rd grade book. The 3rd graders do not know what an adjective is in Turkish! And we try to teach them in English (what an adjective is and how it is used). We need Turkish base first, only then can we design the curriculum, and teach English in parallel with their counterparts that are learned in Turkish." (T_8)

Teacher: Teacher is of course a factor that can either facilitate or debilitate the process of EFL writing. Teacher participants themselves explained that if they get learners do more writing practice, if they make the courses interesting to learners, and if they try to make writing in English a part of learners' life, the learners both get motivated and achieve more success. Otherwise, students cannot achieve the desired level of success in EFL writing. While some of the teacher participants stated that they could be more helpful in teaching writing to students, some others claimed they do as much as they can in guiding them to write:

"Not only the students themselves but we, teachers, have influences on their writing. We can facilitate their learning of writing by endearing writing." (T_8)

"When the teacher (form tutor) does not broaden a kid's horizon or when she does not have him/her read books, then the kid gets difficulty in learning reading, and therefore, writing in English." (T_3)

Insufficient Training Hours: Almost all the participants are of the same opinion in that foreign language training hours are not adequate. Thus, they say they cannot allocate enough time for writing. Some teachers claim that if adequate time was given for the training, students would surely be better in writing in the L2 language since more time would be allocated for writing classes:

"Time is not enough; I mean lesson hours. We cannot finish activities. If I could allot some other time for writing, then maybe we could write more and get over this problem. We have 2 hours a week. Our books are heavily focused on listening. Writing is totally left aside especially in primary school." (T_4)

Cursive Writing: Primary school teachers observed that their students get difficulty in integrating some specific letters that are not in the Turkish alphabet into the words they write in cursive writing. So, it is among the factors that negatively affect EFL writing:

"First of all.. The process of cursive writing has affected them negatively. It's clear that as they haven't been trained for such letters as x, w etc. they get difficulty while trying to write them in cursive writing." (T_10)

(Note: Letters such as x and w are not in the Turkish alphabet)

Material: Participants also mentioned the role of materials in teaching writing effectively. While some of the teachers and instructors who have adequate and appropriate materials stated their satisfaction, the majority who

lack these materials reported this as a significant obstacle before teaching writing. Most teachers complained about inadequate and inappropriate materials that were provided to them for writing classes. Although there were some teachers who were satisfied with the materials, most of them had negative statements about them.

Method: The method that teachers and instructors apply in their classroom has an either facilitating or a complicating effect as well. Both cases were reported as below by the participants:

"We teach most with visuals. They accommodate visually not auditory. I do not do dictation exercises until fourth grade. Therefore, they can write it as soon as they see the picture because they learn it in that way. If you say it, they cannot write, they are not get used to it." (T_5)
"It can be presented as a game. Writing with the smart board banners may come differently to them, more enjoyable. We can make use of them; I use them in fact." (T_4)

Motivation: Primary teacher participants' observation on their students is that they are motivated to learn how to write in another language. In this sense, motivation was found to be more as a facilitating factor in learning EFL writing in primary school. However, as motivation can be in any point between two ends of a continuum, it should be regarded as an intervening factor.

Information Processing Capacity: The fact that primary school students process limited information at a given time makes it difficult for them to think and compose during writing. The following quote from a teacher explains this case:

"They are slow and they make mistakes because they are very younger. They do spelling mistakes as well, particularly 2nd graders. However, when they proceed to 4rd grade, they are getting better. They write faster, and they make fewer mistakes." (T_7)

Family/Culture: Families and the environment in which students live also have influence on their learning of EFL writing. While the students whose families are interested adapt to writing more quickly, those who come from indifferent families have trouble in achieving success.

"I think the background is a bit of education families give their children. The child is successful if he is educated in literacy, science, and so on in family. It starts from the family." (TS_9)
"Families need to be informed about English teaching, and its importance." (TS_2)

It was found that primary school is the level at which students lay the foundation of EFL writing. They suffer from interference, lack of writing experience, lack of transfer of writing skills they had in their L1, and lack of some sub-skills of writing in an L2 in this foundation stage. The causal conditions of EFL writing failure in the foundation stage, the strategies to cope with the problems, and the intervening factors in students' learning of EFL writing can be illustrated smoothly in a figure.

In figure 1 below, the findings for the foundation level were presented in a refined manner as part of the FEDCom Model which was provided at the end of the findings section. The reason why cursive writing is shown in red is that it is found unique to the foundation level.
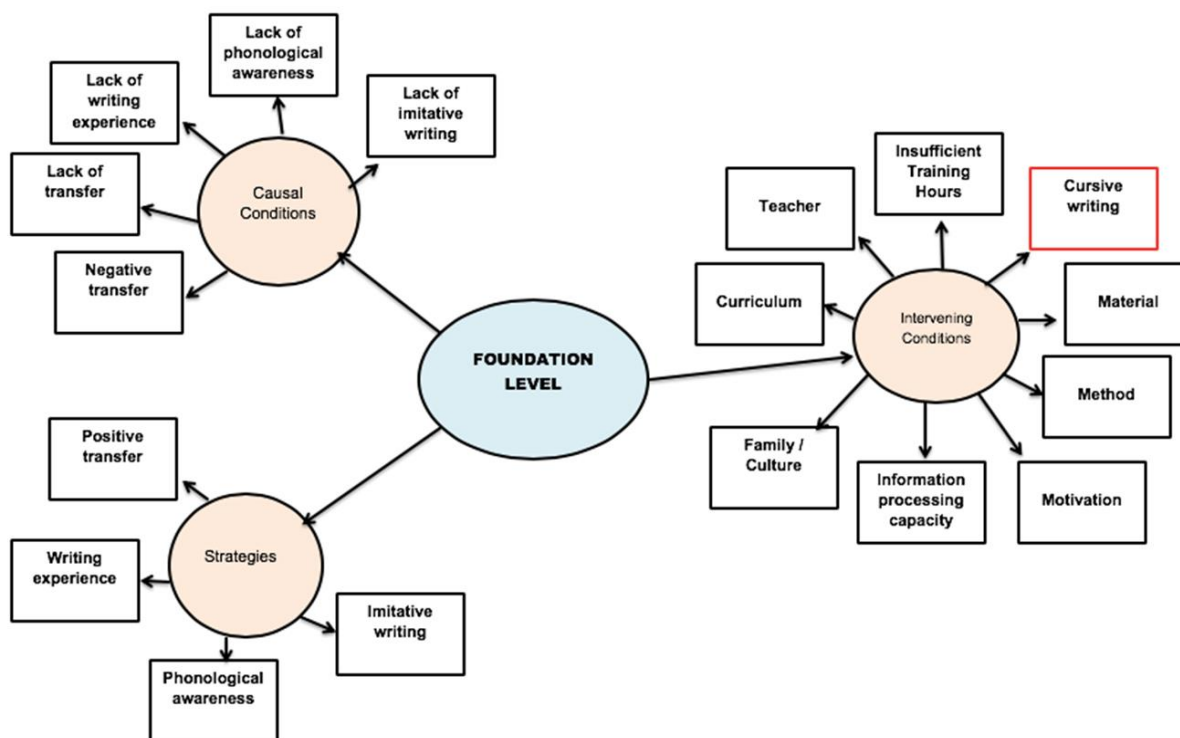
Figure 1. Causal conditions, strategies and intervening conditions of EFL writing problem at *primary school* level

The Difficulties of Learning EFL Writing at Secondary School

Secondary school was considered as the period of expansion for writing. It is deduced from the participants' views that EFL writing, the foundation of which was laid on in primary school, is expanded during this level. For example, a lower secondary school English teacher made the following statement:

"They improve it (their writing) at the end of the 5th grade. The spelling mistakes reduce. We want them to make sentences with the words they have learned so that the vocabulary will be permanent. They learn more about writing when they move towards the 7th and 8th grades." (T_7)

In this section, the findings from secondary school data were presented as an answer to the second parts (i.e. relating to secondary school) of the research questions. At the end of the section, the expansion part of the FEDCom model was illustrated in a figure.

Causal Conditions

It was discovered that the same causal conditions namely negative transfer, lack of transfer of writing skills, lack of writing experience, lack of phonological awareness, and lack of imitative writing are also valid for secondary school students' EFL writing failure. One extra causal condition, lack of grammatical knowledge, which was not seen in the previous stage, was revealed in this one. The secondary school teachers stated that one of the most problematic areas they encounter in EFL writing classes is grammar. Some example expressions of the participants which are representative of the case were provided in the appendix.

Strategies to cope with the problems

From the analysis of the data, it was ascertained that strategies/solutions to cope with the problems in learning EFL writing in secondary schools were nearly the same as those in primary school: positive transfer, increasing the amount of writing experience, raising phonological awareness, achieving imitative writing, and improving grammatical knowledge. Although the names of the strategies are the same as those in the primary school level,

the content and quality of these strategies differ. For instance, while writing basic sentences that students always hear and see can be a "writing experience" in the primary school level, in the secondary school this "writing experience" might be writing diaries including simple sentences just as one of the secondary school teachers expressed:

"In order that the students improve their writing skills, they have to engage in a variety of writing exercises. For instance, they can write what they do during a day in their diaries." (T_13)

Intervening conditions

In addition to the causal conditions of the phenomena, the data revealed eight intervening conditions, the majority of which is almost the same as those detected for the first stage. These are curriculum, teacher, insufficient training hours, material, method, motivation, information processing capacity and family/culture. Although they were named the same as those in the previous stage, they are different from them in terms of content and quality. For instance, *Motivation*, in contrast to the previous stage, was found as a more debilitating factor in secondary school EFL writing. Students want to prepare for LGS (Liselere Geçiş Sınavı-High School Entrance Exam) in which there is multiple-choice English test which has little weight on the overall exam, rather than learning English for communicative purposes.
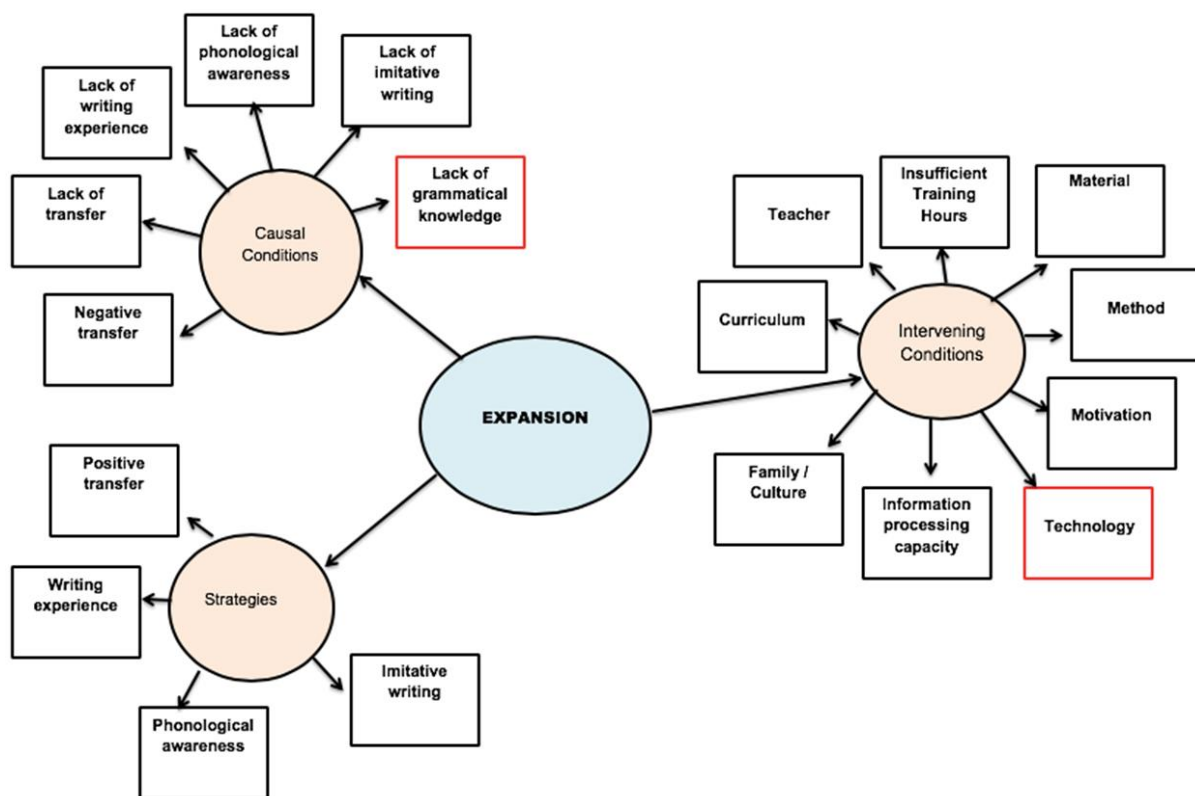
Figure 2. Causal conditions, strategies and intervening conditions of EFL writing problem at *secondary school* level

Technology emerged as another debilitating factor in this level although it was not observed in the data relating to primary school level. The teachers complained that students engage in mobile devices too much and this prevents them from writing. They think that playing with recent technological devices is easier for the students than engaging in any kinds of writing which students see a challenging task:

"They like playing with computer or phone. They are not struggling with writing. There is too much distraction." (T_18)

The conditions that caused failure in EFL writing in secondary school (Expansion) level, the strategies to overcome the challenges and the factors which sometimes facilitate but mostly debilitate learning of EFL writing can be seen in figure 2 above as part of the FEDCom model.

The Difficulties of Learning EFL Writing at High School

High school years were described as *development* period in terms of students' EFL writing. The teacher participants indicated that in this period students move further ahead of the level that they reached at secondary school. The below statement of a high school English teacher is an example among many which pointed this case:

"They are beginning to write simpler. But later on, in conjunction with the subject, the writings are further developed when the examples are given in front of them. They move their writings further ahead the level they achieved before high school." (T_23)

In this section, the findings addressing the third parts (e.g., relating to high school) of the research question regarding the causal and intervening conditions of the challenges that Turkish EFL writing learners encounter in high school were presented. At the end of the section development part of the FEDCom model was illustrated in a figure.

Causal conditions

The majority of the factors that negatively affect high school students' EFL writing are almost the same as those in the previous stages. The causal conditions in high school level are negative transfer, lack of transfer of writing skills, lack of writing experience, lack of grammatical knowledge, limited vocabulary, and lack of productive writing. On one hand, it can be observed that the causal conditions 'lack of phonological awareness" and "lack of imitative writing" which both emerged in the previous levels did not emerge among causal conditions for high school students' underdeveloped EFL writing. On the other hand, two extra causal conditions, "limited vocabulary" and "lack of productive writing", which were not seen in the previous two stages, revealed in this one. The participant teachers from high school stated that students lack the necessary vocabulary for writing. They added that students are not productive enough to express themselves in writing. The below quotes from two different high school teachers show the case clearly:

"Writing is really a skill. In other words, you have to have vocabulary knowledge, and grammatical knowledge to produce something. They get in trouble at this point. The vocabulary of the students is not sufficiently developed." (T_30)
"Speaking and writing are productive skills. You have to produce something, and those who are not so imaginative cannot achieve it. That's a complication." (T_24)

Strategies to cope with the problems

The strategies that should be used to overcome the challenges in writing English in secondary school were positive transfer, enhancing writing experience, improving grammatical knowledge, having rich vocabulary, and productive writing. Different from the ones in the previous stages, "having rich vocabulary" and "productive writing" emerged as the new strategies to cope with the problems in EFL learning at high school level. Teachers emphasized the importance of rich vocabulary and productive writing in learning EFL writing:

"As students' vocabulary [knowledge] increases, their writing gets richer." (T_22)
"By giving an example, you can say, "Here is how it is written; you can do it like this". You might want something like their rewriting something by changing a few words. It is first because of creating the feeling "I can write"." (T_26)

Intervening conditions

Some factors that were emerged from the data were found as having either positive or negative effect on learning EFL writing. The intervening conditions discovered in the data relating to high school level were almost the same as those in the previous levels. However, two new different intervening conditions also emerged in high school data, which are "individual differences" and "change in developmental task".

High school teachers stated that the ability to write in a foreign language change from based on the individual. While some students are really good at EFL writing, some others are really inadequate as if writing was not for them:

"Some people like to work individually. Some of them enjoy group work. There are individual level differences too. Some students may be able to write better than others." (T_22)

In adolescence, individuals form their identity and develop their autonomy (Erikson, 1956; Steinberg, 2005). These changes in developmental tasks are reflected on their writings. For example, a high school teacher made the statement below:

"Another factor is they are in adolescence. They cannot express themselves well. They cannot express them in a planned way. When they write something, they do it very complex and complicated. However, they get happier if we help them plan and guide them. But they also complain about it: "You are limiting us." First, you will not block their autonomy. You will leave them free. It (the thing they write) should be more flexible. You cannot expect students to be autonomous in teacher-dominated class." (T_26)

The conditions that prevented success in EFL writing at high school level (Development), the strategies to overcome the challenges of the issue, and the intervening factors -sometimes facilitating but mostly debilitating learning of EFL writing- were presented as part of the FEDCom Model in the figure below:
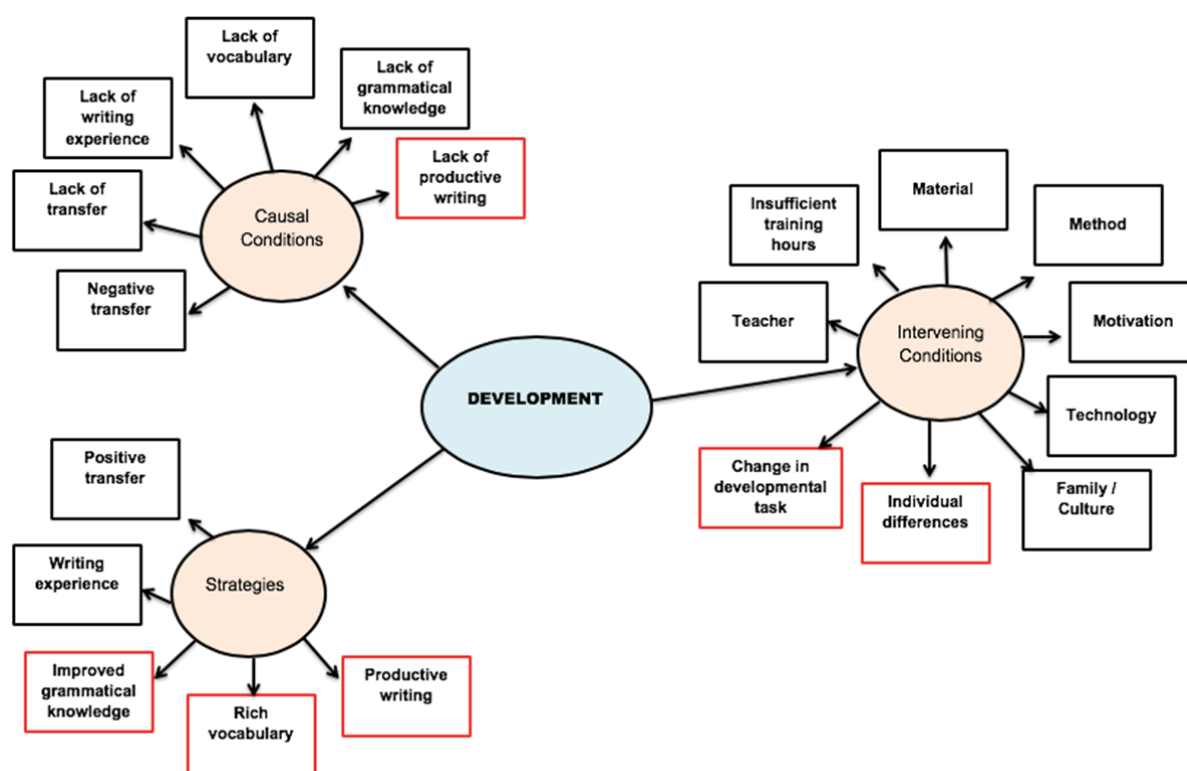
Figure 3. Causal conditions, strategies and intervening conditions of EFL writing problem at *high school* level

The Difficulties of Learning EFL Writing at University

University years were considered as the *completion* level with regard to students' EFL writing. Based in primary school, expanded in secondary school and developed in high school, students' EFL writing is completed at university. Participants in the study also stated that this period was the period in which they nearly completed learning how to write on their own:

"A little more emphasis should be put on writing in pre-university education. It gets a little difficult (late) to acquire writing skill at college. They should acquire it in the previous stages." (T_32)

In this section, the findings from high school data were presented. At the end of the section both the development part and the whole FEDCom Model were illustrated in a figure.

Causal conditions

The causal conditions of students' underdeveloped EFL writing that were discovered at university level were negative transfer, lack of transfer of writing skills, lack of writing experience, limited vocabulary, lack of grammatical knowledge, lack of productive writing, and lack of rhetorical organization. All these factors, except one, were also observed in the previous stages (ie. primary, secondary, and high school). Although almost all these factors have the same titles, they may differ from each other on the basis of content and quality. To exemplify, while writing basic sentences that students always hear and see can be a "writing experience" in the primary school level, the "writing experience" at university can be writing in a variety of genres and writing essays that are rich in vocabulary.

The only different causal condition emerged at this level was "lack of rhetorical organization". The instructors working in the preparation school of a large university observed that their students cannot organize their ideas in writing an essay. One of them commented on the issue as below:

"Generally, they cannot organize sentences, ideas, etc. in a proper and smooth fashion. We observe that they see writing as a task. They do not want to seek ways for improving their abilities on writing organization." (T_31)

Strategies to cope with the problems

In addition to the strategies (i.e. writing experience, positive transfer, improving grammatical knowledge, having rich vocabulary, productive writing) that were found in the previous levels, "learning rhetorical organization" was discovered as another strategy to cope with the problems of EFL writing at university. In order to show that the strategies having the same title with those in the previous ones are different from them, two representative examples were provided below:

Positive Transfer: "I'm trying to raise their metalinguistic awareness. 'See why we are learning bla bla, because of bla bla'. I raise their awareness. The structure of each language is different. You cannot translate a sentence that you set in your mind in mother tongue to foreign language. I say that you have to express it in its simplest form in the target language according to the rules." (T_34)

Productive Writing: "First, you need to write essays together with students. Secondly, you need to have them write in class. Otherwise, it is not all right when you want them to write at home. And they need to write on different types and issues. They have to experience it to produce something original on their own later." (T_39)

Intervening conditions

The factors which either facilitate or debilitate learning EFL writing at university level were determined as teacher, material, method, motivation, technology, family/culture, and individual differences. All of these factors were observed in the previous levels. However, the fact that they have the same names as those in the former levels does not mean they are the same. In fact, all of these elements, both conditions (i.e. causal conditions and intervening conditions) and strategies in each level may differ from each other in terms of content and quality

For example, while *method* as an "intervening condition" in the primary school may mean using visuals in writing classes, method as an intervening condition in university may mean guided writing or free writing. To illustrate, it is useful to highlight this difference by giving two sample examples from the data relating to primary school level and university level respectively:

"We teach most with visuals. They accommodate visually not auditory. I do not do dictation exercises until fourth grade. Therefore, they can write it as soon as they see the picture as they learn it in that way. If you say it, they cannot write, they are not get used to it." (T_5)

"When we first established the prep school, METU had a writing book. And we started with guided writing there. For example, it is written there; I, 25, London, student etc. The student wrote by combining these prompts. Free writing followed this. First guided and instructed writing then free writing. Step by step because it is always recommended in seminars and conferences. Personalization is very important. For example, like and dislike. I am asking immediately; 'what do you like / dislike?'. They learn more easily. They will learn chunks first. I am opposed to starting with free writing." (T_39)

The causal conditions behind the challenges in EFL writing, the strategies to overcome the challenges, and the intervening factors which mostly make learning of EFL writing difficult were given in the figure below:
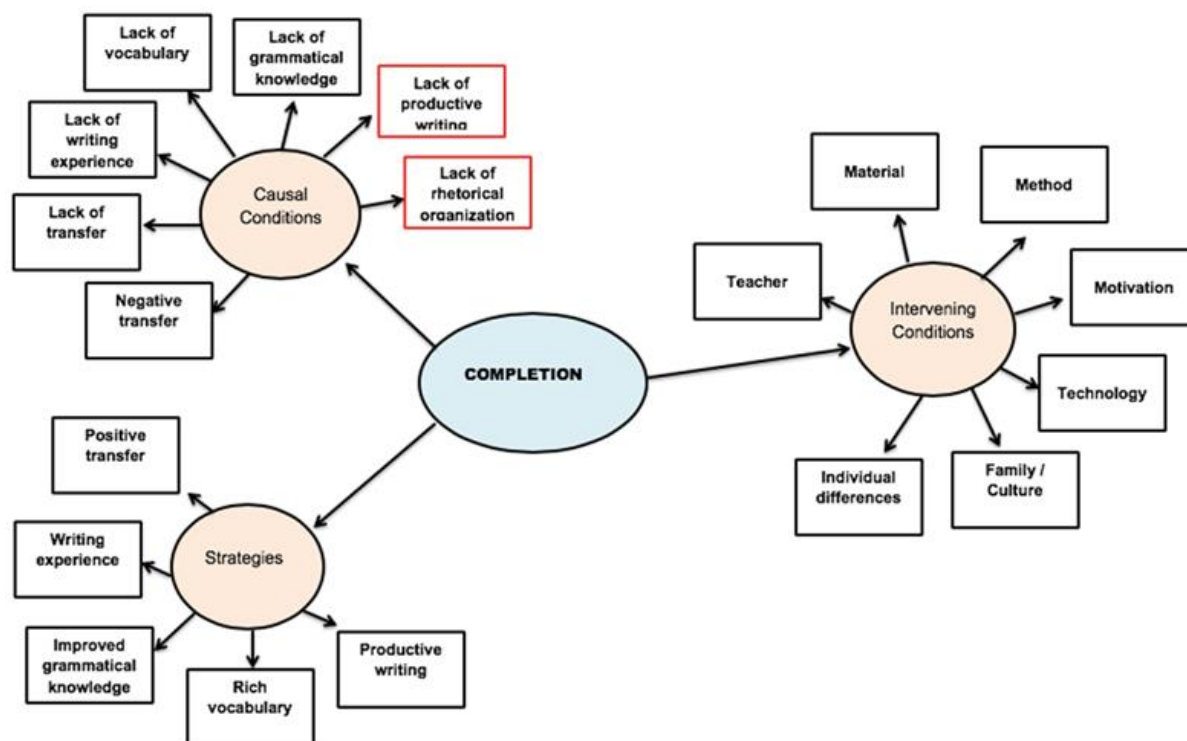

Figure 4. Causal conditions, strategies and intervening conditions of EFL writing problem at *university* level

The findings obtained from the teachers from each stage of education display an integrated structure with the FEDCom model. The expression of the participants supports this argument. The following expression points to the whole of the FEDCom model:

"Language learning in general and writing in particular require continuity. If you say I learned everything in the class and go home and do not practice, you will not achieve success. The research shows that if you repeat the same day no matter which subject you learned; you will know 80 percent of that subject. You can never hold over most particularly in writing. It needs continuity. It should start at primary school, be developed in secondary and high school, and students should be ready to write in a foreign language till university." (T_40)

Results Relating to the Retrospective Interviews with the University Students

The results of the retrospective interviews with university students confirmed the model discussed in the study. All of the causal factors, strategies, and intervening conditions that were emerged from the interviews with the teacher and instructor participants were also seen in the data obtained from the retrospective interviews with the university students.

Some of the statements of the university students that point to the FEDCom model were provided below. Some students gave the answers below to the question concerning their writing lesson experience in primary school, secondary school, high school, and finally in university. These statements can be accepted as an indication of the fact that primary school, secondary school, high school and university are where learning of EFL writing is founded, expanded, developed, and completed respectively.

Foundation:
"What we were doing .. In our elementary school, for example, there were pictures, and we were supposed to write what we saw in those pictures. After you already knew simple words, you could do this. We were writing on word basis." (S_3)

Expansion:
"In 7th and 8th grades, we were doing some studies on what we learned in elementary school. In 6th, 7th, 8th grades, we would write much more than we do in the primary school. We move on to write sentences instead of writing just words." (S_9)

Development:
"We were placed according to our level of English. The ones who reached a certain level in English would join the higher level class." (S_7)

Completion:
"I think we integrate the knowledge and skills on writing in English at the university that we brought with us from high school. Everyone in college must write in their area of expertise and gain independence in writing." (S_2)
The two statements of different students below point to the FEDCom model in a whole:
"Writing should be taught starting from primary school and then gradually be progressed." (S_1)

"I started learning English in third grade. In class four, we gained both grammatical knowledge and other skills. Our vocabulary level rose. When I came out of primary school and began secondary school, we started to write short paragraphs gradually. I move on to the high school I improved my writing a little bit further. I reached the point where I could write an essay in my own field." (S_6)

The findings show that writing is developmental in nature. The data obtained from both teachers and students supported the FEDCom model. The model incorporating conditions of EFL writing at all levels of educations is given in figure 5 below.
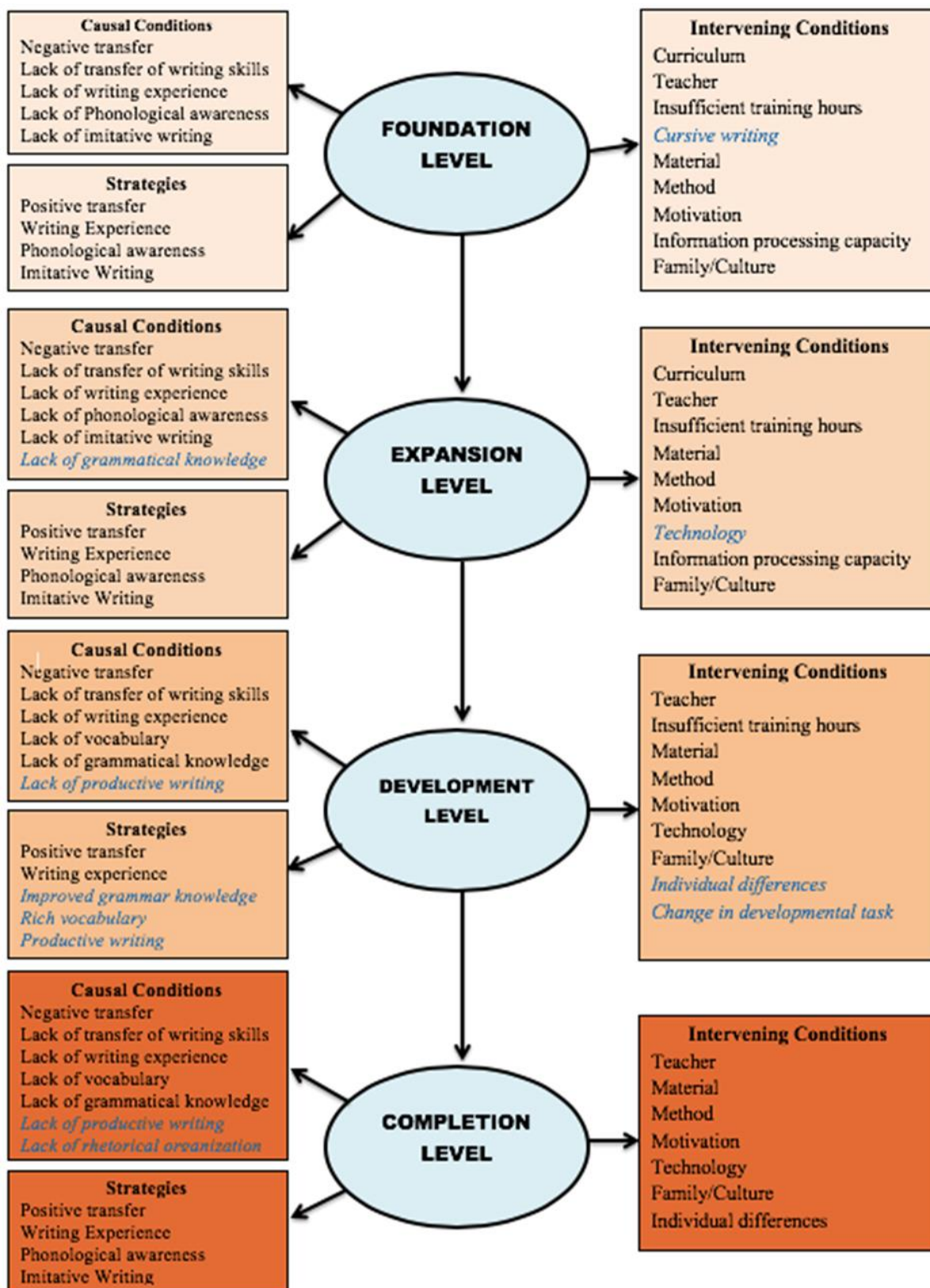
Figure 5. The FEDCom Model of EFL Writing

Discussion and Conclusions

This qualitative study aims to determine the challenges the students face in EFL writing primarily from the perspective of teachers at all levels from primary school to university; secondarily from the perspective of a group of university students retrospectively. It intends to provide insight for the parties who want to put forth solutions to these challenges. The findings of the study reveal causal and intervening conditions for the problems encountered in EFL writing at each stage and strategies to overcome them. The study yields a four-stage FEDCom model, which indicates EFL writing is *based* in primary school, *expanded* in secondary school, *developed* and *completed* in high school and university respectively.

The results of this research have similarities and differences with the studies in the literature on the challenges of learning EFL writing. Studies in the relevant literature have discovered similar challenges in learning EFL writing in different contexts such as the application of inappropriate methods and materials in Ukrainian (Tarnopolsky, 2000) and Iranian (Zoghi & Fakhimie Shokri, 2019) contexts, students' lack of transfer of writing skills due to their insufficient writing experience in L1 in Kuwait context (Al-Zankawi, 2018), negative transfer, phonological unawareness, grammatical problems, limited vocabulary, lexical bundle and crosslinguistic influence (Güngör & Uysal, 2020), and problems in rhetorical organization in Taiwanese university (Chen, 2002) and Indonesian contexts (Hidayati, 2018), punctuation and spelling problems in Saudi Arabian university (Younes & Albalawi, 2005); cohesion and coherence in Palestinian context (Hammad, 2016) and other various contexts (Polio, 2017); the effect of educational policy in Iranian (Zenouzagh, 2018) and Chinese (Zhao & Huang, 2020) contexts; and insufficient time allocation for writing, curriculum, and imposition of western writing pedagogy (Leki, 2001). Moreover, the findings are parallel with the conclusion that writing process is a complicated recursive process rather than a simple linear one (Zamel, 1983). Finally, having found that the culture and negative transfer from the mother tongue impact EFL writing (Güngör & Uysal, 2020), the study affirmed the studies of Kaplan (1966), Ostler (1987), Derakhshan and Karimian (2020), and Yang (2019).

There are other studies that found some other causes behind learning EFL writing, which are different from the ones the current study revealed. These include insufficient writing teacher training, writing teachers' lack of writing experience (Leki, 2001), EFL learners' inability to form their own voice in writing (Alagozlu, 2007; Ramanathan & Kaplan, 1996). On the other hand, the study found out some other factors different from those in the existing studies behind the problems in learning EFL writing. These factors cover lack of imitative writing and cursive writing (mostly in elementary level), family, and technology.

Although there have been many studies focusing on the problems of learning EFL writing (Ahmed, 2010; Al-Zankawi, 2018; Derakhshan & Karimian, 2020; Ezza, 2010; Hammad, 2016; Kaplan, 1966; Leki, 2017; Ostler, 1987; Tarnopolsky, 2000; Younes & Albalawi, 2005; Zenouzagh, 2018), this study differs from those ones in some respects. Firstly, the study is an original one in that it presents the causes behind EFL writing problems, the strategies needed to overcome these problems and the intervening factors in EFL writing by making classification. Second, while the majority of the existing studies focus on tertiary level EFL writing, this study tries to capture the phenomenon in a holistic manner by gathering a large amount of data at all levels from primary school to university. Last but not least, this study addresses learning of EFL writing within a developmental perspective.

The findings of this study extend some of the existing declarations in EFL writing literature (Flower & Hayes, 1981; Perl, 1980). For example, Zamel (1983) implies that EFL writing has a developmental nature in micro-level (planning, collecting data, drafting, revising, rewriting, and editing). Each dimension of the model grounded on this study has developmental nature in its own right as well. For example, in writing at primary school level, students go through some developmental sub-stages such as phonological awareness, learning vocabulary and basic grammar, and imitative writing. In addition, each stage in the model can be considered as a prerequisite development period for the next stage. At this point, this study is different from other studies by contributing to EFL writing as it brings developmental perspective to the field.

The results of this study revealed important findings. First of all, this study shows that EFL writing is a whole set of sub-skills that progress in a developmental process. At this point, it verifies the information in the literature (Brockman & Taylor, 2016; Rinehart & Thomas, 1993). The foundation of EFL writing is laid on in the primary school (Foundation), and this skill is expanded in the secondary school years (Expansion). Then, high school years are noted as the years of development of this skill (Development). Finally, university years are regarded as a process in which the developing structure is completed (Completion). These are considered as a model for how a healthy writing development should take place. However, both the causes stemming from the education system and the individual causes prevent the successful completion of this process. For this reason,

learners remain at a certain stage of FEDCom or fall into a lower-level developmental stage. In other words, problems that cannot be solved continue developmentally.

This study revealed that learning EFL writing has a developmental nature. Findings at this point can also be handled from the perspective of developmental psychology. According to developmental psychology, individuals in certain age groups exhibit certain developmental characteristics. In this sense, it is stated that both language and cognitive development progress through various developmental periods (Santrock, 2006). The current findings show that the fact of developmental periods and developmental tasks is also valid in learning to write. According to the results of this study, causal conditions, strategies and intervening conditions have the same names but they are qualitatively different. Besides, there are some other causal and intervening conditions and strategies that are specific to each level of education. These similarities and differences prove that EFL writing has a developmental nature. *The strategies* in this study have shown that there are developmental writing tasks that must be accomplished by students in certain time periods in EFL writing, just as it is in developmental psychology (Havighurst, 1972), which can be evaluated as the critical tasks of learning EFL writing. On the other hand, *the causal conditions* have shown the problems that will arise if these developmental writing tasks cannot be overcome successfully. The factors that were found to intervene in students' learning of EFL writing, just as there are such factors in the developmental psychology as "scaffolding", "culture", and "language" which were expressed by sociocultural theory (Vygotsky, 1980); and some others as "social transfer", "readiness", "experience" and "maturation" that Piaget (1964) described as the factors affecting cognitive development. These factors were taken as "intervening conditions" in this study.

According to Piaget's cognitive development theory, assimilation and adaptation are two major influences on the process of cognitive development (Flavell, 1963; Piaget, 1964). It may be stated that the *causal conditions* in the FEDCom model of EFL writing in this study in fact correspond to *assimilation* in Piaget's (1964) cognitive development theory. *Strategies* on the other hand can be evaluated as the *adapted factors*. According to Piaget's (1964) theory of cognitive development, individuals or institutions around the individuals accelerate cognitive development of individuals as a means of social transmission. At this point, some of the intervening conditions in this study can be seen as means of social transmission. In addition to all these, in every developmental period, the important role that experience plays in EFL writing is emphasized. Piaget (1964) advocated that development progresses more healthily and rapidly, with increasing experience.

The causal conditions, strategies and intervening conditions may demonstrate different characteristics based on the level although the majority of them are found across all levels. For example, while negative transfer in primary school was generally seen in letter level, it causes syntactic and semantic problems in later levels. In addition, some elements of the strategies and both conditions differed based on the level. For example, elements concerning productive writing, grammar, and vocabulary showed up in later levels. Two most striking examples for such differences can be seen in the examples of "change in developmental task" in intervening conditions of the *development* stage, and of "lack of rhetorical organization" in causal conditions of the *completion* stage. In adolescence individuals form their identity and develop their autonomy (Erikson, 1956; Majchrzak, 2018; Steinberg, 2005); and these changes in developmental tasks are reflected on their writings. In university, on the other hand, students are supposed to write advanced writing, which requires more knowledge of rhetorical organization. These examples and similar others also support the model in that EFL writing have different requirement and is affected by different situations in each developmental stage.

Leki (2001) questions the reason why EFL writing is taught, and she gives such reasons as a) professional purposes (e.g. teaching English) b) using it as a tool for professional development and c) using it for self-exploration regarding teaching EFL writing. However, there can certainly be some others who want to be or will be in a position that will not require the ability to write in an L2 (for example, an ordinary worker at a factory). In this case, it is meaningless to try to teach those students EFL writing. A similar situation applies to EFL writing teaching in Turkey. The questions that Turkish students cannot answer about why they should learn EFL writing may lie behind particularly the factors such as "family/culture" and "individual differences" which are among the intervening conditions of this study. In fact, since professional development or career development progresses parallel to the individual's self-development -it is the reflection of self-development, it can be argued that the answer to this question is that "the task of professional development in Turkey cannot be overcome successfully". Successful fulfillment of individual professional development duties during each developmental period leads to healthy individual career choices and a healthy professional development (Super, 1990). However, profession choice in Turkey generally depends on the placement scores in the university entrance exam. Thus, individuals cannot understand why it is necessary to write in English because they do not know whether they will use it. The reason why they do not know is that they are not directed to the professions earlier according to their interests and abilities.

There is classification of teaching L2 writing pedagogies such as controlled and guided writing, free writing, product writing, process writing and genre-based writing (Ferris & Hedgcock, 2005, p. 4-9; Grabe, 2001). The findings of the current study suggest that all these approaches should continue to be used in teaching EFL writing in the Turkish context. While controlled and guided writing should continue in the beginning levels (i.e. foundation and expansion), the other approaches should be applied based on the pedagogical objectives in later levels (i.e. development and completion). Actually, if we act from the Turkish case, the model that emerged in the work provides policy makers and theoreticians with ideas on teaching EFL writing. At this point, while EFL writing should be mandatory during foundation and expansion periods; it should be optional during development and completion periods in an EFL context similar to Turkey. Those who want to use English in their life and those who consider English as a part of their professional identity need to go through all stages of the model. The reason why it should be obligatory in foundation and expansion stages is that individuals have not yet shifted to a professional field because they cannot distinguish their interests and skills.

Consequently, this research study has presented a holistic and comprehensive perspective on the problems of EFL writing and solutions to these problems in the Turkish context. Necessary regulations, which informed by the findings of the current study, on teaching EFL writing can be done at each education level based on this model. Carrying out similar studies in different communities can help cross-cultural comparisons in the field of EFL writing. The model discussed in this study can also be quantitatively tested through developing an instrument based on these findings as well as on the existing studies. The dimensions discussed in the FEDCom model can be re-evaluated according to the results of quantitative studies. According to such findings, EFL writing development programs can be prepared.

Limitations

The study has three main limitations. Firstly, the findings that emerged from the data could have been could be subjected to measure quantitatively. Secondly, the implications of the problems on students' writing could have been presented, which could strengthen the existing findings. Finally, instead of interviewing with only university students, some other students from the other levels could have been interviewed.

Recommendations

The suggestions for future research can be put forth considering the limitations of the current study. Accordingly, the studies ahead can quantitatively test the findings that emerged out through the qualitative analysis that was done in this study. Moreover, tangible examples from the writings of the students from each level can be provided in a study that will focus on the current issue. In addition, longitudinal studies which focus on the progression of students' writing from at least secondary school to the end of the high school can be planned and carried out. Such studies can reveal more about the developmental nature of writing in a foreign language. Finally, future studies can focus on not only the problems but also possible solutions to these problems.

References

Ahmed, A. H. (2010). Students' problems with cohesion and coherence in EFL essay writing in Egypt: Different perspectives. *Literacy Information and Computer Education Journal (LICEJ)*, *1*(4), 211-221.

Akyel, A., & Kamisli, S. (1996). *Composing in First and Second Languages: Possible Effects of EFL Writing Instruction*. Paper presented at the Balkan Conference on English Language Teaching of the International Association of Teachers of English as a Foreign Language (2nd, Istanbul, Turkey, September 5-7, 1996). Retrieved from https://files.eric.ed.gov/fulltext/ED401719.pdf

Alagozlu, N. (2007). Critical thinking and voice in EFL writing. *Asian EFL journal*, *9*(3), 118-136.

Alharbi, M. A. (2019). Saudi Arabia EFL university students' voice on challenges and solution in learning academic writing. *Indonesian Journal of Applied Linguistics*, 8(3), 576-587.

Al Seyabi, F., & Tuzlukova, V. (2014). Writing problems and strategies: An investigative study in the Omani school and university context. *Asian Journal of Social Sciences & Humanities*, *3*(4), 37-48.

Al-Zankawi, M. (2018). *An Investigation of EFL Writing Strategies and Cohesion of Kuwaiti Undergraduate Students*. Doctoral dissertation, Trinity College Dublin, Ireland

Arslan, R. Ş., & Şahin-Kızıl, A. (2010). How can the use of blog software facilitate the writing process of English language learners?. *Computer Assisted Language Learning*, *23*(3), 183-197.

Barkaoui, K. (2019). What can L2 writers' pausing behaviour tell us about their L2 writing processes. *Studies in Second Language Acquisition*, 41(3), 529-554.

Bennui, P. (2016). A study of L1 intereference in the writing of Thai EFL students. *Malaysian Journal of ELT Research*, 4(1), 72-102

Bozkurt, A., Aydin, B., Taskiran, A., & Koral, E. (2016). Improving creating writing skills of EFL learners through microblogging, *The Online Journal of New Horizons in Education*, 6(3), 88-98.

Brockman, E., & Taylor, M. (2016). Four College-Level Writing Assignments: Text Complexity, Close Reading, and the Five-Paragraph Essay. *Teaching/Writing: The Journal of Writing Teacher Education*, *5*(1), 10.

Buyukyavuz, O., & Cakir, I. (2014). Uncovering the Motivating Factors Behind Writing in English in an EFL Context. *Anthropologist*, *18*(1), 153-163.

Chen, Y. M. (2002). The problems of university EFL writing in Taiwan. *The Korea TESOL Journal*, *5*(1), 59-79.

Choi, Y. H. (2016). Writing strategies in the process of L2 computer-mode academic writing with the use of multiple resources. *English Teaching*, 71(3), 3-28.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Sage Publications, Inc.

Cumming, A. (2001a). ESL/EFL instructors' practices for writing assessment: specific purposes or general purposes?. *Language Testing*, *18*(2), 207-224.

Cumming, A. (2001b). Learning to write in a second language: Two decades of research. *International journal of English studies*, *1*(2), 1-23.

Derakhshan, A., & Karimian Shirejini, R. (2020). An Investigation of the Iranian EFL Learners' Perceptions Towards the Most Common Writing Problems. *SAGE Open*, *10*(2), 2158244020919523.

Dogancay-Aktuna, S. (1998). The spread of English in Turkey and its current sociolinguistic profile. *Journal of multilingual and multicultural Development*, 19(1), 24-39.

Emmons, R. H. (2003). *An Effective Writing Formula for Unsure Writers*. http://www.airpower.au.af.Mil/airchronicles/aureview/1975/sept.oct/emmons.html

Erkan, Y. D., & Saban, A. I. (2011). Writing performance relative to writing apprehension, self-efficacy in writing, and attitudes towards writing: A correlational study in Turkish tertiary-level EFL. *The Asian EFL Journal Quarterly*, *13*(1), 163-191.

Erikson, E. H. (1956). The problem of ego identity. *Journal of the American Psychoanalytic Association*, *4*(1), 56-121.

Ezza, E. S. (2010). Arab EFL learners' writing dilemma at tertiary level. English Language Teaching, 3(4), 33-39

Ferris, D., & Hedgcock, J. S. (2005). Teacher response to student writing: Issues in oral and written feedback. *Teaching ESL composition: Purpose, process and practice*, 184-222.

Flavell, J. H. (1963). *The university series in psychology. The developmental psychology of Jean Piaget*. D Van Nostrand. New York

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College composition and communication*, *32*(4), 365-387.

Ginzberg, E. (1972). Toward a theory of occupational choice: A restatement. *Vocational guidance quarterly*, 20(3), 2-9.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory*. Weidenfield & Nicolson. London

Grabe, W. (2001). Notes toward a theory of second language writing. *On second language writing*, 39-57.

Guba, E.G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries, *Educational Communication and Technology*, 29(2), 75-91.

Güngör, F., & Uysal, H. H. (2020). Lexical bundle use and crosslinguistic influence in academic texts. *Lingua*, 242, 102859.

Hammad E.A. (2016) Palestinian University Students' Problems with EFL Essay Writing in an Instructional Setting. In: Ahmed A., Abouabdelkader H. (Eds), *Teaching EFL Writing in the 21st Century Arab World*. Palgrave Macmillan, London.

Hasan, M. K., & Akhand, M. M. (2010). Approaches to writing in EFL/ESL context: Balancing product and process in writing class at tertiary level. *Journal of NELTA*, 15(1-2), 77-88.

Havighurst, R.J. (1972). *Developmental tasks and education*. New York: David McKay.

Hidayati, K. H. (2018). Teaching writing to EFL learners: An investigation of challenges confronted by Indonesian teachers. *Langkawi: Journal of The Association for Arabic and English*, *4*(1), 21-31.

Hyland, K. (2019). *Second language writing*. Cambridge university press.

Hyland, K., & Hyland, F. (Eds.). (2019). *Feedback in second language writing: Contexts and issues*. Cambridge university press.

İnceçay, G. (2012). Turkey's foreign language policy at primary level: Challenges in practice. *ELT research journal*, 1(1), 53-62.

Kaplan, R. B. (1966). Cultural thought patterns in inter-cultural education. *Language learning*, *16*(1-2), 1-20.

Kirkgoz, Y. (2007). English language teaching in Turkey: Policy changes and their implementations. *RELC journal*, *38*(2), 216-228.

Kırmızı, Ö., & Kırmızı, G. D. (2015). An investigation of L2 learners' writing self-efficacy, writing anxiety and its causes at higher education in Turkey. *International Journal of Higher Education*, *4*(2), p57.

Lee, I. (2007). Assessment for learning: Integrating assessment, teaching, and learning in the ESL/EFL writing classroom. *Canadian modern language review*, 64(1), 199-213.

Leki, I. (2001). Material, educational, and ideological challenges of teaching EFL writing at the turn of the century. *International Journal of English Studies*, *1*(2), 197-209.

Leki, I. (2017). *Undergraduates in a second language: Challenges and complexities of academic literacy development*. Routledge.

Li, M. (2018). Computer-mediated collaborative writing in L2 contexts: An analysis of empirical research. *Computer Assisted Language Learning*, 31(8), 882-904.

Majchrzak, O. (2018). *Learner identity and learner beliefs in EFL writing*. New York, NY: Springer.

McMullen, M. G. (2009). Using language learning strategies to improve the writing skills of Saudi EFL students: Will it really work? *System*, *37*(3), 418-433.

Mermelstein, A. D. (2015). Improving EFL learners' writing through enhanced extensive reading. *Reading in a Foreign Language*, *27*(2), 182.

Mıstık, S. (1994). *The effect of peer feedback on the development of Turkish EFL students' writing proficiency*. Doctoral dissertation, Bilkent University, Turkey

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

Miller, P. H. (2002). *Theories of developmental psychology*. Macmillan.

Moon, J. (2008). L2 children and writing: a neglected skill?. *ELT journal*, 62(4), 398-400.

Muslim, I. M. (2014). Helping EFL students improve their writing. *International Journal of Humanities and Social Science*, *4*(2), 105-112.

Nunan, D. (1999). *Second language teaching and learning*. Boston, MA: Heinle & Heinle.

Ostler, S. E. (1987). A Study of the Contrastive Rhetoric of Arabic, English, Japanese, and Spanish: By Shirley Elaine Ostler. Doctoral dissertation, University of Southern California, The USA

Ozsevik, Z. (2010). *The use of communicative language teaching (CLT): Turkish EFL teachers' perceived difficulties in implementing CLT in Turkey*. Doctoral Dissertation, University of Illinois, The USA

Öztürk, H., & Çeçen, S. (2007). The effects of portfolio keeping on writing anxiety of EFL students. *Journal of Language and Linguistic Studies*, *3*(2).

Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading &Writing Quarterly*, 19(2), 139-158.

Perl, S. (1980). Understanding composing. *College composition and communication*, *31*(4), 363-369.

Piaget, J. (1964). Part I: Cognitive development in children: Piaget development and learning. *Journal of research in science teaching*, *2*(3), 176-186.

Phuket, P. R. N., & Othman, N. B. (2015). Understanding EFL Students' Errors in Writing. *Journal of Education and Practice*, 6(32), 99-106.

Polio, C. (2017). Second language writing development: A research agenda. *Language Teaching*, 50(2), 261-275.

Ramanathan, V. and Kaplan, R.B. (1996). Audience and voice in current composition texts: Some implications for ESL student writers. *Journal of Second Language Writing*, 5(1), 21-34.

Richards, J. C., & Renandya, W. A. (Eds.). (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge university press.

Rinehart, S. D., & Thomas, K. F. (1993). Summarization ability and text recall by novice studiers. *Literacy Research and Instruction*, *32*(4), 24-32.

Santrock, J. W. (2006). *Life-span development (10th ed.).* McGraw-Hill.

Schenck, A. D., & Choi, W. (2015). Improving EFL Writing Through Study of Semantic Concepts in Formulaic Language. *English Language Teaching*, *8*(1), 142.

Sivell, J. (2013). Factors underlying students' appropriate or inappropriate use of scholarly sources in academic writing, and instructors' responses. *Iranian Journal of Language Teaching Research*, 1(2), 65-84.

Solak, E., & Bayar, A. (2015). Current Challenges in English Language Learning in Turkish EFL Context. *Online Submission*, 2(1), 106-115.

Steffensen, M. S., Joag-Dev, C., & Anderson, R. C. (1979). A cross-cultural perspective on reading comprehension. *Reading research quarterly*, 10-29.

Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends in cognitive sciences*, *9*(2), 69-74.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research*. 1998. Thousand Oaks.

Super, D. E. (1990). A life-span, life-space approach to career development. In D. Brown, L. Brooks, & Associates (Eds.), *Career choice and development: Applying contemporary theories to practice* (2nd ed., pp. 197-261). San Fran- cisco: Jossey-Bass.

Susser, B. (1994). Process approaches in ESL/EFL writing instruction. *Journal of Second language writing*, 3(1), 31-47.

Tarnopolsky, O. (2000). Writing English as a foreign language: A report from Ukraine. *Journal of Second Language Writing*, *9*(3), 209-226.

Trotman, W. (2010). *Teacher Oral Feedback on Student Writing: An action research approach towards teacher-student conferences on EFL academic essay writing in a higher education context in Turkey*. Doctoral dissertation, University of Warwick, The UK

Uysal, H. H. (2008). Tracing the culture behind writing: Rhetorical patterns and bidirectional transfer in L1 and L2 essays of Turkish writers in relation to educational context. *Journal of Second Language Writing*, *17*(3), 183-207.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard university press.

Yang, X. (2019). A review of negative language transfer regarding the errors in English writing in Chinese colleges. *Journal of Language Teaching and Research*, 10(3), 603-609.

Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of second language writing*, *15*(3), 179-200.

Yeşilyaprak, B. (2007). *Gelişimsel Rehberlik[Developmental Guidance]*. İstanbul: Morpa Kültür Yayınları Ltd. Şti.

Yeşilyaprak, B. (2012). Mesleki rehberlik ve kariyer danışmanlığında paradigma değişimi ve Türkiye açısından sonuçlar: Geçmişten geleceğe yönelik bir değerlendirme [Paradigm shift in the vocational guidance and career counseling and its results for Turkey: A review from the past to the future]. *Educational Sciences: Theory and Practice*, 12(1), 97-118.

Younes, Z.M., & Albalawi, F.S. (2016). Investigating the Factors Leading to Speaking Difficulties: Both Perspectives of EFL Saudi Learners and Their Teachers. *Arab World English Journal*, 7, 268-287.

Yu, S., & Lee, I. (2016). Peer feedback in second language writing (2005–2014). *Language Teaching*, 49(4), 461-493.

Zamel, V. (1983). The composing processes of advanced ESL students: Six case studies. *TESOL Quarterly*, 17,165-187.

Zenouzagh, Z. M. (2018). Multidimensional analysis of efficacy of multimedia learning in development and sustained development of textuality in EFL writing performances. *Education and Information Technologies*, *23*(6), 2969-2989.

Zhao, C., & Huang, J. (2020). The impact of the scoring system of a large-scale standardized EFL writing assessment on its score variability and reliability: Implications for assessment policy makers. *Studies in Educational Evaluation*, 67, 100911.

Zoghi, M., & Fakhimie Shokri, L. (2019). Effects of Pre-Task Strategic Planning on Iranian Intermediate EFL Learners' Writing Motivation and Accuracy. *Journal of English Language Pedagogy and Practice*, *11*(23), 207-225.

APPENDIX: Examples from the Data Obtained from Teachers / Instructors
    A- Examples from primary school data
Already provided in the main article.
    B- Examples from secondary school data
They improve it (their writing) at the end of 5th grade. The spelling mistakes reduce. We want them to make sentences with the words they have learned so that the vocabulary will be permanent. They learn more about writing when they move towards 7th and 8th grades. (T_7)

Example expressions on causal conditions:
Negative transfer:
The biggest problem of course is that words are pronounced and written differently in these two languages. Maybe, it becomes difficult for the kids to learn both how to write and how to pronounce. The kids get satisfied with the pronunciation only, they can write through that way. (T_10)
Lack of transfer of writing skills:
I can say the number of students who use their mother tongue as it should be in real sense considering their ages is about 30 percent. Mostly there are sentence errors. They do not make proper sentences, they do not make long sentences. (T_11)
Lack of writing experience:
If they do it again at home, that is, if they practice more, the result will be better in writing in English (T_10)
The main problem is; the children are not studying regularly, that's most obvious. They are not studying to improve their vocabulary. (T13)
Lack of phonological awareness:
We are facing problems, of course, because the pronunciation is different from the spelling. We spend most of the lesson to check them out. (T_11)
Lack of imitative writing:
We have students who cannot even write to the board. (T_11)
Lack of grammatical knowledge:
They do not know how to organize words because their grammar knowledge is insufficient (T_17)

Example expressions on strategies:
Positive transfer:
They can write composition at home, as they write (composition) in Turkish. (T_12)
Writing Experience:
In order that the students improve their writing skills, they have to engage in a variety of writing exercises. For instance, they can write what they do during a day in their diaries. (T_13)
Phonological awareness:
I, personally, have them listen to simple English songs so that they get accustomed to English sounds. I also get them exercise by pronouncing and by getting my students pronounce short words on the board. (T_14)
Imitative writing:
We have problems with fifth and sixth graders. Our students at seventh and eighth grades are able to write well what's on the board. (T_17)

Example expressions on intervening conditions:
Curriculum:
This is due to the curriculum. I think the curriculum is very dense. If the kids encounter fewer topics at school, if they encounter fewer words, they may learn more robustly (T_13)
Teacher:
The attitude of the teacher and the willingness of him to have students write are important. I am, for example, having them write quite a bit, I do it especially to get them accustomed to writing. I both have them write and give them homework so that they get used to writing a little faster. (T_11)
Insufficient Training Hours:
That is to say, the number of courses is low. We had 24 hours of English lessons per week while we were preparing for university. The number of lessons is low, so we cannot allot time for writing, speaking etc. with children. (T_11)
Material:
We do not have any listening CD with our textbooks. If it was so, and if it was possible to practice listening with native speakers' voices, it would be more efficient. Thus, children's phonological awareness would increase. (T12)
Method:

They already read and understand the book and translate it into Turkish. I like GTM. We learned with GTM. I am also trying to support the good ones (students) in this way. (T19)

Motivation:
Elementary school students are already eager to learn something. If the teacher teaches them in an amusing way, they already progress in 2, 3, 4 (grades). But now when they move on to 6th grade, the child is starting to prepare for the TEOG exam. They are preparing for the test rather than learning English. Their worry is to get high marks in the test. This is a nuisance. They think that they will learn English in college in the future. They do not think it's easier to learn it now. (T_13)

Technology:
They like playing with computer or phone. They are not struggling with writing. There is too much distraction. (T_18)

Information processing capacity:
We have students who can write complete sentences. They usually come out of their shell at 7th and 8th grades. It's not the in the 5th or 6th grade, in later times. (T_11)

Family/Culture:
It might be effective if the family, the teacher, and the student come together and talk. It is necessary that the student give a promise about it. But the parents are often uninterested. (T_18)

C- Examples from high school data

They are beginning to write simpler. But later on, in conjunction with the subject, the writings are further developed when the examples are given in front of them. They move their writings further ahead the level they achieved before high school. (T_23)

Example expressions on causal conditions:
Negative transfer:
They want to carry the Turkish expression in their head to English literally. For example, they use a word in a way that does not fit the context. For example, s/he says 'he has a big head' to mean that 's/he is hobnailed' as s/he translates koca kafalı[†] literally by separating the words of the idiom. S/he cannot use it properly. (T_27)

Lack of transfer of writing skills:
The absence of composition lesson in Turkish also affects their writing in English. They do not know what to write and how to write it. "We do not even do (write) it in Turkish, you want it from us in English," they say. (T_26)

Lack of writing experience:
They write easily when they know what to write about. But when it comes to some more academic subjects, it becomes harder to write because their level of knowledge is not enough. Reading is a solution for this, but "Is the existing one a reading generation?". It is a question mark. Students cannot write without knowledge. He turns and turns and says the same thing. (T_26)

Lack of Vocabulary:
Writing is really a skill. In other words, you have to have vocabulary knowledge, and grammatical knowledge to produce something. They get in trouble at this point. The vocabulary of the students is not sufficiently developed. (T_30)

Lack of Grammatical Knowledge:
They do many grammatical errors. Although we repeat teaching them, and give them feedback when they do wrong, they still lack some grammar points. (T_21)

Lack of Productive writing:
Speaking and writing are productive skills. You have to produce something, and those who are not so imaginative cannot achieve it. That's a complication. (T_24)

Example expressions on strategies:
Positive Transfer:
First of all, grammar should be given to a certain extent. After that you should move on to the use of English; speaking, reading... How does this happen? Reading is important here. For example, I say to my students: While we study in the class, we study passages. When you get home, take this piece and read it aloud to yourself. So you pronounce because you read aloud. Do you hear what you read? Then you are listening. Now take paper and pen, and summarize as well as you can. So you have done exercises on 3 skills and on how to pronounce. (T_25)

[†] Koca: big; kafa: head

Writing Experience:
Students need to read a lot. Reading a lot is crucial. In fact, it is necessary that we do not separate reading in English and reading in mother tongue. Both are necessary. If they read in both, then they can write well. They should be patient and they should allocate time for it. (T_28)
Grammatical Knowledge:
The stronger the grammar knowledge, the more accurate their writing is. (T_21)
Rich Vocabulary:
As students' vocabulary increases, their writing gets richer. (T_22)
Productive Writing:
By giving an example, you can say, "Here is how it is written, you can do it like this". You might want something like their rewriting something by changing a few words. It is first because of creating the feeling "I can write". (T_26)

Example expressions on intervening conditions:
Teacher:
It is not enough that you learn methodology in college. You start little by little in teaching, you see the reaction of students, and you develop something different. You reinforce what they admire, and you avoid what they do not like. (T_27)
Insufficient Training Hours:
It is required that you give an example of what they are going to write so that they can write something like that. They will write after that. You will evaluate what they write and give feedback. But how many hours of a 4-hour lesson in a week does it take? 4-hour English course in a week is insufficient. (T_22)
Material:
Of course, English is not taught like the native language. We are lack of many things; the books are not engaging. We have a problem with the material. Now we have a smart board, but we do not have the internet access. It's pretty good compared to the old one, but there are no funny activities for kids. There is no CD for the listening pieces. We have to reach them from EBA (Educational Information Network), but we cannot reach there at all. (T_25)
Method:
In general, the more stimulation you put in group work, the more the children become happy. Like in a contest, we need to activate the senses. Visuals, for example. If you want them to write on a white page without looking at and seeing anything, this is hard. You need something to activate their senses. (T_23)
Motivation:
When they move on to 11th-12th grades, the students are starting to prepare for the university entrance exam. They are preparing for they do not want to study English. This is a big problem (T_23)
Technology:
The students have become lazy now. Mobile phones, shortening everything, disappearance of letter… Life has become more practical now. They always want everything short. For example, I have just spoken to one of my friend. "Do not write to the board," the students said to him, that is, they do not want to write. There is a serious problem yes, because they do messaging in short cuts. In the past, we wrote letters, the letters vanished. Writing is disappearing slowly. This is a negative side of technology. Of course, teaching English is also influenced by this situation. The kids cannot write, they cannot interpret, and so they cannot express themselves. (T_25)
Family/Culture:
All families and the authorities think that way: Our children should always proceed in the academia; they all should be university graduates. They should study in science colleges, Anatolian teacher training high schools etc. Students have different abilities, all different. The things they can do is clear. Of course we will open the way, but we will not keep them in the direction we want. We will suggest the students to move towards the most appropriate direction. To occupational high school, to craft high school... Those who have such a talent can at least learn an art instead of being dragged to other places. Not everyone will be a doctor or a lawyer. (T_24)
They do not think it's easier now. It's a bit related to the family of course. Parents' view is the child's view. It's all nested. If the parents support and consolidate, our job gets easier. (T_29)
Individual differences:
Some people like to work individually. Some of them enjoy group work. There are individual level differences too. Some students may be able to write better than others. (T_22)
Change in Developmental Task:
Another factor is they are in adolescence. They cannot express them in a well-coordinated manner. They cannot express them in a planned way. When they write something, they do it very complex and complicated. However, they get happier if we help them plan and guide them. But they also complain about it: "You are limiting us." First, you will not block their autonomy. You will leave them free. It (the thing they write) should be more flexible. You cannot expect students to be autonomous in teacher-dominated class. (T_26)

D- Examples from university data

A little more emphasis should be put on writing in pre-university education. It gets a little difficult (late) to acquire writing skill at college. They should acquire it in the previous stages. (T_32)

Examples expressions on causal conditions
Negative transfer:
They want to translate what they think in Turkish into English directly. As can be seen in the example; 'there are green eyes of my sister' (trying to say 'my sister has green eyes'. Secondly, they want to write very complex sentences in English. However, they do not think that 'my English is different from Turkish'. One is my mother tongue and the other is a foreign language. They need to write simpler things in a foreign language first. They cannot make this simplification. (T_34)
Lack of Transfer:
There is a transfer event from mother tongue to the foreign language. When there are major differences, this causes big problems. Even though there is some similarity, it does not provide a complete positive transfer. Something is always missing. (T_32)
Lack of Experience:
Perhaps they have a lot of writing-related shortcomings in the training they had received until they started the university. (T_37)
Lack of vocabulary:
Having a rich vocabulary means being always one step ahead of others. (T_38)
They write translation of a Turkish word in English regardless of its context just by looking up in the dictionary. For example, when a student wanted to write 'this bridge connects two sides', he wrote 'this bridge connects two collars' since the Turkish word 'yaka' is a homonymous word meaning both 'side' and 'collar'. (T_34)
Lack of Grammatical Knowledge:
We get difficulty in structures; grammar, syntax, word order because the syntaxes of the two languages are very different. (T_39)
Lack of Rhetorical Organization:
Generally they cannot organize sentences, ideas, etc. in a proper and smooth fashion. We observe that they see writing as a task. They do not want to seek ways for improving their abilities on writing organization. (T_31)
Lack of Productive Writing:
But students do not produce. It's just grammar loading. When they arrive at university, we do not observe that they can produce something. They say "I know present simple" but they have difficulty in making sentences. (T_36)

Examples expressions on strategies
Writing Experience:
Students will listen/watch BBC radio/TV, even if they do not understand it. First, they will follow foreign radio-TV channels. The second, they will use an English-English dictionary. They will see through the samples and use them. They will also use the collocations dictionary. They will learn such basic collocations as "get a high mark-low mark; take an exam-have an exam .. " etc. However, they will use paper-based dictionaries not electronic ones. I, having more than 30 years of experience, believe that paper-based dictionaries are more helpful and efficient. (T_39)
Positive Transfer:
I'm trying to raise their metalinguistic awareness. 'See why we are learning it, because of it'. I raise their awareness. The structure of each language is different. You cannot translate a sentence that you set in your mind in mother tongue to foreign language. I say that you have to express it in its simplest form in the target language according to the rules. (T_34)
Improved Grammatical Knowledge:
It's my observation that when a student learns the syntax and other grammar-related issues well, they also improve in writing. They start making sentences that are well-ordered. (T_38)
Rich Vocabulary:
Suppose we started on sentence basis. When you explain what a sentence, a paragraph a topic sentence, a supporting sentence, a concluding sentence etc. they start writing something. We observe the gradual development. (T_33)
Writing Experience:
My students have been writing regularly since the term began. As their writing practices increase, their writing also improves. (T_34)
Productive Writing:

First, you need to write essays together with students. Secondly, you need to have them write in class. Otherwise, it is not all right when you want them to write at home. And they need to write on different types and issues. They have to experience it to produce something original on their own later. (T_39)

Examples expressions on intervening conditions
Teacher:
The facilitating factor can be the relationship between us. I keep saying them the mistakes not as mistakes, but as things that are fairly normal in learning English by way of relieving them. I am tolerant towards them. I give the feedback more positively and I tell the reasons for their mistakes. I tell them that their mistakes are normal, and they will overcome them by studying and writing more. (T_37)
Material:
Not having the appropriate material when you need is a nuisance! We have some materials of course, however, they are not those that we need. (T_32)
Method:
When we first established the prep school, METU had a writing book. And we started with guided writing there. For example, it is written there; I, 25, London, student etc. The student wrote by combining these prompts. Free writing followed this. First guided and instructed writing then free writing. Step by step because it is always recommended in seminars and conferences. Personalization is very important. For example, like and dislike. I am asking immediately; 'what do you like / dislike?'. They learn more easily. They will learn chunks first. I am opposed to starting with free writing. (T_39)
Motivation:
First of all, students are afraid of writing. There is a fear that they are not proficient. Most students think so. Even though these students took courses starting from primary school they feel like this at first. (T_37)
Technology:
There is a situation like this: Whatsapp correspondence and Messenger correspondence really kills language skills. For example, if you see my correspondence with an American friend- who has a master's degree- you think that she is probably a Turkish girl, and I am an American. My language is smoother. Abbreviations, acronyms etc. These things affect writing skill negatively. (T_39)
Family/Culture:
I think the background is a bit of education families give their children. The child is successful if he is educated in literacy, science, and so on in family. It starts from the family rather than the previous levels (until university). (T_39)
Individual differences:
Some people cannot express their feelings; they cannot put them on paper. Some of them cannot do it when they are asked to write their opinions on something specific. (T_35)
There are also differences within the same class. Even some cannot write paragraphs in Turkish. But on the other side, there are also ones who can express themselves well through writing in English. (T_40

Digital Learning Experience in Museums: Cultural Readings in a Virtual Environment

Nuray Mamur[1], Vedat Özsoy [2], İbrahim Karagöz[3]
[1]Pamukkale University
[2]TOBB University of Economics &Technology
[3]Ministry of National Education-Turkey

To cite this article:

Digital Learning Experience in Museums: Cultural Readings in a Virtual Environment

Nuray Mamur[1]*, Vedat Özsoy[2] İbrahim Karagöz[3]
[1]Pamukkale University
[2]TOBB University of Economics &Technology
[3]Ministry of National Education-Turkey

Abstract

This study revolves around primary and secondary school visual arts teachers' cultural and critical readings on their VR museum experience through object and representation. As part of the in-service training and research project, the study was carried out in the provinces of Çanakkale, Erzincan, Kayseri, Diyarbakır, Giresun, Mersin, Denizli, which were selected from seven regions of Turkey within a seven-month period. 508 primary and secondary school visual arts teachers participated in the study. The study is a case study and teachers' readings related to the object and representation in the virtual museum and their VR museum experiences formed the units of analysis in the study. The data of the study was collected through worksheets, pre-evaluation forms, participant diaries and focus group interviews with voluntary teachers and content analysis method was employed to analyse the data. In the study, teachers were motivated to incorporate social and personal contexts into their readings of the object-space-representation. The results denoted that the activity enhanced teachers' learning experiences in terms of personal and professional contexts. Although teachers highlighted that VR museum experiences could not offer much more rich experiences than the actual museum visits, they believed that it could be effective for visual arts classes and be blended with various learning activities given that the virtual museum experience offers the possibility of exploring a museum without a barrier of distance and incite enthusiasm and learning interest through digital technologies (e.g., sound, motion, immersive experience).

Key words: Teacher training, cultural heritage, virtual reality, virtual museum, visual culture in visual arts education

Introduction

With rapid technological advances in recent years, museums have sought new ways to document, conserve and exhibit cultural heritage. In museums, digital technologies, which initially manifested themselves in areas such as recording and tracking of works, began to be used in time to communicate with the society and to make exhibited objects, thematic expressions or digital art productions more immersive and impressive. Further, the fact that museums with educational activities have assumed the responsibility of being a major cultural force in society since the second half of the 20th century and consequently integrated with information technologies paved the way for digital environments that offer personalized experiences. For example, according to 24 Hour Museum's data in 2003 and museums' own annual reports, as early as 2002 the number of virtual visitors to many museums' websites had already overtaken the number of physical visitors on-site (Hawkey, 2004). In recent years, however, different types of mobile apps have provided artists, museum educators, academicians, and visitors the opportunity to come together on a common platform with real-time interaction. To illustrate, British Museum provided a virtual tour of the exhibition titled "*Defining beauty: The Body in Ancient Greek Art*", which presented by British historian and broadcaster Dan Snow through Periscope broadcast, a live-streaming app. Lots of people thus interacted within a virtual exhibition environment. In other words, museums have become places where visitors share their ideas and socialize through social platforms. Recent research on the context of museums and technology (Karatay & Karatay, 2015; Özer, 2016; Karadeniz, 2020) shows that digital technologies encourage different age groups to visit the museum and facilitate their access to the museum with presentations such as exhibitions and games. It is seen that this process has accelerated due to the COVID-19 epidemic that emerged in December 2019 and affected the whole world. The physical closure of museums

---

due to the epidemic has caused museums to shift social communication opportunities to digital platforms. Museums have opened their collections, past and present exhibitions to the online experience, designed online educational activities, games or competitions via social media (Karadeniz, 2020; Akça, 2020; Ertürk, 2020).

As can be seen from the examples, the biggest impact of digital media and developing technologies on museums is undisputedly digital access to museum and art experiences anytime from anywhere where the Internet is available. Yücel (2012) asserts that "Digital technologies have been very effective in the development of "accessibility" and "democratization" discourses, which are crucial in contemporary museology in the context of mass communication at least" (p. 28). This is of vital importance in terms of creating a large-scale change. Because museums potentially became a part of everyday life, abandoning the image of corporate and even authoritarian whereby they offered an opportunity for more dynamic and personalized learning.

The term that is often used when the digitized exhibits of a museum are publicly available, is "virtual museum". Lewis (1996) describes the virtual museum as a collection of digitally recorded images, sound files, text documents, and other data of historical, scientific, or cultural interest that are accessed through electronic media. Amongst the opportunities provided by digital process, one can mention that publishing digital museum collections on the web, 360-degree panoramic virtual tour of collections in more advanced applications, various interfaces that allows the users to easily navigate through the collections, enrichment of high-resolution images with sound, music, motion and multiple narrative. In addition to that, the way we perceive artworks in virtual reality and actual physical environment differ. In this context, there is an ongoing theoretical debate in the literature (Barlas Bozkuş, 2014; Yenisehirlioglu, 2007). The focus of the debate addresses to exhibition of artworks in virtual museums without time and space limitations and the blurry effect which this process creates in the identity of artworks. On the other hand, Utkan Özden (2019) contends that "Whereas it will not replace actual museum, virtual museum experience allows us to read and explore any piece just by sitting in front of a computer" (p. 60).

In recent years, the digitization process of museums has evolved into virtual reality (VR) and augmented reality (AR) applications that attract considerable attention in many fields. Historians, curators, and archaeologists have started using virtual reality technologies to promote a deeper level of perception for historical reconstructions and artworks (Zouboula, Fakides & Tsolakidis, 2008; Hürst, Coninck & Tan, 2016; Bülbül, 2017; Schofield, Beale, Beale, Fell, Hadley, Hook, Murphy, Richards & Thresh, 2018; Zhao, Forte & Kopper, 2018). Virtual reality can offer a number of advantages to museums, offering a way to overcome some common problems like the lack of space or the need of visitors to interact with the exhibits (Lepouras, Charitos, Vassilakis, Charisi & Halatsi, 2001). Virtual reality facilitates exhibition of objects that cannot be exhibited especially because of the limited space or that may be too fragile or valuable to be exhibited. Exhibits can be interactively observed from different viewpoints or even manipulated. It also allows the visitors to explore environments that may be no longer exist today be somehow damaged and in need of reconstruction or cannot be easily experienced, either because they exist at a remote site or because their condition does not allow for their interior to be navigated (Lepouras and et al., 2001).

From the perspective of museum education, it is clear that the Internet provides opportunities for art teachers to work on museum collections. "As technology becomes increasingly integral to our lives, it is imperative for students (and teachers) to critically evaluate electronic data and become savvy consumers of art information on the Internet" (Stone, 2001, p. 83). This quotation is taken from the book dated 2001 and titled "*Using the art museum*". In the last 20 years, technological developments have contributed immensely to the progress of museums in virtual environments, and it urged teachers to implement these developments in their classes. The effect of virtual museum environments on the student knowledge and understanding on a certain work of art is no longer questionable.

This study serves to in-service teacher training program supported by the Scientific and Technical Research Council of Turkey (TUBITAK), which was designed to incorporate visual culture inquiries into visual arts courses. The study attempted to integrate virtual museum environments with critical approaches to visual culture whereby it was intended to contribute to the field of learning "*cultural heritage*" in the visual arts teaching program. Visual culture studies in art education is an approach that requires more critical consciousness, in other words, critical-visual literacy, due to the rapidly increasing of visual information as a result of advanced reimaging technologies and the effect of each visible image on human perception and attitudes.

Visual culture theory, which takes a wider perspective on visual arts education, is not only concerned with the appropriate perception of artworks, but also invites us to question our relationship with all kinds of visual

images (advertisements, YouTube videos, social media posts, shopping malls and etc.), which form much of our everyday interaction or appeal to our senses via eye-catching techniques, our viewpoints towards these visual images as well as questioning the underlying factors of our viewpoints. The visual culture theory is closely related to "critical pedagogy, social reconstruction, critical theory, feminist theory/criticism, multiculturalism, social justice and semiotics" (Mamur, 2019).

Tavin (2003) posits that images trigger individuals' emotions in terms of ethnicity, race, nationality, gender, family life, independence, and citizenship. The visual culture theory in visual art education concentrates on every type of cultural and emotional factors that influence individuals' cultural identity. In this respect, aesthetics experiences emanated from everyday visual images have been regarded more meaningful than human's interaction with museum works. This idea is in direct proportion to the frequency of visiting a museum (once a month, maybe once a year) and the lack of connection with artworks. Yet, the recent rise of virtual reality in museums has minimized accessibility issues, and museums have become a part of our ordinary everyday world. When this situation assessed in the context of Turkey, it can be said that the studies on the visibility of museum collections in the digital environment, is not reach sufficient numbers. According to Akça (2020), developing technology and new media opportunities are not used adequately, especially in exhibiting the collections in museums affiliated to the Ministry of Culture and Tourism. Moreover, it is a fact that digitization in these institutions of cultural memory provides added value to the culture sector (pp. 271-272). Given the accelerating digitalization of Covid19 epidemic period, the museums in Turkey are expected to enter this expedited process. The initiatives of museums towards digital platforms create the opportunity to diversify classroom practices in the context of museum and cultural education through online access. In this vein, this study was intended to integrate museum education with visual culture inquiries. This is yet not the only reason. Another reason is the effects of museum education and visual culture education on creating personal meanings through cultural and internal inquiries and the relationship between museum education and visual culture education. We can highlight three factors for this relationship.

One of them is the significance of intercultural interaction built through images in visual culture education. Many museums in the world have now moved their collections to online environments, and these environments "create a cultural heritage and art platform, thereby promoting intercultural codes on a global scale." (Barlas Bozkuş, 2014, p. 330). Exhibiting cultural heritage and works of art through virtual museums increases intercultural recognition and accelerates the process of commodification of culture. Given especially the online platforms of the world's leading museums in the USA and the UK, it is seen that the promotion of the dominant cultural elements that serve to the tourism industry is prioritized rather than the promotion of local cultural identities (Barlas Bozkuş, 2014, p. 332). The fact that museums become a part of the culture industry makes it possible to associate with the critical questioning dimension of visual culture teaching.

The second factor is that museums offer infinite viewing opportunities that stimulate its visitors' perceptions. As for the design of the space, attractive elements are used. "Several visual culture elements are used to determine what the audience will see, how they will determine the direction, what they should focus on, and why being attractive is important" (Karadeniz, Okvuran, Artar & Çakır İlhan, 2015, p. 210). This is valid not only for physical museums but also for virtual museum environments. In fact, it is seen that sound, light, music, motion, and impressive narrative are employed to create more attractive virtual museum environments.

The third factor is that the social ground in the context of makers and viewers is as important as works of art in visual culture education. In other words, the interpreter's role and socio-cultural context are relevant in the intertextual inquiry methods of visual culture (Anderson, 2003; Duncum, 2002; Tavin, 2000; Wilson, 2002). As Hagberg (2016) points out, "Museums are not full of physical objects. They are full of intentional objects made manifest. This distinction is embedded at a foundational level within our highly-evolved language about these intention-enmeshed artefacts: once one knows the identity of an artefact, one can rarely immediately subsequently inquire into its meaning" (p. 261-293). We can thus imply that both museum education and visual culture studies entail a visual learning environment that stimulates audiences' interpretation strategies.

In that context therefore, it can be said that virtual museums offer new presentation and narrative forms and create new representations with the narrative techniques they present to visitors. According to Leppert (2009), vision and representation mutually work to produce the knowledge. From the perspective of visual culture, "the connection between reality and its representation should be noticed and analyzed" (Türkkan, 2008; Dilli, 2020). What kind of questions can be posed regarding representation in these virtual environments? Özsoy (2019) stresses the priority of "*looking and seeing*" practices in museums and/or exhibition spaces which will ensure that the perception is meaningful and permanent. In other words, the interaction of looking and seeing with the object should be achieved. In museum education, objects play a crucial role in the learning process. "The fact

that the meaning of the object varies, and it can be interpreted differently each time is an important feature of the learning process with the object" (Onur, 2012, p. 230). How this learning process can be realized in virtual environments? This study attempted to read object and representation in the context of virtual museum and to focus the investigation of teachers' virtual experiences. To this end, the study sought to answer the following research questions: *What personal experiences and professional qualifications did visual arts teachers gain regarding the activity of 'Representation: Virtual Museum Experience'*?

Method

The study is a case study which delves into the activity, namely, *"Representation: Virtual Museum Experience"* as part of an in-service teacher training program. This activity focuses on learning cultural heritage in visual art education along with a critical approach to visual culture. As for the implementation of the activity, it involves two VR museum experiences and analysis of this experience through the object and representation. In case studies, the case may involve an individual, a small group, an organization or association" (Creswell, 2013, p. 98). In the study, teachers who attended the in-service training program were regarded as a case. Teachers' evaluations regarding the object and representation in the virtual museum and their VR museum experiences formed the units of analysis in the study. Since the study is concerned with a single analysis unit, "single case-holistic design" was used. "A holistic single case pattern is a study conducted with a single unit of analysis (such as an individual, a program, a school)" (Yıldırım & Şimşek, 2013, p.290).

As part of the in-service training and research Project, the study was carried out in the provinces of Çanakkale, Erzincan, Kayseri, Diyarbakır, Giresun, Mersin, Denizli which were selected from seven regions of Turkey within a seven-month period, between the months of September 2018 and April 2019. 508 primary and secondary visual arts teachers participated in the in-service training program. 284 of the teachers are visual arts teachers, 224 are classroom teachers. Necessary explanations were made to these teachers about the research dimension of in-service training, and focus groups of 5-6 people were formed in each province (except Giresun[†]) who were willing to participate in the study. In this context, focus group interviews were held with a total of 66 teachers, including 31 classroom teachers and 36 visual arts teachers.

Data Collection

This study was realized in the three phases. In the first phase, **a pre-interview form** was submitted to gather preliminary information about teachers. This form was employed to find out whether teachers used virtual museum and visual reality applications in their classes and, if so, how they used. The pre-interview form comprises of three questions.

1. *What do you think virtual reality is about?*
2. *How do you benefit from mobile or digital applications in visual art education?*
3. *In which aspects do you think digital technologies contribute to primary and secondary school students in visual art education?*

Afterwards, the next step included the phases of briefing and implementation. Teachers were informed about digital applications and virtual reality applications in museums thorough virtual museum examples in the world and Turkey. The second phase is the VR museum experience which involves ***Art Plunge***[‡] and ***The Museum of Innocence***[§]. These museums were selected because they have collections that urge teachers to question object, representation, and space in the context of visual culture education.

[†] Focus group interviews were not conducted in Giresun province as the number of participating teachers was limited to two.

[‡] **Art Plunge** is a virtual reality gallery that exhibits famous works of art. The VR headset allows the users to get the feeling of being inside famous paintings and to perceive objects and figures in an immersive virtual environment. By the use of sound and motion, each painting offer an enhanced experience. (https://store.steampowered.com/app/570900/Art_Plunge/).

[§] **The Museum of Innocence** is a museum inspired by Nobel Prize-winning author Orhan Pamuk's novel of the same name. The Museum of Innocence tells the story of Istanbul life between 1950 and 2000 through memories and flashbacks centered around two families – one wealthy, the other lower middle class. The author presents what the novel's characters used, wore, heard, saw, collected and dreamed of. (https://tr.masumiyetmuzesi.org/page/muze)
https://artsandculture.google.com/exhibit/masumi%CC%87yet-m%C3%BCzesi%CC%87/XgJyIqBekvaEKw?hl=tr

Image 1. Art Plunge, A girl reading a letter in front of the open window A, J. Vermeer, 1657
Image 2. The Museum of Innocence, The Pain of Waiting

In the literature, learning in museum is defined as meaning-making from objects and experiences (Falk, 2009; Dierking, 2002; Burnham and Kai-Kee, 2015). Because museums are expected to give "visitors a creative and emotional experiences" (Clutterbuck, 2007, p. 75). The aim here is to read and make sense. To this end, teachers were invited to use **worksheet** whereby they were motivated to think about their VR experience. The content of worksheet is based on fostering the practice of looking and seeing in the context of visual culture through the concept of representation along with Dierking's contextual model of learning in the context of museum education. The notion of *"representation"* is used to demonstrate the creation of meaning in visual culture. The use of the contextual model of learning is recommended for visual art education in museum and gallery and this model includes overlapping contexts. These contexts are as follows: *personal context* which refers to the internal and/or external motivation degree demonstrated by learners, *physical context* which defines clues that help the student understand the work of art and phenomena and lastly, *socio-cultural context* which is related to museum audiences who interact with each other, museum and museum staff (Cited in. Lepouras and et al., 2001). This model is mostly offered for learning from physical museums. However, Özer (2016) researched the effect of this model on learning through the virtual aviation museum and stated that it could be a good model for virtual museums, except for some limitations in the physical context. The content of the worksheet is shown in Figure 1.
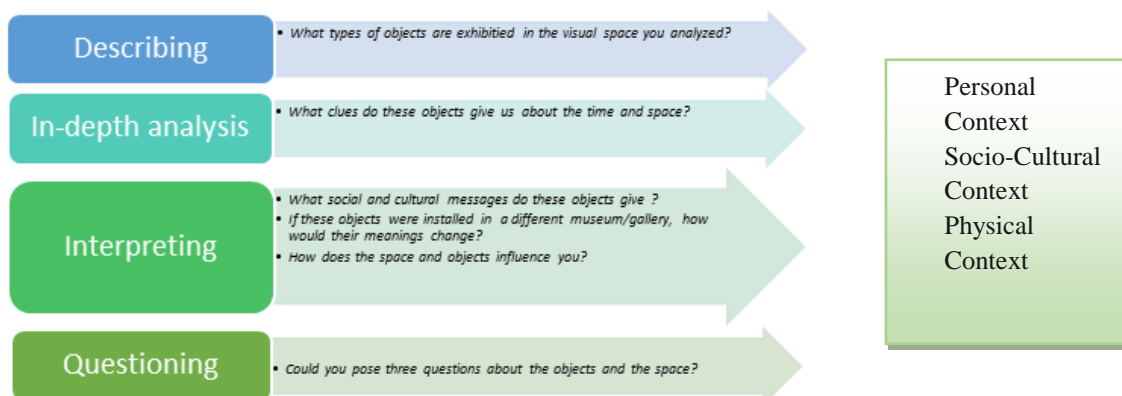
Figure 1. The format of worksheet

After the museum experience, teachers were asked to share their personal experiences in a way that serves to socio-cultural context. Further, **participant diaries** were distributed to each teacher to write their personal experiences. The participant diary consisted of questions that help teachers describe their thoughts, emotions and experiences and included questions that investigated the contribution of the application to the teaching profession and art education. The questions posed are:

1. *What are the contributions of VR learning activity to you?*
2. *What are the contributions of VR learning activity in terms of your teaching profession?*
3. *Which aspects of VR museum experience did make you satisfied and concerned?*
4. *In what ways do you think VR museum applications contribute to primary and secondary school students in the context of visual culture?*

In addition to that, focus group interview with volunteer teachers were conducted in each province as part of the in-service training program. Thus, teachers were asked to share their views on whether the theory of visual culture can be incorporated into art course. In the research, the focus group interview was used to analyse the new ideas of teachers about their classroom practices in depth, based on the education received and experiences. Focus group interviews were preferred because they enabled an intense brainstorming thanks to group interaction compared to individual interviews. According to Yıldırım & Şimşek (2013), "some issues that will not come to mind in individual interviews may come to mind within the framework of the statements of other individuals in group meetings and it may be possible to make additional comments" (p.179). The participants of the focus group interview comprised of 5-6 classroom and visual arts teachers individually from each province. Two interview questions are related to museum experience. Following the interview, these questions were transcribed and regarded as data source.

1. *How do you use the VR museum experience in visual art courses to promote the importance of preservation of cultural heritage?*
2. *Given the VR experience you gained during the learning activity, what do you think you can do for a more effective course?*

All these data collection tools were presented to the project coordinator for checking, and then presented to an expert who has research in the field of visual culture pedagogy and museum education for validity. Following the pilot study in the first city (Çanakkale) the Project was realized; the content of the implementation and data collection tools were revised. To this end, the structure of questions was revised whereby it was intended to use the time more effectively.

Data Analysis

In the study, 1564 pages of data (pre-evaluation form, worksheet and participant diaries) were obtained from participating teachers. These data were classified on a provincial basis and were read by two authors of this study during the research process. Since similar data contents are encountered in each province, 165 pages of data belonging to 55 teachers in total were allocated for in-depth analysis, based on the number of participants in the province in a random manner. Considering that sufficient data saturation was reached after the analysis, the data set in which the analysis was made was considered sufficient by the authors of the study. Table 1 shows the number of participants by province.

Table 1. Distribution of analysis data by provinces

	Çanakkale	Kayseri	Erzincan	Diyarbakır	Giresun	Mersin	Denizli	Total
Class Teacher	3	5	4	5	-	4	4	25
Visual Arts Teacher	5	5	5	5	1	5	4	30
Total	**8**	**10**	**9**	**10**	**1**	**9**	**8**	**55**

In addition to that, a total of 8 pages of data obtained from focus group interviews, which were conducted with voluntary teachers upon the completion of the project, were analyzed. The data obtained from the study was analyzed through content analysis. During data analysis, the first and third writers of the research independently prepared the drafts of similar vocabularies and thematics. Then, draft codes and themes were compared and following the mutual agreement, the second writer of the study submitted the final version to the expert. Following the data analysis, two themes were developed as follows: 1) *learning experiences in virtual museums,* 2) *the evaluation of museum experience in the context of teaching profession.* Table 2 provides information about which data collection tool is employed regarding identification of the codes.

Table 2. The relationship between theme and data collection tools

Theme	Data Collection Tool	Abbreviation Code
1- Learning experiences in virtual museums	Worksheet	WS
	Participant diary	PD
2- The evaluation of virtual museum experience in the context of teaching profession	Pre-evaluation form	PEF
	Participant diary	PD
	Focus group interview	FGI

Findings

Under the theme of **learning in virtual museums,** the following two sub-themes were identified: 1) making sense of space and object and 2) questioning approaches through space and object. Four codes, namely, "making sense of socio-cultural codes", "questioning the relationship between change and continuity", "making sense of the relationship between object and space" and "questioning the perception of space, time and reality" were created in the sub-theme of the making sense of space and object. The relationship between the codes created were given in Figure 2.

Figure 2. The relationship between the codes created in the sub-theme of making sense of space and object.

Regarding the readings on making sense of socio-cultural codes, it is seen that the participants highlighted social status and gender representation in the context of interior space, objects in the space and figure's stance. To illustrate, while the visual arts teacher coded 49 stated, "*In the work of art (Mona Lisa) I examined, the figure's stance, clothing, the objects around give me the impression of social status. It evokes a sense of perfection*" (WS-49, Visual Arts Teacher [VAT]), the classroom teacher coded 1 commented, "*Given that the woman's clothes and the room decoration are quite gorgeous, I get the impression of upper-middle- income and assuming that reading letter I mean she is literate, I get the impression that she has a high level of status*" (WS-1, Classroom Teacher [CT]). On the other hand, the visual arts teacher coded 4 questioned the place of woman in society considering the interior space and the image of woman and remarked: "*Women is at home, closed and passive. These images show that object can change in time, but perceptions don't change*" (WS-4, VAT). It is worth noting that she emphasized the perception of change and continuity. More specifically, by underlining perceptions about women, she noted that although the object changes, perceptions about woman hardly change. A great number of teachers questioned the relationship between change and continuity in the worksheets. For instance, while the visual arts teacher coded 30 stressed object permanence, saying that "*By offering different socio-cultural experiences, objects show us that time is passing but objects are permanent and do not disappear*"(WS-30, VAT), the classroom teacher coded 31 considered object as a reflection of human in the pursuit of comfort , stating that "*I see although centuries pass, the needs of people remain the same, objects only become more contemporary although time goes by and object is only a reflection of human in the pursuit of comfort*"(WS-31, CT). The teacher coded 43 who experienced the Museum of Innocence considered the objects as the witnesses of life stating that *"When I look objects, I see the experience. I see that life is defeated by time, but the object is permanent, these objects are the witnesses of passionate love…"* (WS-43, VAT). Similar expressions regarding the question of change and continuity were recorded in almost every worksheet of the teachers. On the other hand, the participants' perceptions of the change of object-space are different from each other. While some teachers claimed that the exhibition of the same object in a different space would not cause a change in mood, some teachers argued that it would not evoke the same feelings. To illustrate,

> *Its impact power on us can change, but I do not think that its meaning will change (WS-12, VAT)*
> *It does not give the same feeling. The fact that objects are in their own environment increases its effect. It emphasizes the meaning better. (WS-22 CT).*
> *Pulling the object from its environment changes the emotion (WS-11, CT).*

Given the teachers' expressions, they have different opinions about the object and its spatial context. It is seen that some comments or sense-making approaches of the participants rely on experiencing virtual environment. Regarding the code of "questioning the perception of space, time and reality", one teacher stated:

> *I had the feeling that I lived in the space and time that the painting was drawn. It was like everything created an emotional depth like watching the girl reading the letter in that room, feeling the warm wind coming from the open window, the girl's reflection on the mirror (WS-27, VAT).*

As can be seen from the teacher's statement, virtual reality environment influenced teachers' sense-making the object, space, and time. Given teachers' statements, it is evident that virtual reality activated the senses of hearing, touching, and smelling rather than seeing. It is thus can be argued that the practice of reading through the objects in the work of art or space urged teachers to create various meanings. "Creating meaning, making the world meaningful and meaningful interaction with others are important for visual culture practices" (Sarıbaş, 2020, p. 245). According to Hall (2017), "Meaning is also produced whenever we express ourselves in, make use of, consume or appropriate cultural 'things'; that is to say, when we incorporate them in different ways into everyday rituals and practices of daily life and in this way give them value or significance" (p.10).

The sub-theme of questioning approaches through space and object consisted of teachers' questions and approaches based on their experience regarding the work of art or the object in the museum. In other words, teachers' mind was occupied with these questions during their experiences. Consequently, it was observed that teachers developed questions pertaining to the space, objects, virtual reality, personal narrative, and formal aesthetics. Some examples of teachers' questions are categorized and presented in Table 3 below:

Table 3. The Questions Teachers Posed

Code	Some Questions Teachers Posed
Questions regarding the space	*If we change the space in the painting, what changes do you think happen? (WS-15, VAT)* *What kind of taste and smell do you get from this space? (WS-24, VAT)* *What does this space inspire you?? (WS-30, VAT)* *What emotions does this space evoke you? (WS-32, CT)* *How would you call this space?? (WS -2, VAT)*
Questions regarding the object	*How did the letter change the woman's mood? (WS-26, VAT)* *What is the story of the carpet on the table?? (WS-27, VAT)* *What is the relationship between objects? (WS -3, VAT)* *How does objects function in painting? (WS -7, VAT)* *If we changed the objects in the work of art, what changes in the meaning of artwork would occur? (WS-15, VAT)* *What period do objects show us? (WS-8, VAT)* *If you were the artist, how would you design objects and space? (WS-21, VAT)* *What would be the differences in the narration, if more visual and sound effects were added to the objects? (WS-25, VAT)* *If you added new objects to the work of art, what would you add, Why? (WS--41, VAT)* *Which objects in the space do have a different intended purpose today (WS--51, CT)*
Questions regarding VR experience	*If you were really there, what would you do? (WS-17, VAT)* *Could you describe the difference between watching the work of art and being inside? (WS-19, VAT)* *What music and motion effect would you add to this space? In which aspects motion and music do influence the meaning of the work of art? (WS-39, CT)*
Questions Regarding Personal Context	*Is there any section or object that you face the reality in this space? (WS-48, VAT).* *Which object in the museum does say something about you? (WS-41, CT)* *If any museum object you used was exhibited in the future, what thing would you want to exhibit about you? (WS-55, CT)*
Questions Regarding Formal Aesthetics	*If we change the colors, does it change the meaning? (WS-25, VAT)* *How does lighting affect the artwork? (WS-4, VAT)* *How were the integrity created in the artwork? (WS-9, VAT)*

These findings revealed that teachers were mostly inclined to question the object. This might be because of the museums chosen for the activity. Since historical, natural, or artistic objects are generally exhibited in museums, object-based learning is mostly adopted in educational activities. It can be said that the worksheets used for investigating teachers' virtual museum experience led teachers to examine the objects critically and effectively rather than just monitoring it whereby teachers developed different perceptions towards the object. For example, the visual arts teacher coded 2 stated, "This *experience definitely evokes excitement and interest. As I examine the details of the work of art, it makes you feel admiration and excitement"* (WS-2, VAT). The expression of "excitement" here is emancipated from the immersive VR perception of the piece. Yet, the teacher emphasized her increased admiration as she examined the artwork in detail. This might be because virtual museum experience facilitates more careful viewing. Education via museum should be realized through a perspective that stimulates critical thinking skills. In particular, a perspective that focuses on the effect of details on creating meaning is required. Shari Tishman (2018) uses her concept of slow looking to explain it. Slow looking is a way of building knowledge between the work of art and visitor, that is to say, as a mode of learning through observation. There is a need to allocate time for observation. Some of the teachers' comments on observation and learning are as follows:

> *Realizing, investigating, inciting curiosity (PD. 12, CT).*
> *Creating opportunity for cultural encounters (PD. 55, VAT).*
> *Different perspectives, teaching diversity (PD. 49, VAT).*
> *A contribution to learning appetite, critically looking (PD. 36, CT).*

Given teachers' participants diaries, they most frequently highlighted the following descriptions: "realizing, investigating, curiosity, different ways of looking, immersive perception of the space and objects, critical/multidimensional questioning, cultural encounters and learning appetite". It is seen that teachers' VR museum experience in the context of visual culture enabled them to focus on cognitive and emotional learning.

5 sub-themes were developed **under the theme of the evaluation of virtual museum experience in the context of teaching profession**. These are: 1) Preliminary information regarding virtual museum experiences, 2) Evaluations regarding the contribution of virtual museum experience, 3) Evaluations regarding potential problems and challenges 4) Evaluations regarding the contribution of VR museum experience to visual arts course/education and lastly 5) Ideas regarding the classroom practices.

Preliminary information regarding virtual museum experiences was obtained from pre-evaluation forms and entails information with respect to teachers' virtual museum practices. The data obtained in this sub-theme was grouped into the following three categories. Code and themes are presented in Figure 3.
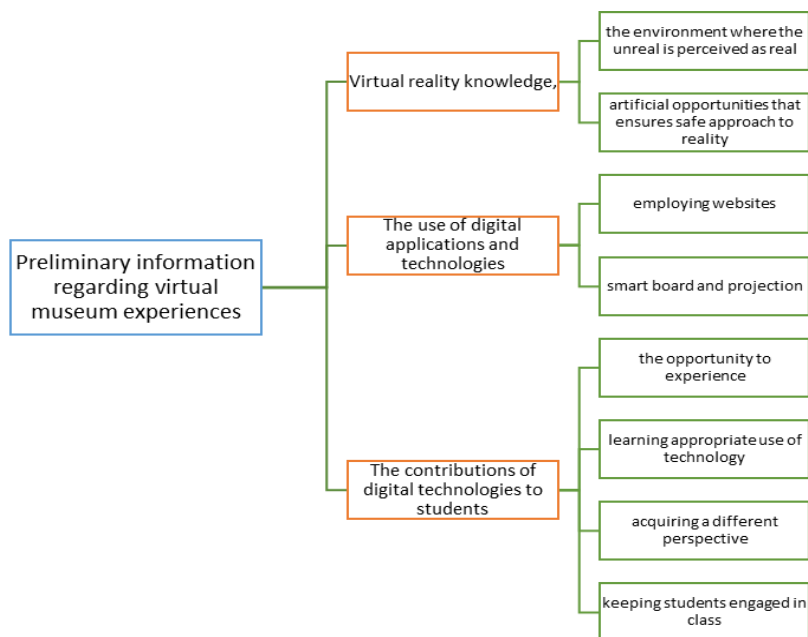

Figure 3. Codes and themes for preliminary information on virtual museum experiences

Teachers mostly described the virtual reality as the experience of spaces that cannot be visited physically and where the reality does not exist. Some teachers, however, reported that visual reality keeps the person from

dangers and allows the user to experience the space. The 5-coded visual arts teacher expressed her opinion in this direction as "Artificial possibilities that bring the real person closer by keeping them away from possible dangers as if they were physically in that environment.

In the category of the use of digital applications and technologies, a great majority of the teachers were found to use digital applications to access to different web pages. Besides, some teachers also stated that they used digital tools available in classes such as smart board and projection. Teachers predominantly expressed their purpose in making use of these technologies as teaching abstract subjects by concretizing.

As for the category of the contribution of digital technologies to students, most of the teachers noted that digital technologies promoted equality of educational opportunities for low-income students. Teachers frequently stated that students could easily and safely access to physical environment or resources that they would not normally have access to. Amongst the teachers who reported that using technology in classroom would lead to more highly engaged students, the teacher coded 41 commented, *"I believe that the goal can be realized more quickly as it draws more attention. Besides, since the age group we reach is very good at technology use, curiosity and interaction will occur."* (PEF-VAT) In addition to that, some of the teachers noted that integrating digital technologies into courses will also help students learn appropriate use of technology.
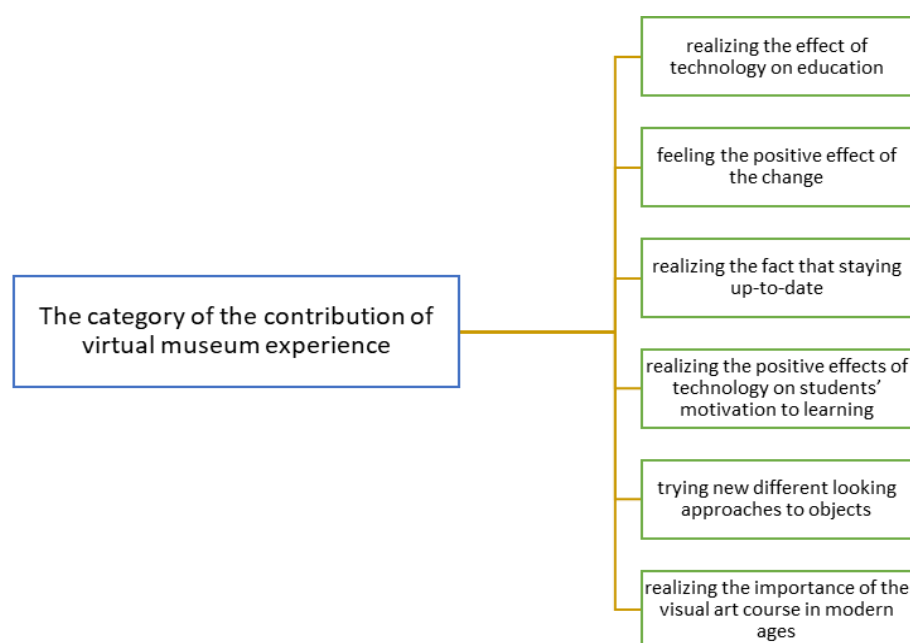

Figure 4. Codes for teachers' evaluations of the virtual museum experience on their contribution to them

Under the category of the contribution of virtual museum experience, 6 codes have emerged as shown in Figure 4. Most teachers emphasized that they gained awareness of the effects of technology on education. For example, the visual arts teacher coded PD.1 stated*, "Looking at the evolution of technology, I see how it impacts on art and thanks to technology, I could move freely in places where it is not possible to go".* The visual arts teacher, yet, coded PD.2 made self-criticism and presented his/her observation saying that *" As employees of Ministry of National Education, I saw how we lagged behind tech opportunities, I face the truth that we fell behind with staying up to date.* The visual arts teacher coded PD.11 stressed the opportunities virtual museum provided to class environment stating that "*I realized that I can make my students do activities such as museum visits or art work analysis without restriction of place, location or time.*" Similar expressions were also identified in the focus group interviews. For example:

> We need to digitize. This is the need of our era. Even if we want, we cannot alienate... However, we should not introduce the digital environment to the students immediately. As teachers, we should be conscious and selective. (Denizli, FGI. VAT-1).

We could thus imply that VR museum experience raised teachers' awareness of digital learning and using technology as a learning tool. It is seen that teachers frequently used the words "realize" and "feel" when they shared their views in respect to the contributions of technology to themselves. They also made self-criticism for

their lack of digital skills. This critical attitude also gave them a refreshed belief concerning their fields or the significance of the visual arts course. Below are some excerpts taken from the comments of the teachers:

> *Visual arts course is not just a practice course, unfortunately, students and my colleagues still perceive it like that. We gradually started to destroy this perception. I believe my course covers many areas of life (PD. 38, VAT).*

> *Once again, I realized how much I love my job. This feeling will reflect positively on my students and my environment. Then, I think I will be a more helpful teacher (PD. 42, VAT).*

Some concerns, however, were raised related to VR museum experience. These concerns were divided into 7 codes under the sub-theme of "*evaluations regarding potential problems and challenges in practice*". The codes are presented in Figure 5.
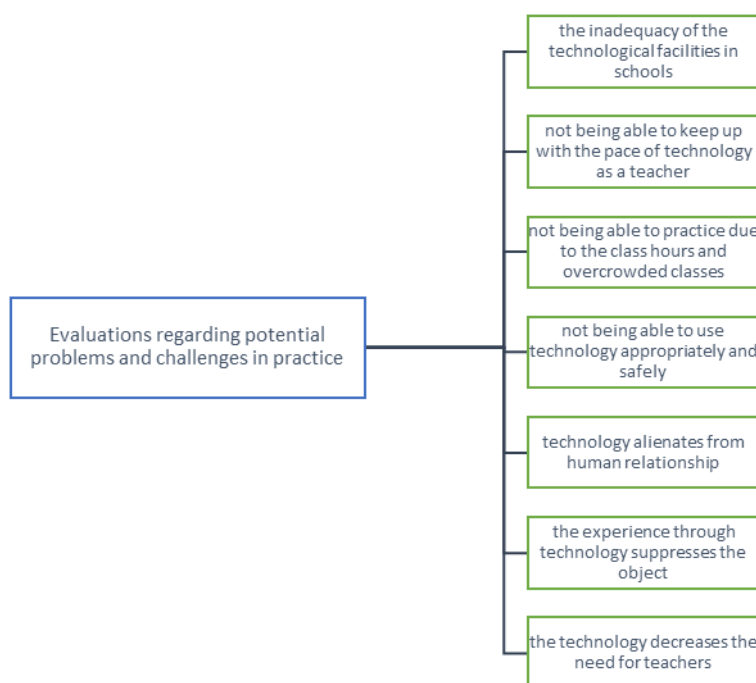

Figure 5. Concerns about virtual museum experiences

Some of the teachers explained their anxieties through their own competencies, the limitations offered by the internet and the inability to control it safely, and the dominant influencing role of technology. For instance, the teacher coded 4 stated, *"My only concern is that there will be no need us in case of heavy use of technology in future?"* (PD. 4, VAT). The teacher coded 5 commented*, "The possibility that technology will suppress hand crafted artworks makes me concerned"* (PD. 5, VAT). The teacher coded 12 made a more objective comment saying that *"My satisfaction and concern are at the same level. Limitless…* (PD. 12, CT). The teacher coded 14 stated, *"I concern that digitizing of art, human and artist and its prevalence will alienate people from socializing, social and human relations"* A great majority of the teachers stressed the challenges that they might encounter during the practice in school environments. Amongst the challenges they identified, there are Internet infrastructure in school, hardware inadequacies in rural schools, and the difficulty of the practice due to limited course hours with overcrowded classrooms. For instance, the teacher coded 35 stated, *"It offers the opportunity to watch and see on-site regardless of time, money and formal procedures. What makes me concerned is internet is only available in cities and inequality of opportunity, rural schools are lacking these opportunities"* (PD. 35, CT*)*. More detailed evaluations were found in focus group interviews with respect to the difficulties in schools. Example quotations from the teachers' statements are presented below:

> *You need to have a sound Internet. Internet does not work properly in schools. Besides, since the child could not go to museum physically and it cannot be like the real atmosphere it has an adverse side (Çanakkale, FGI. VAT-5).*

> *Some psychological effects can be… The child would like to detach from reality (Çanakkale, FG. VAT-2).*

> *Visual realty draws a lot of attention from students. But I will have flying experience when I go to Louvre Museum to see Mona Lisa, I will see Paris, and I will interact with people. Virtual reality is nice, but it cannot replace it (Denizli, FGI. VAT-3).*

As teachers' statements demonstrated, they mostly believed that virtual museum experience could adversely affect the child in terms of detaching from reality. Yet, when it comes to the contribution of VR museum experience to the course and visual art education, a total of 9 codes were determined. These codes are presented in Figure 6.
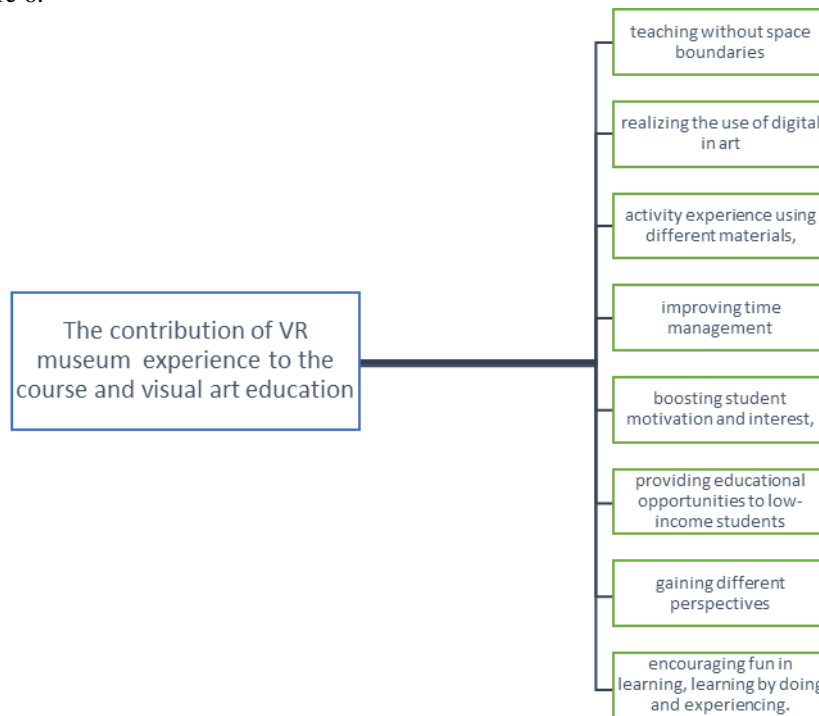

Figure 6. Evaluations on the contributions of the virtual museum experience to visual arts teaching

For example, the classroom teacher coded 14 reported, *"I can incite interest and curiosity in art. Eventually, students follow technological products. If we make them experience that they can use it to access to art, they will be more interested. They have the chance to encounter more branch of art and artworks".* The visual arts teacher coded PD 42 stressed its contribution in terms of accessibility stating that *"Given that students may be stay in the city they live throughout their life, this experience they have, some will have the chance to see these art pieces in the distance, this will be an unprecedented happiness for them"* Teachers coded 29 and 15 mostly focused on the effect of this experience on children.

> Instead of looking and examining a fixed painting, children have the chance to go to museum without changing the time and space, they will like it (PD. 29, VAT).
> It will promote children's interest in art. Their interest on making art and producing (PD. 15, CT).

Similar expressions were observed in the focus group interviews with teachers. Teachers agreed on that it would be useful in terms of practicing the topic learned, permanent learning and critical thinking. Some excerpts of the participants are given below:

> It is tough to teach children the concept of art. But I believe that this immersive experience will allow the practicing of any concept very well (Kayseri, FGI. CT-4).
> Instead of lecturing, you let the children experience it personally, he/she wants to see the real one after experiencing. Even this is a big gain (Erzincan, FGI. VAT-1).
> When I experienced it, I felt the period Leonardo lived. It is important for students in terms of feeling the same, being curious about the history and making more in-depth research (Erzincan, FGI. VAT-2).

During the focus group interviews, teachers were asked to answer what activities they prefer to conduct in line with the experiences they gained after the in-service training and their responses were classified under the theme of *"opinions regarding the in-class activities"*. With respect to the VR museum experience in visual culture education, teachers' responses are as follow: examining the artwork/object in virtual environment, completing the painting or reinterpreting, developing stories or visual narratives about the before and after the moment, building interdisciplinary perspective, planning activities integrated with interactive boards. To illustrate, the following expressions were obtained in the focus group interviews and participant diaries:

> We can ask the students to make a painting like "if you were an artist, how would you make this painting or complete it? (Kayseri, FGI. VAT-1).

> *Writing activities can be conducted. Like which dentist Mona Lisa did she visit before posing, what kind of things did she come across at the street and even drawing activities can be done (Erzincan, FGI. VAT-6).*
> *It can be incorporated into all courses and subjects and it can be integrated with other courses. Different cultures can be introduced (PD 13, CT).*

Although teachers frequently mentioned that they found the experience of a physical visit to the museum more valuable, they expressed that virtual museum experience could be more effective in classes in terms of the opportunity to visit exhibitions without having to travel far, promoting interest and learning appetite by means of technological opportunities (sound, motion, immersive and etc.) and it can be incorporated into various learning activities. Interestingly, the idea of creating interdisciplinary lesson plan was mostly emphasized by classroom teachers.

Discussion and Conclusion

Visual culture studies in art education (Duncum, 2002; Tavin, 2007) are predominantly shaped on aesthetic experiences (Billboards, advertisements, Youtube videos, social media posts, digital games, shopping malls, tourist attractions, etc.) obtained from daily life. It is emphasized that daily aesthetic experiences are more on one's identity and world experiences than the artistic experiences encountered in museums. However, although some arts education researchers (Eisner, 2002; Efland, 2005) are positive towards the creation of an arts teaching program that relates learning with real-life experiences, they express concern that a focus solely on visual culture will hamper the true goals of arts education. However, in recent years, museums have found more places in digital media through new narratives, exhibitions, digital games and even various teaching activities, making it easier for them to penetrate into daily life and meet with the audience. Based on this fact, the two museums experienced with virtual reality glasses were included in a teacher education process created in the context of visual culture inquiries. Thus, it was aimed to present a new perspective to the teachers on how visual culture can be interpreted in the context of museum education.

The contact with the artworks and objects in the virtual museum is not same as the contact in real museum environment. However, virtual museum applications and technologies bring museum collections of an actual museum anywhere in the world to class environments, thereby serving to educational purposes. Learning in virtual museum environments should be constructed in a way that ensures effective learning and empowers students' critical thinking skills as the learning in actual museums do. "In virtual museum activities, the quality of the activity and its educational outputs are closely linked to teachers' planning and implementing of teaching" (Yılmaz, Yıldırım, Filiz & İbrahimoğlu, 2018, p. 34).

It is important to examine what appears in visual culture studies, to question who or what the meaning attributed to it and the created images represent. Associating, questioning and interpreting create individual and social contexts that are attributed to the meaning of that object. (Mamur, 2012, p.2163). This study revealed that the participant teachers questioned the visual culture in the context of the relationship through object, space-time, social cultural codes, permanence and temporariness and they produced meanings. Visual culture in art education involves multidimensional thinking and deeper thinking skills, which is designed through constructivist approach and critical pedagogy. This also can be seen in the theoretical foundation of museum education. "Museum education centers on a holistic approach that emphasizes individual learning characteristics such as cognitive styles, constructivist approach and multiple intelligence theory" (Lepouras and et al., 2001). Indeed, it was found that following their VR museum experience, teachers developed questioning approaches pertaining to space, object, virtual reality, personal context, form and aesthetics and they effectively employed their cognitive and emotional skills. Similar findings were found in a study conducted in Eskişehir Cartoon Museum (Sarıbaş, 2020) on museum, visual culture and representation. In the physical environment of the museum, visual arts teacher candidates made inquiries on visual representation and meaning, and the meanings they created through the cartoons they examined were reflected in their reproduction as cognitive and affective wealth. The important outcomes of the implemented program were shown as the motivation of the students and the context they established with their own lives. On the other hand, it can be said that social interaction is limited because the teachers have a more personalized museum experience. In this study, it was found that teachers' motivation in the process, being influenced by the virtual collection and their opinions about the virtual museum application were positive; however, it has been found that the influence stemming from the physical context provided by a real museum has shifted more to the possibilities offered by technology. Similar results were found in a study conducted by Özer (2016) with the content of a virtual museum, and it was stated that the effect of the physical environment did not occur, and the social context remained in a more limited perspective. The data obtained from reflective tools demonstrate that teachers gained learning experiences through the in-service training project in terms of personal and professional contexts. Teachers reported that they became

aware of the power and capacity of especially VR technologies and virtual museum environments in visual art courses in terms of practicing the course, permanent learning, critical thinking and demonstrating cultural diversity to students by bringing any museum across the world into the classroom environment. In this context, it can be said that when technological tools are combined with well-structured curricula, they direct learners to higher level thinking skills. According to Silberman (2016), technological tools facilitate learning by engaging students in high-level active learning, allowing them to reach information faster and to do many things with that information, allowing them to think critically, analyse topics and practice (Cited in Karadeniz, 2020).

While access to museums and incorporating them into visual culture teaching activity seems to be the strength of this study, most teachers expressed their concerns about not keeping up with the pace of technology or using it properly and safely in the classroom. In addition to that, teachers mentioned that the practice of virtual museum experiences might be challenging due to the limited course hours with overcrowded classrooms and schools lacking Internet infrastructure. Another finding, however, revealed that such educational process increased teachers' willingness to benefit from technology more effectively in art teaching. It was observed that teachers developed some ideas to enrich their own course activities with virtual museum experiences as a result of the experiences they gained during the activity process. Teachers most frequently developed such ideas as examining artworks / objects in the virtual environment, comparing works of different cultures, writing stories about before or after the moment or creating visual narratives, and ensuring interaction between virtual museum environment and other courses. Creating networks of interdisciplinary relationship and cultural comparisons or the opportunity to examine the codes of a different culture through art works might be the most important output of virtual museum examination in the scope of visual culture. Because while a student has the opportunity to encounter cultural museum objects in his/her country, she may not have the opportunity to encounter objects and works of another culture. It can thus be said that virtual museum environments give us the opportunity to realize, understand and interpret cultural codes and intercultural interactions.

Based on the results of the research, it is especially suggested that the awareness and competencies of teachers should be further increased through different in-service training activities in order to diversify the opportunities for remote access to museums in art classes with critical visual reading approaches. This study was carried out as a qualitative research, without the aim of generalization. However, new research can be made on the use of virtual museums in education by considering different learning theories and models. In these studies, it can be studied on models in which both real and virtual museum applications are used together to strengthen the social and physical dimensions of the contextual learning model.

Acknowledgements or Notes

References

Akça, S. (2020). Teknoloji ve bilgi çağında müzeler: Genel bakış [Museums in the age of technology and information: Overview]. *The Journal of Turkish Librarianship, 34*, (2), 263-274.

Anderson, T. (2003). Art education for life. *Journal of Art and Design Education, 22* (1), 58-66.

Barlas Bozkuş, Ş. (2014). Kültür ve sanat iletişimi çerçevesinde Türkiye'de sanal müzelerin gelişimi [The development of virtual museums in Turkey: As perspective of culture and art communication]. *The Journal of Academic Social Science Studies, 26,* 329-344.

Burnham, R. & Kai-Kee, E. (2015). *Müze dersleri: Yorum ve deneyim [Teaching in the art museum: Interpretation as experience].* (Çev./Trans. Aylin Onacak), İstanbul: Koç Üniversitesi Yayınları.

Bülbül, H. (2017). Müze eğitiminde bir seçenek olarak "Second Life" müzeleri [Using second life museums as an option ın education via museum]. *Milli Eğitimi Dergisi, 214*, 167-180.

Clutterbuck, R. (2008). Müzeleri anlamak, eğitim ve müze seminerleri [Understanding museums, education and museum seminars]. In İnci San (Ed.)*, Eğitim ve Müze Seminerleri [Education and Museum Seminars]* (pp. 63-83) Ankara: Kök Yayıncılık.

Creswell, J.W. (2012*). Educational research: Planning, conducting, and evaluating quantitative and qualitative research.* 4th ed. Boston: Pearson Education.

Dierking, L.D. (2002). The role of the context in children's learning from objects and experiences. In Paris, S.G. *Perspectives on object-centered learning in museums* (pp. 3-18). New Jersey: Lawrence Erlbaum Ass.

Dilli, R. (2020). Görsel kültür eğitimi [Visual culture education]. In Vedat Özsoy (Ed), *Kültürel ve eleştirel sorgulama olarak görsel sanatlar öğretimi [Teaching visual arts as cultural and critical inquiry]* (pp. 191-208). Ankara: Pegem Akademi.

Duncum, P. (2002). Theorizing everyday aesthetic experience with contemporary visual culture, *Visual Arts Research, 28*(2), 4-15.

Efland, D. A. (2005). Problems confronting visual culture, *Art Education, 58* (6), 35-40

Eisner, E. (2002). *The arts as the creation of mind,* Yale University Press.

Ertürk, N. (2020). Müzeler karantinada! [Museums in quarantine!]. *Turkish Studies, 15* (4), 389-402. https://dx.doi.org/10.7827/TurkishStudies.44342

Falk, J. H. (2009*). Identity and the museum visitor experience*. Walnut Creek, CA: Left Coast Press.

Hagberg, G. (2016). Word and object: Museums and the matter of meaning. *Royal Institute of Philosophy Supplement, 79*, 261-293.

Hall, S. (2019). *Temsil kültürel temsiller ve anlamlandırma uygulamaları [Representation: Cultural representations and signifying practices]*. (Çev./Trans. İdil Dündar), İstanbul: Pinhan Yayıncılık.

Hawkey, R. (2004). Learning with digital technologies in museums, science centres and galleries. King's College, London.

Hürst, W. Coninck, F. & Tan, X. J. (2016). Complementing Artworks to Create Immersive VR Museum Experiences. Proceedings of the 13th International Conference on Advances in Computer Entertainment Technology, 34, 1–6, November 09-12, 2016, Osaka, Japan, DOI: https://doi.org/10.1145/3001773.3001806

Karadeniz, C., Okvuran, A., Artar, M. & Çakır İlhan, A. (2015). Yeni müzebilim bağlamında müze eğitimine çağdaş yaklaşımlar ve müze eğitimcisi [Contemporary approaches and museum educator within the context of new museology]. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, 48,* (2), 203-226.

Karadeniz, C. (2020). Müzede dijital teknolojilerin kullanımı ve salgın sürecinde dijital katılım [The use of digital technologies in the museum and digital participation through pandemic process]. *Idil,* 70, 975–984. doi: 10.7816/idil-09-70-06

Karatay, A. & Karatay, A. (2015). New technologies in museum information design and augmented reality applications, *Proceedings of International Fine Arts Symposium* (pp.201-209), 12-13 November 2015, Konya, Turkey.

Leppert, R. (2009). *Sanatta anlamın görüntüsü: İmgelerin toplumsal işlevi [Art and the committed eye: The cultural functions of imagery]*. (Çev./Trans. İsmail Türkmen). 2nd ed. İstanbul: Ayrıntı Yayınları.

Lepouras, G. Charitos, D., Vassilakis, C., Charisi, A. & Halatsi, L. (2001). Building a VR museum in a museum, in Richir, S., Richard, P., Taravel, B. (eds.), Proc. 2001 *VRIC Laval Virtual International Conference*, 16-18 May 2001, Laval, France.

Lewis G. (1996), The response of the museum to the web, *Museum discussion list*, Erişim: http://home.dc.lsoft.com/archives/museum-l.html. 15.06.2020.

Mamur, N. (2012). Perceptions concerning visual culture dialogues of visual art pre-service teachers. *Educational Sciences: Theory & Practice 12*(3), 18-26.

Mamur, N. (2019). Öğrenme kuramları ve görsel sanatlar öğretimi [Learning theories and visual arts teaching]. In Vedat Özsoy (Ed.), *Görsel sanatlar öğrenme ve öğretim yaklaşımları* [Visual arts learning and teaching approaches]. (pp.139-152) Ankara: Pegem Akademi.

Onur, B. (2012). *Çağdaş müze eğitim ve gelişim: Müze psikolojisine giriş* [Contemporary museum education and development: An introduction to museum psychology]. Ankara: İmge Kitabevi.

Özer, A. (2016). Sanal müzede öğrenmenin bağlamsal modelinin kullaniminin öğrencilerin akademik başarısı, motivasyonu ve memnuniyet düzeylerine etkisi [Using The Contextual Model of Learning Effect of Academic Achievement, Motivation and Satisfaction Levels of Students in Virtual Museum] (Unpublished doctoral dissertation). Ankara Üniversitesi.

Özsoy, V. (2020). Görsel sanatlar eğitiminde görsel kültürün temelleri. [Basics of visual culture in visual arts education]. In Vedat Özsoy (Ed) *Kültürel ve eleştirel sorgulama olarak görsel sanatlar öğretimi* [Teaching visual arts as cultural and critical inquiry]. (pp. 1-20). Ankara: Pegem Akademi.

Saribaş, S. (2020). Müzede görsel kültür temelli bir uygulama. [An application based on visual culture in the museum]. In Vedat Özsoy (Ed) *Kültürel ve eleştirel sorgulama olarak görsel sanatlar öğretimi [Teaching visual arts as cultural and critical inquiry]*. (pp. 241-261). Ankara: Pegem Akademi.

Schofield, G., Beale, G., Beale, N., Fell, M., Hadley, D., Hook, J., Murphy, D., Richards J. & Thresh, L. (2018). *Viking VR*: *Designing a virtual reality experience for a museum, immersive experiences* içinde. (pp. 805-815) DIS 2018, June 9–13, 2018, Hong Kong.

Stone, D. L. (2001). *Using the art museum,* Massachusetts: Davis Pub.

Tavin, K. (2003). A critical pedagogy of visual culture as art education: Toward a performative inter/textual practice. (Unpublished doctoral dissertation). The Pennylvania State University.

Tavin, K. (2002). Engaging advertisements: Looking for meaning in and through art education. Visual Arts Research, 28(2), 38-47.

Tishman, S. (2018). *Slow looking: the art and pratice of learning through observation,* Routledge.

Türkkan, B. (2008). İlköğretim görsel sanatlar dersi bağlamında görsel kültür çalışmaları: Bir eylem araştırması [Visual culture studies in primary school visual arts course: An action research]. (Unpublished doctoral dissertation). Anadolu University, Eskişehir.

Utkan Özden, H. (2019). Çağdaş sanat eserleri ve anlamları: Çağdaş sanat müzelerinin sanal müze olarak varlığı [Contemporary art works and their meanings: The existence of contemporary art museums as virtual museums.]. *Unimuseum, 2* (2), 59-62.

Wilson, B. (2002). Conflicting forms of cultural literacy: Inside school and beyond schooling. *Journal of Research in Art and Education, 3*, 123-151.

Yenişehiroğlu, F. (2008). Türkiye'de müzecilik bilgi alanının gelişmesi üzerine gözlemler [Observations on the development of the museology field in Turkey]. *Sanat Eğitimi ve Sanal Müze: Geçmişten Geleceğe Türkiye'de Müzecilik Sempozyumu Bildiri Kitabı [Art Education and Virtual Museum: Past to the Future Symposium Proceedings of Museology in Turkey]*, 21-22 May 2007, Ankara (s. 247-257). Ankara: VEKAM Yayınları.

Yıldırım, A. & Simşek, H. (2013). *Sosyal bilimlerde nitel araştırma yöntemleri [Qualitative research methods in the social sciences].* Ankara: Seçkin.

Yılmaz, A., Yıldırım, G., Filiz, F. & İbrahimoğlu, Z. (2018). *Etkileşimli müze ve tarihi çevre eğitimi: Kuramsal temeller ve örnek etkinlikler [Interactive museum and historical environmental education: Theoretical foundations and sample activities].* Ankara: Pegem Akademi.

Yücel, D. (2012). *Yeni medya sanatı ve müze [New media art and museum].* İstanbul: İstanbul Kültür Üniversitesi Yayınları.

Zhao, Y., Forte, M, & Kopper, R. (2018). VR Touch Museum*, IEEE Conference on Virtual Reality and 3D User Interfaces*, 18-22 March, Reutlingen, Germany.

Zouboula, N., Fakides, M. & Tsolakidis, C. (2008). Educational Uses of Virtual Reality: Constructing a VR Museum, *International Conference on Interactice Collaborative Learning Proceedings.* September 24 - 26, 2008 Villach, Austria.

The Relationship between Authentic Leadership and Work Engagement

Ramazan BAŞARAN [1], Erkan KIRAL [1]
[1] Aydın Adnan Menderes University

To cite this article:

The Relationship between Authentic Leadership and Work Engagement[*]

Ramazan Başaran[1][**], Erkan Kıral[1]
[1] Aydın Adnan Menderes University

Abstract

The purpose of this research to examine the relationship between the authentic leadership behaviors of the school administrator and the work engagement of the teachers, according to the perceptions of the teachers of secondary school. The research was designed in the relational survey model and carried out with 300 teachers. Descriptive and inferential statistical techniques were used in the research. As a result of the research, it was found that teachers' general perceptions of authentic leadership of school administrators was relatively high and the highest level on balanced processing. Teachers' general perceptions work engagement was relatively high and the highest level on dedication. Positive, medium and low-level relationships were found between authentic leadership and work engagement. According to teachers, the authentic leadership of school administrators significantly predicted teachers' perceptions of work engagement. Based on these findings it has been proposed to various suggestions have been put forward such as protect and increase the authentic leadership behaviors of school administrators and teachers' work engagement levels.

Key words: Authentic, Leadership, Engagement, Administrator, Teacher.

Introduction

Educational institutions are institutional structures that have taken the biggest responsibility in raising capital, which enables countries to survive and reach better points. Many activities from the acquisition of basic skills of the child to professional shaping are presented to him/her by the school's internal stakeholders in the schools that form the basis of educational institutions. So, schools are the places where structure the foundation of the intellectual capital which is considered as the most important capital of the countries. According to Taymaz, the duties that schools undertake in raising individuals cannot be ignored (2009), and school administrators and teachers who are internal stakeholders of the school should be aware of this responsibility. They should make an effort to improve their school's day by day and take different responsibilities besides their responsibilities. As a matter of fact, many changes from culture to technology in the world have caused different expectations both in internal stakeholders and external stakeholders of the school and put more responsibilities on school administrators to meet these expectations. School administrators should display leadership behaviors in addition to management in order to meet the changing expectations and needs of the school's stakeholders. One of the recently expected leaderships from school administrators is authentic leadership (Kıral, 2018b). In recent years, it has become almost impossible to make schools effective by only managing schools, and it has become a sought-after feature that school administrators exhibit leadership. Authentic leadership is an important form of leadership that school administrators should demonstrate. Authentic leadership includes school administrators to be like themselves with sincere behavior rather than giving them leader impression, their relationship is honest and transparent, and the establishment of a school climate based on trust and ethical principles at school. The trust factor in the focus of authentic leadership is a very important concept as it affects the quality of education (Kıral & Başaran, 2018a). Decreasing trust in school is an important factor leading to deterioration of school climate and poor quality of education (Kıral & Başaran, 2018b).

---

[*] This study is produced from the first author's master's thesis titled "The Relationship Between Authentic Leadership and Work Engagement in Secondary Schools" under the supervision of the second author, and it was presented as an oral paper at VI. International Eurasian Educational Research Congress held at Ankara University between 19 - 22 June 2019.

[**] Corresponding Author: *Ramazan Başaran, basaran2006@gmail.com*

When the origin of the concept of authentic is examined, almost all cultures have "knowing themselves and acting with this consciousness" (Caza, Bagozzi, Woolley, Levy & Caza, 2010). So, people who know themselves and take this into account in their actions can build trust and influence people in their relationships with others. In this interaction, if the individual can leave more influence on the other side than the one left by him/her, he/she can come to the fore as a leader. According to George, the authentic leader; is a person who is aware of his/her limits while using his/her natural abilities and strives to overcome these limits, is followed by his/her followers because he/she knows where to stop and shows disciplined and consistent behaviors (2003). The authentic leader acts principally and does not compromise his/her principles. Knowing the purpose of life, the authentic leader wants to develop both himself/herself and those around him/her and makes an effort for it. According to Avolio, Gardner, Walumbwa, Luthans & May, the authentic leader knows who he/she is, what he/she thinks, how he/she behaves; a hopeful, confident, flexible and high moral character; is the person who can grasp the moral perspectives, values and powers of others (2004). The authentic leader is self-confident and confident, sincere, aware of his/her values and beliefs (Ilies, Morgeson & Nahrgang, 2005). Authentic leader can create a positive climate in the organization by focusing on the followers and influencing their development with the leadership it will exhibit. According to Luthans & Avolio, he/she acts in authentic leadership by putting elements of self-discipline and self-awareness in his/her behaviors and provides self-improvement. In fact, authentic leadership is a leadership approach that emerged under the influence of positive psychology (2003). This approach essentially added authenticity to the phenomenon of leadership and formed leadership actions accordingly. In order for the leader to be considered as authentic, he/she has to exhibit four components, which will create a link between himself/herself and his/her followers; self-awareness, balanced processing, internalized moral perspective and relational transparency. Relational transparency refers to sharing the information and feelings that are appropriate for the situations clearly; internalized morality refers to behave in harmony with high moral standards and the person's value patterns, choices and needs; balanced processing refers to balanced information processing and objectively analyzing the data related to the subject, and the tendency to discover the ideas of others before making a decision; self-awareness refers to one's knowing his/her strengths and weaknesses and the way he/she makes sense of the world (Avolio et al., 2004; Gardner, Avolio, Luthans, May & Walumbwa, 2005). As can be seen, a school administrator who is transparent in his/her relations, knows himself/herself, has adopted universal ethical principles and cares about the employees can affect the educational activities of teachers, which is the most strategic element of education. The school administrator can ensure that teachers, who are the practitioners of the educational process with authentic leadership behavior, passionately fulfill this process and engage in their work.

The origin of work engagement comes from the concept of engagement. Using the concept of engagement for the first time, Kahn (1990) based this concept on the role theory of the individual; it has been investigated as personal engagement and disengagement. Schaufeli & Bakker (2010) expressed engagement in behavior as a manifestation of a particular mental state that occurs when they are physically at work, by giving the employee energy to the job role. Schaufeli, Bakker & Salanova examined work engagement as one of the psychological conditions affecting the individual in the area of positive organizational behavior (2006). In this context, although work engagement is claimed to be an umbrella concept that covers similar theories in the field of organizational behavior (Macey & Schneider, 2008), it is considered as a concept that can be applied empirically, practiced, definitively and distinctively defined in academia (Leither & Bakker, 2010). Work engagement is defined as negatively related to burnout, but as a separate, self-concept, a positive, satisfying mood for the job of the employee, which is characterized by vigor, dedication and absorption. Vigor means that the employee gives all his/her energy to his/her work without getting tired immediately while working and mental endurance; dedication means that the job is meaningful and attractive to the employee, requires struggle, the employee become integrated and embraces it with passion; absorption means that the employee focuses on his/her job, does not see external stimuli, pays all his/her attention, does not care about anything else, and is happy when doing his/her job (Schaufeli, Salanova, Gonzales-Roma & Bakker, 2002). As it is seen, meaningfulness of the work for the employee and working with passion by concentrating on it can make the job more qualified. Teachers engaged in education and training are also expected to be engaged in their jobs. As a matter of fact, teachers, who are engaged in their jobs, can lead to educate more qualified children, who are the assurance of their future. However, this engagement of teachers who are engaged in their jobs can be provided by school administrators who will lead them. It can be expected that teachers will be willing and efficient in fulfilling the functions of the school, they will take responsibility voluntarily, and they will be satisfied with their work in a school environment where authentic leadership behaviors are exhibited by the school administrator. For this reason, the fact that the school administrator, who is accepted as an important element of effective schools, is authentic in leadership behaviors and the effect of this on teachers' work engagement is worth examining. In addition, work engagement is beneficial for all organizations, but it has a special importance especially in educational organizations. In this context, a teacher who feels vigorous, dedicated and absorbed while doing his/her job will make a positive contribution to the quality of education.

When the literature reviewed, it was dedicated that there has been no study conducted about the relationship between authentic leadership behaviors by school administrators for teachers and teachers' work engagement in secondary schools. It is important to determine the authentic leadership level of school administrators and to know to what extent the school administrators' authentic leadership effect teachers' work engagement. Thus, school administrators and policy makers assigning them can take the necessary measures. As a matter of fact, the purpose of this study is to reveal the effect of the authentic leadership behaviors exhibited by the school administrators according to the perceptions of teachers working in secondary schools on the levels of teachers' work engagement.

Research questions
1. What are the authentic leadership behavior levels of school administrators according to the perceptions of secondary school teachers?
2. What are the secondary school teachers' perception levels of work engagement?
3. According to the perceptions of secondary school teachers, is there a significant relationship between the authentic leadership behaviors of school administrators and teachers' perception levels of work engagement?
4. According to the perceptions of secondary school teachers, do the authentic leadership behaviors of school administrators predict teachers' levels of work engagement?

Method

Sample

The relationship between the authentic leadership behaviors exhibited by the school administrators and the teachers' work engagement was examined through relational survey model (Karasar, 2016), which is one of the research models aiming to determine the degree of change between two or more variables.

The target population of the current study is 1059 teachers working in 35 official secondary schools in 2017-2018 academic year in Aydın province Efeler district (National Education Directorate [Aydın NED], 2018). In this research, each secondary school was handled as a cluster with a multi-stage sampling method (Balcı, 2016; Karasar, 2016) based on certain procedures, including a series of stages, and simple random sampling method was used with proportional cluster sampling (Barreiro & Albandoz, 2001). By reaching all 35 secondary schools in which 1059 teachers work, it was ensured that each school is represented in the sample with the number in the proportion represented in the target population. In calculating the sample size, the sample size calculation table was used (Can, 2016). It is assumed that the target population consisting of 1059 secondary school teachers can represent 282 secondary school teachers at the level of $\alpha = .05$ significance and 5% tolerance. It was decided that the sample should be composed of 338 middle school teachers by taking 20% more of the sample calculated due to problems that may occur during the data collection process and returns. 338 volunteer teachers participated in the research but 38 data collection tools (outlier, not filling proper, etc.) were removed from the analysis. Thus, the research was carried out through the data collection tools filled in by 300 participants.

As a result of the analysis of the data in the research; it was found out that 173 (57.7%) of the teachers are female and 127 (42.3%) are male; 264 (88%) are married and 36 (12%) are single. The ages of the teachers vary between 23 and 63 ($\bar{X} = 40.41$). Teachers' year of service in their school varies between 1 and 30 years ($\bar{X} = 5.84$). It was found out that 282 (94%) of the teachers are undergraduate and 18 (6%) are graduate.

Measures

Authentic Leadership Scale

In the current study, Authentic Leadership Scale (ALS), which was developed by Walumbwa, Avolio, Gardner, Wernsing & Peterson (2008) and adapted into Turkish by Kıral (2018a), was used to find out the authentic leadership perceptions of teachers towards school administrators. Adaptation of the scale into Turkish on the sample of secondary school teachers was made by Kıral (2018a). ALS consists of four dimensions: relational transparency, internalized moral perspective, balanced processing and self-awareness. It was determined that as a result of exploratory factor analysis, the scale preserved its original structure with four factors and the first and second fit indices of the scale were appropriate as a result of confirmatory factor analysis. In the reliability analysis, Cronbach Alpha reliability values were found to vary between .60 and .80 for general authentic leadership and the four dimensions of authentic leadership. Due to the fact that the ALS is a relatively new

adaptation, the number of items is less, adapted with a similar sample and the adaptation studies in other studies were carried out in different disciplines other than education, ALS which was adapted by Kıral (2018a) was preferred. The scale consisted of three items for each dimension and a total of 12 items includes 5-point Likert type rating ["I disagree at all" (1), "I totally agree" (5)]. The high scores obtained from the scale indicate that the feature of each dimension is high. In the current study, Cronbach Alpha reliability coefficient was found to vary between .75 and .92 for general authentic leadership and its four dimensions. The fact that these values are in the range of .60 to .80 indicates that the scale is "quite reliable" (Tavsancıl, 2014).

Work Engagement Scale

In the current study, Work Engagement Scale (WES), which was developed by Schaufeli & Baker (2003) and adapted into Turkish by Kıral (2018b), was used to find out the teachers' work engagement level. WES consists of three dimensions: vigor, dedication and absorption. It was determined that as a result of exploratory factor analysis, the scale preserved its original structure with three factors and the first and second fit indices of the scale were appropriate as a result of confirmatory factor analysis. As a result of the adaptation of the scale into Turkish, the internal consistency coefficients were found to vary between .78 and .90 for general work engagement and the three dimensions of work engagement. Due to the fact that the WES is a relatively new adaptation, the number of items is less, adapted with a similar sample and the adaptation studies in other studies were carried out in different disciplines other than education, WES which was adapted by Kıral (2018b) was preferred. The scale consisted of four items for each dimension and a total of 12 items includes 5-point Likert type rating ["I disagree at all" (1), "I totally agree" (5)]. The high scores obtained from the scale indicate that the feature of each dimension is high. In the current study, Cronbach Alpha reliability coefficient was found to vary between .75 and .89 for general work engagement scale and its three dimensions. According to these values, the scale can be said to be "quite reliable" (Tavsancıl, 2014).

Analysis

Descriptive and inferential statistics techniques were used to analyze the data collected in the research. The demographic characteristics of the teachers collected through the "personal information form" were analyzed with descriptive statistics such as frequency and percentages; authentic leadership and work engagement levels analyzed with average and standard deviation. The relationship between authentic leadership behaviors and work engagement perception levels were analyzed with the Pearson Moments Product Correlation; whether teachers' level of work engagement was predicted by perceived school administrators' authentic leadership behaviors was tested by Multiple Regression analysis. It was evaluated as the findings that the data obtained in the study showed a normal distribution due to the fact that measures of central tendency of 300 data determined to be suitable for statistical analysis were close to each other (mean, median and mode), the skewness and kurtosis coefficients were in the range of -1 to +1, and the values obtained when the skewness and kurtosis coefficients are divided by the standard error of skewness and kurtosis, respectively, between -1.96 and +1.96 (Can, 2016). In the current research, statistical analysis of the data was handled in a model that will reveal the effect of independent variables (school administrators 'authentic leadership behaviors towards teachers) on dependent variable (teachers' work engagement). The evaluation ranges of scales were 1.00 - 1.79 very low; 1.80 - 2.59 low; 2.60 - 3.39 medium; 3.40 - 4.19 high and 4.20 - 5.00 very high.

Findings

The findings obtained as a result of the research are given below in accordance with the purpose of the research.

Findings regarding the Authentic Leadership behavior levels of school administrators perceived by teachers

The findings obtained from the statistics regarding the authentic leadership behavior levels of school administrators perceived by secondary school teachers are included in Table 1.

Table1. Descriptive Statistics of Perception Levels of Teachers' Authentic Leadership Behavior Exhibited by School Administrators

Dimension	n	X̄	Sd	Ranking
Balanced processing		3.91	.044	1
Relational transparency		3.90	.042	2
Internalized morality	300	3.81	.041	3
Self-awareness		3.75	.042	4
General Authentic Leadership		3.84	.036	

As can be seen in Table 1, teachers perceive that school administrators have mostly balanced processing then relational transparency, internalized morality and self-awareness, respectively. In general, and in all dimensions, teachers' perceptions about the authentic leadership of school administrators are high.

Findings regarding teachers' perception of Work Engagement

The findings obtained from statistics regarding secondary school teachers' level of perception of work engagement is given in Table 2.

Table 2. Descriptive Statistics of Secondary School Teachers' Average Scores Regarding the Dimensions of Work Engagement Perception Levels

Dimensions	n	X̄	Sd	Ranking
Dedication		4.38	.030	1
Vigor		4.01	.033	2
Absorption	300	3.96	.041	3
General Work Engagement		4.11	.029	

As can be seen in Table 2, teachers have the highest levels of dedication, then vigor and absorption perception, respectively. Teachers' levels of perception of work engagement is relatively high in both general and all dimensions.

Findings regarding the relationship between teachers' perception level of Authentic Leadership behavior exhibited by school administrators and Work Engagement perception level.

Correlation test results in terms of the relationship between teachers' perception level of authentic leadership behavior exhibited by school administrators and work engagement perception level are given in Table 3.

Table3. The Relationship Between Teachers' Perception Level of Authentic Leadership Behavior Exhibited by School Administrators and Work Engagement Perception Level

Dimensions	RT	IM	BP	SA	GAL	V	D	A	GWE
RT	1								
IM	.53**	1							
BP	.73**	.56**	1						
SA	.67**	.53**	.76**	1					
GAL	.86**	.77**	.90**	.87**	1				
V	.36**	.34**	.36**	.32**	.41**	1			
D	.30**	.30**	.32**	.28**	.35**	.63**	1		
A	.19**	.26**	.25**	.28**	.29**	.56**	.55**	1	
GWE	.32**	.35**	.36**	.34**	.41**	.85**	.83**	.86**	1

**p<.01
Note: **RT**: Relational Transparency **IM**: Internalized Morality **BP**: Balanced Processing **SA**: Self-awareness **GAL**: General Authentic Leadership **V**: Vigor **D**: Dedication **A**: Absorption **GWE**: General Work Engagement

When Table 3 is analyzed, it is seen that the relationship of relational transparency with absorption is low; with internalized morality, self-awareness, vigor, dedication, and general work engagement perception is medium; with balanced processing and general authentic leadership is positively high and significant. It is seen that the

relationship of internalized morality perception with absorption is low; with balanced processing, self-awareness, vigor, dedication, and general work engagement perception is medium; with general authentic leadership perception is positively high and significant. It is found that the relationship of teachers' balanced processing perception is low with the perception of absorption; medium with perceptions of vigor, dedication, and general work engagement; self-awareness and general authentic leadership perceptions are positively high and significant. It is observed that the relationship of teachers' self-awareness perception is low with the perceptions of dedication and absorption; medium with perceptions of vigor and general work engagement; positively high level of significant relationships with the perception of general authentic leadership. It is understood that the relationship of teachers' perception of general authentic leadership is low with the perception of absorption; positively medium and significant with vigor, dedication, and general work engagement. It is detected that the relationship of teachers' perception of vigor is medium with perception of dedication and absorption; positively high with general work engagement. It is seen that the relationship of teachers' perception of dedication is medium with perception of absorption; positively high and significant with general work engagement. It is found that the relationship of teachers' perception of absorption is positively high and significant with general work engagement.

Findings related to the prediction of teachers' Work Engagement by Authentic Leadership behaviors presented by school administrators

The results of multiple regression analysis performed to determine whether teachers' perception of vigor is predicted by the sub-dimensions of authentic leadership are given in Table 4.

Table 4. Multiple Regression Analysis Results Regarding Prediction of Vigor Perception.

Variables	B	Standard Error B	β	t	p	Binary r	Partial r
Constant	2.530	.195	-	12.988	.000	-	-
Relational Transparency	.131	.065	.164	2.018	.044	.361	.117
Internalized Morality	.142	.054	.176	2.651	.008	.343	.153
Balanced Processing	.094	.071	.124	1.330	.184	.358	.077
Self-awareness	.015	.068	.019	.226	.822	.317	.013
R=.413; R^2=.170	$F_{(4; 295)}$ = 15.145; p=000						

As can be seen in Table 4, it was found the linear combination of all dimensions of authentic leadership significantly predicted the perception of vigor [R = .41; R^2 = .17; $F_{(4-295)}$ = 15.145; p <.01]. Independent variables explained 17% of the variance related to perception of vigor. When the t-test results related to the significance of the regression coefficients were examined, it was understood that teachers' perception of relational transparency and internalized morality had significant effects on teachers' perception of vigor, but the perceptions of balanced processing and self-awareness did not have any significant effects on teachers' perception of vigor.

The results of multiple regression analysis performed to determine whether teachers' perception of dedication is predicted by the sub-dimensions of authentic leadership are given in Table 5.

Table 5. Multiple Regression Analysis Results Regarding Prediction of Dedication Perception.

Variables	B	Standard Error B	β	t	p	Binary r	Partial r
Constant	3.240	.181	-	17.868	.000	-	-
Relational Transparency	.065	.061	.089	1.070	.286	.295	.062
Internalized Morality	.112	.050	.152	2.238	.026	.297	.129
Balanced Processing	.101	.066	.148	1.538	.125	.318	.089
Self-awareness	.019	.063	.026	.296	.767	.279	.017
R=.355; R^2=.126	$F_{(4; 295)}$ = 10.615; p=000						

As can be seen in Table 5, it was found that the linear combination of all dimensions of authentic leadership significantly predicted the perception of dedication [R = .36; R^2 = .13; $F_{(4-295)}$ = 10.615; p <.01]. Independent variables explained 13% of the variance related to perception of dedication. When the t test results related to the

significance of the regression coefficients were analyzed, it was seen that perception of internalized morality had a significant effect, but the others didn't have significant effects.

The results of multiple regression analysis performed to determine whether teachers' perception of absorption is predicted by the sub-dimensions of authentic leadership are given in Table 6.

Table 6. Multiple Regression Analysis Results Regarding Prediction of Absorption Perception

Variables	B	Standard Error B	β	t	p	Binary r	Partial r
Constant	2.693	.249	-	10.814	.000	-	-
Relational Transparency	-.087	.083	-.088	-1.042	.298	.185	-.061
Internalized Morality	.161	.069	.162	2.339	.020	.262	.135
Balanced Processing	.080	.090	.086	.881	.379	.254	.051
Self-awareness	.181	.087	.187	2.094	.037	.278	.121
R=.316; R^2=.100	$F_{(4;295)}$ = 8.159; p=000						

In Table 6, it was determined that the linear combination of all dimensions of authentic leadership significantly predicted the perception of absorption [R = .32; R^2 = .10; $F_{(4-295)}$ = 8.159; p<.01]. Independent variables explained 10% of the variance related to the perception of absorption. When the t test results related to the significance of the regression coefficients were analyzed, it was seen that the perceptions of internalized morality and self-awareness had significant effects, but perception of relational transparency and balanced processing didn't have significant effects.

The results of multiple regression analysis performed to determine whether teachers' perception of general work engagement is predicted by the sub-dimensions of authentic leadership are given in Table 7.

Table 7. Multiple Regression Analysis Results Regarding Prediction of General Work Engagement Perception

Variables	B	Standard Error B	β	t	p	Binary r	Partial r
Constant	2.821	.173		16.345	.000	-	-
Relational Transparency	.036	.058	.052	.633	.527	.322	.037
Internalized Morality	.138	.048	.193	2.906	.004	.351	.167
Balanced Processing	.092	.063	.137	1.463	.145	.361	.085
Self-awareness	.072	.060	.102	1.196	.233	.343	.069
R=.411; R^2=.169	$F_{(4;295)}$ = 14.990; p=000						

When Table 7 was analyzed, it can be seen that the linear combination of all dimensions of authentic leadership significantly predicted the perception of general work engagement [R = .41; R^2 = .17; $F_{(4-295)}$ = 14.990; p <.01]. Independent variables explained 17% of the variance related to perception of general work engagement. When the t test results regarding the significance of the regression coefficients were examined, it was seen that perception of internalized morality had a significant effect, but the others didn't have significant effect.

Discussion, Conclusion and Recommendations

According to the teachers, school administrators exhibit the mostly balanced processing behavior, then relational transparency, internalized morality, and lastly self-awareness behaviors and their general authentic leadership behaviors were high. Similarly to the results of this research, there are researches with authentic leadership levels of "high" (Çelik, 2015; Keser, 2013; Kıral, 2018a; Örs, 2015; Ünal, 2015) and there are also researches that have different results with authentic leadership levels of "low" (Kulophas, Hallinger, Ruengtrakul & Wongwanich, 2018); medium (Fox, Gong & Attoh, 2015; Gök, 2015; Secove Lopes, 2013; Shapira-Lishchinsky & Tsemach, 2014; Srivastava & Dhar, 2017) and "very high" (Akıncı, 2016; Özden, 2015). It can be said that school administrators exhibit high level of authentic leadership behaviors and this is perceived by teachers in the same way. This is very important for the school climate. In schools where authentic leadership behaviors are exhibited, there are findings that it will increase teachers' extra role behavior (Srivastava & Dhar, 2017), work engagement (Adil & Kamal, 2016; Kulophas et al., 2018; Seco & Lopes, 2013), organizational commitment (Fox et al., 2015), organizational citizenship behaviors (Özden, 2015; Shapira-Lishchinsky and Tsemach, 2014) and life satisfaction (Gök, 2015). The exhibition of authentic

leadership behaviors in the school increases the school's performance by supporting teachers' belief that their school's success will increase (Lester, Meglino & Korsgaard, 2002; Rego, Vitória, Magalhães, Ribeiro & Cunha, 2013). In fact, it can be stated that this situation will contribute to the success of the school and the quality of education. According to the research findings made in Turkey, it can be said that the authentic leadership behaviors of the administrators in educational institutions are perceived at a high and very high level by teachers, but this perception is generally medium or low in studies conducted abroad.

Teachers had mostly the perception of dedication of work engagement, then vigor and absorption, respectively. It was found that teachers' perception level of dedication is "very high", their perception level of vigor and absorption is "high". It can be said that teachers' perception level of general work engagement is "high". The reason for high level of teachers' work engagement can be related to be provided resources they need at school (Salanova, Agut & Peiró, 2005), positive feedback they receive from the school administrator (Barker & Geurts, 2004), find their work meaningful (May, Gilson & Harter, 2004) and receive social support from their environment (Turgut, 2010). High level of teachers' work engagement may affect the school climate positively. As a matter of fact, it can positively reflect the success of the school by providing teachers performance to increase (Salanova et al., 2005). In studies investigating teachers' work engagement, there are many studies showing that teachers' perception of work engagement is at a "high" level (Adekola, 2010; Atçıoğlu, 2018; Hakanen, Bakker & Schaufeli, 2006; Høigaard, Giske & Sundsli, 2012; Kabar, 2017; Karakaya, 2015; Kavgacı, 2014; Kıral, 2018b; Klassen et al., 2012; Sarath & Manikandan, 2014). According to the research findings made in both Turkey and abroad, it can be said that teachers' perception level of work engagement is high.

It has been determined that there is positively high relationship between general authentic leadership and its dimensions and relatively the highest relationship is with balanced processing. There is positively medium and high-level relationship between authentic leadership dimensions and relatively the highest relationship is between self-awareness and balanced processing, the lowest relationship is between self-awareness and internalized mortality. It has been determined that there is positively high relationship between general work engagement and its dimensions, relatively the highest relationship is with absorption. It has been found out that there is positively medium level relationship between work engagement dimensions, relatively the highest relationship is between dedication and vigor dimensions and the lowest relationship is between absorption and dedication. In general, it has been understood that there is positively medium level relationship between authentic leadership and work engagement. This relationship between authentic leadership and work engagement may be due to the teachers' work engagement because of school administrators' authentic leadership behaviors or school administrators' authentic leadership behaviors because of teachers' work engagement. The environment of trust at school with school administrators' authentic leadership behavior (Clapp-Smith, Vogelgesang & Avey, 2009) or being fair, honest and confident of the school administrator may affect the teachers' work engagement (Yukl, 2018). However, teachers' work engagement can lead school administrators to move towards authentic leadership. In the studies that examine the concepts associated with authentic leadership in educational organizations, the relationship of authentic leadership; is medium with psychological capital and work engagement, Adil & Kamal (2016); is medium with psychological capital Feng-I (2016); is medium with organizational commitment and high with trusting the administrator Fox et al. (2015); is medium with academic optimism and work engagement, Kulophas et al. (2018); is medium with work engagement, Seco & Lopes (2013); is low with organizational citizenship and medium with psychological empowerment, Shapira-Lishchinsky and Tsemach (2014); is high with academic optimism and extra role behavior, Srivastava & Dhar (2017); is medium with emotional intelligence, Akıncı (2016); is medium with life satisfaction and low with the support that perceived from the family, Gök (2015); is medium with psychological capital, Kevser (2013), is medium with organizational conflict management strategy, Örs (2015); is medium with organizational citizenship, Özden (2015); is low with teachers' voice, Ünal (2015). In the studies that examine the concepts associated with work engagement in educational organizations, the relationship of work engagement; is medium with transformational leadership and organizational culture, Arifin, Troena, & Djumahir (2014); is positive medium with organizational commitment and negative medium with burnout, Hakanen et al. (2006); is positive medium with job satisfaction, is negative medium with burnout and intention of leaving the job Høigaard et al. (2012); is high with the job satisfaction, is negative medium with the intention of leaving the job, Klassen et al. (2012); is medium with effectiveness of the school, Atçıoğlu (2018); is medium with emotional intelligence, Kabar (2017); is medium with school climate, Karakaya (2015); is medium with self-recovery power, self-efficacy, self-esteem, autonomy, leader-member interaction and trust in manager Kavgacı (2014); is low between the organizational climate and the perception of organizational support, Köse (2015); is medium between school principal supervision styles, institutional support and parent support, Oymak (2015); is low with servant leadership, Sönmez (2014); is positive medium with spiritual leadership, Tan (2015); is low negative with loneliness in business life, Sezen (2014).

According to the teachers, the general authentic leadership behaviors of school administrators that they perceive significantly affected teachers' perception of vigor. 17% of the change of teachers' perception of vigor was due to the authentic leadership behaviors of school administrators. It was found that the authentic leadership exhibited by the school administrators had significant effects on the perception of teachers' relational transparency and internalized morality, but it didn't have significant effects balanced processing and self-awareness dimensions. It can be said that the vigor dimension, which involves the energetic and vigorous teachers, is influenced by authentic leadership behaviors, which include the sincere attitude of the school administrator, establishing an open and honest communication and adhering to ethical principles. It can be thought to be effective on teachers' feeling vigorous that It can be thought that the school principal's manager to express himself clearly and accurately (Ilies et al., 2005) with the effort of providing confidence, which is the most important goal, as well as possible different results. (Tabak, 2012). As a matter of fact, when the communication within the school is transparent and sincere, the teachers may be willing to make efforts in their work and feel more vigorous when they adhere to the ethical principles in the school. Because employees feel more vigorous when they establish trust-based relationships with their managers and devote themselves more to what they do (Hakanen et al., 2006; May et al., 2004; Schaufeli & Bakker, 2004). In fact, Kumor & Israel stated that employees who have transparent relationships in the organization feel more vigorous (2012).

According to the teachers, the general authentic leadership behaviors of school administrators that they perceive significantly affected teachers' perception of dedication. 13% of the change in teachers' perception of dedication was due to the authentic leadership behaviors of school administrators. It has been determined that internalized mortality dimension of the authentic leadership displayed by the school administrators had a significant effect on teachers' perception of dedication, but the dimensions of relational transparency, balanced processing, and self-awareness didn't have significant effects. It can be said that the dedication dimension, which includes teachers' being proud of their work and caring about their work, is influenced by authentic leadership behaviors that include the consistency of the school administrators' behavior with what they think and adhere to ethical principles. It can be said that it is effective for teachers to feel dedicated to their job and to show their values and ethical standards with their behaviors (Luthans & Avolio, 2003). Actually, the fact that the school administrators do what they say and adheres to ethical principles affects the teachers and makes them feel dedicated to their school. According to Hassan and Ahmed, if the manager behaves regarding to ethical principles, the employees are more dedicated and feel more secure (2012).

According to the teachers, the general authentic leadership behaviors that they perceive in school administrators affected teachers' perception of absorption significantly. 10% of the change in teachers' absorption perceptions was due to the authentic leadership behaviors of school administrators. It has been determined that authentic leadership, self-awareness and internalized morality dimensions exhibited by school administrators had significant effects on teachers' perception of absorption, but relational transparency and balanced processing dimensions didn't have a significant effect. It can be said that the absorption dimension, which includes the focus and happy inoccupation of teachers while doing their job, is influenced by authentic leadership behaviors, which include the awareness of the school administrators' own abilities, weaknesses and strengths, core values, beliefs and desires, and adherence to ethical principles. It can be said that school administrators' being aware of their feelings, motivations and desires and showing their trust in this awareness clearly is effective for teachers to feel focused on their work. (Kernis, 2003). As matter of fact, school administrators' self-awareness behaviors based on the internalized moral values help teachers to concentrate on their own work. In fact, employees feel more concentrated if ethical principles are provided in the organization (May et al., 2004).

According to teachers, the general authentic leadership behaviors that they perceive in school administrators significantly affected teachers' perception of work engagement. 17% of the change in teachers' work engagement was due to the authentic leadership behaviors of school administrators. It was found that the authentic leadership exhibited by the school administrators had a significant effect on the teachers' perception of work engagement, but it didn't have significant effect on the perceptions of transparency relations, balanced processing and self-awareness. It can be said that the work engagement, which reflects the positive and satisfying mood and includes vigor, dedication and absorption towards the work of the teacher; is influenced by authentic leadership behaviors that include the loyalty to the moral values and ethical behaviors of the school administrator, without being affected by the pressures of the society. As a matter of fact, there are studies which show that managers displaying authentic leadership behaviors positively affect employees' work engagement (Bamford, Wong & Laschinger, 2013; Giallonardo, Wong & Iwasiw, 2010).

When the literature is analyzed, there is no study found that authentic leadership affects teachers' work engagement; while there are research findings revealing that spiritual leadership (Tan, 2015); servant leadership (Sönmez, 2014); institutional and parent support at school (Oymak, 2015); organizational climate and

organizational support (Köse, 2015); autonomy, leader-member interaction, self-recovery power and self-efficacy variables (Kavgacı, 2014); the emotional intelligence level of the teachers (Kabar, 2017); organizational culture and leadership (Arifin et al., 2014) and multidimensional organizational identification (Emeksiz, 2015) significantly affect teachers' perceptions of work engagement. In addition, work engagement positively affects school climate (Karakaya, 2015); job satisfaction (Emeksiz, 2015); school effectiveness (Atçıoğlu, 2018) and teacher performance (Arifin et al., 2014).

The following suggestions can be made regarding these results. Practices such as school administrators being aware of information about their feelings, motivations and desires, expressing what they want to say, accepting it when they make mistakes and encouraging everyone to say their thoughts, etc. should be performed by creating working groups and workshops at school and this self-awareness training should be given to school administrators simultaneously with teachers who perceive this. Teachers' level of work engagement was above average. In order to be better, it should be ensured that the school management cooperates with the teachers and the expectations and needs of the teachers are met within the bounds of possibility of the school. In order to increase the concentration component that can be explained as the individual's attention to his work and focus on his work while doing his job, measures such as not dividing the lesson of the teacher, reducing the external stimuli, making the physical conditions more suitable for teaching etc., should be taken. According to the findings of the research, it was concluded that the authentic leadership behaviors of the school administrator had a positive effect on teachers' work engagement. In this context, researches can be conducted to examine the relationship between leadership theories developed with new paradigms and teachers' work engagement; authentic leadership and positive organizational behavior, which is thought to affect the quality of education, and other concepts in which teachers' work engagement is related. In addition, the factors affecting teachers' perception of work engagement and the effects of work engagement on teachers can be determined by conducting a qualitative or mixed research.

References

Adekola, B. (2010). Work engagement among secondary school english teachers in Nigeria, *Pakistan Journal of Socila Sciences*, *7(2)*, 51-55.

Adil, A., and Kamal, A. (2016). Impact of psychological capital and authentic leadership on work engagement and job related affective well-being, *Pakistan Journal of Psychological Research*, *31(1),* 1–21.

Akıncı, H. (2016). *The relationship between characterictics of being authentic leader and emotionel intelligence among women directors who work in primary and secondary schools,* (Unpublished Master Thesis). Sabahattin Zaim University, Istanbul, Turkey.

Arifin, F., Troena, E., and Djumahir, M. (2014). Organizational culture, transformational leadership, work engagement and teacher's performance: Test of a model, *International Journal of Education and Research*, *2(1)*, 1–14.

Atçıoğlu, E. (2018). *The relationship between the effectiveness level of schools and the engagement level of employees,* (Unpublished Master Thesis). Sutcu Imam University, Kahramanmaras, Turkey.

Avolio, B. J., Gardner, W. L., Walumbwa, F. O., Luthans, F., and May, D. R. (2004). Unlocking the mask: A look at the process by which authentic leaders impact follower attitudes and behaviors, *Leadership Quarterly*, *15,* 801–823.

Aydın NED. (2018). *Secondary Schools Classroom - Student - Teacher statistics,* Accessed April 6, 2018 from https://aydin.meb.gov.tr/www/istatistikler/icerik/181.

Balcı, A. (2016). *Sosyal bilimlerde araştırma; yöntem, teknik ve ilkeler, [Research in Social Sciences; methods, techniques and principles],* Pegem Akademi: Ankara, Turkey.

Bamford, M., Wong, C. A., and Laschinger, H. (2013). The influence of authentic leadership and areas of worklife on work engagement of registered nurses. *Journal of Nursing Management, 21(3),* 529-540.

Barker, A. B., and Geurts, S. A. E. (2004). Toward a dual-process model of work-home interference, *Work and Occupations*, *31(3)*, 345–366.

Can, A. (2016). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi, [Quantitative data analysis in the scientific research process with SPSS],* Pegem Akademi: Ankara, Turkey.

Caza, A., Bagozzi, R. P., Woolley, L., Levy, L., and Caza, B. B. (2010). Psychological capital and authentic leadership: Measurement, gender, and cultural extension, *Asia-Pacific Journal of Business Administration*, *2(1)*, 53–70.

Clapp-Smith, R., Vogelgesang, G. R., and Avey, J. B. (2009). Authentic leadership and positive psychological capital: The mediating role of trust at the group level of analysis, *Journal of Leadership and Organizational Studies*, *15(3)*, 227–240.

Çelik, M. (2015). *According to the teacher perceptions authentic leadership behaviors of primary school principals* (The sample of Gaziantep Şehitkamil district), (Unpublished Master Thesis). Zirve University, Gaziantep, Turkey.

Emeksiz, E. (2015). *Multidimensional organizational identification as the predictor of the job satisfaction: The mediating role of the work engagement,* (Unpublished Master Thesis), Bahçeşehir University, İstanbul, Turkey.

Feng-I, F. (2016). School principals' authentic leadership and teachers' psychological capital: Teachers' perspectives, *International Education Studies*, *9(10)*, 245-263.

Fox, J., Gong, T., and Attoh, P. (2015). The impact of principal as authentic leader on teacher trust in the K-12 educational context, *Journal of Leadership Studies*, *8(4)*, 6–18.

Gardner, W. L., Avolio, B. J., Luthans, F., May, D. R., and Walumbwa, F. (2005). "Can you see the real me?" A self-based model of authentic leader and follower development, *Leadership Quarterly*, *16(3)*, 343–372.

George, B. (2003). *Authentic leadership: rediscovering the secrets to creating lasting value,* Jossey-Bass: San Francisco.

Giallonardo, L. M., Wong, C. A., and Iwasiw, C. L. (2010). Authentic leadership of preceptors: predictor of new graduate nurses' work engagement and job satisfaction. *Journal of Nursing Management, 18(8),* 993-1003.

Gök, B. K. (2015). *Analysis of the relationship between the elemantary school administrators' authentic leadership features, life satisfaction and perceived social support from the family*, (Unpublished Master Thesis). Sabahattin Zaim University, Istanbul, Turkey.

Hakanen, J. J., Bakker, A. B., and Schaufeli, W. B. (2006). Burnout and work engagement among teachers, *Journal of School Psychology*, *43(6)*, 495–513.

Hassan, A., and Ahmed, F. (2011). Authentic leadership, trust and work engagement, *World Academy of Science Engineering Technology*, *80*, 750–756.

Høigaard, R., Giske, R., and Sundsli, K. (2012). Newly qualified teachers' work engagement and teacher efficacy influences on job satisfaction, burnout, and the intention to quit, *European Journal of Teacher Education*, *35(3)*, 347–357.

Ilies, R., Morgeson, F. P., and Nahrgang, J. D. (2005). Authentic leadership and eudaemonic well-being: Understanding leader-follower outcomes, *Leadership Quarterly*, *16,* 373–394.

Kabar, M. (2017). *Correlations between work engagement levels and emotional intelligence levels of high school teachers*, (Unpublished Master Thesis). Sabahattin Zaim University, Istanbul, Turkey.

Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work, *Academy of Management Journal*, *33(4)*, 692–724.

Karakaya, H. (2015). *The analysis of the relationship between work engagement and school climate in terms of some variables,* (Unpublished Master Thesis). Zirve University, Gaziantep, Turkey.

Karasar, N. (2016). *Bilimsel araştırma yöntemi, [Scientific research method],* Nobel Yayın Dağıtım: Ankara, Turkey.

Kavgacı, H. (2014). *The relationship of primary education institution teachers' work engagement with personal and organizational variables,* (Unpublished doctoral dissertation),Gazi University, Ankara, Turkey.

Kernis, M. H. (2003). Toward a conceptualization of optimal self-esteem, *Psychological Inquiry*, *14*, 1–26.

Keser, S. (2013). *The comparison of elementary school administrators' authentic leadership and psychological capital features,* (Unpublished Master Thesis).Yıldız Technical University, İstanbul, Turkey.

Kıral, E. (2018a). Psychometric properties of authentic leadership scale in Turkey sample: Application in teachers, In: *II. International Education Research and Teacher Education Congress,*September 13-15, 2018, Kuşadası, Turkey.

Kıral, E. (2018b). Psychometric properties of work engagement scale in Turkey sample: Application in teachers, In: *II. International Education Research and Teacher Education Congress,*September 13-15, 2018, Kuşadası, Turkey.

Kıral, E., and Başaran, R. (2018a). Psychometric properties of trust scale in Turkey sample: Application in teachers, In: *9th International Edactional Managment Congress*, December, 01 – 04, 2018, Antalya, Turkey.

Kıral, E., and Başaran, R. (2018b). The relationship between organizational jealousy and trust, In:*II. International Educational Research and Teacher Training Congress*, September, 13-15 Eylül 2018, Kuşadası, Turkey.

Klassen, R. M., Aldhafri, S., Mansfield, C. F., Purwanto, E., Siu, A. F. Y., Wong, M. W. and Woods-Mcconney, A. (2012). Teachers engagement at work: An international validation study, *Journal of Experimental Education*, 80(4), 317–337.

Köse, A. (2015). *The relationship between job engagement, perceived organizational support and organizational climate (sample of Kahramanmaras),* (Unpublished Doctoral Dissertation). Sütçü İmam University, Kahramanmaraş, Turkey.

Kulophas, D., Hallinger, P., Ruengtrakul, A., and Wongwanich, S. (2018). Exploring the effects of authentic leadership on academic optimism and teacher engagement in Thailand, *International Journal of Educational Management*, *32(1)*, 27–45.

Kumor, A., and Israel, D. (2012). Authentic leadership & Work engagement. *The Indian Journal of Industrial Relations, (47)3*, 498-510.

Leither, M. P., and Bakker, A. B. (2010). Work engagement: Introduction, A. B. Bakker ve M. P. Leiter (Ed.), In: *Work engagement: A handbook of essential theory and research*, (ss. 1–9), Taylor and Francis Group: New York.

Lester, S. W., Meglino, B. M., and Korsgaard, A. M. (2002). The antecedents and consequences of group potency : A longitudinal ınvestigation of newly formed work groups, *Academy of Management Journal*, *45(2)*, 352–368.

Luthans, F., and Avolio, B. J. (2003). Authentic leadership: A positive development approach, *Positive Organizational Behavior*, 241–258.

Macey, W. H., and Schneider, B. (2008). The meaning of employee engagement, *Industrial and Organizational Psychology*, *1(1)*, 3–30.

May, D. R., Gilson, R. L., and Harter, L. M. (2004). The psychological conditions of meaningfulness, safety and availability and the engagement of the human spirit at work, *Journal of Occupational and Organizational Psychology*, *77(1)*, 11–37.

Oymak, Ç. (2015). *The relationship between the secondary school principals' supervision styles, organizational and parental support and the teachers' work engagement: The sample of Bala, Haymana, Şereflikoçhisar,* (Unpublished Master Thesis). Gazi University, Ankara, Turkey.

Örs, H. (2015). *Relationship between the authentic leadership and conflict management strategies for school principals,* (Unpublished Master Thesis). Zaim University, İstanbul, Turkey.

Özden, Ö. (2015). *The relationship between the authentic leadership characteristics of school administrators and organizational citizenship behavior of employees,* (Unpublished Master Thesis). Zirve University,Gaziantep, Turkey.

Rego, A., Vitória, A., Magalhães, A., Ribeiro, N., and Cunha, M. P. (2013). Are authentic leaders associated with more virtuous, committed and potent teams?, *Leadership Quarterly*, *24(1)*, 61–79.

Salanova, M., Agut, S., and Peiró, J. M. (2005). Linking organizational resources and work engagement to employee performance and customer loyalty: The mediation of service climate, *Journal of Applied Psychology*, *90,* 1217–1227.

Sarath, P., and Manikandan, K. (2014). Work engagement and work related wellbing of school teachers, *SELP Journal of Social Science*, *5(22)*, 93–100.

Schaufeli, W. B., and Bakker, A. B. (2003). Utrect work engagement scale: Preliminary manual, *Occupational Health Psychology Unit, Utrect University*, Utrect.

Schaufeli, W. B., and Bakker, A. B. (2004). Job demands, job resources, and their relationship with burnout and engagement: A multi-sample study, *Journal of Organizational Behavior*, *25(3),* 293–315.

Schaufeli, W. B., and Bakker, A. B. (2010). Defining and measuring work engagement: Bringing calrity to the concept, A. B. Bakker ve M. P. Leiter (Ed.), In: *Work Engagement: A Handbook of Essential Theory and Research*, (ss. 10–24), Taylor and Francis Group: New York.

Schaufeli, W. B., Bakker, A. B., and Salanova, M. (2006). The measurement of work engagement with a short questionnaire, *Educational and Psychological Measurement*, *66(4)*, 701–716.

Schaufeli, W. B., Martínez, I. M., Pinto, A. M., Salanova, M., and Barker, A. B. (2002). Burnout and engagement in university students a cross-national study, *Journal of Cross-Cultural Psychology*, *33(5)*, 464–481.

Seco, V., and Lopes, M. P. (2013). Calling for authentic leadership : The Moderator role of calling on the relationship between authentic leadership and work engagement, *Open Journal of Leadership*, *2(4)*, 95–102.

Sezen, G. (2014). *Examination of the relationship between teachers' work engagement and loneliness at work levels,* (Unpublished Master Thesis). Sakarya University, Sakarya, Turkey.

Shapira-Lishchinsky, O., and Tsemach, S. (2014). Psychological empowerment as a mediator between teachers' perceptions of authentic leadership and their withdrawal and citizenship behaviors, *Educational Administration Quarterly*, *50(4)*, 675–712.

Sönmez, B. (2014). *Mediation effect of servant leadership behaviors of primary and secondary school principals on the relationship between organizational justice and work engagement (A case study of Gaziantep – Şehitkamil District),* (Unpublished Master Thesis). Zirve University, Gaziantep, Turkey.

Srivastava, A. P., and Dhar, R. L. (2017). Authentic leadership and extra role behavior: A school based integrated model, *Current Psychology*, *1(14)*.

Tabak, A., Polat, M., Coşar, S., and Türköz, T. (2012). The Authentic Leadership Questionnaire:The Study of Reliability and Validity, *Is, GucIndustrıal Relatıons And Human Resources Journal*,*14(4)*, 89–106.

Tan, İ. H. (2015). *The relationship between the elementary and secondary school principals' spiritual leadership behaviors and the teachers' work engagement,* (Unpublished Master Thesis). Zirve University and Sütçü İmam University, Gaziantep, Turkey.

Tavşancıl, E. (2014). *Tutumların ölçülmesi ve SPSS ile veri analizi*, *[Measuring attitudes and data analysis with SPSS],* Nobel Yayın Dağıtım: Ankara, Turkey.

Taymaz, H. (2009). *Okul yönetimi, [School management],* Pegem Akademi: Ankara, Turkey

Turgut, T. (2010). Çalışmaya tutkunluk, [Work engagement], G. İslamoğlu (Ed.), In: *Kurumlarda iyilik de var [There is also goodness in institutions],* Nobel Yayın Dağıtım: Ankara, Turkey

Ünal, M. (2015). *Authentic leadership behaviors of middle school's directors, the relationship between the voice of teachers and organizational identification (The sample of Gaziantep - Şehitkâmil District),* (Unpublished Master Thesis). Zirve University, Gaziantep, Turkey.

Walumbwa, F. O., Avolio, B. J., Gardner, W. L., Wernsing, T. S., and Peterson, S. J. (2008). Authentic leadership: Development and validation of a theory-based measure. *Journal of management*, 34(1), 89-126.

Yukl, G. (2018). *Leadership in Organizations,* Pearson Education, India.