EURASIAN JOURNAL OF EDUCATIONAL RESEARCH

Bimontly Peer-Reviewed Journal, Year: 17 Issue: 69 / 2017
İki Ayda Bir Yayımlanan Hakemli Dergi, Yıl: 17 Sayı: 69 / 2017

EURASIAN JOURNAL OF EDUCATIONAL RESEARCH

Bimontly Peer-Reviewed Journal, Year: 17 Issue: 69 / 2017
İki Ayda Bir Yayımlanan Hakemli Dergi, Yıl: 17 Sayı: 69 / 2017

SAHİBİ VE YAZI İŞLERİ MÜDÜRÜ / Publishing manager
Özer Daşcan
EDITORIAL OFFICE / Yayın Yönetim
Dilek Ertuğrul
Anı Yayıncılık, Kızılırmak Sokak 10/A
06640 Bakanlıklar Ankara, TURKEY
ejer.editor@gmail.com
Tel: +90.312 425 81 50 pbx Fax: +90.312 425 81 11

Printing Date / Basım Tarihi: 20.05.2017
Printing Address / Matbaa Adresi: Sözkesen Mat. İ.O.S. Mat. Sit. 558 Sk. No: 41 Yenimahalle-Ankara
Yayın Türü: Yaygın Süreli
Cover Design / Kapak Tasarımı: Anı Yayıncılık
Typography / Dizgi: **Göksel Çakır**
The ideas published in the journal belong to the authors.
Dergide yayınlanan yazıların tüm sorumluluğu yazarlarına aittir

Eurasian Journal of Educational Research (ISSN 1302-597X) is a bimonthly (6 issues per year) peer-reviewed journal published by Anı Yayıncılık (ISSN 1302-597X) Anı Yayıncılık tarafından yılda altı kez yayımlanan hakemli bir dergidir.

Eurasian Journal of Educational Research **(EJER)** is abstracted and indexed in;

Emerging Sources Citation Index **(ESCI)**

The Education Resources Information Center (**ERIC**)

Social Scisearch,

Journal Citation Reports/ Social Sciences Editon,

Higher Education Research Data Collection **(HERDC)**,

Educational Research Abstracts **(ERA),**

SCOPUS database,

EBSCO Host database,

ULAKBİM national index.

CONTENTS
İçindekiler

REVIEWERS of the 69th ISSUE
69. Sayı Hakemleri

Ali Ersoy

Arda Arıkan

Asiye Toker Gökçe

Berna Güloğlu

Burak Aydın

Çağlayan Dinçer

Duygu Anıl

Ergül Demir

Evrim Çetinkaya Yıldız

Gökhan Atik

Hakan Atılgan

Halil Yurdugül

Jale Çakıroğlu

Kamile Demir

Kürşad Yılmaz

Laura M. Reid Marsk

Lihua Yao

Meral Hakverdi Can

Müge Uluman

Ramazan Yirci

Refika Olgan

Safiye Bilican

Sedef Canbazoğlu

Sezgi Saraç

Şahin Kesici

Şenel Poyrazlı

Uğur Akın

Vesile Alkan

Eurasian Journal of Educational Research 69 (2017) 1-19

Eurasian Journal of Educational Research
www.ejer.com.tr

Eliminate with Created Argument Environment after Evaluated and Categorized Misconceptions in an Ontological Sense*

Aysegul KINIK TOPALSAN[1] Hale BAYRAM[2]

ARTICLE INFO

Article History:

Received: 13 October 2016

Received in revised form: 12 February 2017

Accepted: 19 April 2017

DOI: http://dx.doi.org/10.14689/ejer.2017.69.1

Keywords

ontological categories, argumentation, force and motion, misconceptions

ABSTRACT

Purpose: This study aimed to ascertain misconceptions of students about basic physical concepts in the "Force and Motion" unit of secondary school seventh class curriculum, to eliminate the misconceptions with created argument environment and traditional approaches after evaluated, and categorize these misconceptions in an ontological sense. **Research Methods:** Considered fundamental problems and sub-problems for which answers are sought. A semi-experimental model with pre-test and post-test control groups was utilized. Detected ontological categories were analyzed and discussed for each question located in the "Force and Motion" concept test.

Findings: Before and after applications after physical concepts about the Force and Motion unit were examined and categorized ontologically. 301 examined misconceptions from students in the experimental group arose from placement in the higher category. 150 misconceptions that arose from the placement in the lateral category were identified before application. 252 misconceptions of the 301 misconceptions (83.72%) that arose from placement in the higher category were corrected due to argumentation works that were executed. 128 misconceptions out of 150 (85.33%) misconceptions that were placed in the lateral category were corrected after an argumentation analysis. **Implications for Research and Practice**: Studies such as determination, evaluation, and correction of misconceptions should be performed by using ontological categories. This study shows that the implementation of argumentation works is more successful in the elimination of misconceptions placed in constraint-based interaction-natural, random-event, and matter categories. In this way, learning environments can be designed to be more efficient and infallible.

* This article was presented at the Third International Eurasian Educational Research Congress (EJER, 31-3 June 2016).

[1] Corresponding author: Dr. Istanbul Aydin University, Faculty of Education, e-mail: aysegulkinik@aydin.edu.tr

[2] Prof.Dr. Marmara University, Faculty of Education, e-mail: haleb@marmara.edu.tr

Introduction

Misconceptions are generally defined as concepts structured inadequately or incorrectly by students, apart from concepts scientifically accepted as true, and that were acquired by students by the end of the educational process (Nakhleh, 1992). In the 2000s, misconceptions were treated in philosophical terms, as philosophically-based definitions were being introduced into the literature. Chi and Roscoe (2002) treated misconceptions on an ontological basis, arguing that all concepts and ideas belong to certain ontological categories.

As a term, ontology is defined as the "science of being." One of the simplest definitions of a possible ontology may be "a controllable lexicon." Ontology is concerned with beings and the basic categories to which beings belong (Chi, 2001). An ontological property is one that a being potentially possesses due to the ontological category to which it belongs (Chi, 1997). Concepts are placed in ontological categories according to the ontological properties they possess. The three primarily utilized ontological categories are matter, process, and mental states. Students produce misconceptions when they, for example, place a concept that belongs to the process category into the matter category. Therefore, one should determine the categories in which to place concepts, and, in the case of misplacements, ensure that the concepts in question are re-located to the correct categories by using various educational methods and techniques. This is crucial in order to identify the roots of misconceptions and, thus, eliminate them.

An individual may sometimes perceive concepts differently, apart from scientific situations, and may place them in different categories. Usually, when students do not understand a basic physical concept and place it in the categories that already exist in their mind, they struggle to understand higher level and more complex concepts and learn permanently. Therefore, students should establish bridges in a meaningful way between their intuitive thoughts regarding the events they witness in their lives and the physical topics and concepts (Ayvacı and Devecioğlu, 2002). To meaningfully establish such bridges, students' misconceptions should first be determined, and then eliminated (Ayvacı and Devecioğlu, 2002; Yağbasan and Gülçiçek, 2003; Turgut et al. 2011).

The first stage in eliminating and correcting misconceptions, as well as in planning the relevant teaching process, is to determine conceptual misplacements, assist students to test their own conceptualizations and gain awareness of the possible misconceptions, and enable the learners to acquire the ability of higher level reasoning. The second step is to use, in the teaching process, methods and techniques that would enable students to place concepts in the right categories. In this regard, it is suggested that integrating argumentation, a reasoning activity, into the class environment may be an effective strategy to direct students to conceptual changes (Niaz, Aguilera, Maza and Liendo, 2002; Nussbaum and Sinatra, 2003).

Toulmin's argumentation model is composed of an assertion, the proofs to support the assertion, the reasons that indicate the relationship between the proofs and the assertion, the supportive pre-information that strengthens the reasons, the

qualifiers (restrictions), and finally, the refutations that indicate the situations in which the assertion is invalid (Erduran, Simon and Osborne, 2004). Driver, Newton, and Osborne (2000) suggested that the argumentation-based teaching activities in science classes possess three significant effects: improving conceptual comprehension, research skills, and questioning the validity of scientific knowledge. It is observed that argumentation is quite effective in solving problems in science education. Thus, this study tried to realize conceptual changes with the help of argumentative contexts that included pre-determined misconceptions.

Relevant studies are limited to the identification of misconceptions or the effects of various methods in eliminating misconceptions. It has been observed that the national and international literature include only a few studies that treat misconceptions in ontological terms (Soman, 2000; Özalp, 2008; Özalp and Kahveci, 2011; Şen and Yılmaz, 2012; Sanmarti, Izquierdo and Watson, 1995; Watson, Prieto and Dillon, 1997). These studies only treat misconceptions in ontological terms, but do not propose active methods to eliminate the misconceptions of the identified categories. This study is quite significant in that it ontologically evaluates the concepts regarding the subject of "Force and Motion" and determines the effects of employed argumentations in eliminating misconceptions caused by types of ontological categorizations. As the first study in this capacity, this paper will guide researchers in the subject of eliminating misconceptions that are ontologically determined. This study has treated, in ontological terms, students' misconceptions regarding basic physical concepts that are within the subject of "Force and Motion," such as force, frictional force, work, conservation of energy, mechanical energy, kinetical energy, potential energy, and energy stored in springs. After students' misconceptions in identified subjects were ontologically evaluated and categorized, contexts of argumentations were formed to eliminate the identified misconceptions. Argumentation activities were formulated and implemented based on students' existing misconceptions. This forms the basic stage of this study. Additionally, the extent to which the employed argumentation settings affect the levels of students' use of scientific process skills and increase their achievements at the levels of knowledge and comprehension was revealed.

Method

Research Design

Considering the study's aim, main problem, and sub-questions, it can be said that I used a semi-experimental method with a pre-test and post-test control group design. Dependent variables of the implemented experimental pattern were academic achievement, scientific process skills, and learning concepts. The following pre-tests and post-tests were administered to all participant students in order to determine the effects of two different teaching methods: the Force and Motion Subject Academic Achievement Test (FMAAT) to determine the effect on students' academic achievements, the Force and Motion Subject Concept Test (FMCT) to determine the effect on students' learning concepts, and the Scientific Process Skills

Test (SPST) to determine the effect on students' scientific process skills. The research pattern is indicated in Table 1.

Table 1

Research Design

Group	*Teaching Method Used*	*Pre-tests*	*Post-tests*
Control Group	Traditional Method	SPST, FMAAT, FMCT (n=35)	SPST, FMAAT, FMCT (n=35)
Experimental Group	Argumentation	SPST, FMAAT, FMCT (n=35)	SPST, FMAAT, FMCT (n=35)

Research Sample

The working group of this study was composed of students who attended a foundation university in Istanbul, in the 2012-2013 academic year, in two distinct groups. The working group consisted of 70 teacher candidates (2nd grade, primary school teaching) as 60 female and 10 male students. Working groups were determined based on the results of the pre-tests, and they were placed in two equal size groups with 35 students.

Research Instruments and Procedures

Force and Motion Subject Academic Achievement Test. The Force and Motion Subject Academic Achievement Test was composed of 25 questions to reliably determine whether there were any differences in students' learning levels regarding the Force and Motion subject. In preparing the test, six questions that exhibited a least distinguishing index were determined. These questions were later excluded from the Force and Motion Subject Academic Achievement Test and the investigation continued with the remaining 19 questions. Distinguishing indexes of these 19 questions differed from 0.30 to 0.50. Subsequently, in order to determine the reliability of the Force and Motion Subject Academic Achievement Test that consisted of 25 questions, the Cronbach's Alpha coefficient was calculated, which was found as 0.680. The KR-20 coefficient was also found as 0.833.

Force and Motion Subject Concept Test. The Force and Motion Subject had 17 questions in its finalized version, and seven of the test articles were adopted from the test developed by Ulu (2011) while the researcher formulized the remaining 10 questions by literature survey. To formulize the questions, research was first executed on both domestic and foreign studies on the misconceptions about the concepts of force, frictional force, work, conservation of energy, mechanical energy, kinetical energy, potential energy, and energy stored in springs. The questions were formulized to reveal the cited misconceptions and the further misconceptions based on them. The ontological categories were held as the basis of the question design. The

Cronbach's Alpha value was found as 0.710 and the KR-20 coefficient as 0.704 for the Force and Motion Subject Concept Test.

Scientific Process Skills Test. The Scientific Process Skills Test was applied to the experimental and control groups. The Turkish translation and adaptation of the test was executed by Geban, Aşkar, and Özkan (1992). The multiple-choice test, consisting of 36 questions, measures the following skills: defining variables, formulating hypotheses, operational defining, research design, and data analyses. In his research with 7th grades, Aydoğdu (2006) examined the Scientific Process Skills Test developed by Geban, Aşkar, and Özkan (1992) and excluded some of the articles as they were not compatible with the 8th grade cognitive development level, reducing the number of the articles to 28. For a pilot study, the test with 28 questions was administered to 336 randomly selected students attending nine different primary schools. After the application, the distinguishing indices, difficulties of the articles, and the reliability coefficient of the test were calculated. After the calculation, the questions with a distinguishing index below 0.30 were excluded from the test. Thus, a test with 25 multiple-choice questions and with a reliability of 0.81 was acquired to measure scientific process skills.

Data Analysis

Kolmogorov-Smirnov goodness-of-fit test was used to determine whether the points of the Force and Motion Subject Academic Achievement Test, the Force and Motion Subject Concept Test, and the Scientific Process Skills Test demonstrated normal distribution.

To determine whether there were any differences in subject-related learning levels and concept learning levels in the experimental and control groups before and after the Force and Motion Subject, FMAAT was applied to both groups as a pre-test, and independent group t-test was used to analyze the data obtained.

The answers given to FMCT were qualitatively analyzed. In this analysis, the misconceptions determined in each question of the test were ontologically categorized. Then, ontological category maps were formed, in which the right and wrong ontological categorizations were analyzed, after the pre-test and post-test, by providing frequencies and percentages.

To determine whether there were any differences among the pre-study scientific process skills on the part of the experimental and control groups, SPST was applied to both groups as a pre-test and post-test, and the independent group t-test was used to analyze the total points obtained. To determine whether there were statistically significant differences between the points that the students of both groups obtained in the sub-dimensions of SPST before and after the study, the independent group T-test was applied to the points obtained from the dimensions of defining variables, operational defining, and formulating hypotheses. The Mann-Whitney U Test was applied to the points obtained from the dimensions of research design, and data analyses.

Results

In this part, the findings are examined in two sections. First, the findings will address determining the misconceptions of Force and Motion and the efficiency of argumentations and traditional methods used to eliminate these misconceptions. Then, the findings about argumentations, traditional settings, and the teaching process are treated in terms of their efficiency to eliminate misconceptions that resulted from certain misplacements of concepts in ontological categories.

Having compared the points that the students of the control and experimental groups obtained from the FMAAT pre-test and post-test with the independent t-test, the p value of the pre-test was found as 0.876 ($p>0.05$), and the p value of the post-test as 0.012 ($p<0.05$).

Having compared the points that the students of the control and experimental groups obtained from the SPS pre-test and post-test with the independent t-test, the p value of the posttest was found as 0.000 whereas it was 0.890 for the pre-test. A significant difference was found in favor of the experimental group.

No statistically significant differences were found between the points that the students of the experiment and control groups obtained from the sub-dimensions of the SPST ($p>0.05$). Therefore, there were no differences observed between the scientific process skills that the experimental and control groups had at the beginning of the study. A significant difference was found in the SPST sub-dimensions for the experiment group in the results of the post-test.

No statistically significant differences were found between the points that the control and experimental groups obtained from the FMCT pre-test ($p=0.51$). However, a statistically significant difference was found between the points that the control and experiment groups obtained from the FMCT post-test ($p=0.00$). This result was interpreted as that the applied use of argumentation settings in the lab environment more greatly improved the students' level of learning concepts compared to the traditional understanding in which students carry out the instructions given to them during the lab practices.

Ontological Examination of the Force and Motion Subject Concept Test Misconceptions. The percentages of the students' answers to each question of the FMCT distributed by ontological categories were determined and presented in tables. In addition, toward the aim of the study, the students' misconceptions were examined by dividing them into categories. This process was applied elaborately to the 17 questions of the test. In this section, only the analysis of the first test question is included as an example.

Figure 1

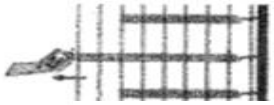


Figure 2

Figure 1. Force and Motion subject concept test, question one.

Question 1. A student compresses a spring in Figure 1 by 10 cm and releases it after a while. Then the student stretches the same spring by 10 cm, as in Figure 2, and releases it after a while. Which of the following judgments is correct?

A) The amount of energy stored in the spring is the same in both cases.

B) No energy is stored in the spring in both cases.

C) More energy is stored in the case given in Figure 1.

D) More energy is stored in the case given in Figure 2.

Which of the following is the reason of your answer in this question?

A) If a spring is compressed or stretched by the same amount, it will have the same amount of energy in both cases.

B) Work is required to store potential energy in the spring. Thus, no energy is stored in the spring in either case.

C) When a spring is compressed and stretched by the same length, it does not retain the same amount of energy. More energy is stored in the compressed spring.

D) When a spring is compressed and stretched by the same length, it does not retain the same amount of energy. More energy is stored in the stretched spring.

E) In my opinion, …….

This question aims to draw attention to the topic of springs and their resilience in the subject of Force and Motion. It facilitated the questioning of the amount of energy stored in springs when compressed or stretched. The students' levels of comprehension regarding the energy stored in springs were evaluated by the evaluation criteria. Their misconceptions regarding the concept were examined in ontological terms, and the sources of misconceptions were determined on an ontological basis. Based on the obtained data, the levels of comprehension on the part of the students of the experimental and control groups in the pre-test and post-test are presented in Table 2 and Table 3, respectively.

Table 2

Students' Levels of Comprehension Regarding the First Question of the Force and Motion Subject Concept Pre-test

Comprehension Level	Experimental group		Control group	
	f	%	f	%
Thorough Comprehension	10	28.57	12	34.29
Misconception	20	57.14	22	62.88
Lack of Comprehension	5	14.29	1	2.86

Table 3

Students' Levels of Comprehension regarding the First Question of the Force and Motion Subject Concept Pre-test

Comprehension Level	Experimental group		Control group	
	F	%	f	%
Thorough Comprehension	27	77.14	24	68.57
Misconception	6	17.14	11	31.43
Lack of Comprehension	2	5.71	0	0

Table 2 indicates that, in the pre-test held before the application, 28.57% of the experimental group and 34.29% of the control group thoroughly comprehended the given concept about the amount of energy stored in springs. It also shows that 57.14% of the experimental group and 62.88% of the control group had a misconception about the given concept, and that 14.29% of the experimental group and 2.86% of the control group did not comprehend the concept investigated in the first question. When we examined the comprehension levels of the students in the same groups regarding the amount of energy stored in springs, we saw that the rate of students with thorough comprehension raised to 77.14% in the experimental group and to 68.57% in the control group, while the percentage of students with misconceptions dropped in a general sense. The table shows that the percentage of the students who could not comprehend the question in the experimental group dropped to 5.71%, while there were no such students in the control group. Another operation performed in the analysis of the first question of the FMCT was to determine the students' misconceptions. Table 4 indicates the misconceptions that the students exhibited in the first question of the FMCT pre-test, and Table 5 indicates the misconceptions that the students exhibited in the first question of the FMCT post-test.

Table 4

Misconceptions in the Answers that the Students Provided for the First Question of the Force and Motion Subject Concept Pre-test

Misconception	Experimental group		Control group	
	f	%	f	%
When a spring is compressed and stretched by the same length, it does not retain the same amount of energy. More energy is stored in the stretched spring.	10	28.57	11	31.43
When a spring is compressed and stretched by the same length, it does not retain the same amount of energy. More energy is stored in the compressed spring.	8	22.88	5	14.29
Work is required to store potential energy in the spring. Thus, no energy is stored in the spring in either case.	2	5.71	6	17.14

Table 5

Misconceptions in the Answers that the Students Provided for the First Question of the Force and Motion Subject Concept Post-test

Misconception	Experimental group		Control group	
	f	%	f	%
When a spring is compressed and stretched by the same length, it does not retain the same amount of energy. More energy is stored in the stretched spring.	2	5.71	3	8.57
When a spring is compressed and stretched by the same length, it does not retain the same amount of energy. More energy is stored in the compressed spring.	3	8.57	6	17.14
Work is required to store potential energy in the spring. Thus, no energy is stored in the spring in either case.	1	2.86	2	5.71

The last operation performed in the analysis of the first question of the FMCT based on the examination of the data given in Table 4 and Table 5 was to examine, in ontological terms, the misconceptions determined in the pre-tests and post-tests. Figure 2 indicates the ontological examination of the misconceptions of the students of the experimental group in the FMCT pre-test and post-test, while Figure 3 displays those of the control group.

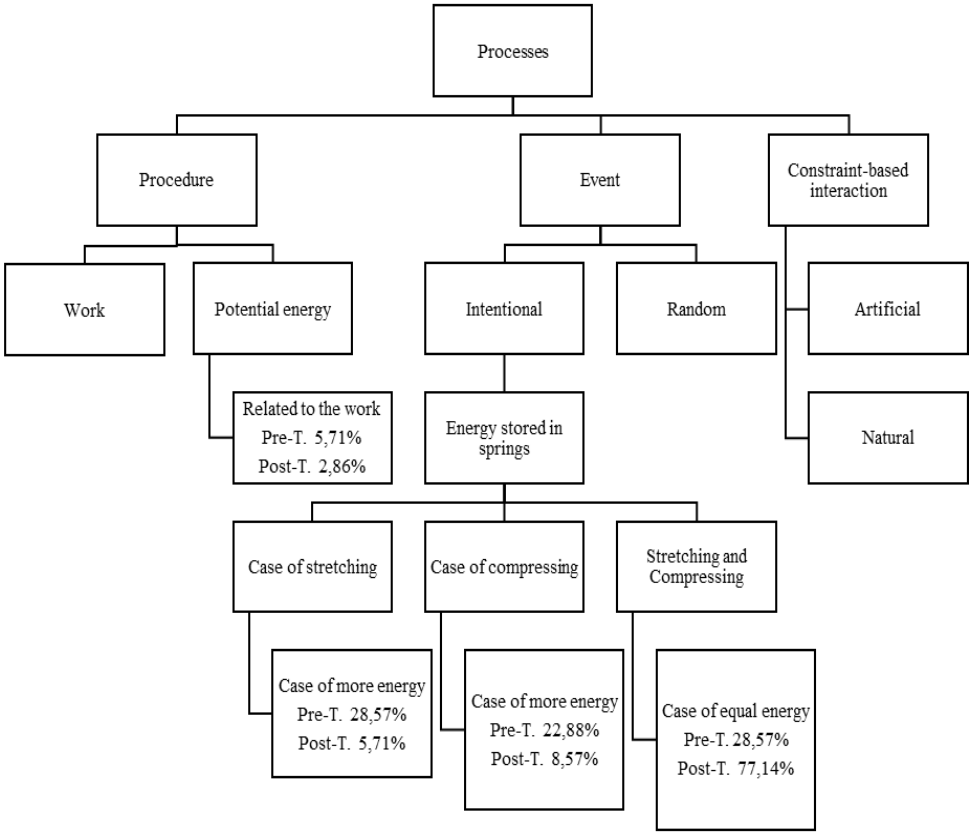


Figure 2. Ontological examination of the misconceptions of the students in the experimental group for the first question of the force and motion subject concept test

Figure 2 shows that the students who correctly answered the first question of the FMCT were those who placed the concept of the amount of energy stored in springs in the category of intentional event, a sub-category of the process category. The rate of these students was 28.57% in the pre-test, while it raised to 77.14% in the post-test.

In this study, we found two different sources, on an ontological basis, for the misconceptions about the energy stored in springs. One concerned the misconceptions that resulted from placing the concept about the amount of energy stored in springs in the categories of "case of compressing" and "case of stretching" that are among the side categories of the mentioned concept. The other concerned the misconceptions that resulted from placing the same concept in the operation category, one of the sub-categories of the process category.

In the misconception that resulted from placing the process category in the operation category, one of the sub-categories of the former, the students stated that it was required to execute a numerical calculatation on the spring for any potential energy to be stored in the stretched or compressed spring.

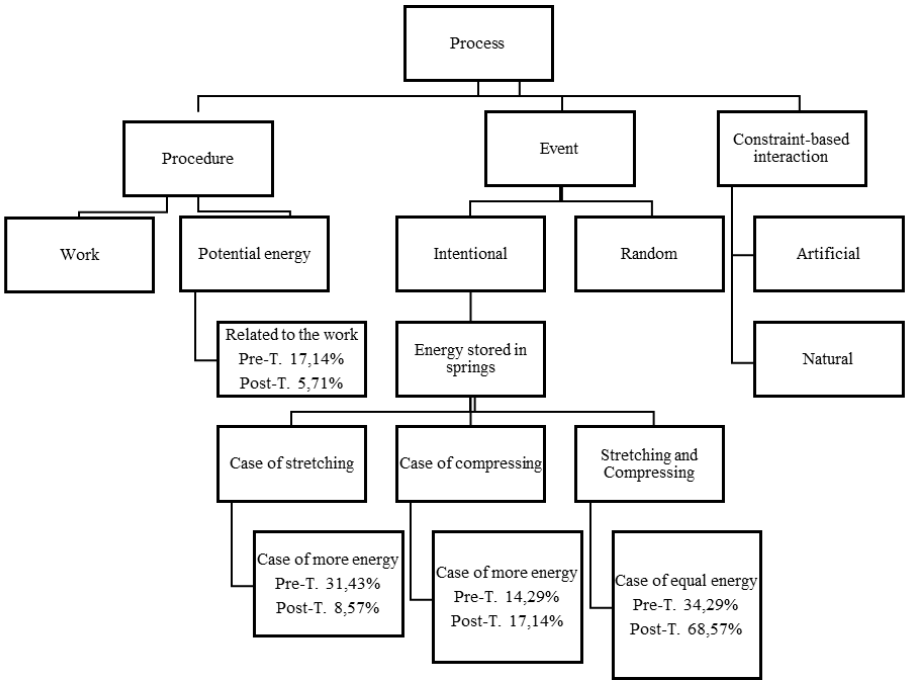


Figure 3. Ontological examination of the misconceptions of the students in the control group for the first question of the force and motion subject concept test

Figure 3 shows that the students who correctly answered the first question of the FMCT were those who placed the concept of the energy stored in springs in the category of intentional event, a sub-category of the process category. The rate of these the students of the control group was 34.29% in the pre-test, while it raised to 68.57% in the post-test. In this study, we found two different sources, on an ontological basis, for the misconceptions about the energy stored in springs. One concerned the misconceptions that resulted from placing the concept about the amount of energy stored in springs in the categories of "case of compressing" and "case of stretching," which are among the side categories of the mentioned concept.

The other concerned the misconceptions that resulted from placing the same concept in the operation category, a sub-category of the process category. In the misconception that resulted from placing the process category in the operation category, a sub-category of the former, the students stated that it was required to execute a numerical calculation on the spring for any potential energy to be stored in the stretched or compressed spring.

Discussion and Conclusion

After comparing the results of the scientific process skills test administered to the control group and the experimental group, a significant difference was found, in terms of the total points, in favor of the experimental group. After examining the results in terms of the sub-dimensions in the scientific process skills, a significant difference was found in all dimensions in favor of the experimental group. We may conclude, in the light of these findings, that the argumentations developed for the questioned concepts are more effective, compared to the activities performed in traditional ways, to enable students to improve the scientific process skills of defining variables, formulating hypotheses, operational defining, research design, data analyses. This conclusion supports the argument that, if students have experiences about scientific processes, these skills will be improved (NRC, 2000).

After comparing the results of the FMAAT post-test administered to the control group and the experimental group, a statistically significant difference was found in favor of the experimental group. Based on this finding, we may conclude that the argumentations developed for questioned concepts are more effective, compared to the activities performed in traditional ways, to increase students' academic achievements. Argumentations, which may easily be incorporated in activities performed in a lab setting, assist students in all areas and create different points of view. In this study, argumentations were used as course material, and, as they enabled the students to take all responsibility for learning, they increased the students' will to learn, allowing them to better internalize the concepts in question. This study indicates the impact of the class for which the number and content of the argumentation were arranged in line with the course of teaching. Studies on argumentations show that students' achievements increase in time (Akkuş et al., 2007). This situation is comparable with the data in the literature.

After comparing the results of the FMCT post-test administered to the control group and the experimental group, a significant difference was found in favor of the experimental group. Based on this finding, we may conclude that the argumentations developed for basic physical concepts are more effective, compared to the activities and experiments performed in traditional ways, to increase students' levels of learning concepts. This conclusion complies with the findings of Kaya (2005); Clark and Sampson (2007); De Vries, Lund, and Baker (2002); Driver et al. (2000); Duschl and Osborne (2002); Niaz et al. (2002); Uluçınar Sağır (2008); Zohar and Nemet (2002); Demirci (2008); Dole and Sinatra (1998); and Nussbaum and Sinatra (2003). Conducted at different levels of primary, secondary, and higher education, these

studies show that course content developed with argumentations increase students' levels of learning concepts more than traditional methods. The most significant suggestion of these studies seems to be that, for conceptual change to be ensured, a convenient learning setting should be prepared in which new concepts can be compared with students' existing concepts, including the formation of deep reflections, relevant argumentations, and counter-argumentations.

In this study, before the application, the students of the experimental group had 301 misconceptions resulting from placement in an upper category and 150 misconceptions resulting from placement in a side category. Out of the 301 misconceptions resulting from placement in an upper category, 252 (83.72%) were eliminated. In addition, out of 150 misconceptions resulting from placing in a side category, 128 (85.33%) were eliminated. This situation reveals the impact of argumentation settings used in the teaching process. The misconceptions that appeared in the upper and side categories were largely eliminated. After examining the upper ontological and side categories, it was observed that the misconceptions placed in the side categories were more frequently eliminated. It was also found that the students acquired new misconceptions because of the argumentation settings and lectures. In this study, three new misconceptions were detected. Even though this kind of study might have been conducted carefully, it may not prevent students from creating new misconceptions. In his doctoral dissertation, Çelik (2010) argued that argumentations may result in similar cases of misconceptions. For conceptual change to be ensured, a convenient learning setting should be prepared in which new concepts can be compared with students' existing concepts, along with the formation of deep reflection, relevant argumentations and counter-argumentations (Dole and Sinatra, 1998; Nussbaum and Sinatra, 2003). The approach based on scientific argumentation may provide a teaching setting convenient for conceptual comprehension and conceptual change, but conceptual confusion may take place during the process, as well.

This study indicates that most of the misconceptions that resulted from the misplacement of the concepts of the sub-categories of the process ontological category, namely those of procedure, intentional event, constraint-based natural interaction, constraint-based artificial interaction, and random event. The misconceptions with the highest rate of occurrence are those that resulted from placement in the categories of procedure and intentional event, which are among the sub-categories of the process category. Slotta and Chi (2006) mentioned how physicists might eliminate strong and stable misconceptions by ontological training and instruct about the categories in which basic physical topics might emerge more intensely. It is seen, in the cited study, that the detected misconceptions, especially regarding the topic of electricity, were concentrated under the process category, and that the concepts were placed in the sub-categories of the process category in several ways due to the concrete examples given by teachers. Similarly, this study has shown that, before the argumentations, the students generally placed the basic physical concepts in question in the sub-categories of the process category, according to their levels of readiness.

It can be said that many of the misconceptions in the control group result from the misplacement of the concepts of the ontological category of the process in its sub-categories, namely those of procedure, intentional event, constraint-based natural interaction, constraint-based artificial interaction, and random event. Of the students in the control group, we determined 318 misconceptions resulted from placement in an upper category, and 131 misconceptions resulted from placement in a side category, all before the application. Out of the 318 misconceptions that resulted from placement in an upper category, 122 (38.36%) were eliminated, and out of the 131 misconceptions that resulted from placement in a side category, 59 (45.03%) were eliminated. This shows that traditional activities performed in lab settings are more effective in eliminating misconceptions that resulted from placement in a side category than from placement in an upper category. In addition, it is observed in the results that traditional activities might cause new misconceptions to be formulated by the students. At the end of the study, 31 new misconceptions were detected.

References

Akkus, R., Gunel, M. & Hand, B. (2007). Comparing an inquiry based approach known as the science writing heuristic to traditional science teaching practices: Are there differences?. *International Journal of Science Education,* 29 (14), 1745-1765.

Aydogdu, B. (2006). *Ilkogretim fen ve teknoloji dersinde bilimsel surec becerilerini etkileyen degiskenlerin incelenmesi* [Investigation of variables affecting scientific surrogacy skills in primary science and technology course]. Yuksek Lisans Tezi, Dokuz Eylul Universitesi Egitim Bilimleri Enstitusu, Izmir.

Ayvacı, H.S & Devecioglu, Y. (2002). Kavram haritasinin fen bilgisi basarısına etkisi [The Impact of the concept map on scientific success]. V. Ulusal Fen Bilimleri ve Matematik Egitimi Kongresi, Ankara

Chi, M.T.H. (1997). Creativity: Shifting across ontological categories flexibly. Ward,T. B., Smith, S. M & Vaid, J. (Eds.), *Conceptual structures and processes: Emergence, discovery and change*, 209-234. American Psychological Association, Washington.

Chi, Y. (2001). Ontology-based curriculum content sequencing system with semantic rules. *Expert Systems with Applications*, 36(4), 7838–7847.

Chi, M. T. H. & Roscoe, R. D. (2002). The processes and challenges of conceptual change. Limon, M. & Mason, L. (Eds), *Reconsidering conceptual change: Issues in theory and practice,* 3-27. Kluwer Academic Publishers.

Clark, D. B. & Sampson, V. (2007). Personally-seeded discussions to scaffold online argumentation. *International Journal of Science Education*, 29 (3), 253-277.

Celik, A.Y. (2010). *Bilimsel tartisma (argumantasyon) esasli ogretim yaklasiminin lise ogrencilerinin kavramsal anlamalari, kimya dersine karsı tutumlari, tartısma isteklilikleri ve kalitesi uzerine etkisinin incelenmesi* [Investigation of effect on Conceptual meanings, attitudes towards chemistry lessons, willingness to discuss and quality with scientific discussion-based approach to teaching].

Doktora Tezi, Gazi Universitesi Egitim Bilimleri Enstitusu Ortaogretim Fen ve Matematik Alanlar Egitimi Anabilim Dali, 122-140

De Vries, E., Lund, K.& Baker, M.J. (2002). Computer-mediated epistemic dialogue: Explanation and argumentation as vehicles for understanding scientific notions. *The Journal of the Learning Sciences*, 11(1), 63-103.

Demirci, N. (2008). *Toulmin'in bilimsel tartisma modeli odakli egitimin kimya ogretmen adaylarinin temel kimya konularini anlama ve tartısma seviyeleri uzerine etkisi* [The impact of Toulmin's focus on the scientific debate model of chemistry teacher candidates on understanding and discussing basic chemistry subjects]. Yuksek Lisans Tezi, Gazi Universitesi, Egitim Bilimleri Enstitusu, Ankara.

Dole, J. A. & Sinatra, G. M. (1998). Reconceptualizing change in the cognitive construction of knowledge. *Educational Psychologist*, 33, 109–128

Driver, R., Newton, P. & Osborne J. (2000). Establishing the norms of scientific argumentation in classrooms. *Inc. Sci. Ed*, 84, 287-312.

Duschl, R. A. & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, 38, 39-72.

Erduran, E., Simon, S. & Osborne, J. (2004). Tapping into argumentation: developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education*, 88 (6), 915-933.

Geban, O., Askar, P. & Ozkan, I. (1992). Effects of computer simulated experiments and problem solving approaches on high school students. *Journal of Educational Research,* 86, 5 - 10.

Nakhleh, M. (1992). Why some students don't learn Chemistry: Chemical misconceptions. *Journal of Chemical Education*, 69(3), 191-196.

National Research Council (NRC). (2000). *Inquiry and the National Science Education Standards.* National Academy Press, Washington, DC

Niaz, M., Aguilera, D., Maza, A. & Liendo, G. (2002). Arguments, contradictions, resistances, and conceptual change in students' understanding of atomic structure. *Science Education*, 86, 505-525.

Nussbaum, E.M. & Sinatra, G. M. (2003). Argument and conceptual engagement. *Contemporary Educational Psychology*, 28, 384-395.

Özalp, D. (2008). *Ilkogretim ve ortaogretim ogrencilerinin maddenin tanecikli yapisi konusundaki kavram yanilgilarinin ontoloji temelinde belirlenmesi* [Diagnostic Assessment of Student Misconceptions about the Particulate Nature of Matter from Ontological Perspective].Yuksek Lisans Tezi, Marmara Universitesi, Istanbul

Özalp, D. & Kahveci, A. (2011). Maddenin tanecikli yapisi ile ilgili iki asamalı taniliyici sorularin ontoloji temelinde gelistirilmesi [Development of two-tier diagnostic items based on ontology in the topic of the particulate nature of matter]. *Milli Egitim Dergisi*, 191, 135–156.

Sanmarti, N., Izquierdo, M. & Watson, J.R. (1995). The substantialisation of properties in pupils' thinking and in the history of science. *Science and Education*, 4, 349-369.

Slotta, J.D. & Chi, M.T.H. (2006). Helping students understand challenging topics in science through ontology training. *Cognition and Instruction*, 24 (2), 261–289.

Soman, S. A. (2000). *Ontological caterization in chemistry: A basis for conceptualchange in chemistry*. Doctorate Thesis, Purdue University, West Lafayette, USA.

Turgut, Ü., Gürbüz, F., Turgut, G. & Açışlı, S. (2011). *Lise 2. sinif fen subesi ogrencilerinin "Kuvvet ve Hareket" konusundaki kavram yanılgıları* [Second-year science sub-school students in high school's misconceptions about force and movement.]. Trakya Universitesi Egitim Fakultesi Dergisi, 1(1), 71-85.

Uluçınar Sağır, Ş. (2008). *Fen bilgisi dersinde bilimsel tartisma odakli ogretimin etkililiginin incelenmesi* [Investigation of the effectiveness of scientific discussion-oriented teaching in science education]. Doktora Tezi, Gazi Universitesi Egitim Bilimleri Enstitusu, Ankara.

Ulu, C. (2011). *Argumantasyon tabanli bilim ogrenme yaklasimina dayali laboratuvar etkinliklerinin 7. sınıf ögrencilerinin kavram ogrenmelerine etkisi: yasamımızdaki elektrik unitesi* [Effects of laboratory activities through the argumentation based inquiry approach on 7th grade students' conceptual learning: electricity in our daily life unit]. Doktora Tezi, Marmara Universitesi Egitim Bilimleri Enstitusu, İstanbul

Yagbasan, R. & Gülcicek, C. (2003). Fen ogretiminde kavram yanilgilarinin karakteristiklerinin tanimlanmasi [Defining the characteristics of conceptual aspects in science teaching]. *Pamukkale Universitesi Egitim Fakultesi Dergisi*, 1 (13), 102-119.

Yılmaz, A.& Sen S. (2012). Erime ve cozunmeyle ilgili kavram yanilgilarinin ontoloji temelinde incelenmesi [Investigation of misconceptions about melting and coaling in ontology basis]. *Amasya Universitesi Egitim Fakultesi Dergisi*, 1(1), 54-72.

Zohar, A. & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, 39(1), 35-6.

Watson, J. R., Prieto, T., & Dillon, J. S. (1997). Consistency in students' explanations aboutcombustion. *Science Education*, 81, 425-444.

Kavram Yanılgılarının Ontolojik Açıdan İncelenmesi ve Bulunan Yanılgıların Oluşturulan Argüman Ortamları ile Giderilmesi

Atıf:

Topalsan-Kınık, A. & Bayram, H. (2017). Eliminating with created argument environment after evaluated and categorized misconceptions in an ontological sense. *Eurasian Journal of Educational Research,* 69, 1-19, DOI: http://dx.doi.org/10.14689/ejer.2017.69.1

Özet

Problem Durumu: Önemli fizik kavramlarını içeren kuvvet ve hareket konusu ile hemen hemen her düzeydeki öğrencide oldukça yüksek oranda kavram yanılgısı olduğu yürütülen araştırmalarla ortaya konulmuştur. Fakat benzer olarak yapılan bu çalışmalarda sadece kavram yanılgıları ortaya çıkarılmış ve yanılgıların nedenleri araştırılmadan farklı yöntem ve tekniklerle giderilmeye çalışılmıştır. Ontolojik kategorilere göre, yanılgıların nedenlerinin ortaya konulduğu değerlendirmenin yapıldığı araştırmalar sınırlı sayıdadır. Bu nedenle yapılan çalışmanın problem cümlesi, yanılgıların nedenlerini tespit etmek ve etkili bir yöntem önermek amacı ile "Kuvvet ve Hareket konusu ile ilgili geliştirilen argüman ortamlarının ve geleneksel olarak yürütülen öğretim sürecinin, Kuvvet ve Hareket konusu ile ilgili tespit edilmiş, ontolojik kategorileştirmeden kaynaklanan kavram yanılgılarını gidermede etkisi nasıldır?" olarak saptanmıştır.

Araştırmanın Amacı: Bu araştırmada, "Kuvvet ve Hareket" konusunda yer alan kuvvet, sürtünme kuvveti, iş, enerjinin korunumu, mekanik enerji, kinetik enerji, potansiyel enerji, yayların depoladığı enerji gibi temel Fizik kavramları ilgili öğrencilerde bulunan kavram yanılgılarını ortaya çıkarmak ve bulunan yanılgıları ontolojik açıdan değerlendirilip, kategorileştirildikten sonra oluşturulan argüman ortamları ve geleneksel olarak uygulanan öğretim süreci ile gidermek amaçlanmıştır.

Araştırmanın Yöntemi: Araştırmanın deseni, temel problemi ve cevap aranan alt problemler dikkate alındığında ön test-son test kontrol gruplu yarı deneysel modeldir. Çalışmada uygulanan deneysel desende, bağımlı değişkenler akademik başarı, bilimsel süreç becerileri ve kavram öğrenme olarak belirlenmiştir. Bu bağımlı değişkenler üzerinde etkisi incelenen bağımsız değişken ise uygulanan öğrenme-öğretme yaklaşımıdır. Ayrıca Kuvvet ve Hareket Konusu Kavram testinde yer alan her bir soru için tespit edilen ontolojik kategoriler derinlemesine analiz edilip tartışılmıştır.

Araştırmanın Bulguları: Uygulamanın ardından deney grubu ile kontrol grubu arasında, bilimsel süreç becerilerinden değişkenleri tanımlama, işlemsel açıklamalar yapma, araştırma tasarlama ile grafiği ve verileri yorumlama boyutlarında deney grubu lehine anlamlı bir fark oluşmuştur.

Yine gerçekleştirilen uygulamanın ardından deney grubu ile kontrol grubu arasında, akademik başarı ve kavram öğrenme düzeyleri açısından deney grubu lehine anlamlı bir fark oluşmuştur.

Yapılan uygulamaların öncesi ve sonrasında Kuvvet ve Hareket Ünitesi ile ilgili belirlenmiş temel Fizik kavramları ontolojik olarak incelenip kategorileştirdikten sonra, deney grubundaki öğrencilerin, uygulamadan önce üst kategoriye yerleştirmeden kaynaklanan 301 kavram yanılgısı, yanal kategoriye yerleştirmeden kaynaklanan 150 kavram yanılgısı tespit edilmiştir. Üst kategoriye yerleştirmeden kaynaklanan bu 301 kavram yanılgısının 252'si yapılan argüman çalışmaları sayesinde giderilmiştir. Üst kategoride giderilen kavram yanılgısının oranına bakıldığında %83,72 olduğu bulunmuştur. Yanal kategoriye yerleştirilen 150 kavram yanılgısının 128'unun da yapılan argüman çalışmaları sonrası giderilmiştir. Yanal kategoride giderilen kavram yanılgısının oranına bakıldığında %85,33 olduğu bulunmuştur. Bu durum öğretim süreci boyunca kullanılan argüman çalışmalarının olumlu etkisini ortaya çıkarmıştır. Üst ve yanal kategoride ortaya çıkan kavram yanılgıları büyük bir oranda ortadan kaldırılmıştır. Üst ontolojik ve yanal kategoriler kendi içerisinde incelendiğinde, özellikle yanal kategoriler içerisine yerleştirilmiş kavram yanılgılarının, yapılan argüman çalışmaları sonrası daha fazla giderildiği görülmüştür. Kontrol grubundaki öğrencilerin, uygulamadan önce üst kategoriye yerleştirmeden kaynaklanan 318 kavram yanılgısı, yanal kategoriye yerleştirmeden kaynaklanan 131 kavram yanılgısı tespit edilmiştir. Üst kategoriye yerleştirmeden kaynaklanan bu 318 kavram yanılgısının 122'si giderilmiştir. Üst kategoride giderilen kavram yanılgısının oranına bakıldığında %38,36 olduğu bulunmuştur. Yanal kategoriye yerleştirilen 131 kavram yanılgısının 59'ü giderilmiştir. Yanal kategoride giderilen kavram yanılgısının oranına bakıldığında %45,03 olduğu bulunmuştur. Bu durum labaratuar ortamında yapılan geleneksel çalışmaların, yanal kategoriye yerleştirilmiş kavram yanılgılarını gidermede, üst kategoriye yerleştirilmiş yanılgılara göre daha başarılı olduğunu göstermektedir. Bunun yanı sıra geleneksel olarak uygulanan çalışmaların öğrencilerde yeni kavram yanılgıları da çıkan sonuçlardan görülmektedir. Yapılan çalışmalar sonrasında 31 yeni kavram yanılgısı ortaya çıkmıştır.

Araştırmanın Sonuçları ve Önerileri: Araştırmada kavram yanılgılarının ontolojik kategorilere göre değerlendirilmesi bu yanılgıların ontolojiye göre hangi nedenlerden dolayı oluştuğunun anlaşılmasını sağlamıştır ayrıca argüman ortamlarının ontolojik olarak tespit edilmiş kavram yanılgılarından sınırlı etkileşim-doğal, rastgele olay ve madde kategorilerindeki yanılgıları gidermediki etkililiği sayısal verilerle ortaya konulmuştur. Bu nedenle aktarılacak konuların bu tür yanılgıları içermesi halinde, argüman ortamları yaratılıp öğrenme ortamı daha etkin ve yanılgısız hale getirilebilir. Yanılgıların nedenlerinin bilinmesi bu ve buna benzer araştırmalar için oldukça önemlidir. Kavram yanılgılarının giderilmesi ancak nedenleri üzerine yoğunlaşıp bunların oluşmalarını engelleyen çalışmaların, öğretim yöntemlerinin, vb. hazırlanmasıyla gerçekleştirilebilir. Bu nedenle ontolojik kategoriler yanılgıların nedenlerinin açığa çıkarılmasını sağladığından çok önemlidir. Kavram yanılgılarının belirlenmesi, değerlendirilmesi ve giderilmesi gibi araştırmalar ontolojik kategorilerden yararlanılarak gerçekleştirilmelidir.

Bunun yanı sıra öğretmenler farklı konularda, farklı argüman teknikleri ile geliştirecekleri çalışmalarla ders içeriklerinin kalitesini daha rahat arttırabilir. Bu

nedenle yurt dışında birçok çalışma ile etkililiği belirlenen bilimsel tartışma modeli öğretmen adaylarına öğretilmeli ve öğretmen adaylarının tartışma becerileri geliştirilmeye çalışılmalıdır. Öğretmenlerin bilimsel tartışma sürecini öğrenmeleri, etkili tartışma yönetebilmeleri için bilimsel tartışma modeli öğretmenlere uygulamalı olarak anlatılmalı ve öğretmenlere bilimsel tartışma etkinlikleri yaptırılmalıdır. Farklı ders içerik ve kazanımlarında geliştirilen argüman çalışmaları bir kitap haline getirilebilirse, öğretmenler süreç içerisinde zorlanmadan argüman çalışmalarını uygulayabilir ve kendilerine uygun çalışmaları, yapılan bu kitapı kaynak alarak daha rahat oluşturabilir. Ayrıca argümanların bilimin doğasının anlaşılmasında, bilimin gelişmesinde, öğrenciler tarafından bilgilerin sorgulanmasında, bilgilerin kalıcı olmasında vb. olumlu etkileri düşünüldüğünde ders kitaplarında argümanlara yer verilmesinin öğrencilere önemli katkılar sağlayacağına inanılmaktadır.

Anahtar Kelimeler: Ontolojik kategoriler, argümantasyon, kuvvet ve hareket, kavram yanılgıları.

Eurasian Journal of Educational Research 69 (2017) 21-36

Eurasian Journal of Educational Research
www.ejer.com.tr

Effects of Content Balancing and Item Selection Method on Ability Estimation in Computerized Adaptive Tests

Alper SAHIN[1] Durmus OZBASI[2]

ARTICLE INFO

Article History:

Received: 14 January 2017

Received in revised form: 09 April 2017

Accepted: 23 April 2017

DOI: http://dx.doi.org/10.14689/ejer.2017.69.2

Keywords

likelihood weighted information
fisher's maximum information
Estimation accuracy

ABSTRACT

Purpose: This study aims to reveal effects of content balancing and item selection method on ability estimation in computerized adaptive tests by comparing Fisher's maximum information (FMI) and likelihood weighted information (LWI) methods. **Research Methods:** Four groups of examinees (250, 500, 750, 1000) and a bank of 500 items with 10 different content domains were generated through Monte Carlo simulations. Examinee ability was estimated by fixing all settings except for the item selection methods mentioned. True and estimated ability (θ) values were compared by dividing examinees into six subgroups. Moreover, the average number of items used was compared. **Findings:** The correlations decreased steadily as examinee θ level increased among all examinee groups when LWI was used. FMI had the same trend with the 250 and 500 examinees. Correlations for 750 examinees decreased as θ level increased as well, but they were somewhat steady with FMI. For 1000 examinees, FMI was not successful in estimating examinee θ accurately after θ subgroup 4. Moreover, when FMI was used, θ estimates had less error than LWI. The figures regarding the average items used indicated that LWI used fewer items in subgroups 1, 2, 3 and that FMI used less items in subgroups 4, 5, and 6. **Implications for Research and Practice:** The findings indicated that when content balancing is put into use, LWI is more suitable to estimate examinee θ for examinees between -3 and 0 and that FMI is more stable when examinee θ is above 0. An item selection algorithm combining these two item selection methods is recommended.

[1] Corresponding Author: Middle East Technical University Northern Cyprus Campus, School of Foreign Languages, Modern Languages Program, alpersahin2@yahoo.com
[2] Canakkale Onsekiz Mart University, TURKEY, dozbasi@comu.edu.tr

Introduction

Traditional paper and pencil tests are on the verge of being outdated due to recent technological advances that affect measurement and evaluation field. In the traditional paper and pencil tests, test takers take all items in a test and spend a considerable amount of time responding to items that are too easy or too difficult for them. Thanks to recent technology and advancements in educational measurement, test takers no longer have to take all items in a test. Rather, they only take the items aligned to their estimated ability (θ) level that is calculated while they are taking the test. This is possible with computerized adaptive tests (CATs). Typically, CATs have some advantages over traditional methods such as providing the test results immediately, reducing the number of items taken by each examinee dramatically, and being more reliable and valid than a conventional test while using fewer items (Hambleton & Swaminathan, 1985; Rudner, 1998; Weissman, 2006; Thompson & Weiss, 2011).

Having an estimate of examinee θ with less error highly depends on putting some sound criteria for item selection, test termination, and θ estimation together in the CAT environment. Item selection method is a very important component of CATs (Choi & Swartz, 2009), as the θ estimation in a CAT environment is conducted in real time according to the responses of the test takers to certain items with known item parameters. Therefore, ensuring that the computer makes the right decision in choosing which item to use next has the utmost influence on θ estimates, which are used for many high-stakes purposes. However, the selection of the appropriate item in the item pool is not an easy process in CATs. It has still been discussed in the literature (Chang & Ying, 1996; Veerkamp & Berger, 1997; van der Linden, 1998; Chen, Ankenmann & Chang, 2000; Cheng & Liou, 2003; Weissmann, 2006).

A successful CAT is based on an item bank composed of items that address a wide range of θ levels. This item bank has its own information function to which each item contributes with its own information function formed according to its item parameters. During a CAT session, items are mainly selected among the ones with the highest information and closest to the location of the estimated θ of the examinee taking the test. As expected, some item selection methods have been proposed by different authors (Kingsbury & Zara, 1989; Lord, 1980; Veerkamp & Berger, 1997; Chang & Ying, 1996) in order to optimize this procedure. However, selection of items in CAT is often dependent on Fisher's maximum information (FMI). FMI mostly uses the maximum likelihood estimate of the θ (Veerkamp & Berger, 1997; Barrada, Olea, Ponsoda & Abad, 2010).

FMI utilizes item information, the conversion of the item characteristic curve, to select items for CATs (Weiss, 1983). Selecting items from an item pool for a multiple-choice test, where the item characteristic curve is defined in three-parameter logistic model (3PLM; will be explained further later), FMI can be calculated using Equation 1 (Embretson and Reise, 2000):

$$I_i[\theta_{m-1}] = \frac{(D_{ai})^2(1-c_i)}{\left[c_i+e^{D_{ai}(\theta_{m-1}-b_i)}\right]\left[1+e^{D_{ai}(\theta_{m-1}-b_i)}\right]^2}\ , \qquad (1)$$

in which,

m = examinee

a_i = item discrimination for item i;

b_i = item difficulty for item i;

c_i = pseudo-chance parameter of item i;

D = scaling constant (mostly used as 1.7)

and in which, c_i is set to 0.00 for two parameter model and a_i to 1.00 (and c_i to 0.00 as well) for one parameter model. The item information for each item in the item bank can be calculated with the formula above. With the help of equation 1, the total item information levels of the items given to one person reaches the maximum (Lord, 1980).

In studies on item selection, FMI or an FMI-based method almost never changes as the performances of newly proposed methods are mostly compared to that of FMI. Although many studies were conducted to develop better alternative item selection methods, their results could yield slight differences or advantages over FMI. According to the current literature, especially when the CAT has more than 20 items, the difference in performance of a newly proposed method and FMI turns out to be trivial (Passos, Berger & Tan, 2007). For example, Chen, Ankenmann and Chang (2000) conducted a simulation study to compare item selection methods of FMI, Fisher interval information, Fisher information with a posterior distribution, Kullback Leibler information (KLI) and KLI with a posterior distribution in terms of test efficiency and ability estimation precision at the beginning of CAT session. In their results, they found that for CATs with more than 10 items, there is no difference between FMI and other selection methods in terms of θ estimation precision. Similarly, Chang and Ying (1996) compared the performance of KLI and FMI in two studies. In the first, they used an item bank of 800 items simulated from a pre-specified uniform distribution, and in the second one they used an item bank of 254 items whose parameters were taken from a National Assessment of Educational Progress reading test. They found that KLI performed slightly better when the test was short. Especially in the second study, the difference was trivial.

Additional studies have reached similar results with negligible differences between FMI and alternative methods for tests with more than 20 items (Barrada, Olea, Ponsoda & Abad, 2009; van Rijn, Eggen, Hemker & Sanders, 2002; Veldkamp, 2003). However, Veerkamp and Berger (1997) suggested a feasible alternative item selection criteria called likelihood weighted information (LWI). In LWI, which was suggested by Veerkamp and Berger (1997) as an alternative to FMI, the information function is formulated as a weighted mean of information function of all possible theta values. The LWI function is defined by Veerkamp and Berger (1997) as:

$$\max_{i \in In} \int_{-\infty}^{\infty} L_n(\theta; x_{n-1}) I_i(\theta) d\theta. \quad (2)$$

in which they define LWI as a product of $L_n(\theta; x_{n-1})$, the likelihood function (I) of the (n-1)th item with a response vector of x_{n-1}.

In their study, Veerkamp and Berger (1997) used two simulated item banks with 200 and 400 items generated in 3PLM. They compared FMI, interval information and LWI for up to 60-item tests. They found that LWI was a good alternative to the FMI. LWI was found to be the only alternative that outperformed FMI in tests over 20 at that time.

There is ample research on the comparison of item selection methods in CATs. However, the current literature lacks further studies considering the recent advances and practical needs of current CAT applications like content balancing. There was no study found in the literature that compared the performance of item selection methods when content balancing was put into use. Moreover, the current literature does not reveal how the examinees with different ability levels are affected from the changes in item selection method and content balancing. The present study addressed these issues by using FMI and LWI as the item selection methods together with content balancing in CAT and sought an answer to the following research question: Does the accuracy of the θ estimation change for examinees with different θ levels depending on the item selection method used when content balancing is put into use?

Method

Research Design

According to the International Council for Science (2004), basic research is defined as experiment- or theory-based research that aims to increase the current information on a topic with indirect concerns about its practicality. The present study is a basic research study, the data of which was generated through Monte-Carlo simulations using SimulCAT (Han, 2012).

Research Sample

As the first step of the item generation process, examinee samples of different sizes (250, 500, 750 and 1000) were generated with a standard normal distribution between -3 and +3. In this way, the true θ levels of these examinees were obtained.

Research Instruments and Procedures

After the generation of the examinee samples, the items in the item bank of the study were generated. For this purpose, a bank of 500 items with equally distributed items in 10 different content domains (with 50 items in each) were generated separately in 3PLM of item response theory (IRT). In 3PLM, each item has item discrimination (a), item difficulty (b) and pseudo-chance (c) parameters. The 3PLM can be shown with equation 3 (Hambleton, Swaminathan & Rogers, 1991):

$$P_{ij}(\theta_j) = c_i + (1 - c_i)\frac{\exp\left[Da_i(\theta_j - b_i)\right]}{1 + \exp\left[Da_i(\theta_j - b_i)\right]}, i = 1..., n \quad (3)$$

in which $P_{ij}(\theta_j)$ can be explained as the probability of a correct response of examinee j to item i on a specific θ level. Moreover, a_i corresponds to estimated a, b_i to estimated b, and c_i to the estimated c parameter for item i.

All item parameters were generated from a uniform distribution with the a parameters ranging between 0 and 1.5, the b parameters ranging between -3 and +3 and the c parameters ranging between 0 and 0.25. The item parameters were generated from a uniform distribution in order to obtain an item bank with more balanced capability of estimating θ in all areas of the θ continuum. The item bank information function of the item bank generated can be viewed in Figure 1.

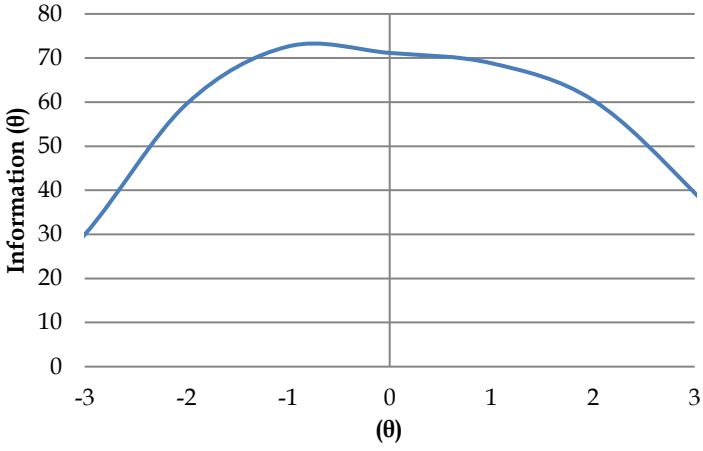


Figure 1. Item bank information function.

Post-Hoc Simulations. Following the generation of examinee and item parameters, five post-hoc simulations were conducted. During these post-hoc simulations, each exam session was set to have at least 10 items and 10% from each content domain. This was done to make sure that the sessions did not terminate with very few items and that there are approximately the same number of items from each content domain in each session. Maximum likelihood estimation was used to estimate examinee abilities in each research condition. Tests were terminated when the standard error of θ estimate was 0.25 and below. No exposure control method was utilized. Moreover, random values between -0.5 and 0.5 were taken as the initial θ estimates of the examinees.

As mentioned earlier, performance of two item selection methods, LWI and FMI methods were compared. This comparison was done with each of four examinee samples, and each research condition was replicated 10 times. In this way, 21 individual θ (including true θ) for each examinee in each examinee sample and a total of 84 scores were obtained. A brief overview of this can be seen in Table 1.

Table 1

A Brief Overview of Scores Obtained Through Simulations

	True ability score for each examinee	LWI (estimated score for each examinee with replications)	FMI (estimated score for each examinee with replications)	Total
250	1	10	10	21
500	1	10	10	21
750	1	10	10	21
1000	1	10	10	21
			Total	84

Data Analysis

Data analysis was handled by investigating the accuracy of θ estimates, conditional on ability subgroups.

The accuracy of θ estimates in each research condition was evaluated by calculating the correlation (r; Gao & Chen, 2005) between the true θ levels of the examinees that were obtained when the examinees were first generated and their estimated θ levels in each research condition and replication. Then, these correlations were averaged to obtain the average correlation of the estimated θ scores for each examinee. Moreover, the mean squared error (MSE; Veerkamp & Berger, 1997; Chang & Ying, 1996) between the true and estimated scores was also calculated using Equation 4:

$$\mathrm{MSE}\left(\widehat{\theta}\right) = \frac{\sum_{j=1}^{N}(\widehat{\theta}_j - \widehat{\theta}_{Ti})^2}{N}, \quad (4)$$

where $\hat{\theta}_j$ is the estimated θ, $\hat{\theta}_{Ti}$ is the true θ for the examinee j in each research condition, and N is the total number of examinees. Apart from the correlations and MSE values, the average numbers of items used in each research condition were also calculated conditional on examinee samples.

Ability Subgroups. Findings were analyzed conditional on examinees' θ level in pre-specified intervals rather than taking all examinees as a whole. This was done to have a deeper understanding of the effects of item selection on the θ estimation for examinees with various θ levels. It is known that examinees with different θ levels are affected differently from variations in CAT methodology (Sahin & Weiss, 2015).

Examinees were divided into subgroups according to their true θ levels with increments of 1.00 standard deviation. For example, examinees with θ levels higher than -2 were put into subgroup 1. Then, examinees between -2 and -1 were put into subgroup 2. Six subgroups were formed as a result of this procedure. Distribution of examinees in each θ group, conditional on examinee samples, can be seen in Figure 2.

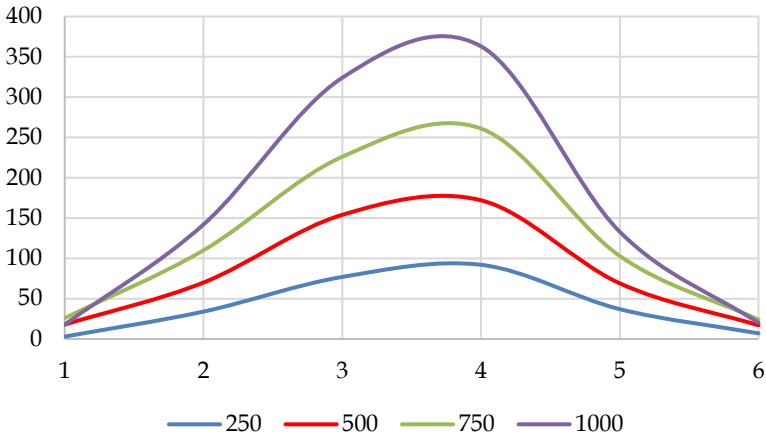


Figure 2. Distribution of examinees in each θ group conditional on the examinee samples.

Results

The average correlation coefficients between true and estimated θ parameters, conditional on the number of test takers and θ groups, are presented in Figure 3. As can be seen in Figure 3, correlations are the highest in group 1, for the students with the lowest θ levels in all examinee samples. The highest correlation (r=0.94) that was obtained with 250 examinees was in group 1 when FMI was used as the item selection method. The lowest correlation obtained in the same group was r=0.26 when LWI was used in group 6.

The highest correlation obtained with 500 examinees was r=0.75 in group 1, when LWI was used. The lowest correlation obtained with the same examinees was r=0.24 in group 5, when LWI was used. When the examinee number increased to 750, the highest correlation was around the same value in group 1, when LWI (r=0.76) and FMI (r=0.75) were used. In addition, the lowest correlation (r=0.22) was obtained from group 6, when LWI was used. The highest correlation obtained when there were 1000 examinees who took the test was in group 1 again with similar values for FMI (r=0.74) and LWI(r=0.75), and the lowest correlation (r=0.19) was obtained in group 6 when the LWI was used.

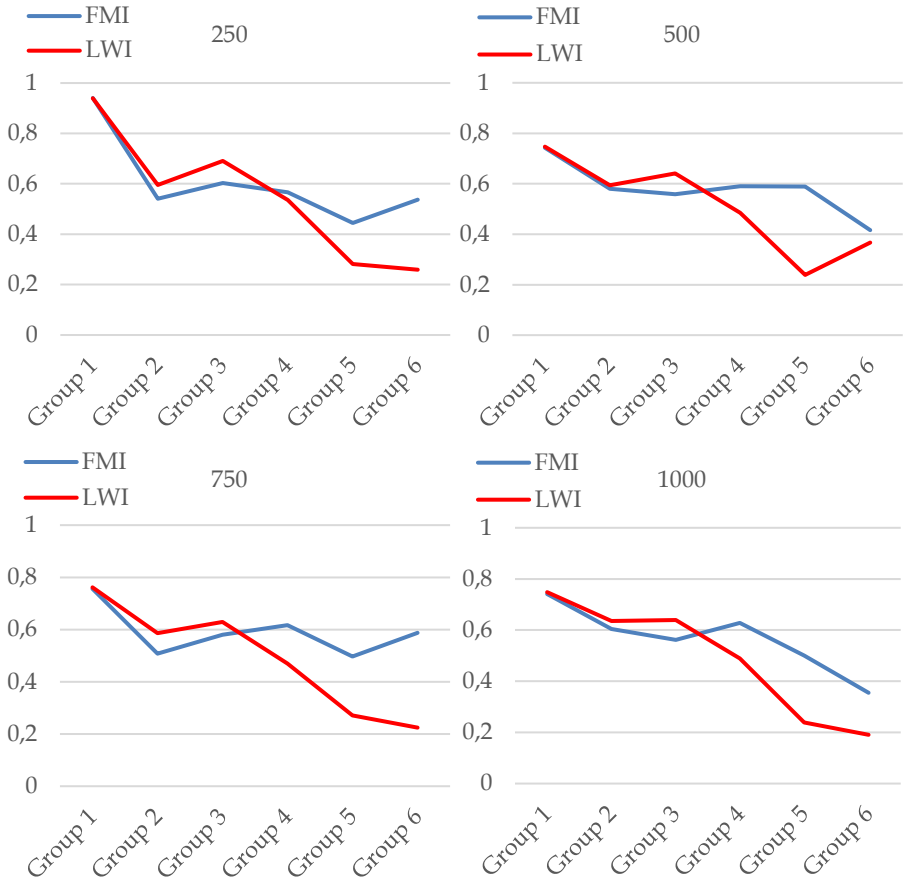


Figure 3. Correlations conditional on number of test takers, item selection method and θ groups.

MSE conditional on number of test takers, item selection method used and θ group of the examinees can be seen in Figure 4. The lowest MSE obtained with 250 examinees was in group 1 (MSE=0.10), when LWI was used. Moreover, MSE=1.11 was the highest MSE value obtained with 250 examinees in group 6, when LWI was used. When the examinee number increased to 500, the lowest MSE was obtained in group 2 (MSE=0.12), when LWI was used. In addition, the highest MSE was obtained in group 6 (MSE=1.22), when LWI was used. In the sample with 750 examinees, the lowest MSE was obtained in group 1 (MSE=0.11), when LWI was used as the item selection method. The highest MSE was in group 6 (MSE=1.35), when LWI was used. In the examinee sample with 1000 examinees, similar results were obtained. The lowest MSE (MSE=0.11) was obtained in group 1 when LWI was used. Group 6 was the one with the highest MSE (MSE=1.27), when LWI was used as the item selection method.

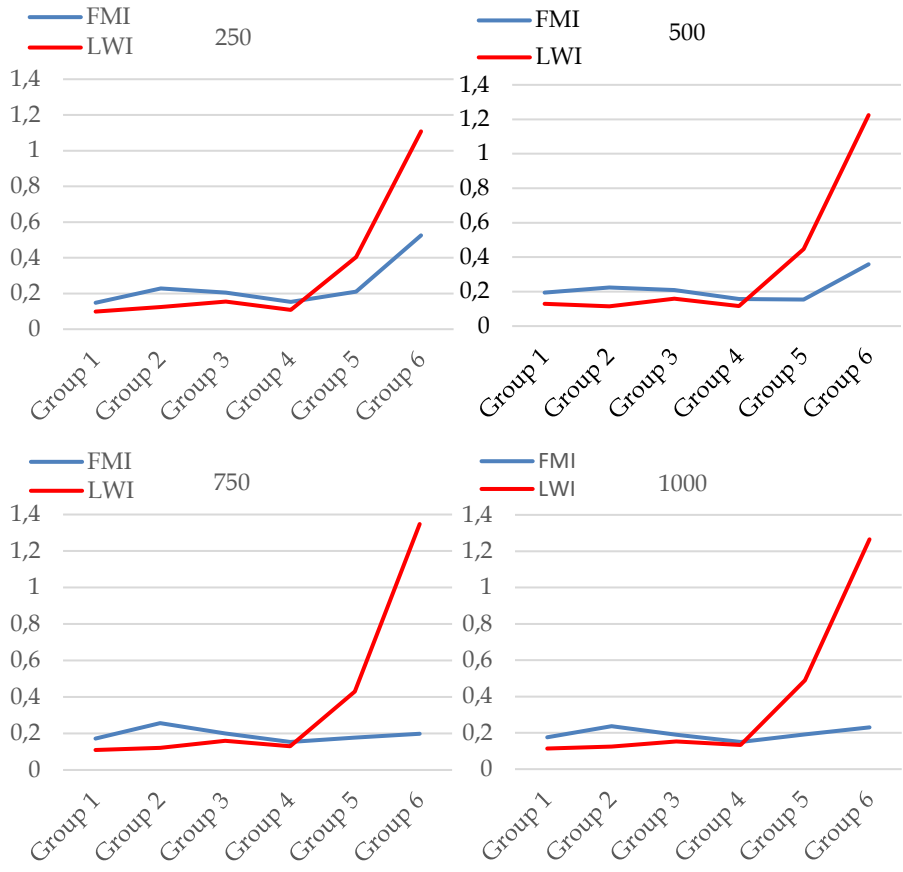


Figure 4. MSE conditional on number of test takers, item selection method and θ groups.

When Figure 5 was analyzed in terms of the average number of items used in each condition of the study, it was seen that 22.56 items were used for group 5, when FMI was used with 250 examinees. The highest average number of items used for the same 250 examinees was 41.77, when LWI was used for examinees in group 6. An average of 31.03 items were used for examinees in group 1 for this examinee sample as well. The highest average number of items used with 500 examinees was 44.55, when LWI was used for examinees in group 6. The lowest average number of items used was 22.78, for group 5 in 500 examinees, when FMI was used. The highest average number of items used for 750 examinees was 45.81, when LWI was used for examinees in group 6. The lowest average number of items used was 22.71 in group 5 of 500 examinees, when FMI was used. Among the 1000 examinees, group 6 got an average 44.1 items, and group 5 got an average of 22.65 items in their sessions.

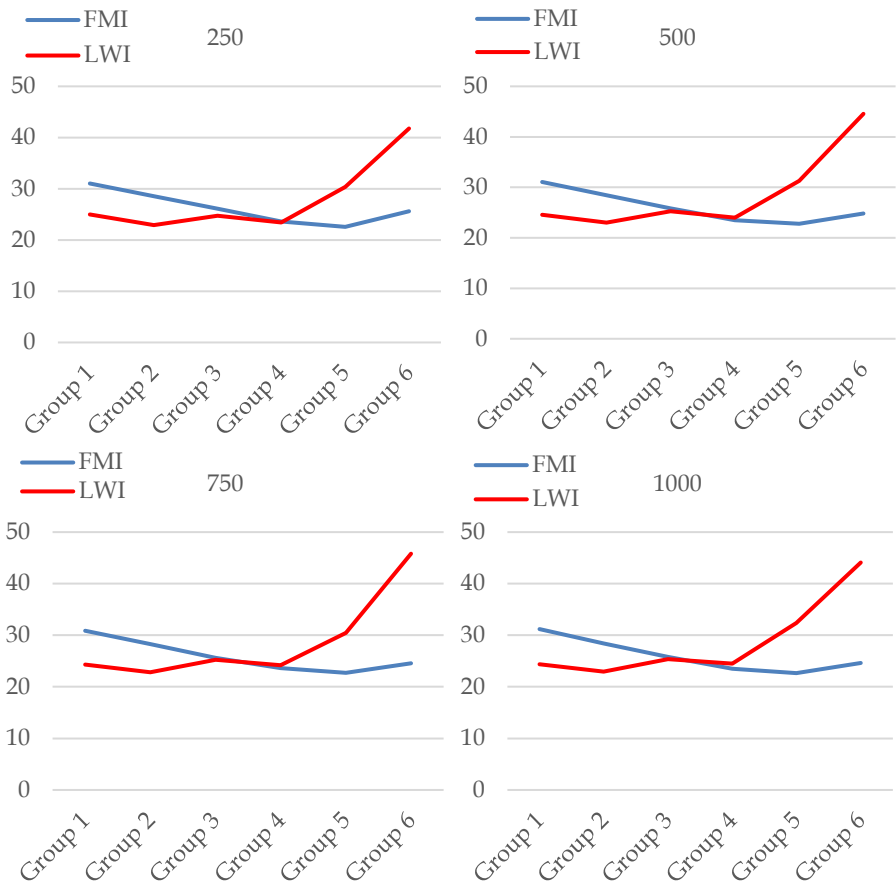


Figure 5. Average number of items used in each condition.

Discussion and Conclusion

The findings regarding the correlations indicated that correlation coefficients decreased steadily as examinee θ level increased from -3 to +3 in all examinee samples when LWI was used as the item selection method. FMI obtained decreasing correlations with 250 and 500 examinees as the examinee level increased. When 750 examinees took the test, correlations were somewhat steady in regard to FMI. When 1000 examinees took the test, FMI was not successful in estimating examinee θ accurately after Group 4. It is interesting to note that LWI is better in estimating the examinee θ levels in θ subgroups 1, 2, and 3. Similarly, FMI outperforms LWI in θ subgroups 4, 5 and 6.

When the figures regarding the MSE are analyzed, parallel conclusions can be drawn. From the figure for MSE, it is visible that there is a dramatic increase in MSE values in subgroup 6 when LWI was used in all conditions. There is also an increase in MSE when FMI was used, but it is somewhat limited compared to LWI. As indicators of estimation accuracy, MSE values indicate that when FMI is used as the item selection method, θ estimates are estimated with less error compared to LWI. Moreover, it is important to note that when the examinee number reached 750, the increase in MSE values when FMI was used became nearly invisible. According to the findings in this regard, as in correlation coefficients, LWI outperforms FMI in θ subgroups 1, 2, 3 and FMI outperforms LWI by having less MSE in θ subgroups 4, 5 and 6. The same rule applies when the average number of items used in all conditions are analyzed.

When all these are put together and interpreted as a whole to answer our research question, it can be said that LWI is more suitable to estimate examinee θ for examinees between -3 and 0 when content balancing is put into use. Moreover, our results also suggest that FMI is more stable when examinee θ is above 0, but it is less accurate in estimating examinee θ when the examinee level is below 0. This is somewhat conflicting with Veerkamp and Berger (1997), who found that LWI might be a sound alternative to FMI. LWI may be a good alternative to FMI when θ estimates are compared as a whole and when content balancing is not put into use; however, when the content balancing is in use and when examinees are divided into θ groups, LWI outperforms FMI only for certain θ subgroups. Therefore, a new item selection algorithm using the LWI method for the examinees with θ levels below 0 and using FMI for examinees with θ levels above 0 might be more beneficial and more robust against possible difficulties that both of these item selection methods experience for certain groups of examinees during CAT administration.

The current study has some limitations. First of all, the current findings are limited to the uses of LWI and FMI when content was balanced in 10 different content areas that comprised 10% of each CAT session. Secondly, although the data analyses were replicated 10 times, because of the nature of the study, the findings may be rather limited to the data generated in this study. Moreover, the item bank generated in the present study had high information in nearly all areas of the θ continuum, so the results may be limited to CATs with similar item banks. Finally, SE(θ) <= 0.25 was used as the test termination criteria. This may be a rather stringent termination criteria for real CAT administrations, and current findings may be limited as such.

The results of this study have caused some questions to emerge, and it is suggested that they be investigated in further detail by follow-up studies. A possible follow-upstudy would investigate the feasibility of using a mixed-method item selection algorithm, as suggested by the findings of the present study, that uses LWI when the examinee level is below 0 and FMI when it is above 0. Moreover, a similar study with real or simulated data that compares the accuracy of the θ estimates when content balancing is and is not used would also be beneficial. Last but not least, a study comparing the performances of LWI and FMI with item banks of different sizes would be highly valuable.

References

Barrada, J. R., Olea, J., Ponsoda, V., & Abad, F. J. (2009). Item selection rules in computerized adaptive testing: Accuracy and security. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *5*, 7-17.

https://dx.doi.org/10.1027/1614-2241.5.1.7

Barrada, J. R., Olea, J., Ponsoda, V., & Abad, F. J. (2010). A method for the comparison of item selection rules in computerized adaptive testing. *Applied Psychological Measurement, 34,* 438-452.

https://dx.doi.org/10.1177/0146621610370152

Chang, H.-H., & Ying, Z. (1996). A global information approach to computerized adaptive testing. *Applied Psychological Measurement*, *20*, 213– 229.

http://dx.doi.org/10.1177/014662169602000303

Chen, S.-Y., Ankenmann, R. D., & Chang, H. H.(2000). A comparison of item selection rules at the early stages of computerized adaptive testing. *Applied Psychological Measurement*, *24*, 241-255.

http://dx.doi.org/10.1177/01466210022031705

Cheng, E.P. & Liou, M. (2003). Computerized adaptive testing using the nearest-neighbors criterion. *Applied Psychological Measurement*, *27*(3), 204-216.

http://dx.doi.org/10.1177/0146621603027003002

Choi, S. W.& Swartz, R. J. (2009). Comparison of CAT item selection criteria for polytomous items. *Applied Psychological Measurement*, *33*, 419-440.

http://dx.doi.org/10.1177/0146621608327801

Embretson, S.E. & Reise, S.P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gao, F., & Chen, L. (2005). Bayesian or non-Bayesian: A comparison study of item parameter estimation in the three parameter logistic model. *Applied Measurement in Education, 18*, 351–380.

http://dx.doi.org/10.1207/s15324818ame1804_2

Hambleton, R. K., & Swaminathan H. (1985). *Item response theory: Principals and applications*. Norwell, MA: Kluwer Nijhof.

Hambleton, R. K., Swaminathan, H., & Rogers, H.J. (1991) *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications.

Han, K. T. (2012). SimulCAT: Windows software for simulating computerized adaptive test administration. *Applied Psychological Measurement, 36*(1), 64-66.

http://dx.doi.org/10.1177/0146621611414407

International Council for Science. (2004). *The value of basic scientific research.* Retrieved March 10, 2017 from http://www.icsu.org/publications/icsu-position-statements/value-scientific-research

Kingsbury, G. G., & Zara, A. R. (1989). Procedures for selecting items for computerized adaptive tests. *Applied Measurement in Education*, *2*, 359–375.

http://dx.doi.org/10.1207/s15324818ame0204_6

Lord, F. M. (1980) *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Passos, V. L., Berger, M. P. F., & Tan, F. E. (2007). Test design optimization in CAT early stage with the nominal response model. *Applied Psychological Measurement*, *31*, 213-232.

http://dx.doi.org/10.1177/ 0146621606291571

Rudner, L. M. (1998). *An on-line, interactive, computer adaptive testing tutorial.* Retrieved February 25, 2017 from

http://echo.edres.org:8080/scripts/cat/catdemo.htm

Sahin, A., & Weiss, D. J. (2015). Effects of calibration sample size and item bank size on ability estimation in computerized adaptive testing. *Educational Sciences: Theory & Practice, 15*, 1585-1595.

http://dx.doi.org/10.12738/estp.2015.6.0102

Thompson, N. A., & Weiss, D.J. (2011). A Framework for the development of computerized adaptive tests. *Practical Assessment, Research & Evaluation*, *16*(1), 1-9. Retrieved from:

http://pareonline.net/getvn.asp?v=16&n=1

van der Linden, W. J. (1998). Bayesian item selection criteria for adaptive testing. *Psychometrika*, *63*, 201-216.

http://dx.doi.org/10.1007/BF02294775

van Rijn, P., Eggen, T. J., Hemker, B. T., & Sanders, P. F. (2002). Evaluation of selection procedures for computerized adaptive testing with polytomous items. *Applied Psychological Measurement, 26*, 393-411. http://dx.doi.org/10.1177/014662102237796

Veerkamp, W. J J. & Berger, M. P. F. (1997). Some new item selection criteria for adaptive testing. *Journal of Educational and Behavioral Statistics*, *22*(2), 203-226. Retrieved from

http://www.jstor.org/stable/1165378

Veldkamp, B. P. (2003). Item selection in polytomous CAT. In H. Yanai, A. Okada, K. Shigemasu, Y. Kano, & J.J. Meulman (Eds.), *New developments in psychometrics* (pp. 207-214). Tokyo, Japan: Springer-Verlag.

Weiss, D. J. (1983). *Final Report: Computer-Based Measurement of Intellectual Capabilities*. University of Minnesota, Department of Psychology. Retrieved from the University of Minnesota Digital Conservancy,

http://hdl.handle.net/11299/135261

Weissman, A. (2006). A Feedback control strategy for enhancing item selection efficiency in computerized adaptive testing. *Applied Psychological Measurement*, *30*(2), 87-99.

http://dx.doi.org /10.1177/0146621605282774

Bilgisayar Ortamında Bireye Uyarlanmış Testlerde İçerik Dengeleme ve Madde Seçme Yönteminin Yetenek Düzeyi Kestirimine Etkileri

Atıf :

Sahin, A. & Ozbasi, D. (2017). Effects of content balancing and item selection method on ability estimation in computerized adaptive testing. *Eurasian Journal of Educational Research*, *69*, 21-36. http://dx.doi.org/10.14689/ejer.2017.69.2

Özet

Problem Durumu: Son yıllardaki teknolojik gelişmelerin ölçme ve değerlendirme alanına katkılarıyla birlikte geleneksel anlamda kağıt kalem testleri artık eski popülerliğini yitirmeye başlamıştır. Gelişen bilgisayar teknolojisi, hem ölçme işleminin süresinin kısalmasını hem de daha geçerli ve güvenilir testlerin işekoşulmasını mümkün hale getirmiştir. Özellikle bireyin yetenek düzeyine uygun sınav sorularıyla karşılaşması zaman ve kullanılan süre açısından önemli bir tasarruf sağlamaktadır. Bu, ancak bilgisayar ortamında bireye uyarlanmış test (BOBUT) uygulaması ile mümkün olabilmektedir. BOBUT uygulaması, başlatma kuralı, madde seçim yöntemi, yetenek kestirimi, içerik dengeleme ve test sonlandırma gibi önemli süreçlerden oluşmaktadır. Bu süreçlerin belki de en önemlisi madde seçim yöntemidir. Bu çalışmada BOBUT uygulamasının en önemli aşamalarından olan madde seçim yöntemleri ele alınmıştır. Alanyazındaki madde seçimine yönelik çalışmalar incelendiğinde madde seçim yöntemlerinin içerik dengeleme (content balancing) kullanıldığında farklı yetenek düzeylerindeki bireylerin örtük puanları üzerinde nasıl bir etki gösterdiğinin halihazırda henüz incelenmediği görülmüştür.

Araştırmanın Amacı: Bu araştırmanın amacı BOBUT uygulamalarında içerik dengeleme kullanıldığında madde seçim yöntemindeki değişikliğin yetenek kestirimine etkisini yaygın olarak kullanılan Fisher'ın en yüksek bilgi (Fisher's maximum information) ve onun önemli bir alternatifi olduğu daha önceki

araştırmalarda tespit edilen ağırlıklandırılmış bilgi oranı (Likelihood weighted information) yöntemlerini kullanmak suretiyle belirlemek ve içerik dengeleme üzerine sonraki dönemlerde yapılacak çalışmalara ışık tutmaktır.

Araştırmanın Yöntemi: Araştırmada kullanılan veriler Monte-Carlo simülasyon yöntemi ile elde edilmiştir. Bu bağlamda, araştırmanın verileri için yetenek düzeyleri -3 ile +3 arasında normal dağılım gösteren 4 farklı büyüklükte 250, 500, 750 ve 1000 birey grupları oluşturulmuştur. Yetenek kestirimlerinde en yüksek olabilirlik kestirim (Maximum likelihood estimation) yöntemi kullanılmıştır. Benzetim ile oluşturulan bireyler bu aşamada elde edilen gerçek yetenek düzeylerine göre altı alt yetenek grubuna ayrılmıştır (Örn. $-3 < \theta < -2$ = grup 1, $-2 < \theta < -1$ = grup 2, … vb.).

Madde havuzu için her birine yönelik 50'şer madde bulunan 10 farklı konu alanında toplam 500 madde benzetim yöntemiyle üretilmiştir. Madde parametreleri a parametresi için 0 ile 1.5, b için -3 ile +3 ve c için ise 0 ile 0.25 arasında sabit (uniform) dağılım gösterecek şekilde üretilmiştir. Birey ve maddelerin elde edilmesi sonrası bir dizi Post-hoc benzetim çalışması gerçekleştirilmiştir. Bu çalışmalar, birey yetenek başlangıç düzeyi -0.5 ile +0.5 aralığında olacak, en kısa test uzunluğu her bir konu alanından %10 oranında madde içerecek şekilde en az 10 madde kullanılacak ve yetenek düzeyi kestirimi standart hata değeri 0.25'ten küçük olduğunda testi sonlandıracak şekilde ayarlanmıştır. Post-hoc benzetimler 10 kez tekrarlanmıştır.

Araştırmanın Bulguları: Farklı madde seçme yöntemleri kullanıldığında, gerçek ve kestirilen yetenek düzeyleri arasındaki korelasyonlar (r) 4 farklı büyüklükteki grup ve bu grupların her birinde 6 farklı yetenek aralığındaki bireyler için ayrı ayrı incelenmiştir. Buna göre 250 kişilik grup için Fisher'ın en yüksek bilgi yöntemi kullanıldığında, gerçek ve kestirilen yetenek düzeyleri arasında en yüksek korelasyon r=0.94 olarak bulunmuştur. En düşük korelasyon (r=0.26) ise madde seçme kuralı olarak ağırlıklandırılmış bilgi fonksiyonu kullanıldığında elde edilmiştir. Sınavı alan kişi sayısı 500'e çıktığında ise en yüksek korelasyon madde seçme kuralı olarak ağırlıklandırılmış bilgi oranı kullanıldığında elde edilmiştir (r=0.75). Kişi sayısı 750'ye çıktığında en yüksek korelasyon katsayıları her iki yöntem için de çok yakın bulunmuştur (r_{fisher}=0.75; $r_{ağırlıklandırılmış}$=0.76). Benzer bir durum, örneklem sayısı 1000'e çıktığında da geçerli olmuş ve benzer en yüksek korelasyonlar elde edilmiştir (r_{fisher}=0.74; $r_{ağırlıklandırılmış}$=0.75).

Farklı birey gruplarında her alt yetenek düzeyi için iki madde seçme kuralı ayrı ayrı uygulandığında elde edilen tahmini yetenek düzeyleri ile bireylerin gerçek yetenek düzeyleri arasındaki ortalama karesel hata (MSE; Mean Squared Error) değerleri karşılaştırılmıştır. Buna göre, en düşük MSE değeri 250 kişilik grupta ağırlıklandırılmış bilgi oranı yöntemi kullanıldığında 1. alt yetenek grubunda elde edilmiştir (MSE=0.10). Yine aynı madde seçme kuralında alt yetenek grubu 6'da ise MSE=1.11 ile diğer yetenek gruplarına göre daha yüksek bir değer almıştır. Birey sayısı 500'e çıktığında, ağırlıklandırılmış bilgi oranı yöntemi kullanıldığında alt yetenek grubu 1 MSE=0.12 ile en düşük değer almıştır. En yüksek MSE ise alt grup 6'da MSE=1.22 olarak hesaplanmıştır. Birey sayısı 750'ye çıktığında ise ağırlıklandırılmış bilgi yöntemi kullanıldığında MSE değeri en düşük alt yetenek

grubu 1'de (MSE=0.11) elde edilmiştir. En yüksek MSE (1.35) ise yine alt grup 6'da elde edilmiştir. Birey sayısı 1000'e çıktığında da benzer sonuçlar elde edilmiştir. En düşük MSE değeri grup 1'de, en yüksek MSE değeri ise yine grup 6'dan elde edilmiştir.

Her iki madde seçme yönteminin kestirim kalitesi kullanılan ortalama madde sayıları açısından da karşılaştırılmıştır. 250 kişinin sınavı aldığı durumda, en fazla sayıda madde, madde seçme kuralı olarak ağırlıklandırılmış bilgi oranı yöntemi kullanıldığında alt yetenek grubu 6'da ortaya çıkmıştır (kullanılan madde sayısı 41.77). En düşük ortalama madde sayısı (31.03) ise alt yetenek grubu 1'den elde edilmiştir. Sınavı alan birey sayısı 500'e çıktığında ise, en yüksek ortalama madde sayısı madde seçme kuralı olarak ağırlıklandırılmış bilgi yöntemi kullanıldığında grup 6'da elde edilirken, en düşük madde sayısı Fisher'ın en yüksek bilgi yöntemi kullanıldığında 5. alt yetenek grubundan elde edilmiştir (22.78). Bu durum sınavı alan birey sayısı 750 ve 1000 olduğunda da değişmemiş, en yüksek ve en düşük ortalama madde uygulanan yetenek aralıkları ve bunlara ait madde seçme kuralları değişmemiştir. Bir başka ifade ile sınavı alan birey grubu 750 ve 1000 olduğunda en yüksek madde kullanımı her iki birey grubunda da madde seçme kuralı olarak ağırlıklandırılmış bilgi oranı yöntemi kullanıldığında grup 6'da sırasıyla ortalama 45.81 ve 44.1 şeklinde elde edilmiştir. En düşük ortalama madde kullanımı ise madde seçme kuralı olarak Fisher'ın en yüksek bilgi yöntemi kullanıldığında grup 5'te sırasıyla 22.71 ve 22.65 şeklinde elde edilmiştir.

Araştırmanın Sonuçları ve Önerileri: Çalışmada elde edilen tüm bulgular göz önüne alındığında, içerik dengeleme kullanıldığında, ağırlıklandırılmış bilgi oranı yönteminin literatürde geçtiği şekliyle Fisher'ın en yüksek bilgi yöntemine aslında tamamen üstünlük sağlamadığı, bu üstünlüğün yetenek değeri -3 ile 0 aralığında olan bireyler için geçerliği olduğu, yetenek düzeyi 0'ın üzerine çıktığı durumlarda ise Fisher'ın en yüksek bilgi yönteminin yetenek kestiriminde daha başarılı olduğu sonucuna varılmıştır. Bu durum 0'dan küçük yetenek düzeylerinde ağırlıklandırılmış bilgi oranı yönteminin, 0'dan büyük yetenek düzeylerinde Fisher'ın en yüksek bilgi yönteminin kullanılmasını sağlayacak bir madde seçme algoritmasının her iki yöntemin de eksiklerini giderebileceğinden hareketle her durumda BOBUT uygulamalarında daha başarılı yetenek düzeyi kestirimleri elde edilmesini sağlayacak böyle bir algoritmanın geliştirilmesi önerilmektedir.

Anahtar Kelimeler: Ağırlıklandırılmış bilgi oranı, fisher'ın en yüksek bilgi yöntemi, kestirim keskinliği.

Eurasian Journal of Educational Research 69 (2017) 37–54

Eurasian Journal of Educational Research
www.ejer.com.tr

The Predictive Role of Interpersonal Sensitivity and Emotional Self-Efficacy on Psychological Resilience Among Young Adults*

Bilge Nuran AYDOGDU[1] Hilal CELIK[2] Halil EKSI[3]

ARTICLE INFO

Article History:

Received: 14 December 2016

Received in revised form: 12 April 2017

Accepted: 02 May 2017

DOI: http://dx.doi.org/10.14689/ejer.2017.69.3

Keywords

psychological resilience, interpersonal sensitivity, emotional self-efficacy, young adults

ABSTRACT

Purpose: In the face of adverse and traumatic events throughout their lives, individuals respond in different ways depending on their degree of resilience, factors of which include their individual resources for coping with those events. This study examined the predictive role of emotional self-efficacy and interpersonal sensitivity on psychological resilience among young adults in order to gain insights into psychological resilience and its protective factors. In particular, its purpose was to examine how perceiving emotions of the self and others, using emotions to facilitate thought, regulating emotions in the self and others, interpersonal awareness, need for approval, separation anxiety, timidity, fragile inner self, and understanding emotions, the emotional self, and others affect perceptions of the future, structural style, social competence, family cohesion, and social resources.

Method: Using the relational screening model, participants were selected via basic random sampling. The sample included volunteers—243 women (73.4%) and 88 men (26.6%)—with a mean age of 21.46 years. The Resilience Scale for Adults, Emotional Self-Efficacy Scale, and Interpersonal Sensitivity Measure were used as measuring instruments. **Findings**: Following simultaneous multiple regression analysis, psychological resilience could be predicted according to emotional self-efficacy and interpersonal sensitivity. **Conclusions and Recommendations:** Similar to earlier research in the field, this study showed that psychological resilience and its aspects can be explained in light of emotional self-efficacy and interpersonal sensitivity. However, since psychological resilience had not heretofore been examined in such detail, this study offers significant contributions to trauma and preventive psychological counselling studies.

*This study was partly presented at the 13th National Psychological Counselling and Guidance Congress in Mersin, 07 – 09 October, 2015

[1] Marmara University, Turkey, bilge.aydogdu@marmara.edu.tr

[2]Corresponding Author: Hilal CELİK, Marmara University, Turkey, hilalcelik@marmara.edu.tr, celikhilal@gmail.com

[3] Marmara University, Turkey, halileksi@marmara.edu.tr

Introduction

Within his or her lifetime, everyone encounters at least one stressful or traumatic event and responds differently. Whereas some become more vulnerable, others appear stronger. According to Fletcher and Sarkar (2013), such diverse responses result from different levels of resilience.

Resilience can refer to "a phenomenon or process reflecting relatively positive adaptation despite the experience of significant adversity or trauma" (Luthar, 2006, p. 742). In slightly different terms, it is a dynamic developmental process that promotes positive adaptation under stressful, adverse, and traumatic circumstances (Masten & Wright, 2010). From three different perspectives, resilience can be a positive outcome despite an individual's high-risk problems, good adaptation under stressful circumstances, and recovery from trauma (Masten, Best, & Garmezy, 1990). Hjemdal, Friborg, Stiles, Rosenvinge, and Martinussen (2006) stated that resilience consists of protective factors or functional mechanisms that promote good outcomes even when a person faces adverse life events. In other words, resilience is composed of protective factors that prompt the greater possibility of a positive outcome. Those protective factors help to prevent negative outcomes by strengthening coping skills and decreasing the effects of risky situations (Rutter, 1990). With the help of protective factors, individuals can generate positive outcomes despite adverse situations (Zautra, Hall, & Murray, 2010).

Although resilience does not prevent adverse life circumstances from emerging, it does provide the necessary tools for functionally coping with them (Hjemdal et. al., 2006). Protective factors of resilience can arise from various dimensions depending on personal and social differences. Those dimensions can be described as emotion regulation, positive emotion (Curtis & Cicchetti, 2007), family relationships (Bowlby, 1982), self-esteem, self-control (Wilson & Agaibi, 2006), and self-efficacy (Bandura, 1993). All of those dimensions have significant impacts on resiliency status because they prompt differences in individuals' judgments, emotions, thoughts, and perceptions related to coping skills (Curtis & Cicchetti, 2007). Accordingly, among other things, self-efficacy can contribute to an individual's level of psychological resilience.

When an individual faces adverse life events or experiences, self-efficacy play a significant role in determining his or her psychological status. *Self-efficacy* has been defined as the "beliefs in one's capabilities to organize and execute the courses of action required to manage prospective situations" (Bandura, 1997, pp. 2). In other words, if people believe in their ability to cope with difficult circumstances, then they can plan their actions accordingly. Self-efficacy is a dynamic process that can change over time (Bandura, 1997) and affects individual motivation, affect, and action (Bandura, 1993). Therefore, how an individual judges his or her capabilities is significant in terms of motivational, affective, and behavioral aspects.

Self-efficacy comes in various types, including academic self-efficacy, social self-efficacy, and emotional self-efficacy (Bandura, 1997). Among them, emotional self-efficacy is a chief focus of this study. Emotional self-efficacy indicates an individual's

beliefs about the transformation of negative emotions in the face of adverse situations (Pool & Qualter, 2011). Emotion is a feeling that accompanies certain thoughts, psychological and biological situations, and tendencies toward actions (Goleman, 1995). Emotional abilities inform people about whether an issue is positive or negative, which in turn forms individuals' attitudes and behaviors. Emotions relay messages to individuals so they can evaluate situations, act on cues from their emotions, and make decisions accordingly (Greenberg, 2002). Although emotion has great importance for humans, the perception of an emotion is more significant than the emotion itself (Goleman, 1995) because how a person feels an emotion can depend on how he or she sees and experiences it. Therefore, the ability to believe in one's emotional competence—in order words, emotional self-efficacy—is important in the context of understanding attitudes and behaviors.

Altogether, self-efficacy can prompt considerable change in the power of an individual's resilience and thus the ability to change his or her resilience mechanisms. In turn, it can help individuals to protect themselves from depression, anxiety arousal (Bandura, 1993), problematic behaviors, addictive behaviors, panic attacks, and phobias (William, 1995), as well as promote health and the immune system (Bandura, 1997).

Another factor that may have an impact on resilience is interpersonal sensitivity. Social settings are essential in peoples' lives, and psychological statuses are affected by interpersonal relationships and social interactions (Aronson & Wilson, 2005). According to Luthar (2006, p. 780), "Resilience rests, fundamentally, on relationships." Positive relationships can enhance psychological resilience and provide external sources for it (Libório & Ungar, 2014). However, relationships can also be a factor of vulnerability when stressful bonding with others emerges. For that reason, the extent to which psychological resilience can undergo changes when relationships are not positive and when sensitivity to social interactions is high should be considered.

Boyce and Parker (1989) defined *interpersonal sensitivity* as a personality trait that leads people to misinterpret others' attitudes and behaviors. *Interpersonal sensitivity* can also refer to fearing others' possible rejection or criticism (Bell & Freeman, 2014). In some research, the term *interpersonal rejection sensitivity* is used instead of *interpersonal sensitivity* to prevent confusion of the concept (Stafford, 2007).

Research has shown that interpersonal sensitivity depends on many factors, one of which is attachment style. Cummings–Robeau, Lopez, and Rice (2009) detected a significant relationship between parental and adult attachment that affects interpersonal sensitivity. In addition, Masten and Wright (2010) demonstrated that people who experience low attachment to parents and friends tend to be more interpersonally sensitive, which prompts a decreased level of resilience. Individuals with a negative self-view and low self-esteem also tend to be more sensitive to interpersonal relations (Otani, Suzuki, Ishii, Matsumoto, & Kamata, 2008). From the other direction, as numerous studies have shown, interpersonal sensitivity can precipitate psychological problems (Bell & Freeman, 2014), including depression,

anxiety disorders, burnout, eating disorders, and social avoidance (Bianchi, Schonfeld, & Laurent, 2015).

As mentioned, facing stressful and traumatic events is an inevitable part of life, and as such, psychological resilience becomes a significant. After all, an individual's ability to cope depends on his or her power of resilience. According to the literature, emotional self-efficacy and interpersonal sensitivity can generate psychological resilience. In response to that knowledge, this study investigates the effects of emotional self-efficacy and interpersonal sensitivity on the psychological resilience of young adults. Its purpose was to examine how perceiving emotions of the self and others, using emotions to facilitate thought, regulating emotions in the self and others, interpersonal awareness, need for approval, separation anxiety, timidity, fragile inner self, and understanding emotions, the emotional self, and others affect perceptions of the future, structural style, social competence, family cohesion, and social resources. In line with that aim, following questions were sought:

1. Is there a significant relation between interpersonal sensitivity and emotional self-efficacy?
2. Is there a significant relation among interpersonal sensitivity, emotional self-efficacy and psychological resilience?
3. Do interpersonal sensitivity and emotional self-efficacy significantly predict psychological resilience of young adults?

Method

Research Design

This research employed the relational screening model, which is used to determine the relationship between two variables or two datasets and the extent to those variables or datasets are related (Cohen, Manion, & Morrison, 2000).

Research Sample

The population of the research included students attending Marmara University in Istanbul, Turkey, during the 2014–2015 academic year. Participants were selected from undergraduate and graduate students studying in 16 faculties of the university by simple random sampling. The names of all faculties were written on slips of paper, which were put in a bag. Eight faculties were selected, after which one department from each faculty was selected in the same way. Selected departments were the Atatürk Education Faculty, Faculty of Technical Education, Faculty of Medicine, Faculty of Arts and Science, Faculty of Business Administration, Faculty of Fine Arts, Faculty of Law, and Faculty of Engineering. Other departments were Department of Guidance and Psychological Counseling, Department of Printery, Department of Medicine, Department of Turkish Language and Literature, Department of Business of Administration, Department of Painting, Department of Law, and Department of Mechanical Engineering.

The sample included 243 women (73.4%) and 88 men (26.6%), all volunteers, with a mean age of 21.46 years (SD = 3.48 years). Simple random sampling was used to

select participants from various departments, including Guidance and Psychological Counseling (13.3%), Teacher Training in Printery (13.6%), Medicine (10%), Turkish Language and Literature (9.7%), Business Administration (16.3%), Painting (11.5%), Law (11.2%), and Mechanical Engineering (14.5%).

Research Instruments and Procedure

Interpersonal Sensitivity Measure. Interpersonal sensitivity was measured with the Interpersonal Sensitivity Measure (IPSM) scale developed by Boyce and Parker (1989). The IPSM is a 36-item, Likert-type questionnaire that assesses pervasive and heightened attention and sensitivity to interactions with others. The scale generates a total score ranging from 36 to 144, with higher scores indicating greater interpersonal sensitivity. The measure has five subscales: interpersonal awareness, need for approval, separation anxiety, fragile inner self, and timidity. The Turkish version was adapted by Erozkan (2005). The IPSM has been found to have high internal consistency (α = .86) and test–retest reliability (r = 0.70). In the Turkish version, Cronbach's alpha coefficients were from .73 to .76 for the subscales and .81 for the whole scale.

Resilience Scale for Adults. The Resilience Scale for Adults was developed by Friborg et al. (2003) and revised by them in 2005. The scale has 33 items in six subscales: structural style, perception of the future, family cohesion, perception of self, social competence, and social resources. The Turkish version of the revised scale was given to two different samples of students and personnel by Basim and Cetin (2011). Test–retest reliabilities of the subscales were from .68 to .81. Cronbach's alpha for the subscales ranged from .66 to .81 for students and from .68 and .79 for personnel. Cronbach's alpha was .86 in both samples for the scale as a whole.

Emotional Self-Efficacy Scale. The Emotional Self-Efficacy Scale was developed by Kirk, Shutte, and Hine (2008). In its original form, the scale consists of 32 Likert-type items addressing perceiving emotions in the self and others (perceive), using emotions to assist thought (assist), understanding emotions and emotional knowledge in the self and others (understand), and regulating emotions in the self and others (regulate). The Turkish version was adapted by Totan, Ikiz, and Karaca (2011). Confirmatory factor analysis showed that the 4-factor structure of the scale was confirmed in a sample of Turkish university students. In that version, Cronbach's alpha was in the range of .70 to .83 for the subscales and .93 for the whole scale. The test–retest reliability was from .65 to .71 for the sub-scales and .62 for the whole scale.

The researchers emailed the departments' instructors to obtain their consent to facilitate the study. After receiving approval, the researchers made arrangements with the instructors and conducted the study in their classes. First, the purpose of the research was explained to the students, and volunteers were recruited to participate. A research assistant was always available to provide assistance to the students and to ensure confidential, independent responses. The participants completed the scales in approximately 40 min.

Data Analysis

This study aimed to investigate the predictive roles of emotional self-efficacy and interpersonal sensitivity on psychological resilience. This main goal of the study was analyzed in two-step process. First the Pearson's correlational analysis was conducted to test relationship between predictor and independent variables. Second, simultaneous multiple regression analysis was performed to address the predictive power of independent variables.

During a simultaneous multiple regression analysis, the assumptions for regression were first examined. The relationship between predictor and independent variables was linear with a normal distribution. In accordance with the principle of multicollinearity, the tolerance value was greater than .20, and the variance inflation factor of the predictor variables did not have a high correlation.

Results

As results of Pearson's correlational analysis of the relationship between dependent and predictor variables, mean and standard deviation values were reported (Table 1). Pearson's correlational analysis showed that psychological resilience had a negative linear relationship with need for approval, separation anxiety, fragile inner self, and interpersonal awareness and a positive linear relationship with using emotions to assist thoughts, perceiving emotions, understanding emotions, and emotion regulation. Perception of the future had a positive linear relationship with using emotions to assist thoughts, understanding emotions, perceiving emotions, and emotion regulation and a negative linear correlation with need for approval, separation anxiety, and interpersonal awareness.

Table 1
*Means, Standard Deviations, and Intercorrelations of the Six Dimensions of Psychological Resilience with Interpersonal Sensitivity and Emotional Self-Efficacy (*n = *331)*

Dimensions	*M*	*SD*	1	2	3	4	5	6	7	8	9
Psychological resilience	126.62	17.91	-.12*	-.18***	-.13*	-.23**	.03	.41***	.51***	.49***	.50***
Structural style	3.51	.85	-.04	-.08	-.08	-.11*	.01	.23***	.22***	.19***	.19***
Perception of the future	3.92	.86	-.16**	-.19***	-.11	-.21***	.01	.22***	.31***	.31***	.29***
Family cohesion	3.84	.77	-.06	-.06	.023	-.14*	.08	.15**	.22***	.24***	.28***
Perception of self	3.71	.74	-.24***	-.29***	-.31***	-.31***	-.07	.46***	.51***	.43***	.40***
Social competence	3.83	.78	.03	-.06	-.07	-.11*	-.03	.38***	.50***	.44***	.46***
Social resources	4.07	.69	-.06	-.09	-.02	-.09	.11*	.26***	.36***	.41***	.44***

Note. 1 = Interpersonal awareness, 2 = Separation anxiety, 3 = Fragile inner self, 4 = Need for approval, 5 = Timidity, 6 = Emotion regulation, 7 = Using emotion to assist thought, 8 = Understanding emotions, 9 = Perceiving emotions
*p < .05, **p < .001, ***p < .000

Family cohesion had positive relationships with perceiving emotions, understanding emotions, using emotions to assist thoughts, and emotion regulation and a negative linear relationship with need for approval. Furthermore, perception of the self positively correlated with using emotions to assist thoughts, emotion regulation, understanding emotions, and perceiving emotions and negatively correlated with need for approval, fragile inner self, separation anxiety, and interpersonal awareness. Another dependent variable, social competence, had positive relationships with using emotions to assist thoughts, perceiving emotions, understanding emotions, and emotion regulation, yet a negative correlation with need for approval. Lastly, social resources positively correlated with perceiving emotions, understanding emotions, using emotions to assist thoughts, and emotion regulation and negatively correlated with timidity (Table 1).

Multiple regression analysis was conducted to determine the best linear combination of interpersonal awareness, separation anxiety, fragile inner self, need for approval, regulating emotions in the self and others, using emotions to assist thought, understanding emotions in the self and others, and perceiving emotions in the self and others for predicting the score of psychological resilience and its protective factors. This combination of variables predicted psychological resilience, with three variables that significantly contributed to the prediction. Using emotions to assist thought contributed the most to predicting psychological resilience; need for approval and understanding emotions and emotional knowledge in the self and others also contributed to that prediction. The adjusted R^2 value was .34, which indicates that 34% of the variance in psychological resilience was explained by the model (Table 2).

Table 2

*Simultaneous Multiple Regression Analysis Summary for Interpersonal Sensitivity and Emotional Self-Efficacy Predicting Psychological Resilience (*N = *331)*

Variable	B	SEB	β
Interpersonal awareness	-1.485	2.682	-.033
Separation anxiety	-3.453	2.661	-.084
Fragile inner self	2.532	1.670	.090
Need for approval	-6.742	2.673	-.173*
Regulating emotions	1.260	1.792	.044
Using emotions to assist thought	7.160	2.089	.259**
Understanding emotions	4.578	2.232	.167*
Perceiving emotions	4.139	2.373	.147
Constant	85.78	7.41	

Note. $R^2 = .34$; $F(8.322) = 22.05$, $p < .000$
*$p < .05$, **$p < .001$, ***$p < .000$

Structural style was not significantly predicted by need for approval, regulating emotions in the self and others, using emotions to assist thought, understanding emotions and emotional knowledge in the self and others, or perceiving emotions in

the self and others (F[5.325] = 4.95, p < .000). The adjusted R^2 value was .06, which indicated that 6% of the variance in structural style was explained by the model.

According to multiple regression, perception of the future was predicted by using emotions to assist thought (β = .187, p < .001) and understanding emotions and emotional knowledge in the self and others (β = .182, p < .05; F[7.323] = 8.71, p < .001). The adjusted R^2 value was .14, meaning that 14% of the variance in perception of the future was explained by the model.

Combinations of variables predicted family cohesion with two variables. Perceiving emotions in the self and others (β = .234, p < .05) contributed the most to predicting family cohesion, although need for approval (β = .129, p < .05) also contributed (F[3.325] = 6.94, p < .001). The adjusted R^2 value was .08, meaning that 8% of the variance in family cohesion was explained by the model.

The combination of variables significantly predicted perception of the self, with four variables that significantly contributed to the prediction, as expected. Using emotions to assist thought contributed the most to predicting perception of the self; regulating emotions in the self and others, understanding emotions and emotional knowledge in the self and others, and separation anxiety also contributed to that prediction (Table 3). The adjusted R^2 value was .38, which means that 38% of the variance in perception of the self was explained by the model.

Social competence was significantly predicted by one variable (F[5.325] = 25.37, p < .000): using emotions to assist thought (β = .318, p < .000). The adjusted R^2 value was .27, meaning that 27% of the variance in social competence was explained by the model.

Social resources were significantly predicted by one variable (F[5.325] = 17.32, p < .000): perceiving emotions in the self and others (β = .310, p < .001. The adjusted R^2 value was .20, which indicates that 20% of the variance in social resources was explained by the model.

Table 3

*Simultaneous Multiple Regression Analysis Summary for Interpersonal Sensitivity and Emotional Self-Efficacy Predicting Perception of the Self (*N = *331)*

Variable	B	SEB	β
Interpersonal awareness	-.155	.107	-.083
Separation anxiety	-.223	.106	-.131*
Fragile inner self	-.076	.067	-.065
Need for approval	-.176	.107	-.110
Regulating emotions	.194	.072	.163**
Using emotions to assist thought	.399	.084	.349***
Understanding emotions	.184	.089	.162*
Perceiving emotions	-.120	.095	-.103
Constant	2.904	.297	

Note: R^2 = .38; F(8.322) = 26.10; p < .000
*p < .05, **p < .001, ***p < .000

Discussion and Conclusion

The present research examined the effects of emotional self-efficacy and interpersonal sensitivity on psychological resilience among young adults. Psychological resilience and its dimensions were analyzed according to each aspect of emotional self-efficacy and interpersonal sensitivity. The results showed that emotional self-efficacy and interpersonal sensitivity significantly predicted psychological resilience.

First, analysis revealed that using emotions to assist thoughts, need for approval, and understanding emotions all predicted psychological resilience the best, in that order. Researchers have posited that emotional intelligence bears significance in psychological resilience (Buyukbayram, Arabaci, Tas, & Varol, 2016; Ozer & Deniz, 2014). Other than emotional intelligence, belief in the capability of emotion is also enhances resilience. In this study, emotional self-efficacy was the most powerful predictor of psychological resilience. According to the findings, two of emotional self-efficacy's dimensions (i.e., using emotions to assist thoughts and understanding emotions) had powerful impacts on psychological resilience. As Schwarzer and Warner (2013) have indicated, self-efficacy makes people more resilient to adverse events. To cope with traumatic experiences, individuals need to believe they have the ability to overcome the situation (Bandura, 1997). Kirk, Schutte, and Hine (2008) have defined *emotional self-efficacy* as the belief in one's ability to transform the negative emotions of negative life experiences. Thus, emotion can be a tool to transform the negative impacts of a certain experience to make oneself more resilient in stressful life events. That finding is consistent with the present research and indicates a positive relationship between emotional self-efficacy and psychological resilience.

Interpersonal sensitivity is another predictive factor of psychological resilience. Individuals become vulnerable to psychological disorders such as depression because of their excessive sensitivity to interpersonal relationships (Boyce, Hickie & Parker, 1991). People with high sensitivity in their social relationships have a greater tendency toward psychological disorders involving somatic symptoms, as well as depression, substance abuse, and Internet addiction (Erozkan, 2011; Yilmaz, Hacihasanoglu, & Cicek, 2006; Herken, Bodur, & Kara, 2000). Moreover, according to Earvolino–Ramirez (2007), interpersonal sensitivity is a protective factor for resilience. One of its dimensions, need for approval, had a significantly powerful impact on psychological resilience. As estimated in that study, higher scores on need for approval indicated lower resilience, due to the avoidance of social relationships and settings because of negative evaluations, humiliation, rejection, and exclusion, among other things. Need for approval from others can affect self-esteem and, depending on the situation, can be a vulnerability or protective factor (Rolf & Johnson, 1990). People who need excessive approval from others tend to accept others' opinions and act accordingly, which creates higher vulnerability and lower resilience. Rutter (1990) explained that positive, healthy relationships with others encourage people's beliefs in their self-capabilities. The results of all of those studies are consistent with the findings of the present study: that a greater need for approval signifies less resilience.

Second, perception of the future was predicted by using emotions to assist thoughts and by understanding emotions. Research has demonstrated that despite adverse life events, people with positive emotions are likely to be goal oriented in their plans (Moskowitz, Folkman & Acree, 2003). The ability to use emotion for cognitive processes and to understand complex emotions contributes to making sense of emotions and acting accordingly (Fredrickson, 2001). LeBlanc, McConnell, and Monteiro (2015) explained that emotions can significantly impact individuals' perceptions of the world and cognitive states, which can shape their decisions and goals. When people believe in their emotional capabilities, their perception of the future as a protective factor becomes stronger, and they view the future optimistically (Fredrickson, 2001).

Third, perception of the self was predicted by using emotions to assist thoughts, regulation of emotions, understanding emotions, and separation anxiety. Repeat emotional achievements during life events make people believe that they can deal with difficult situations in the future (Fredrickson, 2001). Hjemdal et al. (2006) found that emotional stability is importance for personal strength and self-perception. Therefore, being emotionally capable raises people's self-efficacy and self-esteem. The current study's results showed parallels with the findings of Brown and Marshall (2001), which demonstrated that self-esteem and self-perception were highly interrelated with emotions. Another factor predicting perception of self was separation anxiety, which can create vulnerability and increase the likelihood of anxiety and mood disorders (Manicavasagar, Silove & Hadzi–Pavlovic, 1998). Moreover, Prince–Embury and Saklofkse (2013) observed that resilience has relationships to feeling loved, feeling accepted by others, and having healthy interpersonal relationships with peers and adults.

Fourth, regarding family cohesion, results indicated that family cohesion was predicted by perceiving emotions and need for approval, in that order of effect. Family cohesion is a level of mutual emotional bonding among family members and is likely to become stronger when emotional closeness to children is provided (Carruth, Tate, Moffett, & Hill, 1997). Therefore, perceiving emotional cues in one's self and other family members can facilitate bonding among family members. Family members' attitudes, relationships, loyalty, and support for each other fulfill children's emotional needs and need for approval (Hjemdal et al., 2011). When children receive balanced emotional closeness, they can form an identity separate from their family while also feeling togetherness with the family. If they cannot achieve adequate approval and emotional satisfaction, then they may feel sensitive and vulnerable in their need to seek approval, first from family and second from others outside the family (Minuchin, 1975). Individuals who receive balanced emotional closeness can more easily adapt to environments and cope with situations (Metcalf, 2011). Therefore, when children's needs are not properly met, the family becomes a factor of vulnerability, though it might otherwise be a protective factor.

Fifth, the study examined social competence, which was predicted by using emotions to facilitate thought. Social interactions were affected by emotional status. Blair et al. (2015) stated that all social interactions involve emotional ability and that a

connection between them exists. Other studies have shown that negative emotions can decrease social competence increase difficulties in social relationships, and cause social anxiety (Schwartz, Snidman, & Kagan, 1999).

Sixth and lastly, results demonstrated that social resources were predicted by perceiving emotions. In the face of an adverse, traumatic situation, people need to share their emotions with significant others in order to receive support, empathy, and understanding, which contribute to their ability to cope with the events. The ability to perceive emotions in the self and others and receive support from other people increase social resources and strengthen relationships (Kumpfer, 1999). That thinking is consistent with the finding that social resources were predicted by perceiving emotions in the self and others.

Limitations and Recommendations

A few limitations of the study should be stated. First, participants were young adults living in Istanbul, Turkey. The researchers chose Istanbul for its ability to represent the Turkish population, since the city has a cosmopolitan structure. However, current circumstances vary among regions in Turkey, and different family structures, events, and immigration can affect resilience. Because of its geographic position and cultural dynamics (e.g., civil wars in neighboring countries, the impact of internal and external migration, economic and politic instability), people in Turkey have likely encountered more traumatic experiences than their counterparts in other European countries. In Turkey, the in-group mentality is crucial due to the collectivist society. In that regard, when individuals evaluate and infer from their life experiences, their interpersonal relationships and other people's viewpoints play a significant role.

Furthermore, since traditional family experiences are highly common in Turkey, starting from the early periods of an individual's life, a person learns to evaluate and react to an experience based on emotional processes instead of cognitive ones. In that context, interpersonal relationships, emotions, and forms of emotional expression play a substantial role in determining the meaning and importance of a life experience. For that reason, interpersonal relationships and emotions have an essential place in the development of psychological resilience, which is a vital phenomenon for coping with negative life experiences.

Despite those limitations, the study has several strengths. In preventive counselling field, the dimension of psychological resilience had heretofore not been examined in detail. Therefore, the study marks an important attempt to fill that gap. Furthermore, emotional self-efficacy is a developing concept, and existing research on the topic is inadequate. This study showed that emotions constitute a major phenomenon in supporting resiliency factors. For those reasons, the research is considered to have made contributions to trauma studies.

References

Aronson, E., & Wilson, T. (2005). *Social psychology* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*, 117–148. doi: 10.1207/s15326985ep2802_3

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman and Company.

Basim, H. N., & Cetin, F. (2011). Yetiskinler icin psikolojik dayaniklilik olceginin guvenilirlik ve gecerlilik calismasi. [Reliability and validity studies of resilience scale for adults]. *Turk Psikiyatri Dergisi*, *22*(2), 104–114.

Bell, V., & Freeman, D. (2014). A pilot trial of cognitive behavioral therapy for interpersonal sensitivity in individuals with persecutory delusions. *Journal of Behavior Therapy and Experimental Psychiatry*, *45*(4), 441–446. doi:10.1016/j.jbtep.2014.06.001

Bianchi, R., Schonfeld, I. S., & Laurent, E. (2015). Interpersonal rejection sensitivity predicts burnout: A prospective study. *Personality and Individual Differences, 75*, 216–219. doi:10.1016/j.paid.2014.11.043

Blair, B. L., Perry, N. B., O'Brien, M., Calkins, S. D., Keane, S. P., & Shanahan, L. (2015). Identifying developmental cascades among differentiated dimensions of social competence and emotion regulation. *Developmental Psychology, 51*(8), 1062–1073. doi:10.1037/a0039472

Bowlby, J. (1982). *Attachment and loss* (2nd ed.). London: Hogarth.

Boyce, P., Hickie, I., & Parker, G. (1991). Parents, partners or personality? Risk factors for post-natal depression. *Journal of Affective Disorders, 21*(4), 245–255. doi:10.1016/0165-0327(91)90004-c

Boyce, P., & Parker, G. (1989). Development of a scale to measure interpersonal sensitivity. *Australian and New Zealand Journal of Psychiatry, 23*, 341–351.

Brown, J., & Marshall, M. (2001). Self-esteem and emotion: Some thoughts about feelings. *Personality and Social Psychology Bulletin, 27*(5), 575–584. doi:10.1177/0146167201275006

Buyukbayram, A., Arabaci, L. B., Tas, G., & Varol, D. (2016). Ogrenci hemsirelerinin duygusal zekâ ve sosyotropi-otonomi kisilik ozellikleri ile psikolojik dayanikliliklari arasindaki iliski. [The relationship between the emotional intelligence and sociotropy-autonomy: Personality traits of nursing students and their resilience.] *Izmir Katip Celebi Universitesi Saglik Bilimleri Fakultesi Dergisi*, 1(3), 29–37.

Carruth, A. K., Tate, U. S., Moffett, B. S., & Hill, K. (1997). Reciprocity, emotional well-being, and family functioning as determinants of family satisfaction in

caregivers of elderly parents. *Nursing Research, 46*(2), 93–100. doi:10.1097/00006199-199703000-00006

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5 th ed.). London: Routledge- Falmer.

Cummings-Robeau, T. L., Lopez, F. G., & Rice, K. G. (2009). Attachment-related predictors of college students' problems with interpersonal sensitivity and aggression. *Journal of Social and Clinical Psychology*, *28*(3), 364–391. doi:10.1521/jscp.2009.28.3.364

Curtis, W., & Cicchetti, D. (2007). Emotion and resilience: A multilevel investigation of hemispheric electroencephalogram asymmetry and emotion regulation in maltreated and non-maltreated children. *Development and Psychopathology, 19*, 811–811. doi: 10.1017/s0954579407000405

Earvolino-Ramirez, M. (2007). Resilience: A concept analysis. *Nursing Forum, 42*(2), 73–82.

Erozkan, A. (2005). Universite ögrencilerinin kişilerarasi duyarlilik ve depresyon duzeylerinin bazi degiskenlere gore incelenmesi. [Reliability and validity studies of resilience scale for adults.]. *Mugla Universitesi Sosyal Bilimler Enstitusu Dergisi, 14,* 129-155.

Erozkan, A. (2011). Sosyal kayginin kaygi duyarligi, benlik saygisi ve kisilerarasi duyarlık acisindan incelenmesi. [Investigation of social anxiety with regards to anxiety sensitivity, self-esteem, and interpersonal sensitivity.]. *İlkogretim Online, 10*(1), 338-347.

Fletcher, D., & Sarkar, M. (2013). Psychological resilience: A review and critique of definitions, concepts, and theory. *European Psychologist, 18*(1), 12–23. doi:10.1027/1016-9040/a000124

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist, 56*(3), 218–226. doi:10.1037/0003-066X.56.3.218

Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books.

Greenberg, L. (2002). *Emotion-focused therapy: Coaching clients to work through their feelings*. Washington, DC: American Psychological Association.

Herken, H., Bodur, S., & Kara, F. (2000). Universite ogrencisi kizlarda madde kullanimi ile kisilik ve ruhsal belirti iliskisi. [Assessment of the relationship between personality and psychiatric symptoms and substance abuse in the female university students.]. *Klinik Psikiyatri Dergisi, 3*, 40–45.

Hjemdal, O., Friborg, O., Braun, S., Kempenaers, C., Linkowski, P., & Fossion, P. (2011). The resilience scale for adults: Construct validity and measurement in a Belgian sample. *International Journal of Testing, 11*, 53–70. doi:10.1080/15305058.2010.508570

Hjemdal, O., Friborg, O., Stiles, T., Rosenvinge, J., & Martinussen, M. (2006). Resilience predicting psychiatric symptoms: A prospective study of protective factors and their role in adjustment to stressful life events. *Clinical Psychology & Psychotherapy, 13*(3), 194–201. doi: 10.1002/cpp.488

Kirk, B. A., Schutte, N. S., & Hine, D. W. (2008). Development and preliminary validation of an emotional self-efficacy scale. *Personality and Individual Differences*, *45*(5), 432–436. doi:10.1016/j.paid.2008.06.010

Kumpfer, K. L. (1999). Factors and processes contributing to resilience: The resilience framework. In M. D. Glantz & J. L. Johnson (Eds.), *Resilience and development: Positive life adaptations* (pp. 179–224). New York: Kluwer Academic/Plenum.

LeBlanc, V. R., McConnell, M. M., & Monteiro, S. D. (2015). Predictable chaos: A review of the effects of emotions on attention, memory, and decision making. *Advances in Health Sciences Education, 20*(1), 265–282. doi:10.1007/s10459-014-9516-6

Libório, R., & Ungar, M. (2014). Resilience as protagonism: Interpersonal relationships, cultural practices, and personal agency among working adolescents in Brazil. *Journal of Youth Studies, 17*(5), 682–696. doi:10.1080/13676261.2013.834313

Luthar, S. S. (2006). Resilience in development: A synthesis of research across five decades. In D. Cichetti & D. Cohen (Eds.), *Developmental Psychopathology Volume 3: Risk Disorder and Adaptation* (2 nd ed., pp. 696–739). Hoboken, NJ: John Wiley & Sons.

Manicavasagar, V., Silove, D., & Hadzi-Pavlovic, D. (1998). Subpopulations of early separation anxiety: Relevance to risk of adult anxiety disorders. *Journal of Affective Disorders, 48*, 181–190. doi:10.1016/S0165-0327(97)00170-5

Masten, A., Best, K., & Garmezy, N. (1990). Resilience and development: Contributions from the study of children who overcome adversity. *Development and Psychopathology, 2*, 425–425. doi:10.1017/S0954579400005812

Masten, A., & Wright, M. (2010). Resilience over the lifespan: Developmental perspectives on resistance, recovery, and transformation. In J. Reich, A. Zautra, & J. Stuart (Eds.), *Handbook of adult resilience* (pp. 213–237). New York: Guilford Press.

Metcalf, L. (2011). *Marriage and family therapy a practice-oriented approach*. New York, NY: Springer Pub.

Moskowitz, J. T., Folkman, S., & Acree, M. (2003). Do positive psychological states shed light on recovery from bereavement? Findings from a 3-year longitudinal study. *Death Studies*, *27*, 471–500. doi:10.1080/07481180302885

Minuchin, S. (1975). A conceptual model of psychosomatic illness in children. *Archives of General Psychiatry, 32*(8), 1031.doi:10.1001/archpsyc.1975.01760260095008

Otani, K., Suzuki, A., Ishii, G., Matsumoto, Y., & Kamata, M. (2008). Relationship of interpersonal sensitivity with dimensions of the temperament and character inventory in healthy subjects. *Comprehensive Psychiatry*, *49*(2), 184–187. doi:10.1016/j.comppsych.2007.09.004

Ozer, E. & Deniz, M. E. (2014). Universite ogrencilerinin psikolojik saglamlik duzeylerinin duygusal zeka acisindan incelenmesi. [An investigation of university students' resilience level on the view of trait emotional EQ.]. *Ilkögretim Online*, *13*(4), 1240–1248.

Pool, L. D., & Qualter, P. (2011). The dimensional structure of the emotional self-efficacy scale (ESES). *Australian Journal of Psychology, 64*(3), 147–154. doi:10.1111/j.1742-9536.2011.00039.x

Prince-Embury, S., & Saklofske, D. (2013). Translating resilience theory for application: Introduction. In S. Embury & D. Saklofske (Eds.), *Resilience in children, adolescents, and adults: Translating research into practice* (pp. 3–7). New York: Springer Science Business Media.

Rolf, J. & Johnson, J. (1990). Protected or vulnerable: The challenges of AIDS to developmental psychopathology. In J. Rolf, A. Masten, D. Cicchetti, K. Nuechterlein, & S. Weintraub (Eds.), *Risk and protective factors in the development of psychopathology* (pp. 384–404). Cambridge, MA: Cambridge University Press.

Rutter, M. (1990). Psychosocial resilience and protective mechanisms. In J. Rolf, A. Masten, D. Cicchetti, K. Nuechterlein, & S. Weintraub (Eds.), *Risk and protective factors in the development of psychopathology* (pp. 181–214). Cambridge, MA: Cambridge University Press.

Schwartz, C. E., Snidman, N., & Kagan, J. (1999). Adolescent social anxiety as an outcome of inhibited temperament in childhood. *Journal of the American Academy of Child & Adolescent Psychiatry, 38*, 1008–1015. doi:10.1097/00004583-199908000-00017

Schwarzer, R., & Warner, L. M. (2013). Perceived self-efficacy and its relationship to resilience. In S. Prince-Embury & D. H. Saklofske (Eds.), *The Springer series on human exceptionality: Resilience in children, adolescents, and adults: Translating research into practice* (pp. 139-150). Springer New York.

Stafford, L. (2007). Interpersonal rejection sensitivity: Toward exploration of a construct. *Issues in Mental Health Nursing, 28*(4), 359–372. doi:10.1080/01612840701244250

Totan, T., Ikiz, E., & Karaca, R. (2011). Duygusal oz-yeterlik ölceginin Turkceye uyarlanarak tek ve dort faktorlu yapisinin psikometrik ozelliklerinin incelenmesi [Turkish adaptation of the emotional self-efficacy scale: Psychometric properties of the single and four-factor structure.]. *Buca Egitim Fakultesi Dergisi*, *28*, 71–95.

William, S. (1995). Self-efficacy and anxiety and phobic disorders. In J. Maddux (Ed.), *Self-efficacy, adaptation and adjustment: Theory, research, and application* (pp. 69–107). New York: Plenum Press.

Wilson, J. & Agaibi, C., & (2006). The resilient trauma survivor. In J. Wilson (Ed.), *The posttraumatic self-restoring meaning and wholeness to personality* (pp. 369–399). New York, NY: Taylor & Francis Group, LLC.

Yilmaz, S., Hacihasanoğlu, R., & Cicek, Z. (2006). Hemsirelerin genel ruhsal durumlarinin incelenmesi. [The Analysis of the General Psychological Conditions of Nurses.]. *Sted, 15*(6), 92–97.

Zautra, A., Hall, J., & Murray, K. (2010). Resilience: A new definition of health for people and communities. In J. Reich, A. Zautra, & J. Hall (Eds.), *Handbook of adult resilience* (pp. 3–34). New York: Guilford Press.

Genç Yetişkinlerde Psikolojik Dayanıklılığın Duygusal Öz-Yeterlik ve Kişilerarası Duyarlılık Perspektifinden İncelenmesi

Atıf:

Aydogdu, B. N., Celik, H., & Eksi, H. (2017). The predictive role of interpersonal sensitivity and emotional self-efficacy on psychological resilience among young adults. *Eurosian Journal of Educational Research*, *69*, 37-54. http://dx.doi.org/10.14689/ejer.2017.69.3

Özet

Problem Durumu: Yaşamın kaçınılmaz gerçeklerinden biri travmatik yaşantılardır. Faklı gelişim dönemlerinde hemen hemen herkes şu veya bu şekilde travmatik sayılacak düzeyde deneyimlerden geçer. Bu deneyimlere karşı verilen tepkiler, dışsal ve içsel faktörlere göre değişiklik gösterir. Olumsuz deneyimlerin olumlu sonuçlara dönüştürülmesi için sahip olunması gereken koruyucu faktörler, psikolojik dayanıklılık kavramını gün yüzüne çıkarır. Psikolojik dayanıklılık, pek çok faktöre bağlı olarak gelişen bir mekanizmadır. Bu mekanizmayı etkileyen önemli faktörlerden biri, travmatik olaylar sonucunda oluşabilecek olumsuz duyguları dönüştürebilme inancı olarak kabul edilen duygusal öz-yeterliliktir. İnsanların duygularını algılaması, duygularını anlaması, bunları düşünceye yardımcı bir araç olarak kullanması ve duygularını düzenleyebilmesi duygusal öz-yeterliği oluşturan temel bileşenlerdir. Psikolojik dayanıklılık üzerinde etkisi olduğu düşünülen bir diğer faktör ise kişilerarası duyarlılıktır. Kişilerarası duyarlılık diğer kişilerin tutum ve tavırlarının yanlış yorumlamasına neden olan kişilik özelliği olarak tanımlanır. İlişkilerde olumlu ve sağlıklı iletişime sahip olmak, güçlü kişilerarası bağlantıların oluşumuna zemin hazırlar. Bu durum, psikolojik dayanıklılıkları açısından bireylere koruyucu bir faktör olarak kalkan görevi görür. Öte yandan başkalarıyla kurulan olumsuz ve gerginlik unsuru olan bağlar, psikolojik dayanıklılık için risk faktörü olarak kabul edilir. Bu açılardan değerlendirildiğinde, duygusal öz-yeterlik ve

kişilerarası duyarlılığın, psikolojik dayanıklılık üzerinde önemli bir etki gücü olduğu düşünülmektedir.

Amaç: Bu bağlamda bu çalışma, duygusal öz-yeterlik ve kişilerarası duyarlılık özelliklerinin genç yetişkinlerin psikolojik dayanıklılıkları üzerinde bir etki gücüne sahip olup olmadığı ve sahipse bu etkinin derecesini inceleme üzerine yapılandırılmıştır.

Yöntem: Araştırmanın örneklemini 2014-2015 eğitim öğretim döneminde Marmara Üniversitesi'nde öğrenim görmekte olan genç yetişkinlerden oluşturmaktadır. Basit seçkisiz örnekleme yönteminin kullanıldığı çalışmada, yaşları 18-34 arasında değişen *(ss:3.48,* $\overline{X}$ *: 21.26)* 331 genç yetişkine (243 bayan, 88 bay) ulaşılmıştır. İlişkisel tarama modelinin kullanıldığı bu çalışmada araştırmanın amaçlarını test etmek için Eş Zamanlı Çoklu Regresyon analizinden yararlanılmıştır.

Bulgular: Araştırmanın amaçları doğrultusunda yapılan analizler neticesinde duygusal öz-yeterlik (duyguları anlama ve duyguları düşünceye destekleyici olarak kullanma alt boyutları) ile kişilerarası duyarlılığın (onaylanma ihtiyacı alt boyutu) birlikte, psikolojik dayanıklılığın toplam varyansının % 34'ünü açıkladığı sonucuna ulaşılmıştır. Psikolojik dayanıklılığın alt boyutlar bazında ise elde edilen bulgular şu şekildedir: duyguları düşünceye destekleyici olarak kullanma ve duyguları anlama birlikte gelecek algısının toplam varyansının %14'ünü, onaylanma ihtiyacı ve duyguları algılama boyutları birlikte aile uyumunun toplam varyansının %8'ini, ayrılma anksiyetesi, duygu düzenleme, duyguları düşünceye destekleyici olarak kullanma ve duyguları anlama boyutları birlikte kendilik algısının toplam varyansının %38'ini, duyguları düşünceye destekleyici olarak kullanma boyutu tek başına sosyal yeterliliğin toplam varyansının %27'sini ve son olarak duyguları anlama boyutu yine tek başına sosyal kaynakların toplam varyansının %20'sini anlamlı şekilde yordamıştır. Buna karşın psikolojik dayanıklılığın diğer alt boyutu olan yapısal sitil, duygusal öz-yeterlik ve kişilerarası duyarlılık tarafından anlamlı bir şekilde açıklanmamıştır

Sonuç ve Öneriler: Bu çalışma genç yetişkinlerin duygusal öz-yeterlik ve kişilerarası duyarlılıklarının sahip oldukları psikolojik dayanıklılık özellikleri üzerinde belirleyici etkilerinin olduğunu açıkça ortaya koymaktadır. Travmatik deneyim karşısında yaşanan zorluklarla başa çıkabilmek için, insanlar durumun üstesinden gelme yeteneğine sahip olduklarına inanmak zorundadırlar. İnsanlar kişilerarası ilişkilere aşırı duyarlı hale geldikleri zaman çeşitli psikolojik rahatsızlıklar karşısında savunmasız kalırlar. Kişilerarası ilişkilerde beklenilen onaylanma ihtiyacı, düşük düzey dayanıklılığın bir göstergesidir. Bu durum ise olumsuz değerlendirilme, küçük düşürülme, reddedilme, dışlama vb. nedenlerden ötürü sosyal ilişkilerden kaçınma ve uzaklaşmaya yol açar. Araştırmanın önemli bulgularından biri, duyguları düşünceye yardımcı bir araç olarak kullanma ve duyguları anlamanın gelecek algısı üzerindeki belirleyici etkisidir. Duygular kişinin bilişsel algılarını dolaysıyla kararlarını ve hedeflerini belirler, dolayısıyla gelecek algısını şekillendirir. Araştırmanın bir diğer önemli bulgusu duyguları düşünceye yardımcı bir araç olarak kullanma, duygu düzenleme, duyguları anlama ve ayrılma anksiyetesinin birlikte

kendilik algısın üzerinde yordayıcı etkisi olduğu yönündedir. Yaşam olayları karşısında duygusal başarıların tekrarlanması, insanları gelecekte zor durumlarla baş edebileceklerine inanmalarını sağlar. Duygusal açıdan tutarlı ve istikrarlı olmanın kişisel güç ve kendilik algısı üzerinde belirleyici bir rolü vardır. Bu nedenledir ki, duygusal açıdan yetenekli olabilme insanların öz-yeterlik ve benlik saygısı olumlu yönde etki eder. Buna karşın kişilerarası ilişkilerde deneyimlenen ayrılma anksiyetesi benlik algısını zayıflatır. Çünkü ayrılma kaygısı insanlar için savunmasızlık yaratır, kaygı ve duygudurum bozuklukları olasılığını artırır.

Araştırmada elde edilen bir diğer sonuç ise duyguları anlama ve onaylanma ihtiyacının aile uyumu üzerindeki yordayıcı etkisidir. Çocukların ihtiyaçları tam karşılanmadığında, koruyucu bir faktör olması gereken aile ne yazık ki bir savunmasızlık/kırılganlık faktörüne dönüşür.

Çalışmanın bir diğer bulgusu ise duyguları düşünceye yardımcı bir araç olarak kullanmanın sosyal yeterlik üzerindeki yordayıcı etkisidir. Duygular, sosyal etkileşimler üzerinde belirleyicidir. Tüm sosyal etkileşimler duygusal yeteneği içerir ve aralarında güçlü bir bağlantı vardır. Olumsuz duygular sosyal yeterliliğin azalmasına, sosyal ilişkilerde güçlükler yaşanmasına ve sosyal kaygıya yol açar.

Bu çalışmada elde edilen son bulgu ise duyguları anlamanın psikolojik dayanıklılığın alt boyutlarından sosyal kaynaklar üzerinde yordayıcı etkisinin olduğu yönündedir. Olumsuz ve travmatik bir durum karşısında, insanların duygularını başkalarıyla paylaşarak destek alabilmeleri gerekebilir. Kendinin ve başkalarının duyguları algılama ve diğer insanlardan destek alma sosyal kaynakları artırır ve ilişkileri kuvvetlendirir.

Coğrafi konumu nedeniyle savaş ve terör olaylarıyla karşılaşma, sosyo-politik alandaki hızlı ve beklenmedik değişimler ve aile içindeki dalgalanmalar gibi etmenler ülkemizdeki bireylerin travmaya maruz kalma olasılıklarını ne yazık ki güçlendirmektedir. Bu noktada psikolojik dayanıklılığın önemi daha da güçlü olur. Bu çalışma genç yetişkinlerin psikolojik dayanıklılık özelliklerinin duygusal öz-yeterlik ve kişilerarası duyarlılık özelliklerine göre şekillendiği açıkça ortaya koymaktadır. Ne var ki psikolojik dayanıklılık bu bağlamda ele alınarak ayrıntılı bir şekilde incelenmemiştir. Bu çalışma alandaki bu boşluğu bir açıdan doldurabilecek niteliktedir. Ayrıca, duygusal öz-yeterlik gelişmekte olan bir kavramdır ve amprik çalışmalarla desteklenmesine ihtiyaç vardır. Bu çalışma duyguların, psikolojik dayanıklılığı açıklamada temel olgu olduğunu işaret etmekte olup alanda yapılacak yeni çalışmalarla desteklenmesi gerekmektedir.

Anahtar Sözcükler: Psikolojik dayanıklılık, kişilerarası duyarlılık, öz-yeterlik, genç yetişkinler.

Eurasian Journal of Educational Research 69 (2017) 55-74

Eurasian Journal of Educational Research
www.ejer.com.tr

The Effect of Self-Regulated Learning Strategies on Academic Achievement: A Meta-Analysis Study*

Binnur ERGEN[1] Sedat KANADLI[2]

ARTICLE INFO

Article History:

Received: 17 November 2016

Received in revised form: 07 February 2017

Accepted: 14 May 2017

DOI: http://dx.doi.org/10.14689/ejer.2017.69.4

Keywords

self-regulated learning, academic achievement, meta-analysis, moderator analysis

ABSTRACT

Problem Statement: Self-regulated learning strategies (cognitive, metacognitive, resource management, and motivational strategies) influence students' academic achievement, conceptual understanding, and motivation. Reviewing the national literature about self-regulated learning strategies, studies have indicated both significant and insignificant effects on academic achievement; however, no meta-analysis studies have been carried out. **Purpose of Study:** The aim of this study is to calculate the common effect size of empirical and relational studies conducted in Turkey between 2005-2014 that investigated the effect of (or relationship with) self-regulated learning strategies on academic achievement, and to determine whether the common effect size shows a significant difference in terms of course type, self-regulated learning strategy type, school level, and study design. **Method:** A meta-analytical review method was employed to combine the outcome of independent empirical or relational studies. The studies included in this review were collected from the CoHE National Thesis Archive, ULAKBIM, Google Academic, ERIC, and EBSCO databases. A total of 47 studies were assessed in accordance with the inclusion criteria, and 21 studies were included in this study. Cohen's d coefficient was calculated for the effect size in this study. **Findings and Results:** As the heterogeneity among the effect sizes of the studies was high ($Q > \chi^2$, $p < .05$), the common effect size was calculated in accordance with the random effects model. As a result of the meta-analysis, it was determined that self-regulated learning strategies had a "large" effect ($d = 0.859$) on academic achievement. Moreover, the calculated common effect size showed no significant difference according to the type of self-regulated learning strategy, course type, study design, and school level. **Recommendation:** As self-regulated learning strategies exhibit a substantial effect on students' academic achievement, it is recommended that preservice and in-service teachers should learn how to implement these strategies in their lessons to increase their students' performance. For this purpose, professional development programs should be designed for teachers.

* This study was presented at '3rd International Eurasian Educational Research Congress in Mugla 2016.

[1] Mersin University, Conservatoire, TURKEY, binnurergen@mersin.edu.tr

[2] Corresponding Author: Sedat KANADLI, Mersin University, Educational Faculty, TURKEY, skanadli@mersin.edu.tr

Introduction

Technology and knowledge are rapidly improving in today's world. It is important that students acquire knowledge and skills by taking responsibility of self-learning in order to become individuals who learn to learn. They are aware of what and how they have learned, and their deficiency of knowledge and skills while learning to learn, which enables them to self-regulate. As a result, individuals' academic success increases (Zimmerman, 1990) and they acquire skills needed to be a lifelong learner. In this regard, self-regulation is defined as individual-controlled emotions, ideas, and behaviours exhibited to reach particular goals and that occur at different levels and features in each developmental period (Zimmerman, 2001). In other words, self-regulation is a process of influencing, directing, and managing one's own behaviours (Senemoglu, 2005, 231).

Studies about self-regulation, a basic concept of the Social Cognitive Theory, have been undertaken by scholars such as Albert Bandura, Barry Zimmerman, Dale Schunk, Paul Pintrich, and Frank Pajares since the 1980s (Sakiz, 2014). This theory claims that a behaviour is developed not only through experiences, but also through observing and understanding the rewarded or punished behaviours of others as a model (Bandura, 1971). Such a model consists of processes of paying attention to the behaviour, keeping it in mind, willing to reflect on it, and ultimately, performing it (Bandura, 1986). Individuals can observe their own behaviours, compare the behaviours according to their own criteria, and regulate themselves within this period (Senemoglu, 2005, 233). Therefore, the increase in the number of the studies focusing on how students regulate themselves in academic environments has resulted in the concept of self-regulated learning (Dinsmore, Alexander & Loughlin, 2008).

Self-regulated learning is defined as an active and constructive process in which individuals set their own learning goals, regulate their cognition, motivation, and behaviours, and are directed and limited by their own goals and contextual features around (Pintrich, 2000). It helps them get to know themselves, become wise and determinant in their learning-oriented approaches (Zimmerman, 1990). Students getting to know themselves can be viewed as a process that is associated with metacognitive skills, acquiring knowledge with cognitive skills, and obtaining the ability to motivated themselves and manage their environment effectively. For this reason, self-regulated learning model is explained in four categories: cognitive, metacognitive, resource management, and motivational strategies (Pintrich & De Groot, 1990; Pintrich, 1999).

Cognitive strategies are associated with behaviours and cognitive processes students employ during their learning experiences to complete a task or achieve a purpose about an academic subject (Boekaerts, 1996). Cognitive strategies cover sub-strategies of rehearsal, elaboration, and organizational strategies (Pintrich, 2000). Metacognitive strategies involve predicting, planning, monitoring, and evaluation, which help individuals control and regulate their own cognitive processes (Lucangeli

& Cornoldi, 1997). Resource management strategies embody such strategies as controlling and managing one's time and study environment, effort, peer cooperation, and help-seeking (Pintrich, 1999). However, as it is important that students are motivated to apply these strategies, motivational strategies covering intrinsic values, self-efficacy, and test anxiety (Pintrich & De Groot, 1990) stand as the last dimension of self-regulated learning.

Regarding national and international literature about self-regulated learning, many studies have examined this concept in accordance with various variables. Various studies reveal that self-regulated learning enhances students' academic success (Atas, 2009; Camahalan, 2006; Cazan, 2014; Dikbas & Hasirci, 2008; Gulay, 2012; Rowe & Rafferty, 2013), while others argue it has no significant relationship with or effect on academic success (Haslaman & Askar, 2007; Shaine; 2015; Ustun, 2012). However, meta-analysis studies of self-regulated learning are seen only in international literature. Such meta-analysis studies have found that self-regulated learning strategies affect students' academic success (Dignath & Buttner, 2008; Hattie, Biggs & Purdie, 1996), reading comprehension (Chiu, 1998), and motivation (Dignath & Buttner, 2008) at a moderate level (d= 0.50-0.80).

There are many primary studies about self-regulated learning in Turkey, and their various findings conflict with one another. Therefore, these studies must be examined through a meta-analysis to reach a more definitive conclusion. The purpose of the study is to calculate the effect size of studies that have analysed the effect of self-regulated learning on academic success or its relationship with academic success in Turkey, and to reveal whether self-regulated learning strategies differentiate significantly according to study type, school level, study design, and design type. One of the aims of this meta-analysis is for future studies about self-regulated learning in Turkey to use this study as a reference.

The effect of self-regulated learning on academic achievement was examined in this study in terms of five categorical moderators: self-regulated learning strategy, study type, school level, research design, and course type. That the determination of the most effective type of self-regulated strategy, school level, and course on academic achievement are considered to be important for guiding practitioners (teachers or academicians). Similarly, revealing the impact of research design and study type on academic achievement is significant, as it will show researchers the type and design of research that should be executed. The answers to following questions were sought in this regard:

1. What is the effect of teaching based on self-regulated learning on academic achievement?
2. Does the effect of teaching based on self-regulated learning on academic achievement show a significant difference according to self-regulated learning strategies?
3. Does the effect of teaching based on self-regulated learning on academic achievement show a significant difference according to research design, course type, and school level?

Method

Research Design

A meta-analysis method was employed in this study to examine the effect of self-regulated learning on academic achievement. Meta-analysis refers to procedures used to combine the results obtained from individual studies and consists of following (Figure 1) stages (Cooper, 2010, 12):

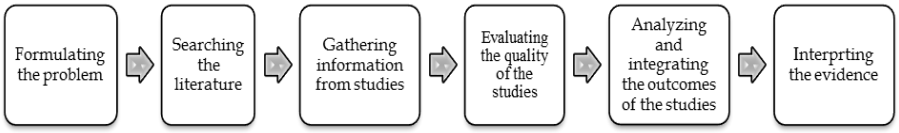


Figure 1. Steps of meta-analysis

Literature Search Procedure

The studies included in this research were obtained from the CoHE National Thesis Center (2015), ULAKBIM (2015), Google Academic (2015), ERIC (2015), and EBSCO (2015) databases. The database search was conducted between February 2015 and May 2015. While searching, these keywords were entered in both Turkish and English: "self-regulating learning", "self-regulated learning", "learning strategies", "learning strategies and academic success", "self-regulated learning and academic achievement", "metacognitive strategies", "metacognition", and "social cognitive theory". Relevant literature was scanned through references of the studies obtained. In total, 115 studies were attained about the literature concerning the effect of self-regulated learning on academic success and the relationship between them. After limiting the studies to those published between 2005-2014 and eliminating duplicated studies, 47 were left.

Study Inclusion and Exclusion Criteria

Quantitative studies published between 2005-2014 and about the effect of self-regulated learning on academic achievement were examined in the context of this study. Inclusion criteria were as follows: *(i)* Must be an article, thesis or assertion carried out in Turkey between 2005-2014 in an empirical and relational design. *(ii)* Must investigate the relationship of self-regulated learning with academic success or its effect on academic success. *(iii)* Empirical studies must have a sample size (N), means score ($\overline{X}$), and standard deviation, while relational studies must have sample size and Pearson correlation coefficient. *(iv)* Studies must employ parametric tests (t-test, F test, Pearson Correlation Coefficient, etc.).

Forty-seven studies examining the effect of self-regulated learning on academic success or its relationship with academic success were identified according to the criteria above. Some of these studies were eliminated, as six were conducted in a qualitative design, nine had limited access, and 11 were published as both a thesis

and an article. As a result, 21 studies (nine empirical and 12 relational) about self-regulated learning were identified to review. A flowchart showing the inclusion process of the studies obtained through the literature review into the meta-analysis is given in Figure 2:

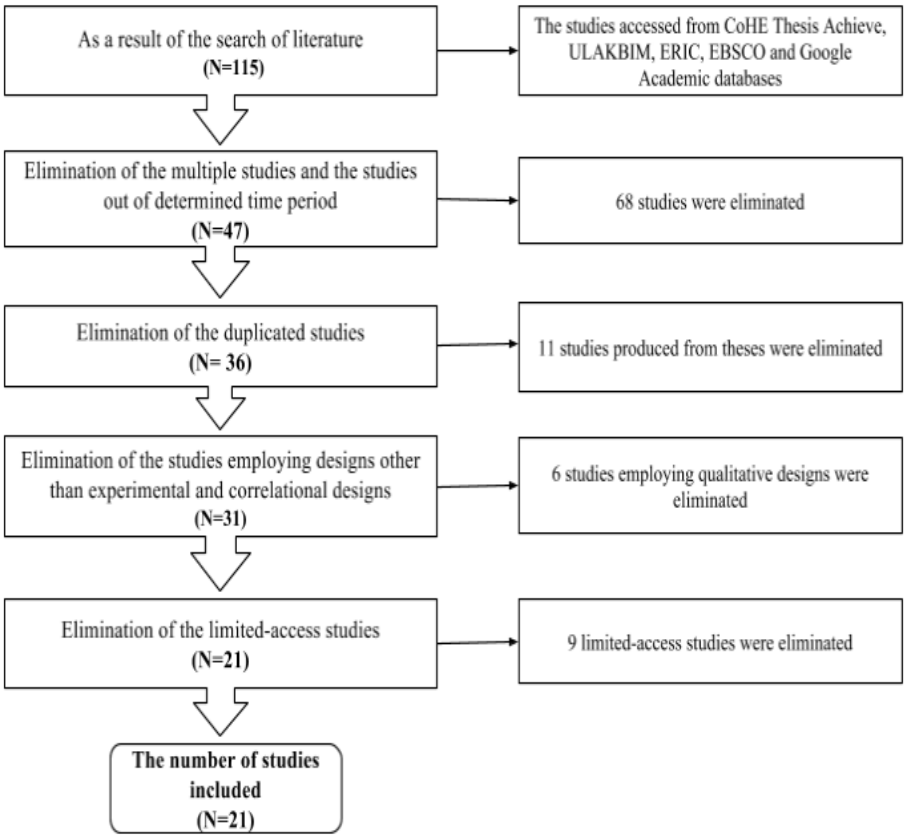


Figure 2. Flowchart of inclusion and exclusion of studies

Coding of Study Characteristics

The studies chosen according to inclusion criteria were coded in terms of their author, date, type, design, and course type. Type of study was coded according to whether it was *a thesis*, *article,* or *proceeding*, while study design was coded as *relational* (if examining the relationship between self-regulated learning and academic success) or *empirical* (if investigating the effect of self-regulated learning on academic success). Course types were placed in four categories: *Science, Language* (Turkish and English), *Social Sciences* (Social Sciences and Teaching Methodology), and *Mathematics*. Rehearsal, elaboration, organisation, and critical thinking were coded as *Cognitive Strategies*; predicting, planning, monitoring, and evaluation as *Metacognitive Strategies*; controlling and managing time and study environment, effort, peer cooperation, and help-seeking as *Resource Management Strategies*; value, expectation

and affective factors as *Motivational Strategies*. Since all the strategy types are employed together in some studies, they were coded as *Self-Regulated Strategies*.

Five studies (22.72%) were chosen at random and given to another coder to calculate inter-coder reliability. An equalisation rate over 80% is accepted as high enough (Miles & Huberman, 1994). After the coding process, inter-coder reliability was found to be 100%.

Data Analytic Strategy

In this study, Cohen's d effect size index defined as the standardised means difference was employed. Cohen's d is calculated by dividing the difference between raw means by standard deviation. According to Cohen (1988), the effect size is accepted as "no effect" if the d-value is up to 0.20, "low" between 0.20-0.50, "moderate" between 0.50-0.80, and "large" over 0.80.

After calculating the effect sizes of individual studies in the meta-analysis method, the effect sizes were combined through a statistical method and the common effect size was calculated. Two models are utilised in calculating common effect size: fixed and random effects models. Although it is disputable which model is to be used, there are two approaches: First one is a test of heterogeneity. This test reveals whether variance observed in effect sizes (Q) significantly differentiates from the variance arising from sampling error (χ^2) (Cooper, 2010, 85). Therefore, the Q-value must be found and compared to the degree of freedom value (df=n-1) in the χ^2 table. If $Q < \chi^2$ ($p>.05$), the effect sizes of studies are interpreted as homogeneous and the combination process is applied according to the fixed effects model. If $Q > \chi^2$ ($p< .05$), the effect size is interpreted as heterogeneous and the random effects model is employed.

Hedges and Pigott (2001) stated that the chi-square test (χ^2) lacks statistical power to measure variance between studies. For that reason, the model to be employed should be determined according to the inference that the researcher wants to conclude (Hedges & Vevea, 1998) and the sampling method of the studies (Borenstein, Hedges, Higgins & Rothstein, 2009, 86). In this study, the random effects model more appropriately fits the purpose of the researcher. Nevertheless, a heterogeneity test was executed, as primary studies were identified through a literature review and generalisation to the universe is an aim of this meta-analysis.

There are sub-groups independent of each other in the studies included in this study. In some of them, the combined effect of self-regulated learning on academic success was reported, while the effect of cognitive, metacognitive, motivational and resource management of self-regulated learning strategies was examined individually in others. Therefore, the studies must be utilised as analysis units instead of sub-groups. The effect size of a study is calculated by combining raw data of sub-groups reported individually (Borenstein et al., 2009, 219), and the common effect size is obtained by the help of combined effect sizes. Later, the studies reporting only the combined effect were excluded from the analysis. Furthermore, whether the common effect size showed a significant difference according to self-

regulated learning strategy types was examined by comparing sub-groups to each other. Moreover, a categorical moderator analysis was applied to reveal whether the common effect size of self-regulated learning on academic success showed a significant difference regarding study design, course type, and school level. Whether the moderator was significant was determined by the significance level of $Q_{between}$ value under the random effects model.

Moderator analysis, funnel plot, Rosenthal's Fail-safe N, and Egger's Regression Intercept tests were executed to reveal the existence of publication bias and its effect on the analysis. Comprehensive Meta-Analysis Software (CMA) 2.0 was utilised in data analysis.

Results

Characteristics of the Included Studies

The sample size of the empirical studies included in this study consists of 770 individuals, while 4583 individuals are included in the relational studies, which makes the total sample size 5353 people. Descriptive features of the studies included in the meta-analysis are given in Table 1.

Table 1

*Frequency Distribution of Studies by Course Type, Study Type, Self-Regulated Learning Strategies, Research Design, and Investigated Variable*s

Variable	Frequency (N)	Percentage (%)
Course type		
Science	4	19.04
Mathematics	8	38.10
Social Sciences	6	28.57
Language	3	14.28
Study type		
Thesis	6	28.57
Article	14	66.67
Assertion	1	4.76
Self-regulated Learning Strategies		
Cognitive	11	40.74
Metacognitive	10	37.03
Resource Management	2	7.41
Motivational	4	14.81
Research design		
Relational design	12	57.14
Empirical design	9	42.86
School level		
Primary	4	19.05
Secondary	8	38.10
High school	1	4.76
Undergraduate	8	38.10

It is seen in Table 1 that 19.04% (*f*=4) of studies were conducted in Science, 38.10% (*f*=8) in Mathematics, 28.57% (*f*=6) in Social Sciences, and 14.28% (*f*=3) in Language courses. Concerning the study type, 28.57% of the studies (*f*=6) were thesis, 66.67% (*f*=15) were articles, and 4.76% (*f*=1) were assertions. With regard to the self-regulated learning strategies, 40.74% (*f*=11) of them were designed in line with cognitive strategies, 37.03% (*f*=10) with metacognitive strategies, 7.41% (*f*=2) with resource management, and 14.81% (*f*=3) with motivational strategies. It was reported that 42.86% (*f*=9) of these studies were empirical while 57.14% (*f*=12) were relational. With regard to the school type, 19.05% (*f*=4) were conducted in primary schools, 38.10% (*f*=8) in secondary schools, 4.76% (*f*=1) in high schools, and 38.10% (*f*=8) in undergraduate programs.

Heterogeneity Test

A heterogeneity test was applied to reveal whether the variance observed in the effect sizes of individual studies demonstrated a significant difference from the variance expected of sampling error, and to determine which model was to be used to combine effect sizes accordingly. Heterogeneity test results are shown in Table 2.

Table 2

Number, Standard Error, Heterogeneity, Effect Size, and Confidence Interval According to Effect Model of Studies

Model	N	Effect Size	Std. Error	95% Interval		Heterogeneity			
				Lower Limit	Upper Limit	Q	df	P	I^2
Fixed Effect	21	0.751	0.017	0.718	0.784	740.77	20	0.00	97.3
Random Effect	21	0.859	0.114	0.636	1.083				

The heterogeneity test results were found to be significant ($p<0.05$), as seen in Table 2. The Q-value was calculated as 740.77, with 20 degrees of freedom (df). This value exceeds the critical value (31.410) of χ^2 with 24 df and confidence intervals of 95%. The I^2 index is 97.30%, which demonstrates a high amount of heterogeneity among the studies. These results reveal that the studies do not share a common effect size; namely, the variance observed in effect size of studies shows a significant difference from the variance of sampling error, and the studies are heterogeneous. As true effect sizes vary from study to study, they should be analysed according to the random effects model; the common effect is the mean of these effects (Borenstein et al., 2009, 76–77).

When the effect sizes of the 21 studies included in this review were combined in accordance with the random effects model, the common effect size was calculated as (d) 0.859 with 0.114 standard error and 95% confidence intervals of 1.083 and 0.636. The value of effect size falls within the "large" interval, according to Cohen's (1988) classification.

Forest Plot

Forest plot is one of the most useful tools to summarise meta-analysis results by visualizing them (Israel and Richter, 2011). The forest plot of the meta-analysis results of the 21 studies included in this review is given below:

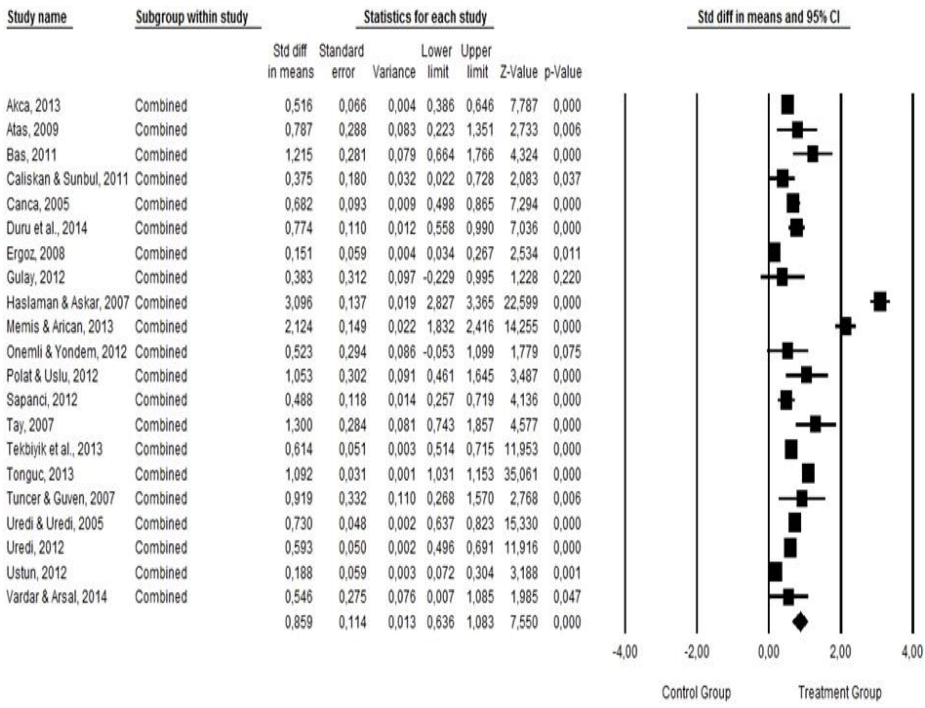


Figure 3. Forest plot of meta-analysis results

When standardised means differences of control and experimental groups are calculated in addition to effect sizes in a 95% confidence interval, the result is seen to be in favour of the experimental group. As a result, 19 of the 21 studies have a significant effect size, while 2 do not. Upon classifying these studies in regard to Cohen's (1988) effect classification, the effect size was found to be "low" in five studies, "moderate" in nine studies, and "large" in seven studies.

Moderator Analysis

Moderator analysis was applied to reveal whether the effect of self-regulated learning on academic achievement showed a significant difference in terms of self-regulated learning strategy, course type, school level, and research design. The results of the moderator analysis can be seen in Table 3.

Table 3

Moderator Analysis

Moderator Name	k	Effect Size	95% CI.		Heterogeneity		
			Lower Lim.	Upper Lim.	Q_b	df	p
Self-regulation strategy	27	0.701	0.548	0.854	2.994	3	0.393
Cognitive	11	0.673	0.432	0.915			
Metacognitive	10	0.897	0.548	1.245			
Management	2	0.818	0.435	1.201			
Motivational	4	0.515	0.204	0.826			
Course type	21	0.639	0.562	0.715	4.182	3	0.242
Language	3	0.682	0.190	1.173			
Science	4	0.618	0.521	0.714			
Mathematics	8	1.098	0.646	1.551			
Social Sciences	6	0.635	0.502	0.769			
School level	20	0.762	0.535	0.990	1.159	2	0.560
Primary	4	1.077	0.159	1.994			
Secondary	8	0.693	0.432	0.955			
Undergraduate	8	0.944	0.408	1.479			
Study design	21	0.825	0.640	1.011	0.521	1	0.470
Experimental	9	0.767	0.522	1.011			
Relational	12	0.905	0.619	1.191			

It is seen in Table 3 that there is no significant difference in the effect size of groups formed according to self-regulation strategy, course type, school level, and study design ($Q_b < \chi^2$; $p > .05$). In other words, the effect of self-regulated learning on academic success does not vary significantly neither according to cognitive, metacognitive, resource management, and motivational strategies, nor according to the courses in which self-regulated learning occurs, be they science, social sciences, mathematics, or language courses. Similarly, the common effect size of the studies shows no significant difference according to whether they are relational or empirical or conducted in primary schools, secondary schools, or undergraduate programs.

Publication Bias

One of the ways to determine the existence of publication bias is that the common effect size of the studies does not show a significant difference according to study type (thesis vs. article). For this purpose, a moderator analysis was executed; it was found that the effect of self-regulated learning strategies on academic success

demonstrates no significant difference regarding publication status under the random effects model (Q_b=0.271, p>0.05). Another way to determine whether there is publication bias is via a funnel plot. When there is no publication bias, the effect sizes of studies included in the analysis will range around the common effect size symmetrically in the funnel plot, while they are expected to pile up very close to each other at the centre or bottom in case of publication bias, depending on the number of lacking studies (Borenstein et al., 2013, 273). A funnel plot for this study is given in Figure 4.

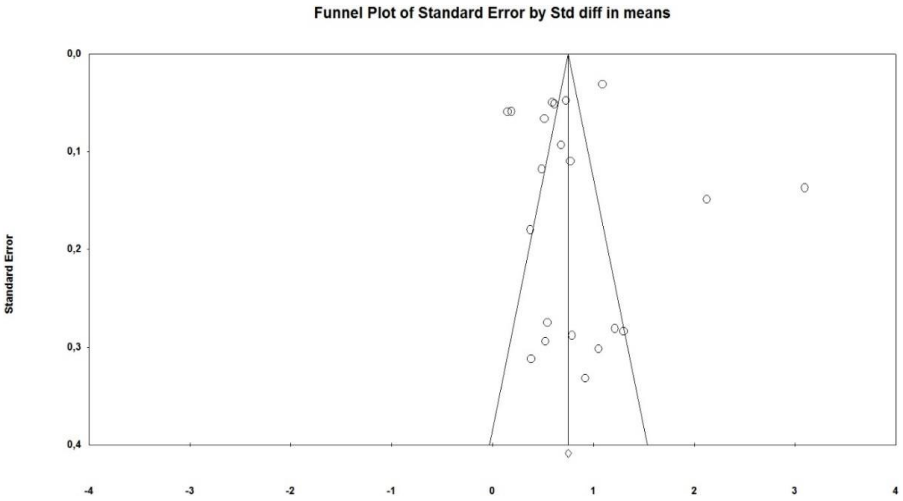


Figure 4. Funnel plot of publication bias

As seen in the funnel plot above, the effect sizes of the studies are generally dispersed at the centre and around the common effect size asymmetrically, which shows a possible existence of publication bias. However, the interpretation of the funnel plot is of the utmost subjectivity (Borenstein et al., 2009, 283). So Egger's Regression Intercept test and Rosenthal's Fail-safe N test were employed to evaluate the amount and impact of publication bias on the results.

If the intercept value (B0) obtained from Egger's regression intercept test does not deviate significantly from zero (p> 0.05), it proves the absence of publication bias, whereas it shows the possible existence of publication bias if the intercept value (B0) significantly deviates from zero (p< 0.05) (Card, 2012, 267). As a result of Egger's regression intercept test, the intercept value (B0) was computed as 0.6996 and the two-tailed p-value as 0.7576. According to these results, it can be interpreted that the common effect size does not result from publication bias since the intercept value did not significantly (p > 0.05) deviate from zero.

Rosenthal's Fail-safe N test (Borenstein et al., 2009, 284) was performed to evaluate whether the observed effect size was strong or if the common effect size resulted from publication bias. This test calculates how many studies with the mean effect of zero need to be added to the analysis to make the p-value non-significant

(Rosenthal, 1979). According to Rosenthal (1979), if 5k+10 (k is the number of the studies) of the studies included in the analysis are needed, the common effect can be said not to result from publication bias. As a result of Rosenthal's Fail-safe N, it was found that 7331 studies with a mean effect of zero would be needed to nullify the common effect size. Considering the number of studies included in the analysis was 21, the threshold of Rosenthal (1979) was computed to be 115 (5*21+10). As the sum of the studies to be added exceedingly outnumbers this threshold, the common effect size can be claimed not to be the outcome of publication bias.

Discussion and Conclusion

In this study as a result of the meta-analysis, a heterogeneity test showed that the individual studies were heterogeneous at a high level *($p < 0.05$, $I^2 = 97.30\%$)*. As effect size varies from study to study, the common effect size should be analysed in regard to the random effects model. The common effect size under the random effects model was calculated as 0.859. The result reveals that the effect of the educational environment designed according to self-regulated learning on academic achievement is "large" *($d=0.859$)*.

The finding that self-regulated learning strategies have a "large" effect on academic achievement shows parallelism with the findings from meta-analysis studies by Hattie, Biggs, and Purdie (1996), Chiu (1998), and Dignath and Buttner (2008). The common effect size in those studies is "moderate", while it is "large" in this study. This can be because relational studies are included in this meta-analysis study in addition to empirical ones. Though there is no significant difference in the common effect size of the studies regarding study designs (empirical vs. relational), the common effect size of empirical studies was found to be "moderate" *($d=0.767$)*, while it was "large" *($d=0.905$)* in relational ones. The result that the common effect size of relational studies is higher than empirical ones can be thought to be the reason for the "large" interval.

It was found in this study that the common effect size of self-regulated learning shows no significant difference according to self-regulated learning strategies (cognitive, metacognitive, motivational and resource management) *($Q_b=2.994$, $p>.05$)*. The common effect size of resource management and metacognitive strategies is "large", while the common effect size of cognitive and motivational strategies is "moderate". Metacognitive strategies help individuals to control and regulate their own cognitive processes (Lucangeli & Cornoldi, 1997); however, resource management strategies enable individuals to manage and monitor their learning environment (Pintrich, 1999). Consequently, individuals make use of their own cognition effectively via metacognitive strategies and benefit from their environment more through resource management strategies to achieve their goals, which will probably increase their academic achievement more than other strategies.

Another result of this meta-analysis review is that the effect of self-regulated learning on academic achievement does not show any significant difference according to course type *($p>0.05$)*. It was found that the effect of self-regulated learning on academic achievement in mathematics is "large", while it is "moderate"

in science, language, and social sciences. The finding as to mathematics corresponds to Dignath and Buttner's (2008) findings; however, unlike this study, Dignath and Buttner (2008) found the common effect size of reading/writing (language) to be "low". Regarding this finding, it can be alleged that self-regulated learning enhances academic success in mathematics more than other courses.

Finally, the effect of self-regulated learning on academic success shows no significant difference in terms of school level. As a result of the meta-analysis, the effect size for primary and undergraduate was found to be "large", while it was "moderate" for secondary school. Accordingly, Chiu (1998) revealed that providing students who exhibit low-level skills with strategy teaching would contribute more to their achievement, in comparison with other students. Therefore, the reason for the high effect size in primary school may be because it requires low-level skills compared to other school levels. The fact that there is no significant difference between school levels shows it would be useful to teach these strategies to all age levels.

Recommendations

It was revealed in this study that metacognitive and resource management strategies have the highest effect size and, thus, it is important that teachers employ metacognitive strategies in learning environments to the increase academic achievement of their students. For this purpose, teachers can be provided with professional development programs about creating self-regulated learning environments as well as requesting teacher candidates to work on improving these skills. Although these programs are useful for all branch teachers, they are especially important for classroom teachers/teacher candidates, as these strategies increase academic success in primary school more than other school levels. In this way, the students educated by these teachers can be lifelong learners, as they will acquire these skills early.

This meta-analysis study revealed that empirical design was employed, at the very least, in studies of self-regulated learning, and that these studies are conducted in language courses. Therefore, further studies may contribute to literature if they are carried out in language courses with an empirical design. In addition, it may be useful for researchers to conduct a meta-analysis study about self-regulated learning to investigate the effect of self-regulated learning on dependent variables such as attitude towards course, retention, self-efficacy, and high-level thinking skills.

Limitations

While calculating the effect of self-regulated learning on academic success, firstly, the sub-strategies of each strategy reported in some of the studies were combined amongst one another and their effect size was calculated. Later, these effect sizes were combined and the common effect size of the study was determined. However, while some of the studies reported the effect of the sub-strategies of each self-regulated learning strategy, some reported the effect of the self-regulated learning strategies, and others reported the effect of self-regulated learning. For instance, the

individual effect of cognitive or metacognitive strategies of self-regulated learning was mentioned in several studies. Therefore, it can be said that this situation can influence the common effect size of self-regulated learning on academic success.

References

(The studies with asterisk indicate that the studies are included in this review).

*Akca, F. (2013). An investigation into the academic success of prospective teachers in terms of learning strategies, learning styles and the locus of control. *Journal of Education and Learning*, *2*(1), 134–146. DOI: http://dx.doi.org/10.5539/jel.v2n1p134

*Atas, I. (2009). *Oz-duzenleyici ogrenme stratejileri kullaniminin ilkogretim okulu 4.sinif ogrencilerinin matematik dersindeki oz-yeterlik algisina ve basarisina etkisi* [The effect of self- regulated learning education strategies on the fourth grade primary students' perception of self-efficacy and success on maths lesson] (Unpublished Master's Dissertation), Gazi University, Ankara. Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi

*Bas, G.(2012). The effect of teaching learning strategies in an English lesson on students' achievement, attitudes, and metacognitive awareness. *Journal of Theoretical Educational Science*, *5*(1), 49–71.

Bandura, A. (1971). *Social learning theory.* New York: General Learning Press.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Boekaerts, M. (1996). Self-regulated learning at the junction of cognition and motivation. *European Psychologist*, *1*(2), 100–112.

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

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

*Canca, D. (2005). *Cinsiyete gore universite ogrencilerinin kullandiklari bilissel ve bilisustu oz duzenleme stratejileri ile akademik basarilari arasindaki iliskinin incelenmesi* [Research of the relationship between university students' cognitive and metacognitive self-regulation strategies and their mathematics achievement according to gender] (Unpublished Master's dissertation), Yildiz Technical University, Istanbul. Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi

Camahalan, F. M. G. (2006). Effects of self-regulated learning on mathematics achievement of selected Southeast Asian children. *Journal of Instructional Psychology*, 33(3), 194–205.

Cazan, A. (April 2014). *Self-regulated learning and academic achievement in the context of online learning environments*. The 10th International Scientific Conference e-Learning and Software for Education, Bucharest.

Chiu, C. W. T. (April 1998). *Synthesizing metacognitive interventions: What training characteristics can improve reading performance?* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd Ed.). Hillsdale, NJ: Erlbaum. Cox

Cooper, H. (2010). *Research synthesis and meta-analysis (A Step by Step Approach) 4th edition*. Thousand Oaks: SAGE Publication.

*Caliskan, M., & Sunbul M.A. (2011). The effects of learning strategies instruction on metacognitive knowledge, using metacognitive skills and academic achievement (Primary Education Sixth Grade Turkish Course Sample) *Educational Sciences: Theory & Practice, 11*(1), 133–153.

Dikbas, Y., & Hasirci, O. (2008).The effects of teaching and using of learning strategies on achievement, retention and attitude of the students. *Journal of Kırsehir Education Faculty (JKEF)*, *9*(2), 69–76.

Dignath, C., & Buttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition Learning, 3,* 231–264.

Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review*, *20*, 391–409. DOI: http://10.1007/s10648-008-9083-6

*Duru, E., Duru, S., & Balkis, M. (2014) Analysis of relationships among burnout, academic achievement, and self-regulation. *Educational Sciences: Theory & Practice, 14*(4), 1263–1284. DOI: http://10.12738/estp.2014.4.2050

*Ergoz, G. (2008). *Oz- duzenleyici ogrenmenin ve guduleyici inanclarin matematik basarisi icinde arastirilmasi* [Investigation of self-regulated learning and motivational beliefs in mathematics achievement*]* (Unpublished Master's Dissertation). Middle East Technical University, Ankara. Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi

*Gulay, A. (2012). *Oz-duzenleyici ogrenmenin 5.sinif ogrencilerinin akademik basarisina ve bilimsel surec becerilerine etkisi* [Effect of self-regulated learning on 5th grade students' academic achievement and scientific process skills] (Unpublished Master's Dissertation) Recep Tayip Erdogan University, Rize. Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi

*Haslaman, T., Askar, P. (2007). Investigating the relationship between self-regulated learning strategies and achievement in a programming course. *Hacettepe University Journal of Education, 32*(2007), 110–122.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research, 66*, 99–136. DOI: http://10.3102/00346543066002099

Hedges, L. V., & Pigott, T. D. (2001). The power of statistical tests in meta-analysis. *Psychological Methods*, *6*, 203–217.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random effects models in meta-analysis. *Psychological Methods*, *3*, 486–504. DOI: http://10.1037/1082-989X.3.4.486

Israel, H., & Richter, R.R. (2011). A guide to understanding meta-analysis. *Journal of Orthopedic & Sports Physical Therapy*, *41*(7), 496–504. DOI: http://10.2519/jospt.2011.3333.

*Kuyumcu Vardar, A., & Arsal Z. (2014). Oz-duzenleme stratejileri ogretiminin ogrencilerin Ingilizce basarilarina ve tutumlarina etkisi [The effect of self-regulated learning strategies instruction on students' English achievement and attitude]. *Journal of Mother Tongue Education*, *2*(3), 32–52.

Lucangeli, D., & Cornoldi, C. (1997). Mathematics and metacognition: What is the nature of the relationship? *Mathematical Cognition*, *2*, 121–139. DOI: http://10.1080/135467997387443

*Memis, A., & Arican, H. (2013). The analysis of 5th grade students' mathematical metacognition levels between the variables gender and achievement. *Karaelmas Journal of Educational Sciences, 1*(2013), 76–93.

Miles, MB. & Huberman, AM. (1994). *Qualitative Data Analysis (2nd edition)*. Thousand Oaks, CA: Sage Publications

*Onemli, M., & Yondem D. Z. (2012). The effect of psychoeducational group training depending on self-regulation on students' motivational strategies and academic achievement. *Educational Sciences: Theory & Practice*, *12*(1), 59–73.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*, 33–40. DOI: http://10.1037//0022-0663.82.1.33

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*, *31*, 459–470. DOI: http://10.1016/S0883-0355(99)00015-4

Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, *92*, 544–555.

*Polat, S., & Uslu, M. (2012). The effect of teaching based on metacognitive strategies on 5th grade students' achievement in science and technology course. *Usak University, The Journal of Social Sciences, 11*, 28–43.

Rowe, F., & Rafferty J. (2013). Instructional design interventions for supporting self-regulated learning: Enhancing academic outcomes in postsecondary e-Learning environments. *MERLOT Journal of Online Learning and Teaching, 9*(4), 590–601.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, *86*, 638–641.

Sakiz, G. (2014). *Oz-duzenleme- ogrenmeden ogretime oz-duzenleme davranislarinin gelisimi, stratejiler ve oneriler* [*Self-regulation- from learning to teaching the development of self-regulation behavior, strategies and suggestions]*. Ankara: Nobel Publishing.

*Sapanci, A. (2012).The relationship of student teachers' epistemological beliefs and metacognitive levels with their academic achievement. *Celal Bayar University the Journal of Social Sciences, 10*(1), 311–331.

*Tay, B. (2007). Ogrenme stratejilerinin hayat bilgisi ve sosyal bilgiler ogretimi dersinde akademik basariya etkisi [The effect of learning strategies on academic achievement in life science and social studies teaching courses]. *Journal of Education and Social Sciences, 173*, 88–102.

*Tekbiyik, A., Camadan, F., & Gulay, A. (2013). Self-regulated strategies as a predictor of academic achievement in science and technology course. *Turkish Studies, 8*(3), 567–582.

*Tonguc, D. (2013). Sekizinci sinif ogrencilerinin motivasyon duzeylerinin ve oz-duzenlemeye dayali ogrenme stratejilerinin matematik basarisini yordama gucu [The predictive power of motivation levels and self-regulated learning strategies on mathematics achievement of eight grade students] (Unpublished Master's Dissertation). Osman Gazi University, Eskisehir. Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi

*Tuncer, K.B., & Guven, B. (2007). Effects of learning strategies on academic achievement, students' recalling level and students' attitudes toward assigned courses. *YYU Journal of Education Faculty, 4*(2), 1–20.

*Uredi, I., & Uredi, L. (2005). The predictive power of 8th grade students' self-regulated strategies and motivational beliefs on mathematic achievement. *Mersin University Journal of the Faculty of Education*, *1*(2), 250–260.

*Uredi, I. (2012). An investigation of the relationship between pre-services teachers' self-regulated strategies, motivational beliefs and academic achievement. *Proceedings of International Conference on Educational Science*, *3*, 2014–2022.

*Ustun, A. (2012). *Cinsiyete gore ogrencilerin kullandiklari bilissel ve bilis ustu oz duzenleme stratejilerinin akademik basarilari uzerindeki etkileri* [The relationship between students' cognitive and metacognitive self-regulation strategies and their academic achievement according to gender] (Unpublished Master's Dissertation). 18 Mart University, Canakkale. Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi

Senemoglu, N. (2005). *Gelisim ogrenme ve ogretim* (12. Baskı) [Development, learning and teaching]. Ankara: Gazi Bookstore.

Shaine, M.H. (2015). The effect of self-regulated learning strategies and self-efficacy on academic achievement of primary school students. *Psychology and Behavioral Sciences*, *4*(3), 107–115.

Yukselturk, E., & Bulut, S. (2009). Gender differences in self-regulated online learning environment. *Educational Technology & Society*, *12*(3), 12–22.

Zimmerman, B.J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3–17.

Zimmerman, B. J. (2001). Theories of self-regulated learning and academic achievement: an overview and analysis. In B. J. Zimmerman, & D. H. Schunk (Eds.), Self-regulated learning and academic achievement: theoretical perspectives (2nd ed.) (pp. 1–38). Mahwah, NJ: Erlbaum.

Öz Düzenlemeli Öğrenme Stratejilerinin Akademik Başarıya Etkisi: Bir Meta Analiz Çalışması

Atıf:

Ergen, B., & Kanadlı, S. (2017). The effect of self-regulated learning strategies on academic achievement: A meta-analysis study. *Eurasian Journal of Educational Research*, *69*, 55-74. http://dx.doi.org/10.14689/ejer.2017.69.4

Özet

Problem durumu: Öğrenme sürecinde öğrencilerin kendi öğrenmelerinin sorumluluğunu almaları beklenmektedir. Bu sorumluluğu alan bireyler kendilerini, kendi belirledikleri ölçütlere göre değerlendirerek eksik bilgi ve becerilerini tamamlayabilirler. Öz düzenlemeli öğrenme olarak adlandırılan bu süreç öğrencilerin akademik başarılarının, kavramsal anlamalarının ve motivasyonlarının arttırılması ve yaşam boyu öğrenen bireyler olmaları açısından önemli görülmektedir. Bireyler kendilerini düzenlerken bilişsel, üst bilişsel, kaynakları yönetme ve motivasyonel olarak adlandırılan çeşitli stratejiler kullanmaktadırlar. Öz düzenlemeli öğrenme stratejileri ile ilgili ulusal alan yazın incelendiğinde, bu

stratejilerin kullanılmasının öğrencilerin akademik başarılarını arttırdığını belirten çalışmaların yanında akademik başarı üzerinde anlamlı bir etkisi ya da akademik başarıyla anlamlı bir ilişkisi olmadığını ortaya koyan çalışmalar görülmektedir. Bu nedenle Türkiye'de öz düzenleme ile ilgili olarak birçok bireysel çalışmanın olması ve bu çalışmalarda çelişkili sonuçların elde edilmesi, bu çalışmaların meta analiz yöntemiyle birleştirerek bir sonuca varma ihtiyacının doğurmuştur.

Amaç: Bu araştırmanın amacı öz-düzenlemeli öğrenmenin akademik başarı üzerindeki etkisini inceleyen ilişkisel ve deneysel çalışmaların meta-analizini yaparak genel etki büyüklüğünü hesaplamak ve akademik başarının öz düzenlemeli öğrenme stratejisine, ders türüne, çalışma türüne, öğretim kademesine ve çalışma desenine göre anlamlı fark gösterip göstermediğini belirlemektir.

Yöntem: Araştırmada öz-düzenlemeli öğrenmenin akademik başarıya etkisinin incelenmesi amacıyla meta-analiz yöntemi kullanılmıştır. Çalışmaların seçiminde Google Akademik arama motoru (2015), TÜBİTAK ULAKBİM DergiPark (2015), YÖK Ulusal Tez Merkezi (2015), ERIC (2015) ve EBSCO (2015) veri tabanları taranarak uygun çalışmalar toplanmıştır. Toplanan çalışmalar şu ölçütlere göre değerlendirilmiştir: *(i)* Türkiye'de 2005-2014 yılları arasında yapılan, deneysel ve ilişkisel desenlerle hazırlanmış makale, tez ya da bildiri olmalıdır. *(ii)* Öz-düzenleyici öğrenmenin akademik başarı ile ilişkisini ya da akademik başarıya etkisini araştırmalıdır. *(iii)* Deneysel çalışmaların örneklem sayısı (N), ortalaması ($\overline{X}$) ve standart sapması (SD); ilişkisel çalışmaların da örneklem büyüklüğü ve Pearson korelasyon katsayısı olmalıdır.*(iv)* Parametrik testleri kullanmalıdır (t testi, F testi vb.). Yapılan değerlendirme sonucunda içerme ölçütlerini karşılayan toplam 21 çalışma analize dâhil edilmiştir. Bu çalışmalar çalışmanın yazarı, çalışmanın tarihi, çalışmanın türü, çalışma deseni, çalışmanın yürütüldüğü ders ve öğretim kademesi ile çalışmada kullanılan öz düzenlemeli öğrenme stratejilerine göre kodlanmıştır. Yapılan kodlamanın derecelendirenler arası güvenirliği %100 olarak hesaplanmıştır. Bu araştırmada etki büyüklüğü indeksi olarak standartlaştırılmış ortalamalar farkı olan Cohen's d kullanılmıştır. Çalışmaların alanyazından toplanmış olması ve evrene genelleme yapılmak istenmesinde dolayı rastgele etkiler modeli uygun olmakla birlikte kullanılacak modelin belirlenmesinde heterojenlik testi yapılmıştır. Alanyazından toplanan çalışmaların kendi içinde alt gruplar içermesinden dolayı her bir çalışma analiz ünitesi olarak kabul edilmiş ve genel etki büyüklüğü buna göre hesaplanmıştır. Bununla birlikte hesaplanan genel etki büyüklüğünün çalışmanın türü, çalışma deseni, çalışmanın yürütüldüğü ders ve öğretim kademesi ile çalışmada kullanılan öz düzenlemeli öğrenme stratejilerine göre anlamlı farklılık gösterip göstermediğini belirlemek için moderatör analizi yapılmıştır. Yayın yanlılığının varlığını belirlemek ve analiz üzerindeki etkisini değerlendirmek için çalışmaların yayınlanma durumuna göre (tez vs. makale) moderatör analizi, huni diyagramı ile Rosenthal'ın Korumalı N'i, ve Egger'in Regresyon Kesişim testi kullanılmıştır. Verilerin analizinde bilgisayar yazılımlarından yararlanılmıştır.

Bulgular: Çalışmaların etki büyüklüklerinin heterojen yapıda ($Q>x^2$, $p< 0.05$) ve çalışmalar arasındaki heterojenliğin ($I^2=97.30$) yüksek miktarda olmasından dolayı genel etki büyüklüğü, rastgele etkiler modeline göre kabul edilmiştir. Rastgele etkiler

modeline göre yapılan meta-analiz sonucunda öz düzenlemeli öğrenmenin akademik başarı üzerinde "geniş" aralıkta (d=0.859) bir etkiye sahip olduğu belirlenmiştir. Yapılan moderatör analizi sonucunda hesaplanan genel etki büyüklüğünün çalışmaların desen türüne, çalışmaların yürütüldüğü ders türüne ve öğretim kademesine, çalışmalarda kullanılan öz düzenlemeli öğretim stratejisine göre anlamlı farklılık göstermediği ($Q_b < \chi^2$, $p>0.05$) belirlenmiştir. Hesaplanan genel etki büyüklüğünün yayın yanlılığının ürünü olup olmadığını belirlemek için yapılan moderatör analizi sonucunda çalışmaların tez veya makale olmasına göre anlamlı farklılık göstermediği belirlenmiştir. Ayrıca Egger'in Regresyon Kesişim testi sonucunda yayın yanlılığı olmadığı ve Rosenthal'ın Korumalı N testi sonucunda genel etki büyüklüğünün oldukça güçlü olduğu belirlenmiştir.

Sonuç ve Öneriler: Bu çalışma sonucunda öz düzenlemeli öğrenme stratejilerin akademik başarıyı önemli ölçüde etkilediği ortaya çıkmıştır. Ayrıca akademik başarının öz düzenlemeli öğrenme stratejilerine, ders türüne, öğretim kademesine, çalışma türü ve çalışma desenine göre anlamlı farklılık gösterdiği sonucuna varılmıştır. Bu sonuçlara göre öz-düzenlemeli öğrenme stratejilerinin tüm derslerde ve her öğretim kademesinde öğrencilerin akademik başarılarının arttırdığı söylenebilir. Bu nedenle öğretmenlerin öğrenme ortamlarında öz düzenlemeli öğrenme stratejileri etkin biçimde kullanmaları, öğrencilerinin akademik başarılarını arttırmaları açısından önemli görülmektedir. Bu amaçla öğretmenlere öz düzenlemeli öğrenme ortamları oluşturmalarına yönelik mesleki gelişim programları hazırlanabileceği gibi öğretmen adaylarının lisans programlarına bu becerileri geliştirmeye yönelik çalışmalar yaptırılabilir. Bununla birlikte bu meta analiz çalışması sonucunda öz düzenlemeli öğretim ile ilgili yapılan çalışmalarda en az deneysel desenin kullanıldığı ve bu çalışmaların en az dil derslerinde (Türkçe, Yabancı Dil) yürütüldüğü belirlenmiştir. Bu nedenle bundan sonra yapılacak bireysel çalışmaların dil derslerinde ve deneysel desen kullanılarak yürütülmesi alanyazına katkı sağlayabilir. Buna ek olarak öz düzenlemeli öğrenme ile ilgili meta analiz çalışması yürütecek araştırmacıların öz düzenlemeli öğrenmenin derse karşı tutum, kalıcılık, öz yeterlik ve üst düzey düşünme becerileri (eleştirel ve yaratıcı düşünme, problem çözme vb.) gibi bağımlı değişkenler üzerindeki etkisini incelemeleri yararlı olabilir. Böylece bu değişkenler üzerindeki genel etki büyüklüğü hesaplanarak hangi tür moderatörlerin bu değişkenleri etkilediği belirlenebilir.

Anahtar Sözcükler: Öz düzenlemeli öğretim stratejisi, Akademik başarı, Meta-analiz, Moderatör analizi

Eurasian Journal of Educational Research 69 (2017) 75-92

Eurasian Journal of Educational Research
www.ejer.com.tr

Analysis of Attainments and Evaluation Questions in Sociology Curriculum according to the SOLO Taxonomy

Fahrettin KORKMAZ[1] Serkan UNSAL[2]

ARTICLE INFO

Article History:

Received: 11 January 2017

Received in revised form: 12 February 2017

Accepted: 26 April 2017

DOI: http://dx.doi.org/10.14689/ejer.2017.69.5

Keywords
sociology course,
curriculum,
analysis of SOLO taxonomy

ABSTRACT

Purpose: This research aims at analyzing the attainments identified in the sociology curriculum for 11th grade implemented by the Ministry of National Education (MoNE) in 2010, and the evaluation questions in the sociology textbook which was taught in the 2016-2017 academic year, based on the Structure of Observed Learning Outcomes (SOLO) taxonomy. **Research Methods:** Document analysis was used in this study. The attainments that constitute the data source of the research were taken from the sociology curriculum for 11th grade published by the Ministry of National Education (MoNE) in 2010. **Findings:** Research findings indicate that 15 out of 60 (25%) attainments in the curriculum could be identified at the level of uni-structural, 14 (23%) multi-structural, 28 (47%) relational, and only 3 (5%) extended abstract. As for the evaluation questions in the textbook, 70 (43%) could be identified as uni-structural, 44 (27%) as multi-structural, 36 (23%) as relational, and 12 (7%) as extended abstract. **Implications for Research and Practice**: When the 11th grade sociology course achievements and assessment questions were examined, it was determined that all thinking levels of the SOLO taxonomy were found at different ratios. It is thought that when the experts prepare the program attainments for the sociology course, the organization of the attainments harmonized by prerequisite considering the principle of progressivity will be of considerable benefit in terms of effectiveness of the program.

[1] Corresponding Author: Ministry of National Education, Gaziantep, korkmaz2725@gmail.com
[2] Kahramanmaras Sutcu Imam University, K. Maras, serkan-unsal09@hotmail.com

Introduction

Sociology is a science that examines social institutions and social relations and interactions within and between groups (Zencirkiran, 2016). According to the American Sociological Association (2013), sociology is a discipline that deals with social life, social changes, social events, and the consequences of human behavior. Giddens and Sutton (2013) define sociology as a discipline that examines societies, social life, and groups. Sociology became an independent discipline in the 19th century as a result of such events as the French Revolution, Renaissance movements, industrialization, and urbanization. The arrival of sociology as an independent discipline in Turkey began with Ziya Gökalp, when he started giving sociology lessons at the Thessaloniki High School of Union and Progress in 1910–1911 (Zabun, 2012). Having a historical background of nearly 100 years, the aim of the sociology course is to provide students with a basic understanding of the society and the social environment they live in, increasing their awareness of the effect of the society on the individual and giving them a sociological perspective (Banoglu & Bas, 2012; Can, 2006).

According to the Ministry of National Education (2009), the sociology curriculum aims to educate individuals who absorb the national and spiritual values of the society in which they live from a sociological point of view, so that they can analyze social and cultural relationships beyond the daily gaze, have the ability to think critically about social life, comprehend social problems, and analyze the conditions that change and transform social life. Nine different sociology curricula with different aims and content have been implemented since 1924; the one in current use was prepared in 2009 according to the constructivist approach (Zabun, 2012).

The Structure of Observed Learning Outcomes (SOLO) taxonomy was developed by John Biggs and Kevin Collis in 1982 (Baghdad, 2013; Biggs & Tang, 2011; Gezer & Ilhan, 2014; Lian & Idris, 2006; Yazıcı, 2013). The learning outcomes are expressions of what students know or can do by the end of the course. They are the assessable outcomes of the education, based on the students' point of view and what is acheivable.

The SOLO taxonomy, as constituted by Biggs and Collins, is the result of the assessment of answers from hundreds of students who have been educated on subjects such as history, mathematics, literacy, geography, computer technology, and foreign languages in a range of schools from primary school to universities. The SOLO taxonomy is an assessment tool that is applied to assess the quality of learning at different school levels in many subject areas (Arı, 2013; Biggs & Collis, 1982; Collis & Biggs, 1979; Yildirim & Baur, 2016). It can be said that the SOLO taxonomy is an important evaluation tool in assessing students' knowledge and skills, examining answers in depth and revealing the quality and structure of answers. Assessment in the SOLO taxonomy is based on the quality and structure of the answers students have given to questions. The answers are analyzed according to certain criteria and the level of learning is then determined (Baghdad, 2013; Biggs & Tang, 2011; Brabrand & Dahl,

2009; Hattie, Biggs, & Purdie, 1996; Konyalihatipoglu, 2016; Musan, 2012; O'Neill & Murphy, 2010).

There are five different thinking levels that are hierarchical in the SOLO taxonomy, and learning outcomes can be evaluated using the five levels as criteria. These structures are: pre-structural, uni-structural, multi-structural, relational, and extended abstract (Biggs, 2011; Biggs & Collis, 1982; Biggs & Tang, 2011; Hattie et al., 1996; O'Neill & Murphy, 2010; Pegg & Dawey, 2012). All these structures reflect the learning quality of a particular chapter or task (Collis & Biggs, 1979). The descriptions of these structural levels are as follows:

Pre-structural: At this level, direct information particles are used (Biggs, 2011). The student uses the knowledge without any understanding, avoids the question, and repeats only the asked question (Biggs & Collis, 1982; Calkins & Cox, 2009; Olsson, 2005; Rooney, 2012). The student uses inappropriate, irrelevant, and erroneous content or methods. The learner has difficulty in understanding the subject and cannot learn anything meaningful. The student might gain scattered information particles but these information particles are disorganized. The student is far from relating to the subject or problem (Arı, 2013). The student has little or no involvement with the assigned task (Biggs & Tang, 2011; Collis & Biggs, 1979; Hattie et al., 1996; Musan, 2012; Pegg & Dawey, 2012). In short, the answer of the student at the pre-structure level is insufficient (Celik, 2007).

Uni-structural: At this level, the student has limited understanding. The student focuses on the question; however, this focus is only related to using the data associated with the question. The student cannot understand the position of the data used in relation to the question within the whole data and the relation of it with the other data. The answers given by the student are limited and incomplete (Baghdad, 2013; Calkins & Cox, 2009; Hattie et al., 1996; Konyalihatipoglu, 2016; O'Neill & Murphy, 2010; Pegg & Dawey, 2012; Yazici, 2013).

Multi-structural: At this level, two or more pieces of information are used (Biggs, 2011) but the student uses the multiple datasets related to the answer without considering the relationship between them; therefore, some inconsistencies can be seen (Biggs & Collis, 1982; Hattie et al., 1996; Light, Calkins, & Cox, 2009; O'Neill & Murphy, 2010; Steel, 2007). The student can deal with various aspects of a topic but cannot establish links (Pegg & Dawey, 2012; Rooney, 2012). At the multi-structural level, the student can quickly understand and address the topics, interpret data in a table, and easily see the relationship between the datasets (Claesgens, Scalise, Wilson, & Stacy, 2009; Collis & Biggs, 1979; Hattie et al., 1996; Lian & Idris, 2006). The student can focus on more than one aspect for the question but cannot interrelate each aspect. For this reason, the student's answers are composed of disconnected pieces of information and there is no relational link between answers (Baghdad, 2013; Biggs & Tang, 2011; Olsson, 2005).

Relational: At this level, the student is able to relate to the topic by taking various aspects of it and knowing how to put the whole together. The student can build an understandable structure and understand that when the trees come together, they will

form a forest. Attribution, buıldıng cause-and-effect relationships, and beıng able to analyze are the characteristics of this level (Brabrand & Dahl, 2009; Claesgens et al., 2009; Olsson, 2005; Pegg & Dawey, 2012). The student uses two or more pieces of information regarding the understandability of the information which is located in the source of information (Biggs, 2011; Biggs & Collis, 1982; Hattie et al., 1996). The student understands how to construct a whole and the relationship between the structures that make up the whole (Arı, 2013; Biggs & Tang, 2011; Hattie et al., 1996; Light et al., 2009; Rooney, 2012).

Extended Abstract: This level is the highest-level thinking pattern. The student can reason by considering abstract features and can make generalizations; this level can represent a new way of thinking (Baghdad, 2013; Celik, 2007; Collis & Biggs, 1979; Musan, 2012; Pegg & Dawey, 2012). The student can perceive the topic from many perspectives, hypothesize, and make generalizations (Biggs, 2011; Brabrand & Dahl, 2009; Claesgens et al., 2009; Light et al., 2009; Olsson, 2005).

When an assessment is made in the SOLO taxonomy, the pre-structural level must be excluded from the thinking level because, at that stage, there is usually no opinion about the topic to learn, or the ideas being proposed are irrelevant (Potter & Kustra, 2012).

The information to be gained for the analysis of the attainments of the Sociology curriculum according to the thinking levels is considered to provide a data source for the curriculum specialists in the stages of observing, designing and organizing them.

In addition, analyzing evaluation questions as well as attainments according to thinking levels allows for a more accurate judgment of the effectiveness of any program. This study aims to provide information as to consistency achieved between the attainments and the evaluation questions, by determining the cognitive levels of both (Ilhan, Oner, Sunkur & Cetin, 2014). In any case, the determination of attainments and learning levels is seen as important in terms of providing information on whether attainments are balanced distributed without intensification or not.

As a result of this research, it is hoped that teachers will be guided to undertake activities appropriate to students' learning levels by ensuring that teachers have knowledge about the different learning levels of attainment. When relevant literature was searched, we did not find any research analyzing the attainments and the evaluation questions of the sociology curriculum according to the SOLO taxonomy. From this standpoint, it is thought that this work will fill a gap in the literature.

The research question is defined as: "How do the attainments in the 11th grade sociology curriculum lesson and the evaluation questions in the sociology textbook disperse according to the SOLO taxonomy?" Within this main research problem regarding 11th grade Sociology lesson:

i) How does the curriculum attainment SOLO taxonomy level of thinking disperse according to the units?

ii) How do the evaluation questions of the course book disperse according to the SOLO taxonomy?

Method

Research Design

The document review method was used in this study, which was prepared according to a qualitative research design. Document review is a method based on analyzing resources that the researcher has collected over a long period of time (Aktas, 2014).

Data Collection and Analysis

The attainments that constitute the data source of the research were taken from the 11th grade sociology curriculum published by the Ministry of National Education (MoNE) Head Council of Education and Morality in 2010. The evaluation questions were obtained from the course book, which was accepted in 2011 and printed in 2016, published on the web site of MoNE. The curriculum of Sociology for Grade 11 includes a total of 60 learning outcomes, and there are 162 evaluation questions in the related course book. The process of analysis for the 60 attainments and 162 evaluation questions according to the SOLO taxonomy was carried out by taking uni-structural, multi-structural, relational, and extended abstract thinking levels as criteria. The indicator verbs and explanations in the teacher's guidebook were used to determine which levels of thinking in the attainment and evaluation questions were equal. It was expected that it would be difficult to show what the level of thinking was for each of the 162 evaluation questions included in the scope of the research. Therefore, sample questions that represent each level of thinking are given below along with the explanation as to why they were coded to a specific thinking level.

Sample assessment questions at the uni-structural level:

i) Philosophy explores what needs to be done when dealing with social problems while ..... deals with what is done.

ii) The most important phenomenon affecting the emergence of Sociology as a science in the 19th century is accepted as .....

When the above sample evaluation questions are examined, it is seen that students were asked to fill in the blanks with suitable words by remembering the terms. Therefore, the two questions were considered as uni-structural.

Sample assessment questions at the multi-structural level:

i) What is the first thing that comes to your mind when someone says "religion"?

ii) Give examples of social events and phenomena.

When these sample evaluation questions are examined, metaphorical thinking and an ability to give examples of an event to express understanding of a concept are expected from the student. Considering that indicator verbs such as listing, explaining, giving examples, describing, and classifying represent the multi-structural level, the above questions were evaluated at that level.

Sample assessment question at the relational level:

What is the relationship between values and norms?

When this sample evaluation question is examined, it is understood that it is necessary for students to be able to see the relationship between the elements in one structure or several different structures in order for them to be able to answer the question. Therefore, this question was evaluated at the level of a relational structure.

Sample assessment question at the extended abstract level:

What are the problems you have observed in your community in the context of the socialization of children in modern society, and what solutions would you offer for these problems?

When this sample evaluation question is examined, it can be seen that the student is asked to develop a proposal for a problem or a situation based on the information he/she has learned. At this level, the student can go beyond the data to run syntheses, reach generalizations, and make predictions. Considering this, the question above is evaluated at the structural level of extended abstract thinking. The thinking levels of the SOLO taxonomy and their corresponding indicator verbs are given in Table 1.

Table 1

SOLO Taxonomy Thinking Levels and Indicator Verbs for These Levels

Uni-Structural	Multi-Structural	Relational	Extended Abstract
Memorize	List	Query	Assume
Define	Explain	Apply	Generalize
Recognize	Report	Outline	Probe
Count	Debate	Differentiate	Design
Draw	Choose	Analyze	Create
Reveal	Calculate	Classify	Judge
Tell	Plan	Compare	Hypothesize
Say	Clarify	Categorize	Evaluate
Express	Make clear	Observe	Prove
Diagnose	Interpret	Summarize	Reflect
Realize	Symbolize	Guess	Apply
Remember	Qualify	Integrate	theory to a
Repeat	Split into main lines	Explain the causes	new field
Mark	Think metaphorically	Evaluate	
Imitate		Apply a given theory to a related field	Guess

Source: Biggs, 2011; Biggs & Tang, 2011; Burnett, 1999; Lian & Idris, 2006; Light et al., 2009.

The evaluation processes of the attainment and evaluation questions, respectively, were independently examined by two researchers according to the SOLO taxonomy. At the end of the evaluation, each item was processed into the suitable level of thinking. In the next process, the markings of the researchers were compared and the percentages of harmony were examined. The reliability of the research is calculated as 92% in the analysis of attainments and 91% in the analysis of evaluation questions. As

a reliability calculation of over 70% is considered reliable for such studies, the results obtained for this study are considered reliable (Miles & Huberman, 1994).

Results

Results Related to the Second Sub-Problem

The second sub-problem of the study was designated as: "How does the 11th grade sociology course SOLO taxonomy thinking levels disperse according to the units?" Findings related to the research problem are presented in Table 2.

Table 2

Dispersal of SOLO Taxonomy Levels by Curriculum Units

Units	Attainments	SOLO Levels			
		US	MS	R	EA
Introduction to Sociology	1. Questions the information that they know about Sociology.			X	
	2. Realizes that the elements forming the society are in interaction.	X			
	3. Embraces the emergence of Sociology as an independent discipline.		X		
	4. Recognizes the methods used in sociological research.	X			
	5. Realizes the contributions of Turkish sociologists to Sociology.	X			
	Total	3	1	1	
Individual and Society	1. Explains the concept of socialization with examples.		X		
	2. Explains the factors affecting the socialization process.		X		
	3. Realizes that socialization is a lifelong process.	X			
	4. Evaluates the effects of socialization on social relations.			X	
	5. Establishes a relationship between social position, status, and role concepts.			X	
	6. Distinguishes social status and social prestige.			X	
	7. Expresses the importance of values and norms in the regulation of social life.	X			
	8. Expresses the functions of social control.	X			
	9. Investigates the causes of social deviance.			X	
	10. Knows the importance of rights and duties arising from status and roles in social life.	X			
	Total	4	2	4	

Table 2 Continue

Units	Attainments	SOLO Levels			
		US	MS	R	EA
Societal Structure	1.Analyzes the elements of social construction.			X	
	2. Distinguishes the types of social interaction.			X	
	3. Learns the structure of the social layer and stratification.		X		
	4. Learns the difference of the stratification of Turkish society from other societies.		X		
	5. Explains the types and causes of social mobility with examples.			X	
	Total		2	3	
Social Change and Development	1. Grasps the phenomenon of "social change."		X		
	2. Evaluates the factors affecting social change.			X	
	3. Realizes that factors affecting social change can change over time.	X			
	4. Evaluates the impact of science, technology and mass media on social change.			X	
	5. Evaluates the effects of modernization on social change.			X	
	6. Interprets the effects of globalization on social change.		X		
	7. Gets to know the elements of social development.	X			
	8. Expresses the importance of social integration.	X			
	9. Analyzes the factors that cause social disintegration.			X	
	10. Develops solutions to address social disintegration.				X
	Total	3	2	4	1
Society and Culture	1. Distinguishes the different meanings of culture.			X	
	2. Analyzes the elements of culture.			X	
	3. Analyzes the place and importance of functions of culture in society.			X	
	4. Evaluates the contributions of culture to social cohesion.			X	
	5. Evaluates concepts related to culture within their relations to each other.			X	
	6. Realizes the importance of their own social culture in the process of acculturation.	X			
	7. Questions the cultural attitudes of the societies and their views on different cultures.			X	
	8. Evaluates the role of their own culture in the intercultural interaction.				X
	Total	1		6	1
Societal Institutions	1. Explains the meaning of the concept of "institution."		X		
	2. Analyzes the functions of social institutions.			X	
	3. Interprets the importance of the family in terms of social life and socialization.		X		
	4. Analyzes the effects of marriage and divorce on the individual and society.			X	
	5. Assesses the conditions required for marriage in terms of the continuity of the family.			X	

Table 2 Continue

Units	Attainments	SOLO Levels			
		US	MS	R	EA
Societal Institutions	6. Compares the family structure in different societies with Turkish family structure.			X	
	7. Evaluates the position of woman in family and society.			X	
	8. Gives examples of the importance that Ataturk gives to women's rights.		X		
	9. Interprets the importance of education in social life.		X		
	10. Learns the importance of education in the process of socialization.		X		
	11. Expresses the ideas of Atatürk about education.	X			
	12. Learns the importance of religion in social life.		X		
	13. Interprets the relation between the concepts of religion and secularism.			X	
	14. Reveals the concept of secularism in the Ataturkist system of thought.			X	
	15. Evaluates the importance of the economy in social life.			X	
	16. Recognizes the basic elements of economy in social life.	X			
	17. Gives examples of Ataturk's views on economic systems.		X		
	18. Gets to know the institution of "politics."	X			
	19. Expresses the basic concepts related to the political institution.	X			
	20. Recognizes and compares the forms of political administration systems.			X	
	21. Evaluates democracy as a form of governance in terms of social life.				X
	22. Gives examples by associating the concepts of citizenship, rights and responsibility in the Ataturkist thought system.			X	
	Total	4	7	10	1
		15	12	28	3

When the SOLO taxonomy is examined on the basis of units in Table 2, it has been determined that 3 out of the 5 attainments in the Introduction to Sociology unit are found to be uni-structural, 1 is multi-structural and 1 is relational; 4 out of the 10 attainments in the Individual and Society unit are uni-structural in nature, 2 are multi-structural and 4 are relational; while 2 out of 5 attainments in the Societal Structure unit are multi-structural and 3 are at the relational structure level. It has also been determined that 3 out of 10 attainments in the Social Change and Development unit are uni-structural, 2 are multi-structural, 4 are relational, and 1 is of an extended abstract level; 1 out of 8 attainments in the Society and Culture unit is uni-structural, 6 are relational, and 1 extended abstract; 4 out of 22 attainments in the Societal

Institutions unit are uni-structural, 7 are multi-structural, 10 relational, and 1 is at the extended abstract level.

Results Related to the Third Sub-Problem

The third sub-problem of the study was designated as: "How do evaluation questions in the 11th grade sociology course book disperse according to the SOLO taxonomy?" Findings related to this research problem are presented in Table 3.

Table 3

Distribution by SOLO Taxonomy on Evaluation Questions in the Sociology Course Book on the Basis of Units

	SOLO Thinking Level				
Units	Number of questions	Uni-structural	Multi-structural	Relational Structure	Extended Abstract
Introduction to Sociology	24	14	7	3	-
Individual and Society	25	5	9	11	-
Societal Structure	25	8	7	9	1
Social Change /Development	20	7	6	3	4
Society and Culture	14	5	4	3	2
Societal Institutions	54	31	11	7	5
Total	162	70	44	36	12

When Table 3 is examined, it is found that 70 of the total 162 questions (43%) are at the uni-structural level, 44 (27%) are at the multi-structural level, 36 (22%) are at the relational level and 12 (8%) are at the extended abstract level. Examining the evaluation questions on the basis of the units, 14 of the 24 evaluation questions in the Introduction to Sociology unit are uni-structural, 7 are multi-structural and 3 are relational; there were not any questions that met the level of extended abstract. Five of the 25 evaluation questions in the Individual and Society unit are found to be uni-structural, 9 multi-structural and 11 relational, and no question rose to the level of extended abstract. Eight of the 25 evaluation questions in the Societal Structure unit are uni-structural, 7 multi-structural, 9 relational and 1 is at the extended abstract level. Seven of the 20 evaluation questions in the Social Change and Development unit are uni-structural, 6 are multi-structural, 3 relational and 4 at the extended abstract level. Five of the 14 evaluation questions in the Society and Culture unit are uni-structural, 4 are multi-structural, 3 relational and 2 are extended abstract. Thirty-one of the 54 evaluation questions in the Societal Institutions unit are uni-structural, 11 are multi-structural, 7 are relational and 5 are extended abstract.

Discussion and Conclusion

There are 60 attainments in the sociology curriculum related to the first sub-problem. Of these achievements, 15 are uni-structural, 14 are multi-structural, 28 are relational in structure and 3 are of the extended abstract structural level.

In conclusion, as this research is the first study to examine SOLO taxonomy for sociology lessons, the results of the research will be compared with other results from different disciplines that examined the SOLO taxonomy, and such similarities and differences will then be discussed. When the results of the study are examined, it can be said that 29 attainments related to the first sub-problem are from superficial learning that includes uni-structural and multi-structural levels, while 31 are from the deep learning that consists of relational and extended abstract levels. In studies related to the SOLO taxonomy, it is generally observed that the one-structural/multi-structural thinking structure is more involved than the relational and extended abstract thinking structure (Celik, 2007; Goktepe & Ozdemir, 2013; Lian & Idris, 2006). As for this study, it was determined that the attainments related to the relational and extended abstract level of thinking are more than others. This situation differs with the results of the abovementioned studies. It is thought that this is caused by the structure of the sociology course. Because sociology is a scientific discipline that studies societal institutions, social relations and interactions between and within groups (Zencirkiran, 2016), this could be the reason why most attainments were at the level of relational thinking.

In the study by Gezer and Ilhan (2015) where they analyzed attainments of a social studies curriculum according to the SOLO taxonomy, approximately half of the attainments in the curriculum correspond to uni-structural and multi-structural levels, while the ratio of the number of attainments reflecting the level of extended abstract thinking was very low. As it is the case in this aforementioned study with only three attainments related to the extended abstract structure, it shows similarity with this study with a very low ratio. In a study conducted by Baghdad and Saban (2014), it was concluded that the majority of those students were below the relational structural level in the study of algebraic thinking skills, according to the SOLO taxonomy. These results contradict the results of this research. In a study by Biber and İncikapi (2016), the knowledge levels of prospective mathematics teachers were mostly uni-structural, multi-structural, and relational in problems related to the topic of functions, and it was determined that there were very few candidate teachers with knowledge at the structural level of the extended abstract. The result of Biber and İncikapi's study (2016) is consistent with the results obtained for the extended abstract structure in this study.

Holmes (2005) trained 28 mathematics teachers in a Web-based training environment and analyzed the results according to SOLO levels. Analysis showed 19% of the math teachers were at a multi-structural level, 30% were at the relational structure level, and only 2.6% met the extended abstract level. The results showed the teachers' thinking levels are highest in the relational structure, while the lowest level appears to be in the extended abstract structure. The research results by Holmes (2005)

parallel the results of the relational structure and the extended abstract structure in this study.

In a study by Milati, Sunardi, and Dyah (2013), reading texts in mathematics textbooks were analyzed according to the SOLO taxonomy, and results showed that 2.3% of the reading texts were uni-structural, 47.3% were multi-structural and 50% were at the level of relational structure. The results of the study conducted by Milati et al. (2013) show parallelism with this study in terms of relational structure and multi-structural being the most prominent thinking structure groups. In another study conducted by Gezer and Ilhan (2016), the achievements of the Citizenship and Democracy Education Course were examined according to the SOLO taxonomy. The results showed most of the attainments (66.67%) were determined at the multi-structural level, a result contradictory to the results of this study. In yet another study, on the analysis of oral communication attainments of 6th, 7th, and 8th grade Turkish course curricula according to the SOLO taxonomy conducted by Kurt (2016), most of the attainments were of the relational and uni-structural level. Results of the aforementioned research showed that most of the attainments were related to the relational structure, which support the results of this research. On the other hand, the results of the same study for the uni-structural level is inconsistent with the results obtained for the relational structure of this research.

A total of 162 evaluation questions in the sociology textbook relating to the third sub-problem have been analyzed. As a result of the analysis, it was determined that 70 of the evaluation questions were at a uni-structural level, 44 were at a multi-structural level, 36 were relational, and 12 were at the extended abstract level. As a result of Gezer and Ilhan's (2015) study, in which they analyzed the evaluation questions of the social studies curriculum course book according to SOLO taxonomy, it was determined that most of the questions were asked to measure at the uni-structural level, followed by multi-structural, relational and extended abstract levels, respectively. These results are in line with the results of this study. Gezer and Ilhan (2014) also analyzed the evaluation questions related to the Citizenship and Democracy Education lesson taught in primary schools and found that the evaluation questions were mostly at the uni-structural level, then multi-structural, and, lastly, relational structure. The results of Gezer and Ilhan (2014) are only in line with the results of the evaluation questions obtained in this study in terms of uni-structural level. It is contradictory, on the other hand, in terms of the order of multi-structural and relational levels. Moreover, while there is no evaluation question for the extended abstract level in Gezer and Ilhan's (2014) study result, there are 12 questions in this research.

It is thought that when experts prepare program attainments for sociology lessons, organizing the attainments to be in harmony with each other on the basis of prerequisites, using the principle of progressivity, would be of considerable benefit in terms of effectiveness of the program. Also, a balanced distribution of the different levels of thinking with regard to the SOLO taxonomy on a unit basis would also contribute to student attainment of different levels of thinking. It is important to make a balanced distribution in the preparation of both the attainments and evaluation questions in terms of both a unit basis and the thinking levels.

Using attainments as a guide in the preparation of evaluation questions will contribute to the consistency between the attainments and the questions. The fact that there is a high level of inconsistency between the attainments and evaluation questions relating to the 11th grade sociology course in terms of uni-structural and relational structural levels suggests that it is important to review both the attainments and the evaluation questions in order to bring up them to an appropriate level.

References

American Sociological Association (ASA), (2013). National Standards for High School Sociology Asa Natıonal Standards for High School Sociology American Sociological Association. 2013. "21st Century Careers with an Undergraduate Degree in Sociology." Second Edition. Washington, DC: American Sociological Association

Ari, A. (2013). Bilissel alan siniflamasinda yenilenmis bloom, solo, fink, dettmer taksonomileri ve uluslararasi alanda taninma durumlari. [Revised bloom, solo, fink, dettmer taxonomies in cognitive area classification and their international recognition cases]. *Usak Universitesi Sosyal Bilimler Dergisi*, (6)2,259-290.

Baghdad, O. (2013). *Ilkogretim 8. Sınıf ögrencilerinin cebirsel düsünme becerilerinin solo taksonomisi ile incelenmesi.* [Investigation of the 8th grade students' algebraic thinking skills with solo taxonomy].Yüksek Lisans Tezi. Eskisehir Osmangazi Universitesi Egitim Bilimleri Enstitüsü, Eskisehir.

Baghdad, O., & Saban, P. A. (2014). Ilkögretim 8. Sınıf ögrencilerinin cebirsel düsünme becerilerinin solo taksonomisi ile incelenmesi. [Investigation of the 8th grade students' algebraic thinking skills with solo taxonomy]. The *Journal of Academic Social Science Studies (JASS*), (26), 473-496.

Banoglu, K., & Bas, Y. (2012). Sosyoloji dersi ögrenci algilarinin derse verilen önem, sosyal kazanim ve ögrenme ortami boyutlari acısından incelenmesi. [An investigation of students' perceptions of their sociology course in terms of the course's importance, their social attainment and the learning environment]. *Egitim ve Bilim,* (37), 164, 31-44.

Biber, A. C., & Incikapi, L. (2016). Problems posed by prospective elementary mathematics teachers in the concept of functions: an analysis based on solo taxonomy. *Mersin Universitesi Egitim Fakültesi Dergisi,* 12(3), 796-809.

Biggs, J. B., & Collis, K. (1982). *Evaluating the Quality of Learning: the SOLO taxonomy.* New York: Academic Pres.

Biggs, J. B. (2011). *Teaching for quality learning at university: What the student does.* McGraw-Hill Education (UK).

Brabrand, C., & Dahl, B. (2009). Using the SOLO taxonomy to analyze competence progression of university science curricula. *Higher Education, 58*(4), 531-549.

Burnett, P.C. (1999). Assessing the Outcomes of Counselling within a Learning Framework. Paper presented at the Annual Conference of American Educational Research Association (Montreal, Quebec, Canada).

Can, S. (2006). Türkiye'de sosyal bilimler üzerine düsünmek. [Thinking about social sciences in Turkey]. Bilim ve Aklin Aydınlıgında Egitim. (Retrieved December 25, 2016 from (http://yayim.meb.gov.tr/dergiler/sayi75/sayi75/sevim%20can.pdf)

Celik, D. (2007). *Ogretmen adaylarinin cebirsel düsünme becerilerinin analitik incelenmesi.* [Analytical examination of the preservice teachers' algebraic thinking skills] Yayınlanmamis Doktora Tezi, Karadeniz Teknik Universitesi, Trabzon.

Claesgens, J., Scalise, K., Wilson, M., & Stacy, A. (2009). Mapping student understandingin chemistry: The perspectives of chemists. *Science Education, 93*(1), 56-85.

Collis, K. F., & Biggs, J. B. (1979). Classroom Examples of Cognitive Development Phenomena: the SOLO Taxonomy. (Retrieved December 12, 2016 from https://eric.ed.gov/?id=ED185055

Gezer, M., & Ilhan, M. (2014). Ilkögretim Vatandaslik ve Demokrasi Egitimi Dersi. (8. Sınıf) kazanimlari ile degerlendirme sorularinin solo taksonomisi'ne göre incelenmesi. [An evaluation of the assessment questions in the textbook and objectives of the 8th grade curriculum citizenship and democracy education course according to solo taxonomy]. *Dogu Cografya Dergisi, 19*(32), 193-208.

Gezer, M., & Ilhan, M. (2015). Sosyal bilgiler dersi ögretim programi kazanimlari ile ders kitabi degerlendirme sorularinin solo taksonomisine göre incelenmesi. [An analyzing of the assessment questions in the textbook and objectives of the curriculum social studies course according to the solo taxonomy]. *Sakarya Universitesi Egitim Fakültesi Dergisi,* (29), 1-25.

Giddens, A., & Sutton, P. W. (2013). *Sociology.* Cambridge, UK: Polity Press.

Goktepe, S., & Ozdemir, A. S. (2013). Ilkögretim matematik ögretmen adaylarının uzamsal görsellestirme becerilerinin SOLO modeli ile incelenmesi. [Examing elementary mathematics teacher landidates' special visualization skills by solo model]. *Kalem Egitim ve Insan Bilimleri Dergisi, 3*(2), 91-146.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of Learning Skills İnterventions on Student Learning: A Meta-Analysis. *Review of Educational Research,* 66, 99-136.

Kurt, A. (2016). Türkce dersi ögretim programı 6, 7 ve 8. sinif sözlü iletisim kazanimlarinin solo taksonomisine göre incelenmesi. [Turkish course education programme, 6, 7 and 8th grade verbal communication gains examining according to the solo taxonomy]. *Bitlis Eren Universitesi Sosyal Bilimler Enstitüsü Dergisi,* 5 (Ek Sayı), 215-228.

Konyalihatipoglu, M.E. (2016). *Ortaokul 7. sinif ögrencilerinin analitik ve bütüncül düsünme stillerinin solo taksonomisi ile incelenmesi.* [A solo taxonomy research on holistic and analytic thinking styles of seventh grade secondary school students]. Yuksek Lisans Tezi. Rize.

Lian, L. H., & Idris, N. (2006). Assessing algebraic solving ability of form four students. *IEJME-Mathematics Education*, 10, 55-76.

Light, G., Calkins, S., & Cox, R. (2009). *Learning and teaching in higher education: Thereflective professional.* California: Sage Publications.

Milati, N., Sunardi, S., & Dyah, N. (2013). Analisis Level Pertanyaan pada Soal Cerita dalam Buku Teks Matematika Penunjang SMK Program Keahlian Teknologi, Kesehatan dan Pertanian Kelas X Terbitan Erlangga Berdasarkan Taksonomi SOLO. *Pancaran Pendidikan*, 2, 83-94.

Miles, M.B. & Huberman, A.M. (1994). *Qualitative Data Analysis: An Expanded. Sourcebook. (2nd Edition).* California: Sage Publications

Musan, M.S. (2012). *Dinamik matematik yazilimi destekli ortamda 8. sinif ögrencilerinin denklem ve esitsizlikleri anlama seviyelerinin solo taksonomisine göre incelenmesi*. [The examınation of 8th grade students' understanding levels of the equation and ineqaulities according to the solo taxonomy in the dynamic mathematics software aided environment]. Yuksek Lisans Tezi. Pamukkale Universitesi Fen Bilimleri Enstitusu Denizli.

O'Neill, G., & Murphy, F. (2010). Guide to taxonomies of learning. UCD Teaching and Learning Resources. (Retrieved December 12, 2016 from http://www.ucd.ie/t4cms/ucdtla0034.pdf

Olsson, T. (2005). *Qualitative Assessment in the Chemical Engineering Curriculum. Myndigheten för nätverk och samarbete inom högre utbildning*. Sweeden: Lund Institute of Technology Lund University.

Pegg, J & Dawey, G. (2012). Interpreting Understading in Geometry: A Synthesis of Two Models. In Lehrer, R., & Chazan, D. (Eds.) *Designing learning environments for developing understanding of geometry and space* (pp. 103-136). New Jersey: Routledge.

Potter, M., & Kustra, E. (2012). A primer on learning outcomes and the SOLO taxonomy. Centre for Teaching and Learning, University of Windsor. (Retrieved January 19, 2016 from www1.uwindsor.cactl/system/files/PRIMER-on-LearningOutcomes.pdf)

Rooney, C. (2012). How am I using inquiry-based learning to improve my practice and to encourage higher order thinking among my students of mathematics? *Educational Journal of Living Theories, 5*, 99-127.

Yazici, N. (2013). *Basarinin olculmesinde solo taksonomiye dayalı hazırlanan rubrik kullaniminin etkisinin karsilastirmali olarak ıncelenmesi.* [A comparative investigating of the effect of rubrics used based on solo taxonomy on the measurement of success]. Yuksek Lisans Tezi. Kahramanmaras Sütcü Imam Universitesi Sosyal Bilimler Enstitusu. Kahramanmaras.

Yildirim, S.G & Baur, S.W. (2016). Development of Learning Taxonomy for an Undergraduate Course in Architectural Engineering Program (Retrieved December 17, 2016 from https://conferences.kstate.edu.aseemidwest2016/files/2016/06/Paper-Yildirim-Baur-1ls8avf.pdf)

Zabun, B. (2012) Turkiye'de sosyoloji dersi ogretim programlarinin amac ve iceriklerinin toplumsal degismeye paralel olarak degisimi. [The changes of the goals and content of sociology curriculums in turkey in parallel with social change]. *GEFAD / GUJGEF*, 32 (1), 157-170.

Zencirkiran, M. (2016). S*osyoloji* [Sociology] (4. Baskı). Bursa: Dora.

Özet

Sosyoloji Dersi Öğretim Programı Kazanımları ve Değerlendirme Sorularının SOLO Taksonomisine Göre Analizi

Atıf:

Korkmaz, F. & Unsal, S. (2017). Analysis of attainments and evaluation questions in sociology curriculum according to the SOLO taxonomy. *Eurasian Journal of Educational Research, 69*, 75-92. http://dx.doi.org/10.14689/ejer.2017.69.5

Problem Durumu: Sosyoloji; toplumsal kurumları, gruplar arasındaki/içindeki sosyal ilişkileri ve etkileşimleri inceleyen bir bilim dalıdır. Sosyoloji toplumsal yaşamı, sosyal değişimleri, sosyal olayları ve insan davranışlarının sonuçlarını konu edinen bir disiplindir. Yine sosyoloji; toplumları, toplumsal hayatı, grupları inceleyen bir bilim dalı olarak tanımlanmaktadır.1924'ten bugüne amaç ve içerikleri birbirinden farklı 9 farklı sosyoloji öğretim programı uygulanmış olup; şu an uygulanmakta olan sosyoloji dersi öğretim programı ise yapılandırmacı yaklaşıma göre 2009 yılında hazırlanmıştır.

SOLO (Structure of the observed Learning outcome) taksonomisi -gözlemlenebilir öğrenme çıktılarının yapısı- John Biggs ve Kevin Collis tarafından 1982 yılında geliştirilmiştir. Öğrenme çıktıları, öğrencilerin dersin sonuna kadar ne yapacaklarını ve yapabileceklerini gösteren ifadelerdir. Sosyoloji Dersi öğretim programındaki kazanımların düşünme seviyelerine göre analizine yönelik olarak edinilecek bilgiler, eğitim programı uzmanlarının kazanımları gözden geçirme, tasarlama ve düzenlemede aşamalarında bir veri kaynağı sağlayacağı düşünülmektedir.

Kazanımların yanında değerlendirme sorularının düşünme seviyelerine göre analiz edilmesi, herhangi bir programın etkililiği hakkında daha doğru yargıda bulunma olanağı sağlar. Çalışma; bu anlayış doğrultusunda kazanımların yanı sıra değerlendirme sorularının da bilişsel düzeyi tespit edilerek; kazanımlar ile değerlendirme soruları arasında tutarlılığa hangi oranda ulaşıldığına yönelik bilgi sunmayı amaçlamaktadır. Diğer taraftan kazanımların ve öğrenme düzeylerinin belirlenmesi; kazanımların sadece belirli bir düzeyde (alt veya üst düzey) yoğunlaşmadan dengeli bir şekilde dağılıp dağılmadığı konusunda da bir bilgi sunması açısından önemli görülmektedir.

Bu araştırma sonucunda; öğretmenlerin kazanımların öğrenme düzeyleri hususunda bilgi sahibi olmaları sağlanarak; öğrenme düzeylerine uygun etkinlik yapmaları konusunda kılavuzluk yapacağı düşünülmektedir. İlgili literatür tarandığında Sosyoloji dersinin kazanımlarını ve değerlendirme sorularını SOLO taksonomisine göre analiz eden herhangi bir araştırmaya rastlanmamıştır. Bu açıdan, yapılan bu çalışmanın literatürdeki boşluğu dolduracağı düşünülmektedir.

Araştırmanın Amacı: Araştırmanın problem cümlesi "11. sınıf Sosyoloji Dersi öğretim programında yer alan kazanımlar ve sosyoloji ders kitabında yer alan değerlendirme soruları SOLO taksonomisine göre nasıl bir dağılım göstermektedir?" şeklinde ifade edilmiştir. Bu ana problem cümlesi bağlamında 11. sınıf Sosyoloji Dersi;

1- Öğretim programı kazanımları SOLO taksonomisi düşünme seviyeleri ünitelere göre nasıl bir dağılım göstermektedir?

2- Ders kitabındaki değerlendirme soruları SOLO taksonomisine göre nasıl bir dağılım göstermektedir?

Yöntem: Nitel araştırma desenine göre hazırlanmış bu çalışmada yöntem olarak doküman inceleme yöntemi kullanılmıştır. Doküman incelemesi, araştırmacının üzerinde çalıştığı konuya göre ulaştığı kaynakları geniş bir zaman diliminde analiz etmeye ve incelemeye dayanan bir yöntemdir. Araştırmanın veri kaynağını oluşturan kazanımlar MEB Talim Terbiye Kurulu Başkanlığı tarafından 2010 yılında yayımlanan 11. sınıf Sosyoloji Dersi Öğretim Programı'ndan alınmıştır. Değerlendirme soruları ise MEB tarafından 2011 yılında kabul edilip, 2016 yılında basılan ve söz konusu bakanlığın internet sitesinde yayımlanan ders kitabından elde edilmiştir. 60 kazanım ve 162 değerlendirme sorusunun SOLO taksonomisine göre analiz süreci tek yönlü yapı, çok yönlü yapı, ilişkisel yapı ve soyutlanmış yapı düşünme seviyeleri kriter alınarak gerçekleştirilmiştir. Kazanım ve değerlendirme sorularının hangi düşünme seviyesine denk geldiğini belirlemek için gösterge fiillerden ve öğretmen kılavuz kitabındaki açıklamalardan yararlanılmıştır. Kazanım ve değerlendirme sorularının değerlendirme süreci iki araştırmacı tarafından önce kazanımlar daha sonra ise değerlendirme soruları ele alınarak her bir araştırmacı tarafından bağımsız olarak SOLO taksonomisine göre incelenmiş; hangi düşünme seviyesine denk geliyorsa karşısına işlenmiştir. Daha sonraki süreçte araştırmacıların işaretlemeleri karşılaştırılarak uyum yüzdelerine bakılmıştır. Hesaplama sonucunda, araştırmacılar arasında 11. sınıf Sosyoloji Dersi kazanımlarına ilişkin 5 kazanımda; değerlendirme sorularında ise 14 soruda görüş ayrılığı olduğu görülmüştür. Görüş ayrılığının

yaşandığı kazanım ve değerlendirme soruları için üçüncü uzmanının görüşüne başvurulmuştur. Uyum güvenirlik hesaplanmasında; [Güvenirlik = Görüş Birliği / (Görüş Birliği + Görüş Ayrılığı)] formülü kullanılmıştır.

Bulgular: Araştırmanın birinci alt problemine ilişkin elde edilen bulgulara göre; SOLO taksonomisi üniteler bazında incelendiğinde sosyolojiye giriş ünitesinde yer alan 5 kazanımın 3'ünün tekli yapı, 1'inin çoklu yapı, 1'inin ise ilişkisel yapıda olduğu; birey ve toplum ünitesinde yer alan 10 kazanımın 4'ünün tekli yapı, 2'sinin çoklu yapıda, 4'ünün ise ilişkisel yapıda olduğu; toplumsal yapı ünitesinde yer alan 5 kazanımın 2'sinin çoklu yapı, 3'ünün ise ilişkisel yapıda olduğu; toplumsal değişme ve gelişme ünitesinde yer alan 10 kazanımın 3'ünün tekli yapı, 2'sinin çoklu yapı, 4'ünün ilişkisel yapı, 1'inin ise soyutlanmış yapıda olduğu tespit edilmiştir. Yine toplum ve kültür ünitesinde yer alan 8 kazanımın 1'inin tekli yapı, 6'sının ilişkisel yapı, 1'inin ise soyutlanmış yapıda olduğu; toplumsal kurumlar ünitesinde yer alan 22 kazanımın 4'ünün tekli yapı, 7'sinin çoklu yapı, 10'unun ilişkisel yapı, 1'inin ise soyutlanmış yapıda yer aldığı tespit edilmiştir. Araştırmanın ikinci alt problemine ilişkin elde edilen bulgulara göre; sosyoloji ders kitabında toplam 162 değerlendirme sorusunun 70'i (%43) tek yönlü yapı, 44'ü (%27) çok yönlü yapı, 36'sı (%22) ilişkisel yapı, 12'si (8) ise soyutlanmış yapı düzeyinde olduğu tespit edilmiştir.

Araştırmanın Sonuçları ve Önerileri: Araştırma sonuçları kendi içerisinde değerlendirildiğinde toplam 60 kazanımdan 15 (%25) kazanımın tek yönlü yapıda olduğu tespit edilirken; değerlendirme soruları içerinde ise tek yönlü yapıya yönelik 70'i (%43) sorunun yer aldığı tespit edilmiştir. Bu sonuçtan yola çıkarak tek yönlü yapı açısından kazanım sayısı ile değerlendirme soruları arasında önemli bir fark olmadığı söylenebilir. İlişkisel yapıya yönelik öğretim programında 28 (%46) kazanım mevcut iken; değerlendirme soruları içerisinde 36 (%22) sorunun ilişkisel yapıda olduğu tespit edilmiştir. Öğretim programları içerinde kazanımlar içerinde 3 (%5) soru soyutlanmış yapıda bulunurken; değerlendirme soruları içerinde ise 12 (%7) soru soyutlanmış yapıda yer almaktadır. Genel olarak 11. sınıf Sosyoloji Dersi öğretim programındaki kazanımlarla, değerlendirme soruları arasında tek yönlü yapı ve ilişkisel yapıya yönelik bir tutarsızlıktan söz edilebilirken; çok yönlü yapı ve soyutlanmış yapıya yönelik ise bir tutarlılık olduğu söylenebilir. Üniteler bazında SOLO taksonomisine yönelik farklı düşünme düzeylerinin dengeli bir şekilde dağılımının sağlanması öğrencilerin farklı şekillerde düşünme seviyesinin kazanımına katkı sağlayacaktır. Diğer taraftan hem kazanımların hem de değerlendirme sorularının hazırlanmasında gerek üniteler bazında gerekse de düşünme seviyeleri açısından dengeli dağılımın yapılmasının önemli olduğu düşünülmektedir. Değerlendirme sorularının hazırlanmasında kazanımların rehber olarak kullanılması kazanımlar ile sorular arasında tutarlılığın sağlanmasına katkı sağlayacaktır.

Anahtar Kelimeler: Sosyoloji Dersi, Öğretim programı, SOLO taksonomisi analizi.

Eurasian Journal of Educational Research 69 (2017) 93-112

Eurasian Journal of Educational Research
www.ejer.com.tr

Personal Factors Predicting College Student Success

Gokcen AYDIN[1]

ARTICLE INFO

Article History:
Received: 07 May 2015
Received in revised form: 18 March 2017
Accepted: 12 April 2017
DOI: http://dx.doi.org/10.14689/ejer.2017.69.6

Keywords
effective learning
college learning
student success
personal variables

ABSTRACT

Purpose: With the changing perspective in modern education systems, success means more than grades and includes emotional, social, cognitive, and academic development. The aim of this study was to investigate the role of personal factors (academic self-efficacy, organization and attention to study, time utilization, classroom communication, stress and emotional components, student involvement with college life) in predicting student success. **Method:** Three hundred and seventeen college students participated in the study, and a demographic information form and the College Learning Effectiveness Inventory (CLEI) were used. A correlational research design was utilized for data analysis. **Findings:** The results indicate that personal variables significantly predicted student success, $\Delta R^2 = .16$, ΔF (6, 310) =10.16, $p<$ .05, and that 16% of the total variance was explained with the model. Among the personal variables of effective learning, stress and time pressure and classroom communication were found to be significant predictors of success. **Implications for Research and Practice**: The findings indicate that students who communicate better and feel more stressed in the classroom reached a higher level of achievement in college learning. The results suggest that activities that increase student communication in the class should be given priority in the classroom environment. In addition, instructors and university counselors should pay attention to the positive relationship between stress and academic success, as a balanced level of stress should not always be feared during studies. For further research, the CLEI should be used with college students in all grades rather than preparatory students to investigate college students' profiles about personal factors.

[1] Corresponding Author: Gokcen AYDIN, Center for Advancing Learning and Teaching, Middle East Technical University, Turkey, gokcenaydn@gmail.com

Introduction

Higher education is important in today's world by means of its contribution to well-qualified graduates, personal development, and economic, scientific, and technological advancements (Nguyen, 2011). The relevance of higher education underlines the issue of student success or college achievement, which is a complex term to define due to multiple personal and institutional factors affecting college success (Mills, Heyworth, Rosenwax & Carr, 2009; Toutkoushian & Smart, 2001). College success is defined by Kim, Newton, Downey and Benton (2010) as "acceptable grade averages, retention toward a degree and attainment of productive life skills" (p. 112). In recent years, the perspective of education systems has undergone a change from "How should we teach students?" to "How should we help students learn?" to develop and maintain the academic achievement of students. It means that student success is related to student and faculty responsibility as well as a campus-wide responsibility (Hunter, 2006).

Studies in recent years have indicated that success means more than grades. Thus, success has been defined in several ways, including whole student development and having many dimensions beyond cognitive and academic factors (Hunter, 2006). According to Pritchard and Wilson (2003), emotional and social factors are crucial in relation to student success, and there is a strong impact of psychological variables on students' academic achievement through students' adjustment to college. Pike and Kuh (2005) emphasized the value of students' behaviors, attitudes, expectations, and their engagement into college life to measure student success. However, Finn and Rock (1997) argued that success means graduating from the institution on time with good grades. In this regard, it is entirely crucial to notice that, even though the emphasis is placed on attaining high grades to signify success, recent years have introduced the idea of considering faculty-student face-to-face interaction as a supplier of an increase in academic success (Crisp, Baker, Griffin, Lunsford & Pifer, 2017).

A vast number of studies have been conducted to examine student success. Kuh (2006) proposed that student success is formed by pre-college experiences (student academic background and readiness for college); student engagement (studying skills, involvement in social life and campus environment); and post-college outcomes (grades and job-related issues). Expectancy-value theory attaches importance to motivation as being a crucial component of academic achievement. According to this theory, motivation is the direct source of expectations for achieving success (Wigfield, 1994). As the level of expectation increases regarding academic tasks, students' motivation increases as well, and they, intentionally and willingly, commit themselves to achieve the desired goals. Similar to this sense, Tinto's academic and social integration model underlines the prevalence of engagement in the new college environment. The more students commit themselves to college, the more they display retention and achievement (Tinto & Pusser, 2006). Furthermore, achievement goal theory is another theory of student success in which defining goals is emphasized because of a leading level of higher achievement (Canfield & Zastavker, 2010).

Research has focused on student success by taking various perspectives into consideration. Kim et al. (2010) proposed that factors affecting student success can be categorized under three types of variables. The first one is previous success in high school (academic) (Wolfe & Johnson, 1995). The second category is constructed by demographics or socio-economic status that are found to influence student achievement and are difficult to change by the students themselves. For example, regarding gender difference in college success, although some of the studies stated that there was no significant difference (Campbell & Fuqua, 2008; Peterson, 2009), others indicated that female students were more successful than male students in the first year of college (Adams, Marsh, Irons, & Carlson, 2010; Mills et al., 2009). The third one is formed by the factors such as perceptions, actions, attitudes, and values that individuals can control and change (Forsyth & Schlenker, 1977). Personal factors are due to many kinds of individual differences. Every individual has the capacity to influence, enhance, or shape his/her own life. Therefore, for the sake of success, personal variables, also known as "psychosocial factors," have taken the attention of researchers. According to Newton, Kim, Wilcox, and Beemer (2008), these factors include attitudes, motivation, usage of campus resources, study approach, etc. (Newton et al., 2008, p. 4). Even though there are several types of personal factors, according to Newton et al. (2008), academic self-efficacy and confidence, strategic organization and study approach, time utilization, stress and emotional components, student involvement with college life and motivation are among the most outstanding personal factors that affect college student success. In this sense, Newton et al. (2008) developed College Learning Effectiveness Inventory, consisting of these six subscales to measure personal variables in college learning. Considering their emphasis on naming these personal variables as the most powerful influences over success, this study was designed to place importance into personal factors and study them separately.

To begin, academic self-efficacy refers to the competency level of students in achieving academic responsibilities, such as tests or homework (Schunk, 1991). Investigating academic self-efficacy is not only a concern for college learning, but also for high school education (Peguero & Shaffer, 2015). A vast number of studies have indicated a positive relationship between academic self-efficacy and college grades (Bong, 2001; Linnenbrink & Pintrich, 2002). As another variable of personal factors, the significance of study habits and attitudes has been underlined to increase college students' success (Crede & Kuncel, 2008). In addition to time management efforts, the ability to organize tasks and concentrate on their studies also affects students' success (Pauk & Owens, 2011). The literature has also advocated the view that the ability to organize time and responsibilities can increase students' level of success. However, when students do not have organizational skills, they tend to view themselves as failures.

Time management, as another personal variable that leads to success, is defined as "the ability to effectively organize your time and responsibilities in order to get most out of your day" (Combs, 2007, p. 74). Combs (2007) also mentioned that time management during the college years creates a difference between students who are

successful in achieving their goals and those who regret being unaware of the critical aspect of managing time. Balduf (2009) also emphasized that a lack of time management leads to underachievement. Furthermore, colleges should be regarded not only as places to provide education, but also as places in which students can develop their social lives. Research has indicated that college students who become a member of social activities engage in the college life and experience the advantages of being social around campus (Reason, Terenzini & Domingo, 2006). Along with activities inside the classroom, evidence shows that attending extracurricular activities forms another part of the personal factors that increase the satisfaction and engagement of students, influencing student success (Pascarella & Terenzini, 2005).

Apart from other personal variables, student achievement is also based on students' satisfaction level regarding the college. Satisfaction is explained by means of various parts of campus life, such as satisfaction about faculty, program quality, college activities and environment, and overall satisfaction about life (Klein, Kuh, Chun, Hamilton & Shavelson, 2005; Pascarella & Terenzini, 2005). Literature has supported the view that students who are emotionally satisfied with the college attend courses at a higher rate and achieve more academic success (Pascarella & Terenzini, 2005). Moreover, Decker, Dona, and Christenson (2007) argued that good relationships between students and faculty members exhibit a greater influence on emotional functioning than academic achievement. Accordingly, learning and teaching happens through open communication (Nurzali & Khairu'l, 2009; Wall, 2007). The increase in sharing and learning happens more frequently by means of communication between student-student and student-teacher. As a crucial factor influencing achievement, student-faculty relationships display a fundamental effect if the faculty is attentive to students' achievements (Komarraju, Musulkin & Bhattacharya, 2010). Similarly, student-student relationships exert a prominent influence over success. Rubin, Graham, and Mignerey (1990) mentioned that students who feel relaxed when communicating demonstrate a higher level of achievement in terms of GPAs at the end of their college education, as openness in communication creates good relationships, which lead to increased academic success. Therefore, classroom communication is viewed as another personal variable that affects student success.

The present study is expected to make a valuable contribution to the literature by investigating the association between personal variables and academic success. Firstly, the study was conducted with a different sample than suggested by Newton et al. (2008), that is, students who were at the beginning of their college education. In addition, the results of the study might contribute to previous studies indicating significant personal factors as well as other affective variables of effective learning (Aydin, 2012) and priorities of prevention facilities, since beginning students would provide valuable information for the development and improvement of services offered by psychological counseling centers. Moreover, an adapted inventory measuring several personal factors can be a practical inventory tool for psychological counselors at university counseling centers while discovering the personal variables that might affect academic success. Furthermore, students might use the inventory

by themselves to gain insight about their strengths and weaknesses. Additionally, the study might be crucial for psychological counselors working with low-achieving first-year students because personal variables, which students are able to exert some control to change by themselves, can be used to increase the academic success of students. Finally, there is a lack of literature concerning first-year college students' success in the Turkish population. Even though there are studies about undergraduate students' success, the literature lacks information about the personal factors that affect academic success of students during their first year of college. Therefore, the present study may offer different perspectives for further research because the role of personal factors is investigated separately by means of a newly adapted instrument measuring effectiveness of college learning. Although there was a previous study investigating personal and affective variables of success in college students (Aydin, 2012), it is necessary to analyze the predictive role of personal variables separately to get a clear picture over the concept of effective learning in college. Considering all the related information regarding the impact of personal factors, the present study aimed at answering the following research question: What is the role of personal factors (academic self-efficacy, organization and attention to study, stress and time pressure, involvement in college activities, emotional satisfaction and classroom communication) in predicting students' academic success?

Method

Research Design

In relation with the purpose of this study, a correlation research design was utilized. The personal variables affecting student success were independent variables, and each personal variable was a subscale of the college learning effectiveness inventory. The dependent variable (students' academic success) was gathered via English proficiency exam scores.

Research Sample

A demographic data form and Turkish version of the College Learning Effectiveness Inventory (CLEI) were administered to language preparatory students of a state university in Turkey. The convenient sampling method was used, and the participants were volunteer students from the pre-intermediate and intermediate English levels at the preparatory school. A total of 317 students participated in the study (50.8% female, 49.2% male). The mean age of the participants was 19.55 with a standard deviation of 2.28, and their ages ranged between 18 and 28. Among the participants, 174 (54.9%) of respondents were pre-intermediate level students and 143 (45.1%) of participants were intermediate level students. The demographic information is indicated in Table 1.

Table 1

Demographic Information of the Participants

Variables	*n*	%
Gender		
Female	161	50.8
Male	156	49.2
English Proficiency Level		
Pre-Intermediate	174	54.9
Intermediate	143	45.1

The language preparatory school, from where participants of the study were selected, aims to provide basic language skills for students whose level of English is below proficiency level during the first year of university. The department functions as a language preparatory school, offering English courses through two semesters. The courses are based on reading, writing, listening, and speaking skills to prepare students for the English medium of instruction in undergraduate study. Students' academic success was obtained via English Proficiency Exam scores given to students at all levels at the end of the year. The exam includes tests consisting of standard grammar and vocabulary, reading comprehension, and listening and writing sections.

Research Instruments and Procedures

Two instruments were used for data collection: The demographic information form was designed by the researcher, including gender, age, and English proficiency level at preparatory school; and the College Learning Effectiveness Inventory (CLEI) developed and revised by Newton et al. (2008) to measure personal factors influencing college student success. The inventory consists of 51 items on a 5-point Likert scale ("1. Never" to "5. Always") and six subscales: Academic Self-efficacy, Organization and Attention to Study, Stress and Time Pressure, Involvement in College Activity, Emotional Satisfaction, and Class Communication. Sample items from the scale are "30. I make study goals and keep up with them," and "36. I feel there are so many things to get done each week that I am stressed." For the scales, high scores indicate expectations to be successful in achieving goals; effective organizational planning; managing pressures, such as procrastination; engaging in activities; encouragement; and active communication with friends and faculty. The Cronbach Alpha levels of scales were .87 (ASE), .81 (OAS), .77 (STP), .81 (ICA), .72 (ES), and .68 (CC) (Newton et al., 2008). A summed score can be calculated for each subscale representing a different personal variable. Similarly, the Cronbach alpha values of the scale with the Iranian sample were between .68 and .79 (Saeed, 2014).

The inventory was adapted to Turkish by Aydin (2012) with a Confirmatory Factor Analysis (CFA) to examine how well the model fit the Turkish population. CFA results confirmed the six-factor structure of the Turkish version of the College Learning Effectiveness Inventory and indicated a significant chi-square value χ^2 (1051) =1957.84, p<.05 with a mediocre fit. However, Items 12 ("I am discouraged

with how I am treated by my instructors.") and 41 ("My friends have good study habits.") were found as insignificant and omitted in the data analysis. In the end, the scale consisted of 49 items in the Turkish adaptation. For internal consistency, Cronbach's alpha level of the total scale was found as .88, and the reliability of the subscales for the Turkish adaptation were .75 (ASE), .79 (OAS), .68 (STP), .73 (ICA), .62 (ES), and .61 (CC). Although the internal consistency value was low for some subscales, the researcher preferred to use this inventory because it exhibited a high quality with different subscales measuring the most prominent factors of personal variables, and there were no learning effectiveness inventories that measured all these psychosocial factors together in Turkish.

Academic achievement scores were gathered via The English Proficiency Exam (EPE) scores administered by the language preparatory school. The data for this study were gathered from preparatory students via the demographic information form and the Turkish version of the College Learning Effectiveness Inventory with an explanation of the current study. Prior to data collection, permission was granted from the Human Subjects Ethics Committee. The scales were administered in approximately 15 minutes via paper-pencil format during class hours with the permission of instructors. The data were collected in two weeks, and the English Proficiency Exam results of each student were collected from the administration at the end of the semester.

Data Analysis

In the study, descriptive statistics (means, standard deviations, and correlation coefficients among variables) and a Multiple Regression Analysis were utilized to analyze the data. Prior to running the analysis, a missing value analysis was conducted and all the necessary assumptions of multiple regression analysis (normality, multicollinearity, homoscedasticity, independence of residuals, and outliers) were checked (Tabachnick & Fidell, 2007).

The normality assumption of the residuals was checked through the histogram and normal probability plot of residuals. According to Tabachnick and Fidell (2007), the distribution of the histogram should not be too peaked or too flat. The histogram (Figure 1) of residuals showed an approximately normal distribution. Moreover, the normal P-P plots (Figure 2) indicated no serious deviation from the straight line.

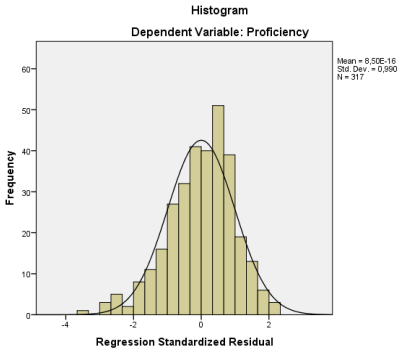


Figure 1

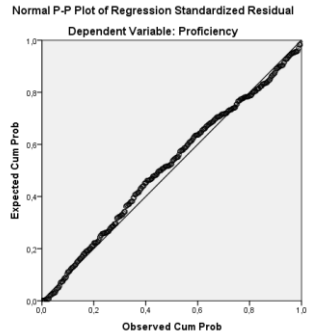


Figure 2

The variance inflation factor (VIF) and tolerance values were checked for the multicollinearity assumption testing high correlations among the independent variables. According to Field (2005), VIF values should be less than 10, and the values of tolerance should be more than .20. Tolerance and VIF value requirements were satisfied, as tolerance changed from .42 to .83 and VIF values from 1.21 to 2.63. Then, scatter plots were checked for the assumption of homoscedasticity, and there seemed to be no violation of assumption as a result of any pattern of scores indicated by randomly spread scatterplots. Moreover, the independence of residuals assumption was checked via the Durbin-Watson value, which should be between 1 and 3 (Tabachnick & Fidell, 2007). It was found as 1.58; therefore, the independence of residuals assumption was satisfied in the current study. Finally, Cook's distances, Leverage, and Mahalanobis distance were checked for assumption relevancy; none of them showed any violation, as Cook's distance was observed as .03; leverage values were within the standards, which should be between 0 and 1; and Mahalanobis distance indicated no outliers, as all the values were lower than critical χ^2 (Field, 2005).

Results

In line with the aim of the study, which was to investigate the role of personal factors in predicting student success, the findings are presented in two sections: the relevant descriptive statistics (means and standard deviations), and the findings of the Multiple Linear Regression Analysis [the effect size (adjusted R^2) of the overall regression model, the associated significance test value (p), and the individual contribution of each predictor (β)]. Descriptive statistics of model showed that the mean of the English Proficiency Exam is 62.88 with a standard deviation of 10.51. Among personal variables, the mean of academic self-efficacy is 57.41 with a standard deviation of 6.47; organization and attention to study *M*=25.27, *SD*=5.76; stress and time press *M*=18.78, *SD*=4.23; involvement in college activities *M*=27.52, *SD*=5.53; emotional satisfaction *M*=21.89, *SD*=3.58; and classroom communication *M*=21.32, *SD*=3.45. The highest mean was academic self-efficacy and the lowest was stress and time press. Then, Pearson's correlation coefficients between independent

variables were analyzed and the level of correlations was suitable for the required limits as indicated in Table 2.

Table 2

Means and Standard Deviations of Subscales and Correlations Between Variables

Variables	*M*	*SD*	*1*	*2*	*3*	*4*	*5*	*6*	*7*
Proficiency	62.88	10.51							
Academic Self-Efficacy	57.41	6.47	.21	1.00	.63	.46	.30	.63	.41
Organization and Attention to Study	25.27	5.76	.13	.63	1.00	.47	.08	.55	.29
Stress and Time Press	18.78	4.23	.36	.46	.49	1.00	.05	.46	.41
Involvement in College Activities	27.52	5.53	.12	.30	.08	.05	1.00	.32	.28
Emotional Satisfaction	21.89	3.58	.26	.63	.55	.46	.32	1.00	.40
Classroom Communication	21.32	3.45	.28	.41	.27	.41	.28	.40	1.00

In terms of correlations, all independent variables were positively correlated with academic success. While the highest correlations were related to stress and time press (r = .36, p<.05) and classroom communication (r = .28, p<.05), correlations for involvement in college activities (r = .12, p<.05) and organization and attention to study (r = .13, p<.05) dimensions were rather low.

Finally, a Multiple Linear Regression Analysis was conducted to examine the relationship between personal factors (academic self-efficacy (ASE), organization and attention to study (OAS), stress and time press (STP), involvement in college activities (ICA), emotional satisfaction (ES) and class communication (CC)) and students' academic success. The results of the analysis are summarized in Table 3.

Table 3

Results of the Multiple Regression Predicting Academic Success

Predictors	*B*	*SE*	*β*	*T*	*P*	*ΔR²*	*ΔF*
Model						*.16**	*10.16*
Academic Self-Efficacy	*.02*	*.13*	*.01*	*.13*	*.09*		
Organization and Attention to Study	*-.21*	*.13*	*-.11*	*-1.59*	*.11*		
Stress and Time Press	*.75*	*.16*	*.30*	*4.70**	*.00*		
Involvement in College Activities	*.08*	*.11*	*.04*	*.76*	*.45*		
Emotional Satisfaction	*.31*	*.21*	*.11*	*1.45*	*.15*		
Class Communication	*.39*	*.19*	*.13*	*2.08**	*.04*		

Note. Dependent Variable = English Proficiency Exam. *p<.05

The regression model was significant as shown in Table 2. Overall, 16% of the variance of the scores can be accounted by the personal factor variables in predicting student success. Results of the multiple regression analysis indicated that personal variables significantly predicted student success, $\Delta R^2 = .16$, ΔF (6, 310) =10.16, $p < .05$. The stress and time press ($\beta = .30$, $p < .05$) and class communication ($\beta = .13$, $p < .05$) appeared as significant predictors of student success. However, academic self-efficacy ($\beta = .01$, *ns*), organization and attention to study ($\beta = -.11$, *ns*), involvement in college activities ($\beta = .04$, *ns*), and emotional satisfaction ($\beta = .11$, *ns*) were not found to be significant predictors in the model.

Discussion and Conclusion

The purpose of the current study was to examine the role of personal variables in predicting students' academic success. The results showed that 16% of the variance in the criterion variable was explained by the model through personal variables. Furthermore, stress and time pressure and classroom communication appeared as significant predictors of success. Contrary to expectations, academic self-efficacy, organization and attention to study, involvement with college activities, and emotional satisfaction were not found as significant predictors.

The variance explained by the model showed that while the model was significant in terms of predicting academic success of students, the explained variance was not good enough. Consequently, the results should be discussed with caution. First, there might be some problems related to the inventory. The Cronbach Alpha of the Turkish version was similar to the original scale, and they were rather low. Similar to the current study, Leung, Ng, and Chan (2011) found the reliability of the subscales to be a little bit lower, as .71 (ASE), .40 (OAS), .43 (STP), .73 (ICA), .45 (ES), and .43 (CC) with college students. This finding may indicate that the inventory might have some problems within itself, or there could be some cultural aspects that cannot be reflected in adaptation studies into other languages.

According to the results, classroom communication, including asking questions in a relaxed classroom environment, offering new ideas, and sustaining good relations with peers and faculty members are positively correlated with academic achievement among the personal variables. This finding was in line with the literature stating that students who had better communication skills found it easy to express themselves and behaved in a more relaxing manner, which led to higher achievement (Reason et al., 2006; Rubin et al., 1990). In addition, the study indicated surprising results for the second personal predictor, which was stress and time pressure. Although the literature stated the adverse effect of stress over student academic success (Alzaeem, Sulaiman & Wasif Gillani, 2010; Bland, Melton, Welle & Bigham, 2012), the findings of the present study showed that stress was a significant positive predictor of students' academic success. It means a higher level of stress results in higher achievement. This finding was partially consistent with the literature underlying the benefit of an appropriate rate of stress for motivation and performance (Cahir & Moris, 1991) and perceived stress resulting in academic success (Jepson & Forrest, 2006). According to Heikkila, Niemivirta, Nieminen and

Lonka (2011), stress can help people survive in critical situations and should not always be considered as negative. Furthermore, the study provided an interesting result to the literature: when stress is balanced and kept at a tolerable level, it might bring success, or when students experience time pressure, they may learn more effectively. Overall, it can be inferred that an appropriate level of stress might be necessary for higher achievement.

The findings of the study did not reveal the other personal variables, such as academic self-efficacy, organization and attention to study, involvement in college activities, and emotional satisfaction, to be significant predictors of academic success in the model, contrary to the literature (Komarraju et al., 2010; Kuh, Kinzie, Schuh, & Whitt, 2005). The study provided surprising results in terms of the insignificant relationship between academic self-efficacy and academic success. Whether students had the belief that they could succeed or not did not correlate to their academic success. Although the positive influence of self-efficacy on academic achievement was underlined in various research studies (Landis, Altman & Cavin, 2007; Linnenbrink & Pintrich, 2002; Margolis, 2005; Zimmerman, 2000), this study provided adverse results compared to previous research. The reason may be because preparatory students participated in the current study; that is, their academic self-efficacy should be considered from the view of language learning self-efficacy and, specifically, their English self-concept should have been studied, as well. As the students were freshmen who recently started college life, they might not have finished the adjustment process. Furthermore, it was found that academic self-efficacy can be influenced by other variables (Peguero & Shaffer, 2015). Additionally, the findings of the present study indicated that students' academic success was not correlated with organization and attention to study. This finding was very surprising because the literature pointed out underachievement as a result of inadequate knowledge of how to study (Balduf, 2009), and that students should pay attention to their studies and responsibilities, concentrate better while studying, and be organized in order to be successful (Pauk & Owens, 2011).

The current study did not provide significant results in terms of the relationship between involvement in college activities and academic success. As another indicator of personal variables, attending college activities was not associated with students' academic success. In the literature, it is highly possible to find a great amount of research supporting that, as the proportion of participation in college activities increased, college students became more successful (Kuh et al., 2005; Pascarella & Terenzini, 2005; Reason et al., 2006). Similar to the findings of the current study, Aitken (1982) found involvement in extracurricular activities as an insignificant predictor of academic success. It is crucial to mention that Aitken (1982) highlighted that the impact of involvement in college activities could be seen in the second year; that is, the first year might not reflect student involvement. This result might be valid for the current study, as well. As the preparatory school building is not close to the center of the university, where most of the extracurricular activities take place, the language preparatory school students might not have been informed about the possibilities and activities around campus, or how to participate in the activities. In

addition, they may also be struggling with other variables, such as loneliness, homesickness, etc.

The personal variable of emotional satisfaction covers the interest of faculty in students' academic success, enjoying the courses and university, instructors' behaviors, and feeling satisfied about future career plans. Unlike the literature (Pritchard & Wilson, 2003; Pascarella & Terenzini, 2005), emotional satisfaction was not found as a significant predictor of academic success in the present study. Literature has stated that when students feel cared for by faculty, they believe in their capacity to achieve more and increase academic self-efficacy (Komarraju et al., 2010). Hence, according to Decker et al. (2007), this belief should be considered as a better indicator of emotional satisfaction than academic success. The reason of the insignificance of emotional satisfaction as a predictor of success might be understandable considering Umbach and Wawrzynski's (2005) suggestion. According to the researchers, students might be prone to ask support not from faculty, but from other sources like friends or family.

Finally, as a newly adapted instrument was utilized within the study, the findings related to the instrument should also be discussed. It was found that reliability results were lower than the original scale, but similar to psychometric properties of the scale adapted into other languages (e.g., Saeed, 2014). One reason might be the nature of the sample that participated in the study. As preparatory school students are the newcomers at the university, they might have not acquired effective learning strategies at college yet, due to the lack of experience and having another major concern, namely learning a foreign language, which requires other competences such as reading, listening, speaking, etc. That is, the scale might not be valid for this group of participants, or there may be some problems related to the Turkish version of the scale. In addition, another reason may be some concerns related to item reflection in the original scale, because the findings of the study were similar to the study that was conducted with the same type of sample, i.e., preparatory school students (Aydin, 2012). The College Learning Effectiveness Inventory was adapted and used for the first time in that study, and although there were other significant affective variables in the study, among personal variables, communication and stress were found as the significant predictors, similar to this study. Due to attaining nearly the same results in both studies, it can be recommended that the instrument should be used with university students in other grades rather than preparatory students in order to attain a clearer picture about personal variables of success regarding concerns for item reflection.

The results of this study can provide some information regarding personal factors related to the academic success of preparatory students. As language preparatory school takes place during the first year of college, possible personal factors can be investigated as the starting point. It is expected that the findings may offer valuable information to language preparatory school administrators, instructors, and university counseling centers who provide psychological help to students. The positive influence of classroom communication over students' academic success might suggest that, when students have a good relationship with each other, feel relaxed while asking questions, and contribute with different ideas to the topic, they

are more likely to attain academic success. Instructors at preparatory school can use this information to create a classroom or outside environment in which students can work and study together with their friends and provide a place to work collaboratively for activities such as group projects, role plays, or other performance studies. In addition, this finding can shed a light for faculty not to use direct instruction methods, as students learn better when they interact among themselves.

University counseling centers can view the invaluable finding of stress and time pressure as a predictor of student success into consideration while preparing activities. The impact of stress can be included in student seminars about stress and academic studies so that students do not put so much pressure on themselves when they feel stressful about their courses or academic studies. Contrary to the literature, the positive relationship between stress and academic success contributes to the literature in a way that perceived stress can influence academic success in a positive way, as well. Instructors at preparatory school may employ this finding, that a balanced level of stress is not always something to fear during studies. Not only the instructors, but also faculty at universities can be informed about the possible positive effect of stress.

In conclusion, in the present study, a newly adapted instrument into Turkish, the CLEI, was used to measure the personal factors of effective learning. The reliability results of the scale demonstrated that the results could have been influenced by the grade level of the participants. Contrary to the hypothesis in the beginning, collecting information about personal variables that influence effective learning might not reflect accurate results when the participants have not yet adapted to university life. Therefore, for further research, the inventory should be used with other grade levels (such as juniors and seniors) in different departments as well as at different universities to explore information about effective learning, as undergraduate students may provide more valuable information related to their experiences at university more than preparatory students who are freshmen and have less experience both academically and socially. As the very similar study indicated nearly the same results, there should be extra studies regarding this inventory. The CLEI can be used to reveal students' profiles about personal factors (organization and attention to study, stress and time press, classroom communication, etc.) that impact their achievements at colleges since it includes a wide range of psycho-social factors.

The current study had also some limitations. The first one was related to generalizability of the results because only the language preparatory students enrolled in a state university participated in this study. Therefore, the results cannot be generalized to the entire population of college students. Although the sample was chosen on purpose that the students were freshmen at the college, enrollment in the language learning preparatory school caused other variables to be taken into consideration. Moreover, the self-reporting nature of the study is another limitation, as the students may have not provided accurate responses to the items. Finally, although the inventory was composed of different psycho-social variables, the low reliability of subscales might have indicated possible problems related to the inventory. Therefore, the inventory should be used by taking all these aspects into consideration.

References

Adams, N.J., Marsh, G., Irons, E.J., & Carlson, N.L. (2010). First time university students: Predictors of success. *National Social Science Journal, 36*(1), 1-7.

Aitken, N.D. (1982). College student performance, satisfaction and retention. *Journal of Higher Education, 53*(1), 32-50. doi:10.2307/1981537

Alzaeem, A.Y., Sulaiman, S.A., & Wasif Gillani, S. (2010). Assessment of the validity and reliability for a newly developed stress in academic life scale (sals) for pharmacy undergraduates. *International Journal of Collaborative Research on Internal Medicine & Public Health, 2*(7), 239-256.

Aydin, G. (2012). *The role of English Proficiency level, personal and affective factors predicting language preparatory school students' academic success.* Unpublished Master's Thesis, Middle East Technical University, Ankara, Turkey.

Balduf, M. (2009). Underachievement among college students. *Journal of Advanced Academics, 20* (2), 274-294. doi:10.1177/1932202X0902000204

Bland, H.W., Melton, B.F., Welle, P., & Bigham, L. (2012). Stress tolerance: New challenges for Millennial college students. *College Student Journal, 46*(2), 362-375.

Bong, M. (2001). Role of self-efficacy and task-value in predicting college students' course performance and future enrollment intentions. *Contemporary Educational Psychology, 26*(4), 553–570. doi:10.1006/ceps.2000.1048

Cahir, N., & Moris, R. (1991). The psychology student stress questionnaire. *Journal of Clinical Psychology, 47*(3), 414-417. doi:10.1002/1097-4679(199105)47:3<414::AID-JCLP2270470314>3.0.CO;2-M

Canfield, C., & Zastavker, Y.V. (2010). *Achievement Goal Theory: A framework for Implementing group work and open-ended problem solving*. 40th ASEE/IEEE Frontiers in Education Conference, Washington, DC.

Campbell, K.C., & Fuqua, D.R. (2008). Factors predictive of student completion in a collegiate honors program. *Journal of College Student Retention, 10*(2), 129. doi:10.2190/CS.10.2.b

Combs, P. (2007*). Major in Success: Make college easier, fire up your dreams, and get a great job*. Ten Speed Press: New York.

Credé, M., & Kuncel, N.R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science, 3*(6), 425-453. doi: 10.1111/j.1745-6924.2008.00089.x

Crisp, G., Baker, V. L., Griffin, K. A., Lunsford, L. G., & Pifer, M. J. (2017). Mentoring undergraduate students. *ASHE Higher Education Report*, *43*(1), 7-103.

Decker, M.D., Dona, P.D., & Christenson, S.L. (2007). Behaviorally at risk African American students: The importance of student-teacher relationships for student outcomes. *Journal of School Psychology, 45*(1), 83-109. doi:10.1016/j.jsp.2006.09.004.

Field, A. (2005). *Discovering statistics using SPSS*. (2nd ed.) London: Sage.

Finn, J.D., & Rock, D.A. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology, 82*, 221-234.

Forsyth, D.R., & Schlenker, L. (1977). Attributing the causes of group performance: Effects of performance quality, task importance and future testing. *Journal of Personality, 45*, 220-236. doi:10.1111/j.1467-6494.1977.tb00148.x

Heikkila, A., Niemivirta, M., Nieminen, J., & Lonka, K. (2011). Interrelations among university students' approaches to learning, regulation of learning, and cognitive and attributional strategies: a person oriented approach. *Higher Education, 61,* 513–529. doi: 10.1007/s10734-010-9346-2

Hunter, M.S. (2006). Fostering student learning and success through first-year programs. *Associations of American Colleges & Universities*, 4-7.

Jepson, E., & Forrest, S. (2006). Individual contributory factors in teacher stress: The role of achievement striving and occupational commitment. *Educational Psychology, 76*, 183-197. doi: 10.1348/000709905X37299

Kim, E., Newton, F.B., Downey, R.G., & Benton, S.L. (2010). Personal factors impacting college student success: Constructing College Learning Effectiveness Inventory (CLEI). *College Student Journal, 44*(1), 112-125.

Klein, S.P., Kuh, G., Chun, M., Hamilton, L., & Shavelson, R. (2005). An approach to measuring cognitive outcomes across higher education institutions. *Research in Higher Education, 46*(3), 251-276. doi:10.1007/s11162-004-1640-3

Komarraju, M., Musulkin, S., & Bhattacharya, G. (2010). Role of student-faculty interactions in developing college students' academic self-concept, motivation, and achievement. *Journal of College Student Development, 51*(3), 332-342. doi: 10.1353/csd.0.0137

Kuh, G.D. (2006). *The National Survey of Student Engagement: Conceptual framework and overview of psychometric properties.* Indiana University Center for Postsecondary Research and Planning Report. Retrieved from http://nsse.iub.edu/pdf/conceptual_framework_2003.pdf

Kuh, G.D., Kinzie, J., Schuh, J.H., & Whitt, E.J. (2005) Never let it rest: Lessons about student success from high-performing colleges and universities. *Change: The Magazine of Higher Learning, 37* (4), 44–51.

Landis, B.D., Altman, J.D., & Cavin, J.D. (2007). Underpinnings of academic success: Effective study skills use as a function of academic locus of control and self-efficacy. *Psi Chi Journal of Undergraduate Research, 12*(3), 126-130.

Leung, C., Ng, C.W.R., & Chan, P.O.E. (2011). Can co-curricular activities enhance the learning effectiveness of students?: An application to the sub-degree students in Hong Kong. *International Journal of Teaching and Learning in Higher Education*, *23*, 329-341.

Linnenbrink, E.A., & Pintrich, P.R. (2002). Motivation as an enabler for academic success. *School Psychology Review, 31*, 313-327.

Margolis, H. (2005). Increasing struggling learners' self-efficacy: What tutors can do and say. *Mentoring and Tutoring, 13*(2), 221-238. doi:10.1080/13611260500105675

Mills, C., Heyworth, J., Rosenwax, L., & Carr, S. (2009). Factors associated with the academic success of first year health science students. *Advances in Health Science Education, 14*(2), 205-217. doi: 10.1007/s10459-008-9103-9

Newton, F.B., Kim, E., Wilcox, D., & Beemer, N. (2008). *Administration and scoring manual for the college learning effectiveness inventory* (CLEI). Kansas State University, Manhattan.

Nguyen, T.K.Q. (2011). Globalization and higher education in Vietnam. *Journal of Interdisciplinary Studies, 23*(1), 117-136.

Nurzali, I. & Khairu'l, N.I. (2009). The effects of classroom communication on students' academic performance at the International Islamic University Malaysia. *Unitar E-Journal, 5*(1), 37-49.

Pascarella, E.T., & Terenzini, P.T. (2005). *How college affects students: A third decade of research*. San Francisco: Jossey-Bass.

Pauk, W., & Owens, R.J.Q. (2011). *How to study in college*. (10th ed.) Boston, MA: Wadsworth Cengage Learning.

Peguero, A. A., & Shaffer, K. A. (2015). Academic self-efficacy, dropping out, and the significance of inequality. *Sociological Spectrum, 35*(1), 46-64.

Peterson, V.M. (2009). Predictors of academic success in first semester baccalaureate nursing students. *Social Behavior and Personality, 37*(3), 411-418. doi:10.2224/sbp.2009.37.3.411

Pike, G.R., & Kuh, G.D. (2005). A typology of student engagement for American colleges and universities. *Research in Higher Education, 46*(2), 185-209. doi:10.1007/s11162-004-1599-0

Pritchard, M.E., & Wilson, G.S. (2003). Using emotional and social factors to predict student success. *Journal of College Student Development, 44*, 18-28. doi:10.1353/csd.2003.0008

Reason, R.D., Terenzini, P.T., & Domingo, R.J. (2006). First things first: Developing academic competence in the first year of college. *Research in Higher Education, 47*(2), 149-175. doi:10.1007/s11162-005-8884-4

Rubin, R.B., Graham, E.E., & Mignerey, J.T. (1990). A longitudinal study of college students' communication competence. *Communication Education, 39*, 1-14. doi:10.1080/03634529009378783

Saeed, M. (2014). Psychometric properties of College Learning Effectiveness Inventory (CLEI). *Journal of Psychological Models and Methods, 3*(14), 91-105.

Schunk, D.H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, *26*, 207-231. doi:10.1080/00461520.1991.9653133

Tabachnick, B.G., & Fidell, L.S. (2007). *Using multivariate statistics*. (5th ed.) Boston: Pearson.

Tinto, V., & Pusser, B. (2006). *Moving from theory to action: Building a model of institutional action for student success*. National Postsecondary Education Cooperative. Retrieved from http://nces.ed.gov/npec/papers.asp

Toutkoushian, R.K., & Smart, J.C. (2001). Do institutional characteristics affect student gains from college. *The Review of Higher Education, 25*(1), 39-61. doi:10.1353/rhe.2001.0017

Umbach, P.D., & Wawrzynski, M.R. (2005). Faculty do matter: The role of college faculty in student learning and engagement. *Research in Higher Education, 46*(2), 153-183. doi: 10.1007/s11162-004-1598-1

Wall, K. (2007). Gesture and its role in classroom communication: An issue for the personalised learning agenda. *Education Review, 19*(2), 32-39.

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review, 6*(1), 49-78. doi: 10.1007/BF02209024

Wolfe, R.N., & Johnson, S.D. (1995). Personality as a predictor of college performance. *Educational & Psychological Measurement, 55,* 177-185. doi:10.1177/0013164495055002002

Zimmerman, B.J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*, 82–91. doi:10.1006/ceps.1999.101

Üniversite Öğrencilerinin Başarısını Yordayan Kişisel Faktörler

Atıf:

Aydın, G. (2017). Personal factors predicting college student success. *Eurasian Journal of Educational Research*, *69*, 93-112. http://dx.doi.org/10.14689/ejer.2017.69.6

Özet

Problem Durumu: Günümüzde yükseköğrenim derecesine sahip olmak yalnızca seçkin mezunlar açısından değil aynı zamanda ekonomi, bilim, teknoloji ve kişisel gelişim için önemlidir. Yükseköğretim kurumları, en önemli misyonlarından biri olarak öğrenci başarısının altını çizmektedir. Farklı teoriler başarı söz konusu olduğunda akademik ve sosyal uyum, beklenti ve motivasyon, hedefleri belirleme gibi noktalara vurgu yapmaktadır. Öğrenci başarı pek çok farklı biçimde tanımlanmasına rağmen, modern eğitim sistemindeki değişen bakış açısıyla birlikte, başarının notlardan çok daha öte bir anlam ifade ettiği görülmektedir. Bu da, öğrencinin bütün olarak gelişiminin, bilişsel ve akademik boyutlarının yanı sıra duygusal ve sosyal boyutlarının da olduğuna işaret etmektedir. Yani öğrenci başarısı üniversite öncesi deneyimler ve üniversite yaşamına katılım gibi pek çok unsurla şekillenmektedir. İlgili alan yazın, diğer etkenlerin yanı sıra, akademik öz-yeterlilik, çalışmaya dikkatini verme ve organize olma, sınıf iletişimi, stres ve zaman baskısı, duygusal etkenler ve öğrencilerin üniversite yaşamına katılımları gibi öğrencilerin kendi kontrolleri altında olan kişisel değişkenlerin öğrenci başarısını etkileyen en önemli faktörler olduğuna işaret etmektedir. Kişisel faktörler bireysel farklılıklardan oluşmaktadır ve her birey kendi yaşamını geliştirme ya da değiştirme kapasitesine sahiptir. Bu nedenle, başarı söz konusu olduğunda "psiko-sosyal faktörler" olarak da adlandırılan kişisel faktörler araştırmacıların odak noktası olmaktadır. İlgili alan yazında etkili öğrenme üzerindeki kişisel faktörleri belirlemede kullanılabilecek Türkçe bir ölçme aracının olmaması sebebiyle geliştirilen kapsamlı ölçme araçlarının uyarlanarak yeni örneklemde kullanılması ve ayrıca etkili öğrenmeyi üzerindeki faktörlerin belirlenerek bu kapsamda yapılabilecek önleyici rehberlik hizmetleri geliştirilmesi açısından bu çalışma önem taşımaktadır.

Araştırmanın Amacı: Bu çalışmanın amacı, üniversitede etkili öğrenme üzerindeki kişisel faktörlerin (akademik öz-yeterlilik, çalışmaya dikkatini verme ve organize olma, sınıf iletişimi, stres ve zaman baskısı, duygusal etkenler ve öğrencilerin üniversite yaşamına katılımları) öğrenci başarısını yordamadaki rolünü araştırmaktır.

Araştırmanın Yöntemi: Bu çalışmada, ilişkisel tarama modeli kullanılmıştır. Çalışmanın örneklemini Türkiye'de bir devlet üniversitesinde öğrenim gören 317 İngilizce Hazırlık Okulu öğrencisi (161 kadın, 156 erkek) oluşturmaktadır. Veri toplama aracı olarak demografik bilgi formu ve Üniversitede Etkili Öğrenme Envanteri kullanılmıştır. Demografik bilgi formu, yaş, cinsiyet, dil seviyesi gibi bilgilerden oluşmaktadır. Üniversitede Etkili Öğrenme Envanterinin Türkçe formu

5'li dereceleme ölçeği (1. Asla, 5. Her zaman) üzerinde 49 madde ve 6 alt boyuttan oluşmaktadır. Bu alt boyutlar, akademik öz-yeterlilik, çalışmaya dikkatini verme ve organize olma, stres ve zaman baskısı, öğrencilerin üniversite yaşamına katılımları, duygusal etkenler ve sınıf iletişimidir. Ölçekte sorulan sorulara verilen yüksek puanlar başarıyla ilgili beklentinin ve hedeflerin olduğunu, etkili planlama yapılabildiğini, erteleme gibi akademik baskılarla baş edilebildiğini, kampüsteki etkinliklere katılımını ve öğretim elemanı ve öğrencilerle iyi bir iletişimi göstermektedir. Ölçeğin alt boyutlarının güvenirlik katsayıları .87 ila .68 arasındadır.

Araştırmanın Bulguları: Çalışmanın amacına uygun olarak yapılan Çoklu Regresyon Analizi sonuçları kişisel değişkenlerin öğrenci başarısını anlamlı şekilde yordadığını göstermektedir, $\Delta R^2 = .16$, ΔF (6, 310) =10.16, $p<.05$. Sonuçlar, bu modelin toplam varyansın % 16'sını açıkladığını göstermiştir. Stres ve zaman baskısı ($\beta = .30$, $p< .05$) ve sınıf iletişimi ($\beta= .13$, $p< .05$) öğrenci başarısını anlamlı şekilde yordarken, akademik öz-yeterlilik, çalışmaya dikkatini verme ve organize olma, duygusal doyum ve öğrencilerin üniversite yaşamına katılımları başarıyı anlamlı ölçüde yordamamıştır. Üniversitede etkili öğrenme üzerindeki kişisel faktörler arasında stres ve zaman baskısı ile sınıf iletişiminin yer aldığı görülmektedir ve iki değişkenin de başarıyla pozitif yönde ilişkili olduğu bulunmuştur.

Araştırmanın Sonuçları ve Önerileri: Bu çalışmaya göre, sınıfla iletişim, stres ve zaman baskısı öğrenci başarısını yordayan değişkenlerdir. Diğer bir deyişle, sınıf içi iletişimi iyi olan ve çalışmaları sırasında stres ve zaman baskısı hisseden öğrencilerin üniversitede etkili öğrenmede daha başarılı olduğu bulunmuştur. Ancak akademik öz-yeterlilik, çalışmaya dikkatini verme ve organize olma, öğrencilerin üniversite yaşamına katılımları başarıyı yordamamaktadır. Seçilen örneklemin üniversitenin hazırlık okulunda öğrenim gören, yani üniversitede henüz ilk yılını geçiren öğrencilerden oluşması, bu durumun kaynağı olarak düşünülebilir. Yani öğrencilerin henüz üniversite yaşamına alışamadıklarından, bu yeni yaşama etkin bir şekilde katılamadıkları ve üniversitede nasıl çalışacaklarına dair ders çalışma becerilerini henüz geliştirememiş oldukları düşünülebilir. Öğretim elemanları bu sonuç ışığında, derslerinde öğrencilerin sınıf içi iletişimi geliştirebilecekleri rol oynama, takım çalışmaları gibi etkinliklere yer verebilirler. Araştırmanın sonucu ayrıca, düz anlatım tekniği kullanan öğretim elemanlarına kullandıkları bu yöntemin öğrenmede etkili olmayabileceği yönünde bir bilgi verebilir. Ayrıca öğretim elemanlarına ve üniversite psikolojik danışmanlarına, stres ve akademik başarı arasındaki pozitif ilişkiye dikkat etmeleri önerilebilir çünkü bazen dengeli bir stres, çalışmalar sırasında her zaman korkulacak bir durum olarak karşımıza çıkmayabilir. Öğrencilere yönelik hazırlanan akademik başarı ve stres hakkındaki seminerlerde ilgili bu sonuçlardan bahsedilerek, öğrencilere kendi üzerlerinde kurdukları baskıyı azaltma yönünde yardımcı olunabilir. Bunun yanı sıra, Türkçeye uyarlanan Üniversitede Etkili Öğrenme Envanterinin yapılan güvenirlik hesaplamaları ve bu çalışma sonucunda elde edilen bulgular, ölçme aracıyla ilgili bazı sorunların olabileceğine ve dikkatle yaklaşılması gerektiğine vurgu yapmaktadır. Gelecekteki araştırmalar açısından, hazırlık sınıfından ziyade birinci, ikinci ya da son sınıf öğrencileri seçilerek üniversiteye uyumunu sağlamış bir örneklem üzerinde çalışılması önerilebilir. Bu doğrultuda,

uyarlanan ölçme aracı, üniversitenin ilk yılı ve dil öğrenme odaklı bir grubunu yansıtan hazırlık öğrencileri dışında farklı sınıf düzeyleri ve farklı üniversitelerde de kullanılarak ilgili alan yazına katkıda bulunulabilir. Ölçme aracı, öğrencilerin etkili öğrenmelerini etkileyen kişisel faktörler açısından oldukça zengin değişkenleri bir arada bulundurması açısından kullanışlı ve yeni bir ölçme aracı olarak düşünülebilir ancak ölçeğin kendi içinde ya da Türkçesi üzerinde bazı sorunlar olabileceği göz ardı edilmemelidir. Son olarak, farklı örneklemlerde kullanılarak uygun sonuçlar elde edilmesi halinde, üniversite psikolojik danışma merkezleri tarafından öğrenci başarısını etkileyebilecek değişkenleri bulmada kullanılabilecek pratik bir envanter olarak görülebilir.

Anahtar sözcükler: Etkili öğrenme, öğrenci başarısı, üniversite öğrenimi, kişisel değişkenler.

Eurasian Journal of Educational Research 69 (2017) 113-134

Eurasian Journal of Educational Research
www.ejer.com.tr

Using Bona Adaptation to Improve Accent Defects as a Voice Training Method *

Kivanc AYCAN[1]

ARTİCLE İNFO

Article History:

Received: 09 March 2016

Received in revised form: 29 July 2016

Accepted: 10 April 2017

DOI: http://dx.doi.org/10.14689/ejer.2017.69.7

Keywords

Bona, Voice Training, intonation, MAXQDA

ABSTRACT

Purpose: In this research, it is observed that if solfeggio syllables, consonants, and vowels are spoken properly, voice intensity (accent), duration, pitch (high pitch-low pitch) and intonation (the ability to carry a musical voice) related to proper pitch level. In this study, it is observed that rhythmic structures do not form without pronouncing the syllables in solfeggio, and, therefore, a voice training system based on bona exercises has been developed.

Research Methods: This research is an action research because it deals with the methods of designing individual lessons. This study examines the vocal exercises that adapted bona (rhythmic articulation exercises) and was performed during the lesson; therefore, it is an action research and is subjective. **Findings:** Bona exercises have an important place in solfeggio training. Solmization syllables have been used especially in bona and solfeggio exercises since they ease solfeggio. Eventually, all of the participants agreed that bona as an individual voice lesson method is helpful to understanding solfeggio principles. **Implications for Research and Practice**: The question of how PVT classes can be of use in solfeggio classes has never been answered. It is deduced from this study that the correct intonation or tone can be obtained by pronouncing the solmization syllables with the principle of correct voice production, using the intensity of the consonants, which affects the intensity, duration, and pitch of the voice. Within this scope, PVT exercises can play a huge role in musical training programs. Otherwise, it will not be of benefit other than drawing attention to the beauty of one's voice. The usage of breathing and bona exercises with solmization syllables during the Personal Voice Training course was useful. Students have found that reading the notes according to their intensities and doing breathing exercises contributes considerably to pronouncing the consonants and finding the tone.

*This study was partly presented at the 2rd İnternational Eurasian Educational Research Congress in Ankara, 8 June – 10 June, 2015

[1] Erciyes University, TURKEY, kivancaycan@hotmail.com

Introduction

It is observed that the main point of origin of the formation of musical notation is the obligation to give information on composition and interpretation. In order for music to be permanent, repeated and thus transferred, a notation was required. Guido d'Arezzo met this requirement in the 10^{th} century by adding three lines to the one-lined staff, adding the notes into a four-lined system using a hymn memorializing the Baptist Saint John and using the first six syllables of this hymn by referring to each note of this six-voiced chord (Hexachord) respectively as *ut, re, la, mi, fa, sol, la* (*Ut Quent laxis/Resonare fibris/Mira gestorum/Famuli tuorum/Solve polluti/Labii reatum*). These syllables were used afterward in some monasteries since they ease solfeggio. However, this notation was subjective, and half pitches formed everywhere because the notes were not fixed in a certain line or space. The clef enabled the fixation of the voices. First, the bass clef was placed on the top line, the place where the half-pitch (fa-mi) has been determined. Then, with the help of the C-clef, the half pitch (do-si) was fixed to the same point, helping to upraise the voices for a fifth, thus providing a stretch of one and a half octaves. Before this invention, melodies were spread around by word of mouth and learned through memorization. Before then, a type of memory support had been provided by *Neuma* (sign) writing. This big invention provided the grounds for multi-vocal music to spread and helped European music to gain an advantage over Eastern music (Aktuze, 2003; Mimaroğlu, 1995). Given the fact that *Neuma* writing did not signify the altitude of the voices but defined only the flow of certain patterns in time (i.e., their rhythmic structures), it may also be regarded as the ancestor of the subsequently developed bona studies (Yazan, 2007).

Today, a system developed by Guido d'Arezzo, called solfeggio, is used worldwide in the beginning of musical education to help students to read and understand music. However, when starting solfeggio practices, it is observed that, instead of teaching the rhythmic patterns, musical voices are usually introduced first. The books derived from the need for teaching the way to read notes with their rhythmic values are limited to the ones written by P. Bona and by C. Pedrone and F. Fontaine, the latter two of whom add to Bona's original work. Music teachers use various methods to teach students how to read notes with their rhythmic values. The most common of these methods is "bona."

The term bona, perceived in Turkish musical education as "reading the notes with their rhythmic values" or "rhythm studies without melody," inherits this name from its creator. In Bona's book, contrary to the solfeggio practices in which melodies are vocalized, the notes are read only by their rhythmic values. The purpose of this method is to prepare the student to read music. In this method, the notes and their values, table of intervals, table of comparison of the expansion of the treble, C and bass clefs and 119 pieces composed by Bona, 24 of which are reviewed Pedrone, are listed in order. Furthermore, there are also pieces written for working on reading treble, C and bass clefs numerated with different numbers together (Sari, 2010). It is more appropriate for the students in the institutions of amateur or professional education to practice reading the notes with their rhythmic values in

order to acknowledge the notes. However, it takes place in only one part of the books on solfeggio, since practicing by reading the notes with their rhythmic values is accepted to be a part of solfeggio exercises. These exercises take place only in the last part in the preface to the solfeggio books. For example, in the program of the Turkish High School of Fine Arts and Sports (Komisyon, 2012), priority is mostly attached to the technical terms (interval notion, values of the notes) and solfeggio exercises under the title of musical signs and terms. In addition, it is conferred that bona practices are applied with the aim of reading the notes with their rhythmic values. However, the difference between these exercises and solfeggio are not sufficiently scrutinized because the notions of voice intervals and the values of the notes are discussed in the bona exercises and can only be of meaning if it is applied.

Usually, however, it is conferred that the bona exercises are not perceived as a practice of reading the notes with their rhythmic values. For this reason, it is observed that the rhythmic patterns and musical terms and applied bona practices are segregated and tend to be memorized with solfeggio. The biggest struggle in the solfeggio practices, which we can describe in short as singing the musical melodies with the names of the notes, is acknowledging the names of the notes and perceiving their values. Before giving examples of the notes from an instrument, especially from a piano to a beginner of musical education, the names and values of the notes must be taught in an applied version. However, this training is usually cut short, and the emphasis is given to the perception of musical voices because the exercises performed are seen as rhythmical exercises without melody (Sari, 2010). For this reason, at the end of solfeggio training, it is observed that students work harder on musical voices instead of taking an interest in musical rhythm patterns. This situation generates the perception and teaching of bona exercises as rhythmical exercises without melody rather than as reading the notes with their rhythmic values, resulting in students' antipathy during these exercises.

The six-voiced chord (Hexachord), also called an *Aretinian syllable,* provides easement in solfeggio and takes its current form with the addition of the syllable "Si." However, today, there is the impression that in order for the syllables to provide an easement in solfeggio, they first need to be spoken. This impression comes from the fact that only the shrillness-gravity or the pitches of the voices are emphasized in solfeggio training. Student views have confirmed this in the study on this matter.

When the alphabetical sounds in Italian (a, b, c, d, e, f, g, h, i, l, m, n, o, p, q, r, s, t, u, v, z) (Tanis, 1974) are examined, it is found that *Aretinian syllables* include the Italian sounds i, e, a, o and u. However, the vowel "U" has lost its place among the sounds in the syllables after the transformation of the syllable "Ut" to the syllable "Do." It is conferred that most of the consonants of the Italian language (b, c, h, n, p, q, v, z) are not used since the first syllables of the text of a hymn are preferred in solfeggio. When it comes to singing, this fact, while providing easement in speaking with the syllables of solfeggio, causes a defect in the unspoken sounds and syllables (i.e., b, c, h, n, p, q, v, z and u). This is the exact point where the importance of bona occurs. In these exercises, which are seen as rhythm exercises without melody but

which help students to read the notes with their rhythmic values, students acknowledge the notes and try to speak them.

Voice training is an important area of application, whether in solfeggio or bona exercises because the subject of voice training is the realization of musical sounds with the use of the human voice. Reading the musical writing with the right sounds and rhythmic patterns in voice training is of importance. Italian voice trainers in the 17^{th} century understood the importance of this and composed pieces without lyrics in order for their students to develop the skill of using their voices and to learn the art of glamorization. Currently, these pieces are called vocal exercises. Each vocal "*ricercare*" (featured or unfeatured compositions predisposed to voice usage) is related to the exercises performed to sing basic polyphony composed of two parts. In this era, *ricercare*s were frequently published, whereas new solfeggios were scarce. The discovery of the Paris Conservatoire in the 18^{th} century regarding exercise in solfeggio changed the teaching plan substantially. Solfeggio had been systematically developed by musicians during the 19^{th} century. The French's interest in Italian methods of training caused solfeggio to spread all around Paris (Sadie, 1980). The origin of the exercises in personal voice training and vocal lessons are based on these methods. However, today, the vocal trainers lack this information and teach the exercises as they had learned in the mentor system. Since the performed exercises are based on memorization, the liaison between solfeggio and the song to be sung is not formed. For instance, it is explained in the preface of the exercise book, "Concone," named after Giuseppe Concone and including his compositions of pieces without lyrics, that the first 25 exercises were intended to be performed in solfeggio. Each note, which is related to its Italian name (do, re, mi, fa, sol, la, si), had to be pronounced and spoken separately, and each tone had to be applied equally throughout and clean in pitch, with the accent of the sound and with the right intonation. As observed in this example, the relation between solfeggio and voice exercises is obvious. Training in singing is mostly applied in amateur and professional musical training and includes, in accordance with the necessities of voice training, imposing upon individuals the right behaviors on how to achieve accurate, beautiful and effective singing while covering the teaching of different songs (Toreyin, 1998). Using this definition, it is possible to assert that the first step in singing a song is to speak the notes with an accent on the consonants of their names; the second step is to reach the right intonation (right pitch, intensity, and duration of the sound, especially in the consonants) by accenting (speaking sounds at the right intensity and duration) the consonants of the syllables in the lyrics and the words. It is thereby possible to conclude that it is the only way to develop the skill of accurate, beautiful and effective singing.

Purpose of the Study

This study originated from the problems that students in personal voice training lessons faced while singing the chosen songs according to the vocal groups. Specifically, students had difficulty with solfeggio while working on the pieces. The

source of these problems is the fact that students could not spend enough time on the solfeggio and its applications because they had received only brief preparations for the fine arts faculty, music department. In the beginning stage and during the development stages of the application, solutions to these problems were attempted while observations were made, and exercises were performed regarding these observations. In standard personal voice, education and vocal training lessons, solfeggio and its application are not frequently used because the backgrounds of the students are accepted as sufficient. However, solfeggio practices comprised a considerable percentage of the lessons because the group of students in this study did not have the adequate background. This study is in accordance with the logic of action research since it determines the application process (Simsek & Yildirim, 2013) of an applicator with respect to the problems occurring in the lesson and their solutions. Since action research is process-oriented (Simsek & Yildirim, 2013), data were collected about the problem, starting with the researcher's Ph.D. process.

In the study prepared with these thoughts in mind, this study intended to help students speak the consonants in solfeggio syllables by introducing rhythmic structures through bona exercises because it is observed that the rhythmic structures are formed properly if the consonants are pronounced. Furthermore, an exercise called repetition of the melody was applied in order to make the unspoken sounds and syllables in solfeggio spoken. In this exercise, the used sounds and especially the non-used ones in solfeggio, such as b, c, h, n, p, q, v, z and u, are combined. For example, syllables such as na, sa, se, si, va, za, ze, etc. have been used. Today, it is confirmed that musical training containing physical exercise is more effective than traditional music education in improving auditorial skills (Ozmentes, 2005). Bona of the notes and melody repetition exercises are studied step by step in this study. Similar studies have been performed while completing authors Ph.D., the researcher conducted similar studies in applying breathing exercises (Aycan, 2012).

Method

Research Design

This study examines the vocal exercises adapted from bona (i.e., rhythmic articulation exercises) performed during a lesson; therefore, it is an action research and is subjective.

Research Sample

In this study, the views of 20 students enrolled in a personal voice training program at Erciyes University, Faculty of Fine Arts and Music Department were collected. Interviews were first recorded and then a developed interview form was applied. In addition, notes were taken on a developed observation form called the individual class observation report.

Research Instruments and Procedures

Due to the sheer volume of the collected data, context analysis was conducted with *MAXQDA plus 11* which is supported by Erciyes University, BAP agency.

Data Analysis

The interview form contents were analyzed with the *MAXQDA plus 11* program.

Validity and Reliability

The observation form questionnaires are presented with the quantitative research expertise from Erciyes University. The collected data are then coded with quantitative research expertise.

Results

Figure 1 describes nineteen of the 20 students participating in this study found the usage of breathing and bona exercises with solmization syllables during the Personal Voice Training (PVT) courses useful. One student responded that he did not have an opinion on this matter.

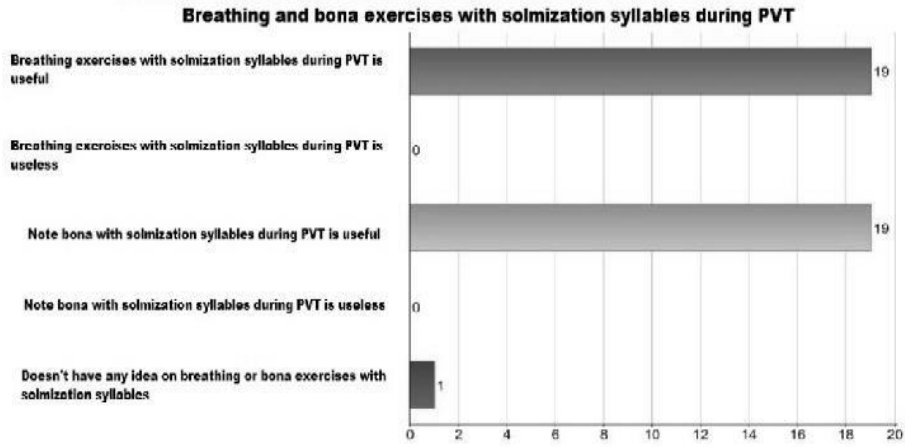


Figure 1. Breathing and bona exercises relationship with the solmization syllables

Figure 2 describes the most frequently performed exercises were options a, b and c, and the next most frequent were options h, d, e, f, and g. According to this pie chart, the students in the study chose exercise option "a" the most (20.3%) and exercise option "g" the least (2.5%). It was observed that the students mostly worked on their postures in front of mirrors.

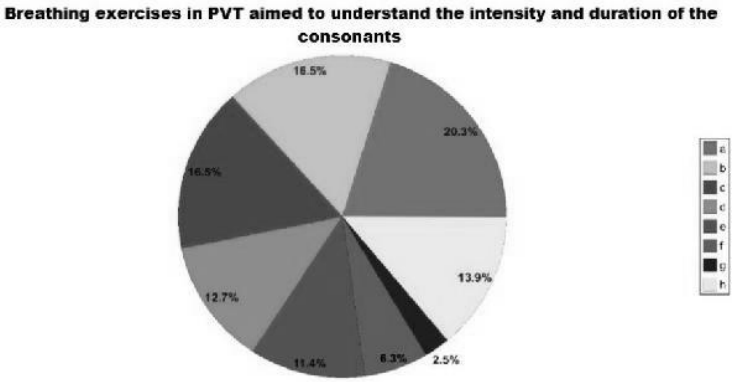


Figure 2. The intensity and duration of the consonants relationship with the breathing exercises

Figure 3 describes it is deduced that the students who participated in the study benefited the most from exercise "c" (31%) and then "b", "d" (26.2%) and "a" (16.7%). While observing the current situation, it is possible to assert that the students benefited nearly the same amount from exercises "b", "c" and "d."

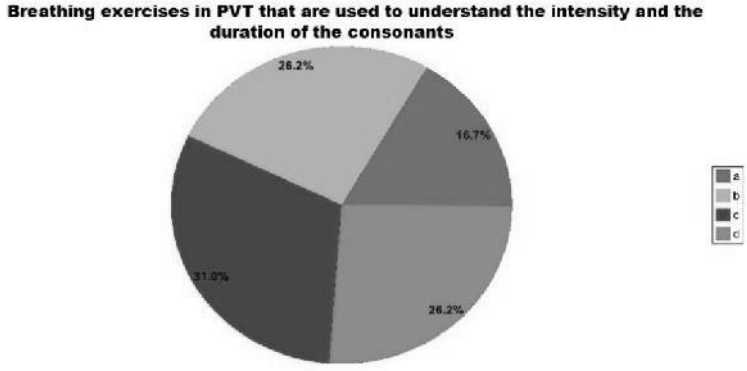


Figure 3. The intensity and duration of the consonants relationship with the voice exercises

Figure 4 describes the students performing in this study chose the following options in descending order of frequency: "a" (32.3%), "b" (30.6%), "c" (29%) and "d" (8.1%). In conclusion, students preferred the bona exercises, given that "a" had the most bona exercises, "b" had fewer, "c" had even fewer and "d" had the fewest.

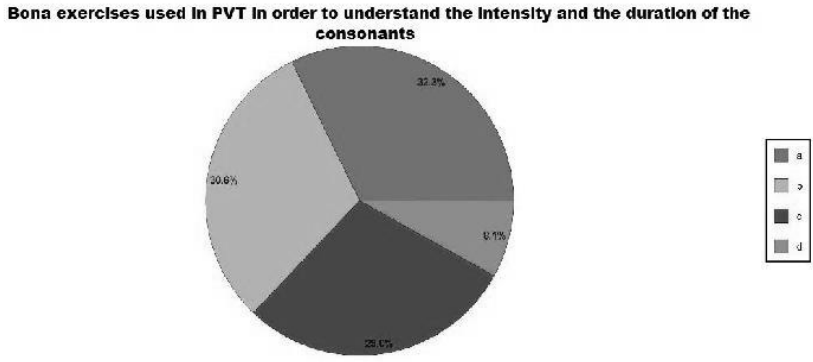


Figure 4. The intensity and duration of the consonants relationship with the bona exercises

Figure 5 describes all 20 students who participated in this study stated that the bona exercises of the notes and words applied in PVT positively affected the duration of their speech (slower or faster than usual speech). Nineteen of the students affirmed that t he bona exercises increased the comprehensibility of the speech (speaking the consonants in the syllables, words, and sentences) and that these exercises helped with the solfeggio course. Nine of the students stated that the exercises helped with their personal instrument course. Two of the students explained that bona exercises helped with the musical theories course and therefore increased their self-confidence in their solfeggio practices. One student said that bona exercises eased his rhythm decipherment, and another student stated that these exercises helped with the chorus lessons.

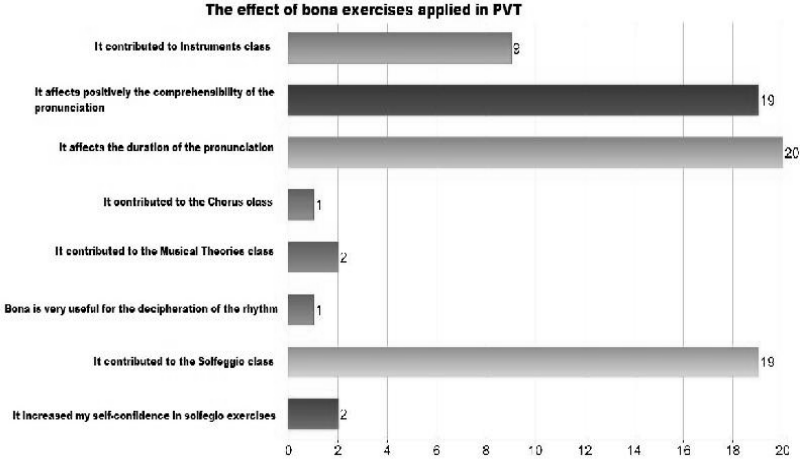


Figure 5. The effect of bona exercises applied in PVT

Figure 6 describes nineteen out of the 20 students who participated in this study stated that there is a relation between bona exercises of the notes and the words applied in the PVT courses and the accentuation of the consonants. One student did not declare an opinion.

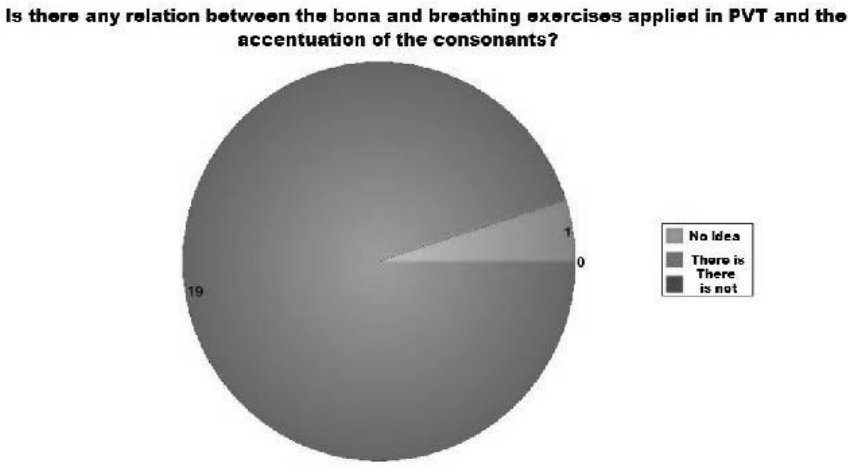


Figure 6. Bona and Breathing exercises relationship with the accentuation of the consonants

Figure 7 describes fourteen of the 20 students who participated in this study stated that being able to sing voice training pieces flawlessly with the support of word bona increased their self-confidence. Twelve students expressed that word bona helps them to sing without being dependent on the piano. Ten students indicated that they could observe the right posture in front of the mirror, and 9 of them saw themselves breathing in and out with the right method in front of the mirror. Seven students stated that the word bona helped them to speak the syllables and words more correctly and assisted them with memorizing the melody. Five of the participants affirmed that bona improved their musical auditory skills and prevented them from singing off tone; these 5 participants also explained that they could observe their mimicry in front of the mirror. Three students stated that the exercises enabled them to discover themselves and increased their efficiency in the PVT classes. Two students stated that bona helped with the vocalization of their voices and their appropriate timbres and enables the teacher to control whether the students had practiced.

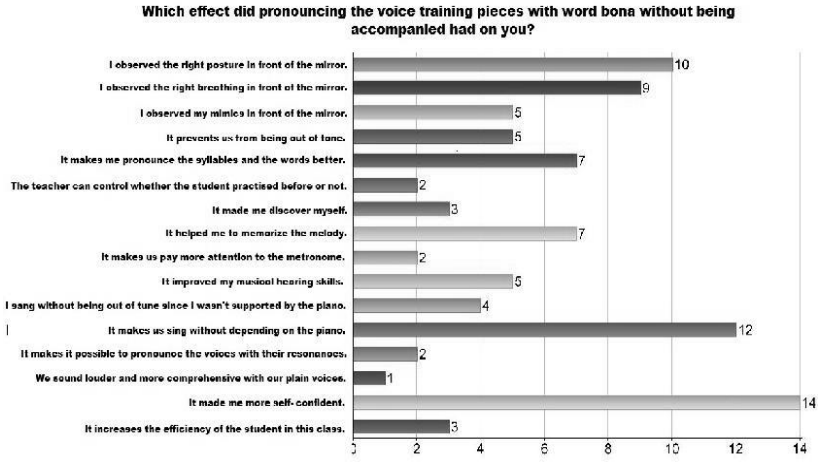


Figure 7. The effect of pronouncing the voice training pieces with word bona without being accompanied

Fifteen of the first and second-year students who entered the Music Department in the 2012-2013 and 2013-2014 academic years were mostly trained in amateur music training. The other 5 students graduated from Anatolian Fine Arts High School (A.F.A.H), where they received professional music training. Six of the 20 students stated that before entering the F.F.A Music Department, they were not trained in any matter relating to solfeggio and its theories in their amateur or professional institutions. It was observed that the 15 students who were in amateur institutions were trained only for short periods, such as 1-3 months, and only 1 student in an amateur institution attended courses for 5 months. Regarding the qualification of the studies during these periods, it is determined that 13 students out of the 20 were not trained with a bona book, and 12 of them had never worked on dictation. In short, the majority of the students affirmed that they were not trained with bona on the theories of solfeggio, and they were not trained in dictation, i.e., the written application of solfeggio. It is determined that the application side of solfeggio education is limited to the solfeggio books of the Muammer Sun-İ and Lavinyak series. The only students who worked on the books in the Lavinyak series are the graduates of A.F.A.H. In addition, 19 of the 20 students participating in this study found it useful to be trained in theories and the application of solfeggio before entering the F.F.A. Music Department; however, 12 of the 15 students who were trained in solfeggio and its theories in amateur education state that since they were not trained in bona and dictation, they could not perceive theoretical and practical music education as a whole (bona, solfeggio and dictation training), and they found the training that they had to be lacking, limited and insufficient when compared to the department's training.

Students who entered the F.F.A. Music Department and found their education "[*sic*] insufficient" stated that during the solfeggio courses, which are intended to teach solfeggio and its theories, bona had to be the priority. Nevertheless, it was not.

Instead, they learned solfeggio and were left by themselves to study. Eleven of the students also affirmed that solfeggio practices should be performed with a group, as doing so increases the memorization of the melodies in their minds. Moreover, in a group, they can ask other group members for help when they are unable to succeed in an application. It is deductible that the students had to work together because the training was insufficient. However, since a group of 25 would be too crowded, the students thought that the practice would be more of use if it were performed in a group of 3-5 or 10-15 people at maximum. They indicated that the teachers could not handle individual problems due to the size of the class is 25 people. Seven participants are of the opinion that the class had to be a private, one-on-one lesson, stating that the efficiency of the class diminishes proportionally with the size of the class. Six students have stated that *voice parameters* (*intensity, duration, accentuation, pitch, and intonation, with intonation being the pronunciation of the voice in accordance with each pitch of the sounds*) are not considered important to achieving the intonation (the harmonious action of a singer with an instrument, vocalizing the voice completely). According to these 6 students, only the pitch of the voice in these parameters is important. The students who think that the aim is only to get the right pitch state that the course keeps relying on memorization. The students confirm that information about the causes of disharmony in intonation is usually given, but individualistic solutions are not being found. While 5 students did not comment on this matter, 9 of them stated that voice parameters were considered important to reaching the correct intonation. The students of the negative opinion gave the serious criticism that only general information is given and that there is no effort toward individual solutions.

When the students are asked about the time they spend on the theoretical parts, especially the studies of solfeggio, it is observed that 15 of them spend around 2-5 min to 15 min on bona exercises. Two of them stated that they did these exercises only during their voice training courses, and 1 student expressed that he did bona exercises in each solfeggio exercise. When the individual views of the students are examined, it is observed that they spend very little time on bona exercises, and their only concern is to perform the right solfeggio and have the right pitch. Their concern over having the right pitch is due to the negative effects of the solfeggio course. Earlier, bona exercises were not performed by one of the participants, but now that same student does bona exercises for about 15 minutes before starting solfeggio. Another student stated that he did not do any bona exercise except for during the personal voice training class.

Suggestions

Unique opinions from the students are as follows:

i. A student stated that when they are vigilant of their intonation while doing solfeggio, as they observed in the voice training class, the bona exercise they did made the melody and rhythmical structure of the solfeggio clearer.

ii. A student stated that after learning how to apply the bona exercises in the voice training class, he first thought of bona while working on each musical note.

iii. A student stated that the subject of bona only occurred in classes such as solfeggio and FMT, and it appeared only as a description; the applied studies were made only for voice training class.

iv. A student stated that when he applied bona exercise as he learned in the voice training class before working on solfeggio, he reached the right intonation, and his voice was vocalized close to its pitch value. He stated that doing word bona in a musical piece affects the nuance and intonation.

Nineteen students out of the 20 who participated in the study found the usage of breathing and bona exercises with solmization syllables (*do, re, mi, fa, sol, la* and *si*) during the Personal Voice Training course useful. One student did not comment on this subject. Students have found that reading the notes according to their intensities and doing breathing exercises contributes considerably to pronouncing the consonants and finding the tone. Additionally, they have stated that they benefited from the pronunciation of the consonants of the solmization syllables according to their intensity and duration in their personal voice training classes and that they have done similar exercises in their solfeggio class. Unique opinions of the students are as follows:

i. Examining note bona and breathing exercises together is very important. It has been possible not to swallow the letters in the pieces with the exercises, for example, whereby the consonant "f" is accentuated while singing the note "Fa."

ii. While performing solfeggio exercise in the PVT class, emphasizing the consonants, such as pronouncing the note "do" as "ddo" by expressing the first consonant better, would create a clearer sound. Through this technique, the singer is speaking rather than aiming the pitch (shrillness-gravity) of the voice, and this aids in the creation of intonation. The purpose of the solfeggio courses is only the pitch of the voice and the right intonation; therefore, there is a relation between the duration and the accentuation of the consonants of the notes. For instance, we cut short the syllable or the consonant in a quarter note, then we say it as an eighth note.

iii. Note bona and word bona applied in PVT class are very useful for solfeggio.

Breathing exercise is very important for the intensity of a note and the right intonation. "Since I play the side-blown flute, the breathing exercises developed by bona exercises have helped me a lot," 1 student reported. This same student explained that the breathing exercises which are related to note bona increases the intensity of the voice and lengthens the duration. Bona and breathing exercises applied in class will be irreplaceable for the voice training class since note and word bona exercises help the words to be pronounced correctly.

In this study, it is determined that exercising by accentuating the consonants of the names of the notes, i.e., the solmization syllables, during the applications which are inspired by bona exercises causes the voices to sound accordingly to the voice intervals of the people (tenor, soprano, mezzo-soprano, baritone, and bass/alto) such as in octave, quartet or quintet intervals. In other words, it is found that bona exercises, seen as exercises without melody/tune, could sound with melody/tune if

the consonants were pronounced. The sounds that appear not to have melody resonate from 1 octave below one's voice interval if the consonants of the syllables in the notes are accentuated. The students who participated in this study have also confirmed that they pronounce the sounds more easily in their solfeggio exercises if they are accompanied by the notes played on the piano in their octaves. They have explained that in situations where only the notes are played or sang, they have difficulty. These difficulties are linked to whether they accentuated the solmization syllables during solfeggio exercises. During the Personal Voice Training (PVT) lesson, in the breathing exercises aiming at the distinction among the intensity and the duration of the consonants, students are mostly (53.5%):

i. Putting their hands on the end of their ribs and their abdomens;

ii. Putting their palms on the sideline of their pants in a parallel position to their bodies, feeling their shoulder blades slightly stretched backwards, standing with their feet open and aligned with their shoulders, putting one foot a little forward and smiling (the edges of the lips are turned up, the lower jaw is a little tense and the tongue is on the back of the teeth in the lower jaw) standing upright;

iii. Standing up, putting their hands in front of them, palms facing down, shoulder blades lowered and frowning (edges of their lips declined or parallel to the floor, loose lower jaw and tongue, hunching their backs;

iv. Observing the difference between standing upright and hunching their backs in front of the mirror;

v. Breathing through the nose on the 4 strikes of a 44 period or breathing through the mouth on the ½ strike of a 44 period, pronouncing the consonants **"s"** or **"f"** as in the Turkish **"hiss"** or **"hiff"**;

vi. Noticing the difference – preferably while practicing in front of a mirror – of intensity and duration between pronouncing the consonants **"s"** or **"f"** on the ½ strike of a 44 period as in the Turkish **"hiss"** or **"hiff"** and **"hiss"** or **"hiff"** while standing with their feet open and aligned with their shoulders, rising a little on their big toes and breathing in through the mouth; and

vii. Performing their exercises in walking tempo, as 1 step breathing in, 1 step breathing out, watching the duration and the intensity of the consonants "s" or "f," pronouncing them as **"hiss"** or **"hiff"**.

During the Personal Voice Training (PVT) class's voice exercises aiming at the distinction among the intensity and the duration of the consonants, students are largely benefiting (by 31%) from performing the following exercises:

i. Working on resonance with double, triple, quartet and quintet voice intervals during voice resonance studies as "*Himm*" or "*Buzzling*" in the beginning or at the end of the PVT classes; and

ii. Singing the melodies in the exercise book by Concone by watching the intensity of the consonants and trying them with words such as **"Filiz, Deniz, Fikret, etc."**

and doing the voice exercises with double, triple, quartet and quintet voice intervals using their own names or other names.

During the Personal Voice Training (PVT) class's bona exercises aiming at the distinction among the intensity and the duration of the consonants, most students (91.9%) stated that they:

i. Performed the bona exercise (note or word bona) aimed to work the diaphragm once or twice a week for 10-15 minutes according to the length and the level of difficulty of the piece, clapping while standing or sitting with their hands aligned with their jaws and shoulders, holding their elbows around 3-4 cm away from their ribs;

ii. Performed note or word bona exercises once or twice a week, placing their hands close to their knees, parallel to the line of their pants while sitting, clapping on their knees at each syllable; and

iii. Customized the breathing exercises during the note or word bona exercises in the voice training pieces by changing the sounds **s-f-s,** which are hard to pronounce, to **s-l-s, s-d-s,** and **s-r-s.** Tried to pronounce the consonant "L" as in the Turkish "*liği*", the consonant "D" as in the Turkish "*diği*", the consonant "R" as in the Turkish "*riği*" and comparing them with "*liği*", "*diği*" and "*riği*," respectively, to see the difference of intensity. When the data of the individual class observation report filled by the students are examined, it is seen that they did not do the note and word bona exercises regularly for their voice training class. The reason is students prefer to memorize solfeggio in a group rather than perform exercises individually. The graduate student from a Korean High School had exercised bona individually ever since he first learned about bona exercises in his voice training class, so his opinions are of value. This student states, "After understanding how to do the bona exercise, I saw that my intonation was correct and clear, and I sang each song feeling its tempo during solfeggio classes. Furthermore, I breathed easily and learned how to use it efficiently. Additionally, with the help of bona exercises, I breathe and use my breath more easily in my exercises with a trumpet with which I was trained in high school and still continue in my individual practices, and I understand better the rhythmical structures of the pieces. Therefore, bona exercises have contributed to my training a lot."

Students' unique opinions on how effective the bona exercises are when applied in their PVT classes are as follows.

i. The note and bona exercises I learned about during by voice training class have corrected some of my mistakes in the duration and comprehensibility of my pronunciation in the pieces I worked on so far.

ii. When I did not pronounce the consonants and swallowed the letters, I could not pronounce the sounds that I easily could by repelling (sounding from the nasal cavity and throat). Sounds were not clear and solid, but these exercises helped to solve problems with the voices sounding more natural.

iii. Pronunciation of the consonants in the syllables, words, and sentences makes it possible to pronounce more easily the lyrics of the songs in Turkish and even more so in a foreign language.

iv. Applied bona exercises contribute to the musical theories class, and therefore we become more self-confident in our solfeggio classes.

Students' unique opinions on the bona and breathing exercises that were applied in PVT classes are as follows.

i. Breathing exercises help us to pronounce the syllables more correctly in the word bona and make the words clearer, correct and comprehensive. Doing note and word bona in PVT class with different syllables that are not used in solfeggio, such as ***"na, nu, no, no, ta, çe, çi, su, so…,"*** contributes to the correct usage of the muscles of the jaw and tongue and to pronounce the pieces with the correct breath (at the correct pitch), clean and comprehensive.

ii. There is a relation between the note and word bona since they both aim to pronounce the consonants and the vowels and, therefore, to develop the right breathing technique and musical hearing.

iii. We work with both breath and the accent of the voice while doing note and word bona exercises. We practice **note and word bona** together, repeating the melodies in word bona with different syllables, such as ***"na, nu, no, no, ta, çe, çi, su, so…,"*** which are not used in solfeggio, and we repeat the melodies by using the syllables in the notes, such as ***"do, re, mi…."***

iv. Bona and breathing exercises need to be performed together since they are related. Furthermore, we aim to pronounce the syllables by first working on note bona and then doing solfeggio with the syllables, such as ***"na, nu, no…"***

v. Co-performance of these two exercises eliminates tempo problems and contributes to the comprehension of the rhythmical structure of the pieces.

Students' unique opinions on the benefit of singing the pieces with the support of word bona without being accompanied and in PVT class are as follows.

i. Since the exercises that we do with the accompaniment of the piano cause us to memorize, we might forget even though we find the right voice at that moment. With this exercise, the sounds are remembered easily, so it is useful.

ii. Singing without the company of the piano is hard but the word bona makes it possible to sing without piano accompaniment. Because bona exercise helps all the voices to be intonated, all the words are pronounced and the rhythms are correct and on time.

iii. Word bona might be the solution to the problems encountered in the speaking parts of the lyrics. This is why bona is an effective method for every student to sing in a healthy way.

iv. When I'm accompanied by the piano, I hide behind its sound, I hold my voice behind, but when I am singing with my plain voice, I am louder.

The opinions of the students who think that "Bona" exercises must be the subject of a separate class are as follows.

i. With the help of a 4-hour-per-week bona class, the difficulties in reading notes (rhythmical and tempo difficulties) would be eliminated. In order for the class to be more efficient, the class has to gather individual groups of a maximum of 10 people and has to progress from easy to hard.

ii. This class is especially necessary in the first and second year in order to acknowledge the notes and to read them. We go to the heart while reading the 64^{th} or 32^{nd} notes in the repertoire classes (especially in the folk music repertoire class). We cannot succeed, including me. Anything to be told to students in these classes has to be taught with the most detailed explanation. Teachers have to observe the students in cooperation.

iii. There is no need for a separate class or a class with the name "Bona." These exercises can be performed in some parts of the PVT classes, but they have to be explained individually.

iv. Students who think that there is no need to teach bona as a separate class state that if the Fundamental Music Theories (FMT) class and solfeggio class are performed efficiently, there won't be any problem. They expressed that teachers have to teach bona by using bona books at least in the first semester of the first year of FMT classes. They also affirmed that bona exercises could also be placed in solfeggio classes by doubling the hours of those classes.

Discussion and Conclusion

After this study, the author offers the following suggestions. The students of this study who entered the Faculty of Fine Arts (F.A.F.) Music Department in the academic years of 2012-2013 and 2013-2014 comprised 15 students who had mostly amateur musical education and 5 students who claimed that the education they had in the professional institutions was not sufficient. This leads us to think that the students who had any type of previous musical education somewhere else had to attend a preparatory class. In the book (Sevgi, 1994), it is stated that the aptitude tests that the students take before entering music departments measure the inclination of the students to have this education, but these tests do not necessarily show that the candidates are ready to receive specialty training. It is expressed that the aptitude test will not be enough to measure the knowledge of the candidate students who previously had not reached their goals in music lessons. However, the insufficiencies and differences among the backgrounds of the students who will receive specialty training can be eliminated by offering a lesson (a preconditioned lesson prepared to give fundamental information) called Musical Hearing –

Reading (ear training and solfeggio), which teaches the language of music, and by opening a preparatory class for the field of expertise lessons.

Uçan (1994) identifies the solution to this issue as an implementation that will make the formation system for music teachers much more solid, consistent and functional if a "Preparatory Class for Music" was opened for the "talented but unprepared" students coming from the high schools apart from the Anatolian Fine Arts High School. Sevgi (1994) underlines the necessity to define the exact timing of each class in this music program correctly in order for the relations to form in a healthy way between fields of expertise lessons.

As observed in this study, the two fundamental classes of musical training, Ear Training and Solfeggio and Private Voice Training (PVT), should be planned in a parallel way in order to support each other. However, bona (rhythmic articulation exercise), which is covered in solfeggio classes, is lacking, and this lack inhibits student progress in the PVT class, which relies on fundamental solfeggio skills.

The question of how PVT classes can be of use in solfeggio classes has never been answered. It is deduced from this study that the correct intonation or tone can be obtained by pronouncing the solfeggio syllables with the principle of correct voice production, using the intensity of the consonants, which affects the intensity, duration and the pitch of the voice. Within this scope, PVT exercises can play a huge role in the musical training program. Otherwise, it will not be of benefit other than drawing attention to the beauty of one's voice. As applied in the Ph.D. studies of Aycan (2012), the accentuation defects in Turkish pronunciation of the students of the Turkish Department of Erciyes University Faculty of Education have changed due to the right breathing and accentuation (the amplitude, height, and pitch of the sound) after programmed voice training.

In this study, students place great importance in the *individuality of bona* and *training exercises*. The fact that bona is not performed individually in solfeggio classes and that personal solution is not found for these exercises in the personal voice training classes negatively affects the students. The fact that the students do not do their exercises regularly is due to not knowing for sure in which area they are going to specialize. It is observed that students frequently mix the notions of private voice training with vocal training, which needs a specialty. This is due to the fact that vocal training, which is oriented to a certain song program (*Opera, lied, popular music* and *jazz*), and *private voice training*, in which fundamental skills for vocal training are given, are not yet distinguished.

The basic objectives of private voice training are explained by Giuseppe Concone in his exercise book, "Concone," as follows: "The first 25 pieces are composed with the intent to be performed as solfeggio. Each note which is related to their Italian names (*do, re, mi, fa, sol, la, si*) has to be pronounced separately, and while breathing out each tone has to be applied with equality, purity, with the intensity of the voice and preciseness of intonation" (Concone, 2014). Concone lays down the relation between solfeggio exercises and voice exercises with the need to pronounce the names of the notes with special attention to the duration, pitch, intensity, and intonation of the

voices. From this perspective, PVT class is the most important supporter of the solfeggio class. However, as expressed by the views of the students in this study, solfeggio, bona and voice training exercises have to be covered simultaneously. Therefore, future studies should apply solfeggio classes to voice training exercises that are supported by bona exercises.

The nonexistence of a musical training program that plans training focused on being a specialist affects the implementation of the classes at the right time as well as the classes' efficiency. Uçan (1994) states that there were four majors in the Gazi Music Teaching program between the years 1994 and 1998, listed as "*Musical Training,*" "*Education on Musical Theories,*" "*Voice Training*" and "*Instrument Training*." He also states that the master's level major, "*Musical Training,*" is a composition of these four majors. These majors under the undergraduate program are not certified by different diplomas and are integrated into the program of music teaching. Planned lessons and their subject matter can be placed within a logic frame among these majors, which have existed up until 1998. Today, however, everything about musical training is covered, and thus specialization is not possible in the organization, such as the Major of Musical Training within the Fine Arts Education Department. This is why topics are not covered in detail. The fact that Erciyes University Fine Arts Faculty, Music Department has only two majors, the Major of Music and the Major of Instrument, inhibits training based on specialization. The PVT class exists in the Major of Music. The classes in the Program of Music are the same as the ones in the Music Teaching departments, except for the formation classes. The credits of the classes were similar to the music teaching programs before the *Bologna Process*.

This shows that a new common program needs to be prepared within the specialty training of the music department program and the education sciences department. The *PVT, solfeggio* and *bona* (*rhythmic articulation*) exercises class prepared within this scope might be useful. However, in order for these preparations to be implemented more efficiently, these exercises could be given in the preparatory class within fundamental music education.

Uçan (1994) indicates that accepting students to the undergraduate program of music training happens under the coordinator OSYM (Student Selection and Placement Center) via the Central Aptitude Test (MOZYES), and the procedures for placing students in the program have been done for the first time with the double staged MOZYES in 1994. It is also stated that the "Exam for Expertise in Medicine" (TUS) can be applied with the OSYM as the coordinator to accept students to into masters and Ph.D. programs.

The quality of the students in the Erciyes University Fine Arts Faculty, Music Department could be increased with the help of this central selection system. The students who are placed within a central system could then reach their objectives more easily after defining what kind of a specialist could be trained with a qualified program in the music department.

In the future, the students' breath and vocal expertise need to be analyzed in experimental studies. The researcher has been planning to analyze intonation and accent among the personal voice training students. If the solmization syllables can be analyzed with vocal analysis programs, such as the PRAAT etc., the intonation and accent differences can be demonstrated with precision.

Acknowledgements

We acknowledge to Assoc. Prof. Dr. Mustafa Ozturk.

References

Aktuze, I. (2003). *Ansiklopedik muzik sozlugu: Muzigi anlamak* [Encyclopedic Music Dictionary: Understanding Music] (Vol. 102). Pan Yayincilik.

Aycan, K. (2012). *Ses Egitimi Yontemlerinin Turkce Konusma Egitimindeki Vurgu Kusurlarinin Duzeltilmesine Etkisi*[The Effects of Vocal Tranining Methods on İmprovement of Accent Defects İn Speech Education]. (PhD), Gazi Universitesi, Egitim Bilimleri Enstitusu.

Bona, P. (1997). *Metodo completo per la divisione Bona*. Milano: Casa Ricordi.

Concone, G. (2014). Fifty Lessons for High and Medium Voice, Op. 9. *Vocal Exercises: The Ultimate Collection* [CD-ROM]: Cd Sheet Music.

Celak, I. (2011). San ve koro egitimi programlarinda kullanilan solfej metotlari ve repertuarindaki olculer konusunda bir saptama ve model onerisi [Solfege methods used in singing and choir education programs and a suggestion and model suggestion on the measures in the repertoire]. *Ankara University, Journal of Faculty of Educational Sciences, 44*(1), 131-145.

Kilic, M. A. (2013). Ses terapisi olarak vurgu yontemi [The accent method as a voice therapy]. *Turkiye Klinikleri J E.N.T. Special Topics, 6*(2).

Komisyon. (2012). Muziksel İsitme, okuma ve yazma-10 [Musical Hearing, Reading, and Writing-10] E. Demirbatir (Ed.) *Guzel sanatlar ve spor liseleri muzik alani.*

Metin, E., C.. (2005, 14-16 Nisan). *Ogretmen Yetistiren Egitim Fakultelerinde Bireysel Ses Egitimi Dersi Iceriginin Incelenmesi* [Investigation of Individual Voice Training Course Contents in Teacher Training Faculties]. Paper presented at the Muzik Sempozyumu, Kayseri.

Mimaroglu, I. (1995). *Muzik tarihi* [Music history]. Istanbul: Varlik Yayinlari.

Ozaltunoglu, O. (2003). *Solfej Ogretim Yontemleri* [Solfa Teaching Methods]. (Yuksek Lisans Tezi). Dokuz Eylul Universitesi, Egitim Bilimleri Enstitusu.

Ozdemir, G. (2012). *Muziksel Okuma (Solfej) Performans Testi Tasarimi* [Musical Reading (Solfege) Performance Test Design]. (Doktora Tezi). Mehmet Akif Ersoy Universitesi, Egitim Bilimleri Enstitusu. (322262)

Ozmentes, G. (2005). *Dalcroze Eurhythmics Ogretiminin Muziksel Beceriler, Muzik Dersine İliskin Tutumlar ve Muzik Yetenegine Iliskin Ozguven Uzerindeki Etkileri* [The Effects of Dalcroze Eurhythmics Instruction on Musical Abilities, Attitudes Towards Music Lessons and Self-Confidence About Musical Skills]. (Doktora Tezi). Dokuz Eylul Universitesi, İzmir.

Sabol, J. W., Lee, L., & Stemple, J. C. (1995). The value of vocal function exercises in the practice regimen of singers. *Journal of Voice, 9*(1), 27-36.

Simsek, G. (2007). *San icin piyano eslikli album* [Piano accompaniment album for professional singers]. Ankara: Muzik Egitimi Yayinlari.

Toreyin, A. M. (2008). *Ses egitimi temel kavramlar- ilkeler- yontemler*. [Voice training basic concepts-principles-methods]. Ankara: Sozkesen Matbaacilik.

Yazan, I. E. (2007). *Konservatuvar Sarkicilik Lisans Programlarinda Solfej Egitiminde İzlenen Kaynak ve Yontemlerin Analizi.* [Analysis of sources and methods in Solfege Training in Conservatory Singing Undergraduate Programs] (Yuksek Lisans Tezi). T.C. Baskent Universitesi, Sosyal Bilimler Enstitusu. (211030)

Bona Uyarlamalarının Vurgu Kusurlarının Düzeltilmesi Amacıyla Bir Ses Eğitimi Yöntemi Olarak Kullanımı

Atıf:

Aycan, K. (2017). Using bona adaptation to improve accent defects as a voice training method. *Eurasian Journal of Educational Research*, *69*, 113-134. http://dx.doi.org/10.14689/ejer.2017.69.7

Özet

Problem Durumu: Bona çalışmaları solfej eğitimi içerisinde önemli bir yere sahiptir. Özellikle bona çalışmalarında ve solfej uygulamalarında kullanılan 11.yy.da rahip Guido di Arezzo'nun bir ilahiden çıkardığı Aretinian heceleri ve bu hecelere daha sonra eklemlenen "si" hecesi solfeji kolaylaştırdığı için kullanılmıştır. Fakat günümüzde gerek ses eğitiminde gerek solfej derslerinde bu hecelerin ünsüzlerinin ve ünlülerinin doğru konuşulmasının gerekliliği üzerinde durulmamaktadır.

Araştırmanın Amacı: Bu çalışmada, bahsi geçen heceler konuşulursa sesin şiddeti, süresi, perdesi ve tonlamasında olumlu değişiklikler olduğu gözlemlenmiştir.

Araştırmanın Yöntemi: Çalışma bu ders içerisinde yapılan uygulamaları ele aldığı için bir eylem araştırmasıdır. Bu sebeple de nitel bir nitelik taşımaktadır. Çalışmada Erciyes Üniversitesi, Güzel Sanatlar Fakültesi, Müzik Bölümünden bireysel ses eğitimi dersini alan 20 öğrencinin görüşlerine başvurulmuştur. Görüşmeler önce

kamera kaydına alınmış ardından yapılandırılmış bir görüşme formu uygulanmıştır. Ayrıca çalışmalar uygulanırken bireysel ders izleme tutanağına da notlar alınmıştır. Toplanan verilerin yoğunluğundan dolayı içerik analizi MAXQDA + 11 ile yapılmıştır.

Araştırmanın Bulguları: Özellikle bireylerin bir çalgı eşliğinde ritmik yapıya uyarak çalgıyla uyumlu şekilde, doğru tonlamayla sesleri çıkartabildikleri görülmüştür. Bu heceler konuşulmadığında solfej dersinde bireylerin basari sağlamakta zorlandıkları görülmüştür. Daha solfej dersinde kullanılan bu heceleri bilmeyen ve konuşamayan bireyler bireysel ses eğitimi dersinde de bu sebeple uygulamaları yapamamaktadır. Dolayısıyla çoğunluğu bona ve solfej uygulamalarından temellenen bireysel ses eğitimi dersinde de bireysel çalışmaları yapmakta zorlanan bireylerin basari düzeyleri düşmektedir. Bu çalışmada solfejdeki heceleri konuşmadan ritmik yapının ortaya çıkmadığı görülmüş bu sebeple bona çalışmaları esas alınarak bir ses eğitimi yöntemi geliştirilip, uygulanmıştır.

Araştırmanın Sonuç ve Önerileri: Bu çalışma ile doğru ses üretme prensibiyle solfej hecelerini konuşarak doğru tonlamanın ya da tonun sesin şiddet, süre, perdesini etkileyen unsuz şiddeti ile elde edilebileceği görülmüştür. Bu kapsamda verilecek bir müzik eğitimi programında B.S.E. çalışmaları etkin bir rol alabilir. Yoksa kişideki ses güzelliğini ortaya koymanın ötesinde bir fayda sağlayamaz. Elde edilen bulgulara göre çalışmaya katılan öğrenciler genel olarak bonanın bir ses eğitimi yöntemi olarak kullanılmasında fayda görmüşlerdir. Öğrenciler nota isimlerini şiddetlerine göre okuyup, bununla ilgili nefes çalışmaları yapmanın ünsüzleri konuşmakta ve tona girmekte fazlasıyla katkı sağladığını belirmişlerdir. Ayrıca solfeji kolaylaştırmak için kullanılan solmizasyon hecelerinde ünsüzlerin şiddet ve süresine göre vurgulanarak konuşulmasının bireysel ses eğitimi derslerinde faydalarını gördüklerini, solfej dersinde de bu şekilde bir çalışma yaptıklarını belirtmişlerdir. Nota bonası ve nefes çalışmalarının birlikte ele alınması önemlidir, örneğin "Fa" notasını söylerken "f" unsuzunun vurgulandığı çalışmalar ile parçalarda harfleri yutmadan söylemek mümkün olmaktadır. Bu çalışmada öğrenciler uygulanacak bona ve ses eğitimi çalışmalarının bireyselliğine vurgu yapmaktadırlar. Özellikle solfej dersinde bona çalışmalarına bireysel olarak değinilememesi, bireysel ses eğitimi dersinde ise bu çalışmalara kişisel çözümler getirilmesi öğrenciler üzerinde olumlu etki yaratmıştır. Öğrencilerin çalışmaları düzenli yapmamaları ise müzik eğitimi açısından hangi konuda uzmanlaşacaklarını bilmemelerinden kaynaklanmaktadır. Öğrencilerin bireysel ses eğitimini uzmanlık gerektiren şan eğitimi ile sıklıkla karıştırdıkları görülmüştür. Bunun sebebi belli bir şarkı müfredatına dönük (Opera, lied, popüler müzik ve jazz) eğitimi verilen san eğitiminin temel becerilerin kazandırıldığı bireysel ses eğitimi dersiyle ayrışmamasındandır. Bireysel ses eğitimi dersinin temel hedeflerini ise Giuseppe Concone kendi ismiyle anılan "Concone" egzersiz kitabında şu şekilde açıklamaktadır: "İlk 25 alıştırma solfej olarak söylenilmesi niyetiyle bestelenmiştir. İtalyanca isimleriyle ilişkilendirilen (do,re,mi,fa,sol,la,si) her bir notanın telaffuz edilmesi, söylenilmesi ve soluk verirken her bir tonun eşit/suresince, temiz/ perde değerinde, sesin şiddetiyle ve doğru bir tonlamayla uygulanması gerekmektedir". Aslında G. Concone solfej çalışmaları ve ses

egzersizlerinin ilişkisini nota isimlerinin seslerinin suresi, perdesi, şiddeti ve tonlamasını gözeterek konuşma ve dolayısıyla söyleme gerekliliği ile ortaya koymaktadır. Bu şekilde ele alındığında BSE dersi solfej dersinin en önemli destekçisidir. Ancak bu çalışmada öğrenci görüşleri ile ifade edildiğinde solfej, bona ve ses eğitimi çalışmaları paralel yürütülmesi gereklidir. Dolayısıyla ilerdeki çalışmalarda uygulanan bona uygulaması destekli ses eğitimi çalışmalarının solfej dersleriyle birlikte uygulanmasında yarar görülmektedir.

Anahtar Kelimeler: Bona, Ses Eğitimi, Entonasyon, MAXQDA

Eurasian Journal of Educational Research 69 (2017) 135-158

Eurasian Journal of Educational Research
www.ejer.com.tr

Inventory of Motive of Preference for Conventional Paper-and-Pencil Tests: A Study of Validity and Reliability*

Mehmet Taha ESER[1] Nuri DOGAN[2]

ARTICLE INFORMATION

Article History:

Received: 08 December 2016

Received in revised form: 18 March 2017

Accepted: 04 May 2017

DOI:http://dx.doi.org/10.14689/ejer.2017.69.8

Keywords

Choice of examination type
content validity
exploratory factor analysis
confirmatory factor analysis

ABSTRACT

Purpose: The objective of this study is to develop the Inventory of Motive of Preference for Conventional Paper-And-Pencil Tests and to evaluate students' motives for preferring written tests, short-answer tests, true/false tests or multiple-choice tests. This will add a measurement tool to the literature with valid and reliable results to help determine why students prefer certain exam types and their level of preference. **Research Methods:** In this study, a screening research design was employed during the data collection and the analysis phases. **Findings:** Cronbach's alpha coefficients were calculated for reliability and it was concluded that the inventory was reliable. First, the exploratory factor analysis was conducted; this was followed by a second confirmatory factor analysis and finally a content validity study to determine the construct validity. A total of 14 items, including 11 items according to the results of the exploratory factor analysis, 1 item based on expert opinion and 2 items according to the results of the confirmatory factor analysis were removed from the survey form of the inventory, resulting in a final form containing 20 items. It was observed that the content validity values of each item in every subtest were sufficient. **Implications for Research and Practice**: The study results showed that this inventory was an appropriate instrument for evaluating high school students' preference for paper-and-pencil tests. An inventory developed under the scope of this study may be used to determine the factors predicting the examination type preference levels of students by using different samples. These results may be used when deciding the actions to be taken.

* The present study is based primarily on the master thesis of Mehmet Taha ESER at Hacettepe University, supervised by Nuri DOĞAN, entitled "Factors Affecting Students' Preferences for the Type of Exam."This study was partly presented at the 3rd International Eurasian Educational Research Congress in Muğla, 31 May – 03 June, 2016.

[1] Corresponding Author: Department of Education, Union of Municipalities of Turkey, Ankara, Turkey, tahaeser@gmail.com

[2] Hacettepe University, Ankara, Turkey, nuridogan2004@gmail.com

Introduction

Evaluation is a process of judging based on the comparison of results obtained from a measurement process of criteria (Turgut, 1997). Evaluation usually takes place when the learning process ends, and it is carried out independently from teaching (Gülbahar and Büyüköztürk, 2008). However, methods for evaluating students should be helpful in providing information and feedback on what is learned by students at what level, what they face during the learning process, and how they prepare for exams (Gülbahar and Büyüköztürk, 2008; Birenbaum, 1997;Struyven, Dochy and Janssens, 2005). In the Turkish educational system, usually the post-examination choices of students are considered and discussed. The grade points of students from a large-scale examination are used to allow them to choose their university and department. These large-scale examinations consist of multiple-choice tests, yet the examination type choices of students are never taken into consideration. Students are compressed into a single model and only given multiple-choice test items.

In most traditional methods, student achievement is typically evaluated using mainly written exams, short answer tests, true/false tests and multiple-choice tests (Turgut, 1988; Atılgan, Kan and Dogan, 2009; Gelbal and Kelecioglu, 2007). The classroom and out-of-classroom behaviours of students are followed by using conventional paper-and-pencil tests. Their performance is examined and students are evaluated invarious aspects of the subject. As teachers are used to it, they prefer the traditional paper-and-pencil tests as a measurement tool (Gelbal and Kelecioglu, 2007).

Considering the qualities of the exam types, we see that exams have different advantages and disadvantages. The most significant advantage for multiple-choice, true/false and short-answer tests is that they are quick and easy to score. Written tests offer students an opportunity to demonstrate their knowledge, skills and abilities in a variety of ways. Multiple-choice tests take time and skill to construct; true/false tests encourage guessing; short-answer tests encourage students to memorize terms and details; and written tests require extensive time to grade. Some of these advantages work in the students' favour and some have a positive effect on the validity and reliability of the measurement results (Zoller, 1994). While some researchers and implementers have theoretically mentioned the positive effects of the exam types, there is relatively little research regarding the advantages and disadvantages of the exam types from the eyes of the students (Zoller and Ben-Chaim, 1998; Zoler and Ben-Chaim, 1990).

The initial studies focused on the type of examination chosen by students and whether these choices varied based on gender (Grandt, 1987; Zoller and Ben-Chaim, 1990). The majority of studies since 1994 used the Assessment Preference Inventory developed by Birenbaum (1994, 1997, 2007). Studies after this date mainly reviewed the relations between the learning-related features of students and their assessment preferences. These studies placed emphasis on learning-related qualities, such as assessment preference choices, learning strategies, motivation strategies, learning approaches, study strategies and academic achievement. The findings revealed that there are strong relations between the assessment preference choices of students and their learning-related qualities and emphasized the importance of considering their assessment preferences during the education process (Birenbaum 1997, 2003, 2007;

Biggs, 2003; Struyven, Dochy and Janssens, 2005; Wilson and Fowler, 2005; Birenbaum and Rosenau 2006; Watering, Gijbels, Dochy and Rijt, 2008).

There are various studies on assessments in the literature, particularly for teachers (Cavanagh, 2006; Cooney, Sanchez & Ice, 2001; Kyriakides, 1997; Miller, 2004; Motsoeneng, 2005; Saxe, Franke, Gearhart, Howard & Crockett, 1997; Sherin & Drake, 2009; Uchiyama, 2004, 2005); however, the number of studies on students, particularly in higher education, is limited (Ben-Chaim & Zoller, 1997; Birenbaum & Feldman, 1998; Struyven et. al., 2005 and Zeidner, 1987). These studies indicate that the assessment preferences may vary based on the education, departments and gender (Beller and Gafni, 2000; Ben Chaim & Zoller, 1997; Birenbaum & Feldman, 1998; Birenbaum, 1997; Brown & Hirschfeld, 2007; Bryant, 2001; Büyüköztürk & Gülbahar, 2010; Struyven et al., 2005; Watering et al., 2008; Zoller & Ben-Chaim, 1990). In this sense, the determination of assessment preference of students studying at the education faculties may be considered as an important factor to reflect their viewpoints on education, and to increase the quality of teaching and provide effectiveness in the program.

When we examined the relevant literature in Turkey, we found very few studies which attempted to determine the examination types of students (Gülbahar and Büyüköztürk, 2008; Bal, 2012; Bal, 2012). It was considered necessary to contribute to the field by developing "The Inventory of Motive of Preference for the Conventional Paper-and-Pencil Tests" (IMP-PAPT) as there was scant research to determine the reason for students' preference of an examination type.

Bal (2012) conducted research on the measurement and assessment preferences of prospective classroom teachers in mathematics.The study used the Assessment Preference Scale (APS) tool for the data collection which was developed by Birenbaum (1994) for university students and adapted forthe Turkish culture by Gülbahar and Büyüköztürk (2008). The Assessment Preference Scale used in the study includes mixed types of questions and intends to determine the level of preference of the assessment types in an integrated way, and not to determine the specific assessment type against certain conditions. However, *IMP-PAPT* developed within the scope of thisstudy, doesnot include mixed types of questions and this inventory provides detailed information on the type of assessment preferred under certain conditions. This study is a scale development study, rather than a scale adaptation study. Scale adaptation studies are more limited in terms of time, budget, and in making an international assessment in a cultural sense. They are also limited in researchers' knowledge ofscale development and any literature that has a strong validity and reliability value in relation to the relevant measurement results in the literature (Hambleton and Patsula, 1999). Taking into account the factors mentioned above, a scale development study on the subject has been carried out.

Purpose of the Study

The objective of this study is to develop *IMP-PAPT*for evaluating the motives of students to prefer written tests, short-answer tests, true/false tests and multiple-choice tests. This will add to the literature a measurement tool with valid and reliable measurement results to help determine the motives of students to prefer written tests, short-answer tests, true/false tests and multiple-choice tests and the level of of

preference for these exams. On the other hand, this study will provide teachers with information on the factors affecting the students' preference of examination types and the way these factors affect the examination preference level. Depending on the results, teachers may increase their efforts to develop measurement toold according to the certain qualities of students when they draft examinations to measure the student achievement. It is believed that the factors the teachers pay attention to in the test development process will reflect positively on students, thereby minimizing thenegative effects of tests on students.

In this study, we want to explore which assessment formats are preferred and how students perceive rather conventional assessment formats. Furthermore, we want to investigate the role of perceptions of assessment in the learning process. It is thought that having information about students' preferences for evaluation types will help students become knowledgeable about test anxiety and trait anxiety, as well as identify student learning strategies and learning styles. At the same time, the scale developed within the scope of this study can be used in studies where the factors affecting students' preferences regarding the types of evaluation are to be determined.

Method

Research Design

This study used the screening model. The studies on the screening model by Cohen, Manion and Morrison (2007) indicate that this is an ideal research methodfor studies on variables requiring a wide sample, such as preference and attitude.

Research Sample

The population of the study consisted of the 9th and 12th grade students studying in the central districts of the Bartin province. The exploratory factor analysis (EFA) was used in a study group of 100 student volunteers. The confirmatory factor analysis (CFA) was conducted on the data collected from 783 student volunteers consisting of 485 girls from various high schools (Bartin Davut Firincioglu Anatolian High School, Köksal Toptan Anatolian High School, Bartin Science High School, Bartin Religious Vocational High School) who studied in the Bartin province and completed and agreed to the research application. The 12th year students study in different fields, which are classified as numerical, verbal, equal weight and language. The size of the study group was considered sufficient for both types of analysis (Klein, 1994; Byrne, 1998). The Davis technique was used in the content validity study; and in this context, meant that opinions were received from 12 experts in the field of assessment and educational evaluation who are competent in the related field.

Many studies, which were inspired by Gardner's AMTB, were conducted in the field. Some of them focused on instrumental and integrative orientations for learning. In the Chinese EFL context, Xiong, 2010 investigated motivational differences among middle school students and observed that they had both instrumental and integrative motivation for learning English. In the Iranian EFL context, studies examined learners' motivational orientations and reported high instrumental motivation among foreign language learners (Hashemi and Hadavi, 2014; Vaezi, 2008). In the Turkish context,

some studies supported that finding (Bektas-Cetinkaya, 2012; Koseoglu, 2013; Ozturk and Gurbuz, 2013). All studies indicated the dominance of instrumental motivation among EFL students.

Research Instrument and Procedure

IMP-PAPT drafted by Eser (2011) was created to reveal the motives of preference on the examination types, such as written, short answer, true/false and multiple-choice and to measure the level of preference of these examinations by students. The survey form of the inventory consisted of 34 items. In this study, both exploratory and confirmatory factor analyses were used. CFA and EFA are, in fact, two stages of a whole process and cannot be effectively separated. If the researcher can use these two methods together, the research will achieve a deeper degree of understanding. Anderson and Gerbing suggested that during the procedure of proposing a theory, it is better to establish a model by EFA and verify the model or modify the model by CFA (Anderson & Gerbing, 1990). EFA provides concepts of the hypothesis and calculating tools, which are an important basis and guarantee for the establishment theory in CFA. It is uncertain if anyone in EFA or CFA is omitted in factor analysis (Hu ve Li, 2015).The final form of *IMP-PAPT* consisted of 20 items. Fourteen items were removed from the initial scale, i.e., 1 item by the expert opinion view method, 11 by exploratory factor analysis, and 2 items by confirmatory factor analysis. When writing the items, the motives of preference of student were considered to be the qualities of examinations that were found to be important with respect to validity, reliability and usefulness. Students were asked to state their preference level on the examination types of written, short answer, true/false and multiple-choice. In the process of preparing the inventory, views and feedback were taken from three PhD students and one associate professor, all of whom are experts in the field of measurement tools.

The scoring of the inventory was based on the following: For me, the responses given to the items are not correct=1, partly correct=2 and totally correct=3. When scoring the items, separate scoring was made for each examination type. Points given for each item indicate the level of preference of individuals while the total points indicate the preference level of the concerned examination by individuals. The examination preference levels of individuals indicate a value between one and three, as they were obtained by taking averages. The values closer to three indicate a higher preference level and show that generally a high point is obtained from the motives of preference for the concerned examination. The points of individuals closer to one indicate lower preference level and show that generally a low point is obtained from the motives of preference for the concerned examination.

Results

Results of Exploratory Factor Analysis

The exploratory factor analysis was applied to the items on each subtest to determine the number of dimensions of the subtests in the inventory. As a result of the analysis, the factor loads for the written examination subtest were found to be between 0,32 and 0,69;those for the short answer examination were between 0,32 and 0,68;those for the true/false subtest were between 0,42 and 0,64; and those for the multiple-choice subtest were between 0,31 and 0,66. According to Tabachnick and Fidell (2001), the

factor load value of each item should be 0,32 or higher. Therefore, the factor load lower limit was accepted at 0,32 when deciding the items to remain in the scale. The KMO values for subtests were between 0,71 and 0,75. It was decided that the data number was sufficient for the factor analysis according to the KMO values results. In addition, the Bartlett test results for all tests were found to be significant at a level of 0,01. This result was considered to be proof that the factor analysis could be applied to the data.

When we look at the eigenvalues of the written examination subtest, seven factors were found with eigenvalues higher than one. The variance disclosed by the first factor (eigenvalue 5,806) was found to be 26,392% while the variance disclosed by the second factor (eigenvalue 2,233) was 10,151%. The factors consisting of all components on the written examination subtest were found to explain 65,303% of the total variance. When we look at the eigenvalues of the short-answer examination subtest, eight factors were found with eigenvalues higher than one. The variance disclosed by the first factor (eigenvalue 5,133) was found to be 23,332% while the variance disclosed by the second factor (eigenvalue 1,815) was 8,249%. The factors consisting of all components on the short-answer examination subtest explained 67,231% of the total variance. When we look at the eigenvalues of the true/false examination subtest, eight factors had eigenvalues higher than one. The variance disclosed by the first factor (eigenvalue 5,338) was found to be 24.265%, while the variance disclosed by the second factor (eigenvalue 1,713) was 7,784%. The factors consisting of all components on the true/false examination subtest explained 66,763% of the total variance. When we look at the eigenvalues of the multiple-choice examination subtest, six factors were found with eigenvalue higher than one. The variance disclosed by the first factor (eigenvalue 5,377) was found to be 24.439%, while the variance disclosed by the second factor (eigenvalue 1,839) was 8,359%. The factors consisting of all components on the short-answer examination subtest explained 57,924% of the total variance.

The factor loads and scree plots on the four subtests were examined and a majority of the items in each subtest was collected under a single dimension (Appendix 1, Appendix 2, Appendix 3, Appendix 4). Depending on the factor analysis results, items that are not included in the first dimension and do not have sufficient factor load to be included in any dimension or those that have high or similar factor load in multiple dimensions were removed from the subtests. After evaluating this, it was deemed appropriate to remove 11 items from the test for all subtests (items 13, 15, 16, 17, 23, 26, 27, 30, 31, 32, 34). Experts agreed on the fact that the fourth item was not suitable for the inventory, and, as a result, the fourth item was removed from all subtests regardless of its statistical values.

In conclusion, it was determined that each subtest was one-dimensional and the practice was continued with 22 items taking into consideration the factor loads, eigenvalues, disclosed variance values and scree plots. An inventory was prepared for the motives of preference using four subtests: written examination, short-answer test, true/false test and multiple-choice test. Subsequently, the correlation values between the corrected test points (obtained by subtracting the correlated item from the total point) and item points were checked in order to determine Cronbach's alpha's internal consistency reliability and item discriminating power.

The Pearson correlation of the test and item points for the written examination scale varied between 0,217 and 0,606; the short-answer test scale varied between 0,217 and

0,598; the true/false test scale varied between 0,215 and 0,532; and the multiple-choice test scale varied between 0,236 and 0,571 (Table 1). Since we paid attention to keep the same items for the four subtest types, each item with a test-item correlation of less than 0,20 for any subtest was removed regardless of the test-item correlation level in the subtests (Ebel, 1979, Field, 2009).

Table 1

Item-Test Correlations and Cronbach's Alpha Values for The Written, Short-Answer, True/False and Multiple-Choice Tests.

Item	Written examination	Short answer test	True/false test	Multiple-choice test
1	.392	.456	.471	.315
2	.232	.344	.366	.335
3	.284	.297	.385	.245
4	.554	.447	.532	.523
5	.594	.510	.527	.524
6	.498	.481	.494	.513
7	.606	.598	.497	.518
8	.401	.282	.466	.385
9	.480	.408	.465	.417
10	.395	.434	.417	.571
11	.485	.426	.367	.384
12	.217	.261	.242	.239
13	.418	.217	.233	.253
14	.380	.218	.225	.276
15	.452	.398	.413	.431
16	.393	.438	.426	.488
17	.466	.458	.387	.486
18	.343	.408	.482	.487
19	.577	.512	.506	.470
20	.228	.266	.215	.236
21	.501	.433	.426	.422
22	.560	.398	.396	.479
Cronbach's Alpha	.856	.831	.838	.838

When we looked at the Cronbach's alpha internal consistency coefficients for the points from four subtests on 22 items, we found that these coefficients varied between 0,831 and 0,856. These values are high and the measurement results are sufficiently

reliable. At the same time, the reliability values of the subtest scores are similar and very close to each other with respect to homogeneity.

The factor loads given in Table 1 relate only to the EFA results. Since the EFA was conducted with100 students, and the sample is small, the factor load wastaken as the lower limit of 0.20.

In Table 1, the averages of the item discrimination indices are shown. The mean of the item discrimination indices is 0.39 for the written test, 0.36 for the short answer test, 0.37 for the true/false test, and 0.34 for the multiple-choice test. The subscales are sufficiently distinguished as the average discrimination values for the subtests are over 0.30.

Results of the Confirmatory Factor Analysis

Each subtest was applied to 783 individuals for the confirmatory factor analysis that was planned to test the construct validity of the subtests. The confirmatory factor analyses included the testing of single dimensionality of the subtests as a model. As the second and fourth items caused autocorrelation in some items during the confirmatory factor analysis, these items were removed from the subtests. The confirmatory factor analyses were done after removing the two items. Table 2 includes the model concordance indicators obtained after the confirmatory factor analysis.

Table 2

Model Concordance Indicators According to the Confirmatory Factor Analysis on the Subtests of the Inventory of Motive of Preference for Examinations

Subtest	Chi-square/ 2	GFI/AGFI	NFI	NNFI	CFI	RMSEA	RMR	SRMR
Written	4,81	0,96 / 0,94	0,99	0,99	0,99	0,079	0,032	0,065
Short-Answer	4,52	0,96 /0,96	0,99	0,99	0,99	0,067	0,028	0,057
True/false	4,01	0,97 / 0,96	0,99	0,99	0,99	0,062	0,027	0,052
Multiple-choice	4,01	0,97 / 0,96	0,99	0,99	0,99	0,062	0,028	0,052

Looking at the confirmatory factor analysis result in Table 2, we can state that there is sufficient evidence on the one dimensionality of each subtest. The chi-square statistics in the literature show a lack of index fit (Stapleton, 1997). Therefore, a small chi-square value indicates that the model is fit for the observed structure and vice versa. That is, a big chi-square value indicates that the model does not sufficiently explain the structure. However, as the chi-square statistic is a sum statistics, it will be as high as the number of variants. Therefore, the use of chi-square/degree of freedom might be recommended (Dogan and Basokcu, 2010). Having a chi-square/degree of freedom lower than five indicates that the model fits and a value lower than three indicates that the model has a very good fit (Byrene, 1998). Having chi-square/degree

of freedom values between three and five in the study indicates that the one-dimensional models created for the subtests are fit for the observed structures.

A goodness of fit index is usually a measurement of the variance and covariance amount disclosed by the model. The coefficient of determination calculated in the multiple regression can be interpreted as R^2. The closer the value of the goodness of fit index, the better the fit of the model for the data (Dogan and Basokcu, 2010). For the goodness of fit indices, the values between 0,90-0,95 indicate an acceptable fit; values above 0,95 indicate a high fit (Dickey, 1996; Stapleton, 1997; Byrne, 1998). The values in Table 2 show that the fit indices other than RMR and SRMR are larger than 0,95. The GFI/AGFI, NFI, NNFI and CFI values indicated that the measurement tool had a high fit. Particularly, having the index value of Root Mean Square Error of Approximation (RMSEA) between 0,08-0,05 shows that the model is acceptable, and a value lower than 0,05 shows that the model is good. Particularly, a good fit is indicated by an index value of the Root Mean Square Error of Approximation (RMSEA) closer to 0,00 (Du Toit and Du Toit, 2001). In our study, the RMSEA values lower than 0,08 indicate an acceptable fit. A good fit is also indicated by the fact the RMR and SRMR values are ≤ 0.08, as these two values are indicators of lack of fit (Jöreskog and Sörbom, 1993). A high fit is proven by the fact that the RMR value, which is an indicator of lack of fit, is between 0,027 and 0,032 for each subtest, while the SRMR values are observed to be lower than 0,08 by varying between 0,052 and 0,065. Considering and interpreting all values together provides a verification of the one dimensionality structure of the subtests. The path graph of the confirmatory factor analysis for the subtests is given in the appendices (Appendix5, Appendix 6, Appendix 7, Appendix 8).

Results of Content Validity

For each item in the subtest composing the assessment tool, opinions were received from 12experts in the field of assessment and evaluation in education. In the determination of content validity related to items, the Davis technique (1992) was used. Considering the requirement that a minimum of three experts use the Davis Technique, this number was met as we received opinions from seven experts in terms of content validity. The surveys related to content validity were conducted with the remaining items after the items having a negative effect on content validity were excluded from the test. Using the Davis technique each item related to the subtests were evaluated as 1=*not relevant*, 2=*somewhat relevant*, 3=*quite relevant*, 4=*highly relevant*. When determining the content validity index for each item, the number of experts choosing the option (3) or (4) was divided by the total number of experts to obtain content validity index and 0,80 was determined as the standard value for CVI's (Davis, 1992).

The content validity indexes of the items forming the assessment tool varied between 0,86 and 1 for written examinations, short answer tests, true/false test and multiple-choice tests. Considering that the limit value for the Davis technique is 0,80, the content validity values of each item in every subtest was sufficient.

Discussion and Conclusion

In this study, a scale was developed to determine the levels of high school students regarding their motives of preference for paper-and-pencil tests. The relevant

literature was reviewed to develop the draft scale and then the scale was applied to the high school students. Cronbach's alpha coefficients were calculated for reliability and it was concluded that the inventory was reliable. First the exploratory factor analysis and then the confirmatory factor analysis were conducted to determine the structure validity. A total of 14 items were removed from the survey, including 11 items according to the results of the exploratory factor analysis, 1 item by expert opinion and 2 items according to the results of the confirmatory factor analysis, leaving 20 items in the final form.

The Assessment Preference Scale,developed by Birenbaum (1994) for university students and adapted forthe Turkish culture by Gülbahar and Büyüköztürk (2008) contains similar objectives to the inventory developed in thepresent study and this scale was used in a majority of similar studies (Gülbahar and Büyüköztürk, 2008; Bal, 2012; Birenbaum, 1994; Birenbaum, 1996; Birenbaum, 1997). Further studies may be recommended to examine the criteria validity studyofthe level of relations between the inventory developed in the present study and the Assessment Preference Scale.

The subtests of the inventory developed by the study consist of four traditional examinations: written, short-answer, true/false and multiple-choice test. Future studies may include different types of traditional examinations and the research may revise the scale or develop an inventory of motives of preference for the examination type created by the complementary measurement approach. The inventory developed under the scope of this study may be used to determine the factors predicting the examination type preference levels of students by using different samples. These results may be used when deciding the actions to be done and tools to be used in the assessment process by determining the examination type preferences of the students.

The Assessment Preference Scale used in the study includes mixed types of questions and intends to determine the level of preference of the assessment types in an integrated way, rather than determine aspecific assessment type against certain conditions. However, *IMP-PAPT*developed within the scope of the study does not include mixed type of questions and this inventory provides detailed information on the type of assessment preferred undercertain conditions. As mentioned earlier, this study is a scale development study. Therefore, in order to avoid the difficulties such as limited time, low budget, a language and culture adapted from a different language and culture, a detailed plan was made prior tothe study. As a result, it will be useful for the researchers to make a detailed plan before the scale development studies are carried out.

References

Anderson, J. C. & Gerbing, D.W. (1990). Structural equation modeling in practice: A review and recommended two-step procedure. Psychology Bulletin, V103:411-423.

Atilgan, H., Kan, A., Dogan, N. (2009). *Egitimde Olcme ve Degerlendirme*. Edit. H. Atilgan (3. Basim): Ani Yayincilik.

Bal, A. P. (2012). Sinif Ogretmenligi Ogretmen Adaylarinin Matematik Dersine İliskin Olcme-Degerlendirme Tercihleri. Turk Egitim Bilimleri Dergisi, 10(3), 459-479.

Bal, A. P. (2012). Ogrencilerin Matematik Dersine İliskin Degerlendirme Tercihleri. Selcuk Universitesi Sosyal Bilimler Enstitusu Dergisi, 27, 59-72.

Beller, M., & Gafni, N. (2000). Can item format (multiple choice vs. open ended) account for gender differences in mathematics achievement? Sex Roles, 42, 1-21.

Ben-Chaim, D., & Zoller, U. (1997). "Examination-type preferences of secondary school students and their teachers in the science disciplines", Instructional Science, 25(5), 347–367.

Birenbaum, M. (1994). Toward Adaptive Assessment - The Student's Angle. Studies in Educational Evaluation, 20, 239-255.

Birenbaum, M. & Feldman, R. A. (1998). Relationships between learning patterns and attitudes towards two assessment formats, Educational Research, 40(1), 90–97.

Birenbaum, M. ve Gutvirzt, Y. (1995). On The Relationship Between Assessment Preferences, Cognitive Style, Motivation and Learning Strategies. Paper presented at the 11th conference of the Israeli Research Association. Jerusalem. The Hebrew University.

Birenanbaum, M. (1996). Alternatives in Assessment of Achievements, Learning Process and Prior Knowledge. USA: Kluwer Academic Publishers.

Birenbaum, M. (1997). Assessment Preferences and Their Relationship to Learning Strategies and Orientations. Higher Education. 33, 71-84.

Birenbaum, M. (2003). New insights into learning and teaching and their implications for assessment, in Segers, M., Dochy, F. and Cascallar, E. (eds.), Optimizing New Methods of Assessment: In Search of Qualities and Standards. Dordrecht, The Netherlands: Kluwer, pp. 13-36.

Birenbaum, M. ve Rosenau, S. (2006). Assessment preferences, learning orientations, and learning strategies of pre-service and in-service teachers. Journal of Education for Teaching, 32(2), 213-225.

Birenbaum, M. (2007). Assessment and instruction preferences and their relationship

Brown, G. T. L. & Hirschfeld, G. H. F. (2007). Students' conceptions of assessment and mathematics achievement: Evidence for the power of self-regulation. Australian Journal of Educational and Developmental Psychology, 7, 63-74.

Buyukozturk, S. ve Gulbahar, Y. (2010). "Assessment Preferences of Higher Education Students", Eurasian Journal of Educational Research, 41, 55-72.

Byrne, B. M. (1998). Structural equation modeling with lisrel, prelis and simplis: Basic concepts, aplications and programming. Mahwah, NJ: Lawrence Erlbaum Associates.

Cavanagh, M. (2006). Mathematics teachers and working mathematicaly: Responses to curriculum change. 2016. Retrieved from http://www.merga.net.au/publications/counter.php?pub=pub_conf&id=289

Cohen, L., Manion, L. & Morrison, K. (2007) Research Methods in Education. 6th edn. London: Routledge.

Cooney, T. J.,Sanchez, W. B. & Ice, N. F. (2001). Interpreting teachers' movement toward reform in mathematics. The Mathematics Educator, 11(1), 10-14.

Davis, L. L. (1992). Instrument review: Getting the most from your panel of experts. Applied Nursing Research, 5, 194–197.

Dickey, D (1996), Testing The Fit of Our Models of Psychological Dynamics Using Confirmatory Methods: An Introductory Primer. (Advances in Social Science Methodology, 4 icinde. Editor: Bruce Thompson). London: JAI press Ltd.

Du Toit, M. ve Du Toit, S. (2001). Interactive Lisrel: User's guide. Lincolnwood: Scientific Software International Inc.

Dogan, N. ve Basokcu T. O. (2010). "İstatistik Tutum Olcegi icin Uygulanan Faktor Analizi ve Asamali Kumeleme Analizi Sonuclarinin Karsilastirilmasi", Egitimde ve Psikolojide Olcme ve Degerlendirme Dergisi, C. 1, S. 2, s. 65-71.

Ebel, R. L. (1979). Essentials of educational measurement. Englewood Cliffs, N.J.: Prentice-Hall.

Eser, M. T. (2011*).Ogrencilerin sinav turu tercih nedenlerini etkileyen bazi faktorlerin incelenmesi.* Yayimlanmamis yuksek lisans tezi, Hacettepe Universitesi Sosyal Bilimler Enstitusu.

Everitt, B. S. (2002). The Cambridge dictionary of statistics (2nd ed). Cambridge; New York: Cambridge University Press.

Field, A. (2009). Discovering statistics using SPSS (Third Ed.). London: SAGE.

Gelbal, S., & Kelecioglu, H. (2007). Teachers' proficiency perceptions of about the measurement and evaluation techniques and the problems they confront. *Hacettepe University Journal of Education*, 33, 135-145.

Grandt, J. (1987). Characteristics of Examinees Who Leave Questions Unanswered on The GRE General Test Rights-Only Scoring. ETS Research Report 87- 83, Princton, NJ: Educational Testing Service.

Gulbahar, Y. ve Buyukozturk, S. (2008). "Degerlendirme Tercihlerin Olceginin Turkceye Uyarlanmasi", Hacettepe Universitesi Egitim Fakultesi, 35, 148-161.

Hambleton, R.K. ve 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.

Ho, R. (2006). Handbook of univariate and multivariate data analysis and interpretation with SPSS. Florida: Chapman ve Hall/CRC.

Hu, Z., Li, J. (2015). The integration of efa and cfa: One method evaluating the construct validity. Global Journal of Human-Social Science, 15 (6).

Joreskog, K. G., & Sorbom, D. (1993). LISREL 8: structural equation modeling with the simplis command language. Lincolnwood: Scientific Software International, Inc.

Kline, P. (1994). An easy guide to factor analysis. New York, NY: Routledge.

Kyriakides, L. (1997). Primary teacher's perceptions of policy for curriculum reform in mathematics. Educational Research and Evaluation, 3(3), 214-242.

Miller, T. (2004). Assessment in practicegrade 9 academic and applied mathematics. (Master Thesis). Queen's University, Kingston, Ontario, Canada.

Motsoeneng, K. G. (2005). The attitude of teacher and parents and learners involved in primary and intermediate schools in the Thabo Mofutsanyane district regarding assessment reform in education. Master Thesis, Bloemfontein University, Mofutsanyane Thabo.

Oren, S. F., Ormanci, U., Evrekli, E. (2014). Ogretmen Adaylarinin Tercih Ettikleri Alternatif Olcme-Degerlendirme Yaklasimlari İle Bu Yaklasimlara İliskin Oz-yeterlilikleri. Egitim ve Bilim Dergisi, 39, 173.

Saxe, G. B., Franke, M. L., Gearhart, M., Howard, S. ve Crockett, M. (1997). Teachers' shifting assessment practices in the context of educational reform in mathematics. CSE Technical Report 471, CRESST University of California, Los Angeles.

Sherin, M. G., & Drake, C. (2009). Curriculum strategy framework: investigating patterns in teachers' use of a reform-based elementary mathematics curriculum. Journal of Curriculum Studies, 41(4), 467-500.

Stapleton, C. D. (1997). Basic concepts and procedures of confirmatory factor analysis. Educational Research Association, Reports-Evaluative (142), Speeches / Meeting Papers (150)

Struyven, K., Dochy, F. & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: a review. Assessment & Evaluation in Higher Education, 30 (4), 325- 341.

Tabachnick, B.G., ve Fidell, L.S. (2001). Using Multivariate Statistics. (4th ed), Allyn& Bacon, Boston,

Turgut, M.F. (1997). *Egitimde olcme ve degerlendirme metodlari*. Ankara: Gul Yayinevi.

Turgut, M. F. (1988). *Egitimde Olcme ve Degerlendirme Metotlari*. Ankara: Saydam Matbaacilik, Altinci Baski.

Uchiyama, M. K. (2004). Teachers use of formative assessment in middle school reform based mathematics classrooms. PhD Dissertation, University of Colorado Boulder, Colorado.

Uchiyama, M. K. (2005). Teachers' use of formative assessment. Annual meeting of the american educational research association, Colorado State University, www.aera.net adresinden 13 Nisan 2005 tarihinde alinmistir.

Watering, G. V., Gijbels, D., Dochy, F. ve Rijt, J. V. (2008). Students' assessment preferences, perceptions of assessment and their relationships to study results. High Education, 56, 645-658.

Wilson, K., & Fowler, J. (2005). Assessing the impact of learning environments on students' approaches to learning: Comparing conventional and action learning designs. Assessment and Evaluation in Higher Education, 30(1), 87–101.

Worthington, R.,& Whittaker, T. (2006). Scale development research: A content analysis and recommendations for best practices. Counseling Psychologist, 34, 806-838. doi:10.1177/0011000006288127

Zoller, U. (1994). The examination where the student asks the questions. *School Science and Mathematics* 94(7):347–349

Zoller, U. & Ben-Chaim, D. (1988). Interaction between examination type, anxiety state, and academic achievement in college science: an action-oriented research. *Journal of Research in Science Teaching 26(2): 65-77.*

Zoller, U. & Ben-Chaim, D. (1990). Gender differences in examination-type preferences, test anxiety, and academic achievements in college science education-A case study. *Science Education* 74(6): 597–608.

Geleneksel Kağıt-Kalem Testleri İçin Tercih Nedenleri Envanteri: Geçerlik ve Güvenirlik Çalışması

Atıf:

Eser, M. T. & Dogan, N. (2017). Inventory of motive of preference for conventional paper-and-pencil tests: A study of validity and reliability. *Eurasian Journal of Educational Research, 69, 135-158. http://dx.doi.org/10.14689/ejer.2017.69.8*

Özet

Problem Durumu: Bireylerin birbirlerinden farklı olmadığı fikri daha çok 20.yüzyıl inanışıdır. Bu fikir büyük olasılıkla Batı dünyasında gelişen "demokrasi" fikrine bağlıdır. Bu inanışa göre, en basit tanımlama ile insanlar birbirlerine eşit ise birbirlerinin aynısı olmalıdırlar. Ancak, yapılan araştırmalar sonucunda, her bireyin farklı karakter özellikleri, farklı zeka seviyeleri ve fiziksel yapıları ile oldukça özel bir donanıma sahip olduğu ortaya çıkmıştır. Bu yaklaşıma göre öğretmenlerin kendi sınıflarında daha başarılı sonuçlar almaları için öğrencilerinin karakterlerini, karakterlerini etkileyen etkenleri, öğrencilerin öğrenme modellerini ve öğrenme modellerini etkileyen etkenleri çok iyi bilmeleri ve göz önünde bulundurmaları gerekir.

Öğretim ve değerlendirme süreçlerinin daha da yakınlaştığı ve etkileşim içerisinde bulunduğu modern eğitim sistemlerinde, öğrencilerin değerlendirme süreci üzerindeki algıları ve değerlendirme yöntemleri seçimlerinin eğitim süreci ve öğrenimi boyunca dikkate alınması gerekir. Öğrencilerin başarıları belirlenirken uygulanan geleneksel kağıt kalem testleri; yazılı sınavlar, kısa cevaplı testler, doğru yanlış testleri, çoktan seçmeli testler, performans görevleri, portfolyo vb.'dir. Öğrencilerin bu geleneksel kağıt kalem testleri konusunda görüşlerini almak, öğretmenlere öğrenci başarısını belirlemede geri besleme ve öğrencilerin öğrenme süreçleri konusunda bilgi edinilmesi gerekmektedir. Bu çalışma öğrencilerin değerlendirme süreçleri üzerindeki algılarının önemini ve değerlendirme yöntemlerinin seçimlerini göz önüne alarak gerçekleştirilmiştir.

Araştırmanın Amacı: Araştırmanın amacı, öğrencilerin yazılı, kısa cevaplı, doğru-yanlış ve çoktan seçmeli testleri tercih etme nedenlerini değerlendirmeye ilişkin *''Geleneksel Kağıt Kalem Testleri İçin Tercih Nedenleri Envanteri''* geliştirerek, literatüre öğrencilerin bu sınav türlerini tercih etme nedenleri ile bu sınavları tercih düzeylerini tespit etmeye yardımcı olacak ölçme sonuçlarının geçerliği ve güvenirliği sağlanmış bir ölçme aracı kazandırılacağı düşünülmektedir. Elde edilen sonuçlara bağlı olarak öğretmenler öğrenci başarısını ölçmek amacıyla sınav hazırlarken öğrencilerin belirli özelliklerine göre ölçme aracı geliştirme çabasını arttırabilirler. Öğretmenlerin test geliştirme sürecinde dikkat edeceği faktörler öğrencilere olumlu bir şekilde yansıyacağı, testlerin öğrenciler üzerinde oluşturduğu olumsuz etkilerin en aza indirileceği düşünülmektedir.

Araştırmanın Yöntemi: 100 lise öğrencisinin oluşturduğu bir örneklemden elde edilen envanter ile ilgili veri setine ilişkin faktör analizi sonuçlarına göre; alt ölçekler için elde edilen faktör yükleri 0,32 ile 0,69 arasında değişmektedir. Alt ölçekler için KMO

değerleri 0,71 ile 0,75 arasında bulunmuştur. KMO değeri sonuçlarına göre veri sayısının faktör analizi için yeterli sayıda olduğuna karar verilmiştir. Tüm alt ölçekler için Bartlett testi sonuçları 0,01 düzeyinde manidar bulunmuştur. Bu sonuç, veri setinin faktör analizine uygun olduğunun bir işaretidir. Dört alt ölçeğe ilişkin faktör yükleri ve yamaç- birikinti grafikleri incelenmiş ve birinci boyutta yer almayan, herhangi bir boyutta yer alması için faktör yükü yetersiz olan veya birden fazla boyutta faktör yükü yüksek olan 11 maddenin envanterden çıkartılması uygun görülmüştür. Uzmanlar 4. maddenin envanter için uygun olmadığını bildirmişler ve 4. madde envanterden çıkartılmıştır. Sonuç olarak her bir ölçeğin tek boyutlu olduğuna karar verilmiş ve uygulamaya 22 madde ile devam edilmiştir. Her bir alt ölçek için iç tutarlığı görmek açısından Cronbach Alfa iç tutarlık katsayıları incelenmiş ve iç tutarlık katsayılarının 0,831 ile 0,856 arasında değiştiği gözlemlenmiştir. Bu değerler ölçeklerin kabul edilebilir güvenirliklere sahip olduğunu göstermektedir.

783 kişiye yapılan ikinci uygulama sonucuna doğrulayıcı faktör analizi uygulanmış; 2. ve 4. maddelerin diğer maddelerle otokorelasyona girdiği gözlemlenmiş ve bu maddelerin atılması uygun görülmüştür.

Doğrulayıcı faktör analizine ilişkin sonuçlar için X^2/sd' nin 5'ten küçük olması modelin uyum iyiliğine sahip olduğunun göstergesidir (Byrne, 1998). RMR değerlerinin 0,05' ten küçük olması mükemmel uyuma, SRMR değerlerinin 0,05 ile 0,08 arasında olması ise iyi uyuma işarettir. GFI/AGFI, NFI, NNFI, CFI değerleri ölçme aracının yüksek uyum verdiğini gösteren değerler almıştır. RMSEA değerlerinin 0,10' dan küçük olması kabul edilebilir bir uyumun göstergesidir. Bütün değerler bir arada ele alınıp yorumlanacak olursa; alt testlerin tek boyutluluk yapısına ilişkin doğrulamanın yeterince güvenilir biçimde sağlandığı söylenebilir (X^2/sd: 4,01-6,54; GFI: 0,96-0,97; AGFI: 0,94-0,96; NFI: 0,99; NNFI: 0,99; RMSEA: 0,062-0,084; RMR: 0,027-0,032; SRMR: 0,052-0,065). Araştırma kapsamında son olarak, kapsam geçerliği çalışması yürütülmüştür. Kapsam geçerliği anlamında ölçme aracını meydana getiren her bir alt testi oluşturan maddeler için, konu alanında yeterli donanım ve bilgiye sahip, çalışmanın öneminin farkında olan 12 eğitimde ölçme ve değerlendirme uzmanının görüşleri alınmıştır. Maddelere ilişkin kapsam geçerlik oranları belirlenirken Davis tekniği kullanılmıştır. Geliştirilen ölçme aracını meydana getiren maddelere ilişkin kapsam geçerlik indekslerinin yazılı sınav, kısa cevaplı test, doğru-yanlış testi ve çoktan seçmeli test için 0,86 ile 1 arasında değiştiği gözlemlenmiştir. Davis tekniği için sınır değerin 0,80 olduğu göz önünde bulundurulduğunda, maddelerin her bir alt testteki kapsam geçerlik değerlerinin yeterli düzeyde olduğu söylenebilir.

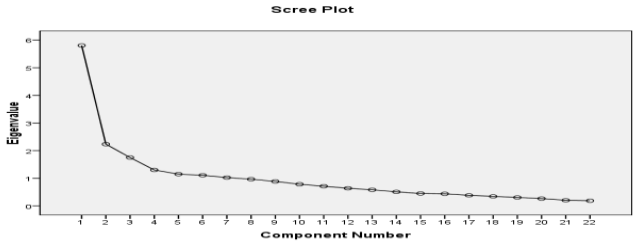
Araştırmanın Bulguları: Bu çalışma sonucunda, öğrencilerin geleneksel kağıt kalem testleri konusunda tercihlerinin belirlenmesine yönelik olan GKKT-TNE geliştirilmiştir. Envanter, 2 bölümden meydana gelmektedir. Envanterin ilk bölümünde demografik bilgilerin yer aldığı 4 madde, ikinci bölümünde ise 3'lü derecelendirilmiş 20 madde yer almaktadır.

Araştırmanın Sonuç ve Önerileri: Araştırma sonuçları, geliştirilen ölçeğin, lise öğrencilerinin kağıt ve kalem testlerine ilişkin tercih sebeplerini değerlendirmek için uygun bir araç olduğu görülmektedir. Bu çalışma kapsamında geliştirilen envanter, öğrencilerin ilgili sınavlara ilişkin sınav türü tercih seviyelerini farklı örnekler

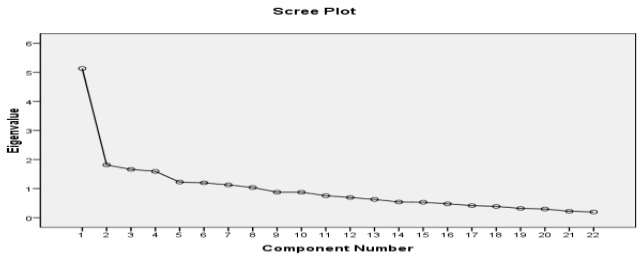
kullanarak tahmin eden faktörleri belirlemek için kullanılabilir. Bu sonuçlar, öğrencilerin sınav türü tercihlerini belirleyerek değerlendirme sürecinde gerçekleştirilecek eylemleri ve araçları belirlerken kullanılabilir.

Anahtar Kelimeler: Sınav türü tercihi, kapsam geçerliği, açımlayıcı faktör analizi, doğrulayıcı faktör analizi.

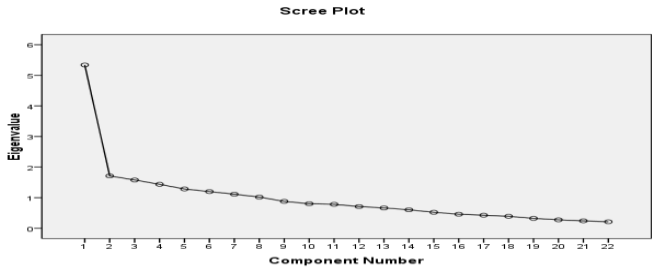
Appendix1**.** Scree Plot of the Written Examination Subtest



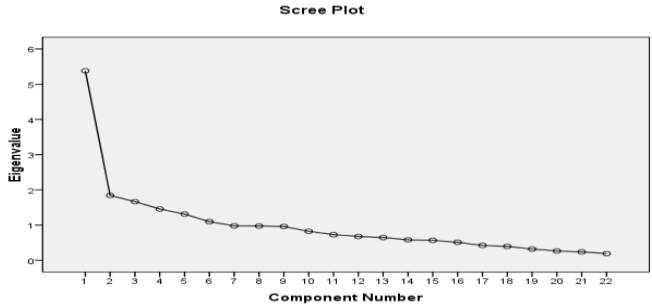
Appendix 2. Scree Plot of the Short Answer Examination Subtest



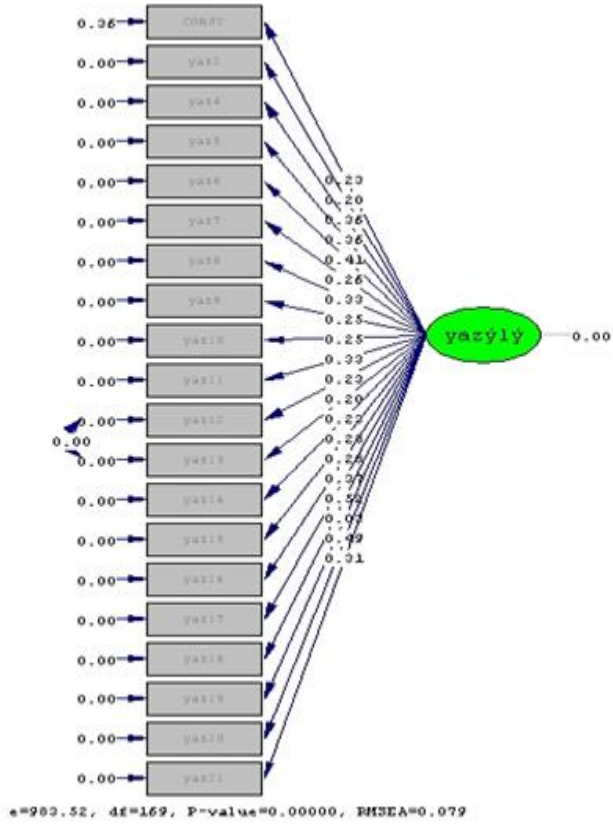
Appendix 3. Scree Plot of the True/false Examination Subtest



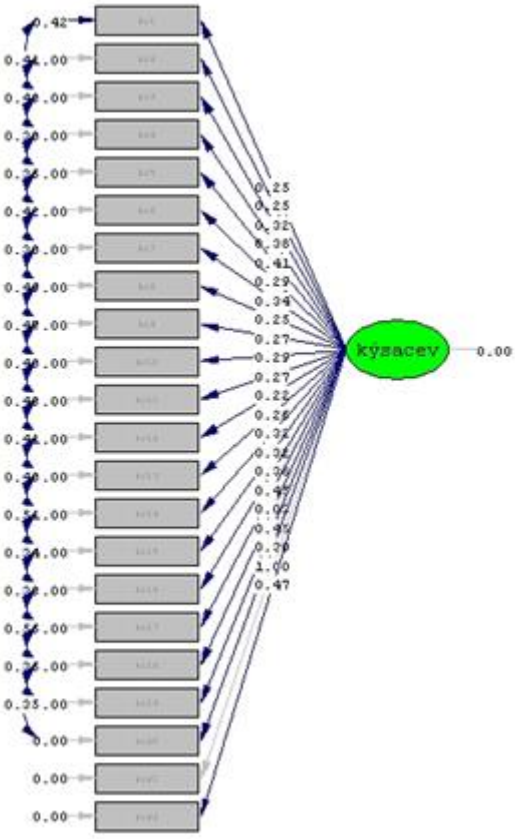
Appendix 4. Scree Plot of the Multiple-choice Examination Subtest



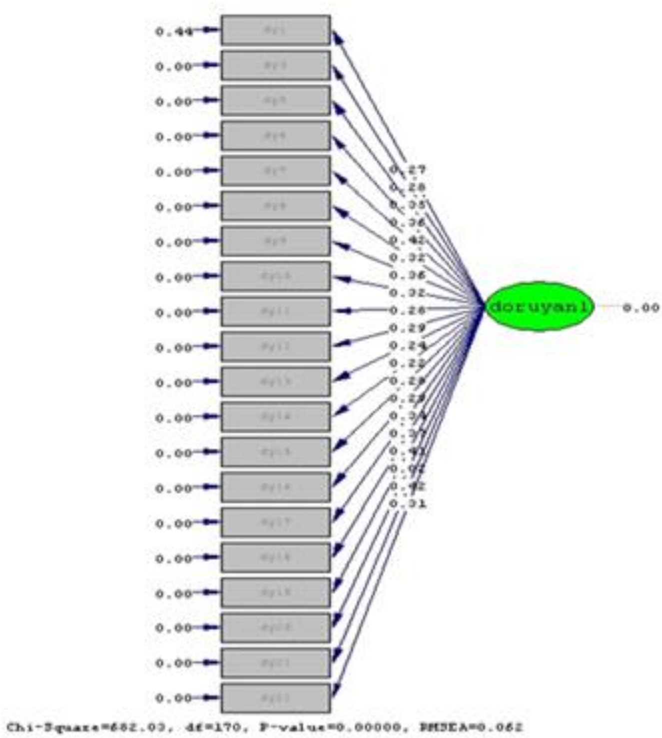
Appendix 5. Path Graph of the Confirmatory Factory Analysis of the Written Examination Subtest



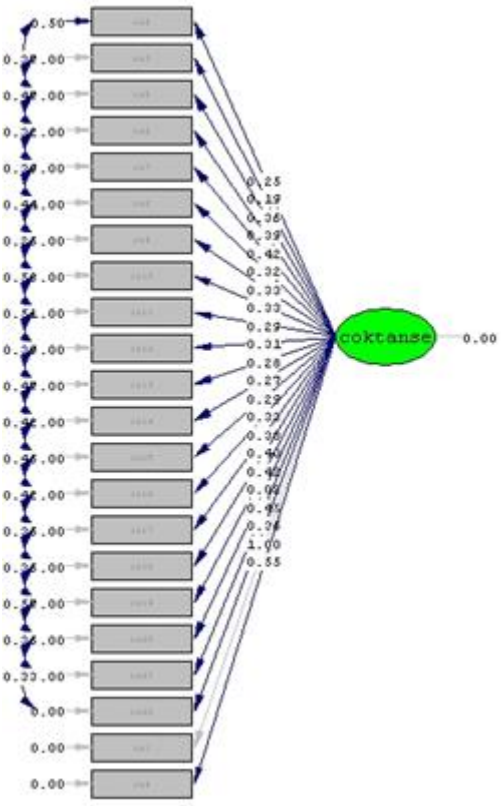
Appendix 6. Path Graph of the Confirmatory Factory Analysis of the Short-Answer Examination Subtest



Appendix 7.Path Graph of the Confirmatory Factory Analysis of the True/false Examination Subtest



Appendix 8. Path Graph of the Confirmatory Factory Analysis of the Multiple-choice Examination Subtest



Appendix 9. The Inventory of Motive of Preference for the Written, Short Answer, True/false and Multiple-choice Tests

Gender:
Grade:
Education Level of Mother:
Education Level of Father:

(1) True for me
(2) Partly true for me
(3) Totally true for me

Dear Students,
Please read the following items and mark the gap under the code with (x) indicating one of the judgments shown on the top right corner. Thank you for participating in our study.

In these examinations	**Examination Types**											
	Written Examinations			Short Answer Tests			True/false Tests			Multiple-choice Tests		
	1	2	3	1	2	3	1	2	3	1	2	3
1) I can cheat easily.												
2) I can easily let others copy from me.												
3) I have a chance in turning the wheel.												
4) I don't need to learn the subjects by heart.												
5) My preparation doesn't take time.												
6) My preparation is easy.												
7) I don't get nervous.												
8) It is easy for me.												
9) I use time efficiently.												

10) I don't have a problem focusing.													
11) I succeed.													
12) I don't feel the need to ask any questions to the teacher.													
13) I feel it is necessary to copy													
14) I don't panic if I finish early.													
15) I think the questions are difficult.													
16) I think the time to answer is not sufficient.													
17) It provides more correct results.													
18) It suits my learning style better.													
19) I easily express what I want to say.													
20) I can predict the score.													
21) I don't feel obliged to study.													
22) Reading is enough to study.													
23) Writing is enough to study.													
24) I finish answering quickly.													
25) It doesn't make me panic.													
26) I cannot be sure about my answer.													
27) I don't feel obliged to express myself/my thoughts.													
28) I get bored.													

29) I feel comfortable.													
30) I have a headache.													
31) I feel bad.													
32) I find it difficult.													
33) I trust in my response.													
34) I want to finish and get out quickly.													

Note: Bold statements are final inventory items.

Eurasian Journal of Educational Research 69 (2017) 159-175

Eurasian Journal of Educational Research
www.ejer.com.tr

A Study of Mathematical Content Provided in Illustrated Children's Books *

Melek Merve YILMAZ GENC[1] Aysegul AKINCI COSGUN [2] Sengul PALA[3]

ARTICLE INFO

Article History:

Received: 11 January 2017

Received in revised form: 21 March 2017

Accepted: 24 April 2017

DOI:http://dx.doi.org/10.14689/ejer.2017.69.9

Keywords

Early childhood, mathematics, NCTM, Picture story books.

ABSTRACT

Purpose: Early childhood is of critical importance in terms of cognitive, affective and physical development. Undoubtedly, a substantially stimulating environment and opportunities offered to children, as well as appropriate educational materials, have an impact on their development. The object of this study is to investigate the mathematical content included in illustrated children's books prepared for pre-school children. **Research Methods:** The research was planned as a descriptive study, and criterion sampling method was used in creating the study group. One hundred seventy-four illustrated children's books offered for sale by bookstores in Adana and Ankara provinces and included by academicians in their private libraries were surveyed. As a result of the survey, 52 illustrated children's books that met the criteria constituted the research study group. A document analysis was conducted to collect research data. Thereunder, a "book review form" developed by the researchers was used to investigate the mathematical content of the books. **Implications for Research and Practice**: As a result of the study, it was discovered that the books investigated most frequently contained the numbers and operations category, followed by geometry, measurement, algebra, data analysis and probability, respectively. Furthermore, the books were found not to contain specific categories or use of different methods and tools in calculations, distinction, algebraic symbols and mathematical models, quantitative changes, symmetry and translation movements, repeated measures, data collection, data organization, data visualization, data recognition, and understanding and applying basic probability concepts. Such results suggest that the illustrated children's books published in our county fail to satisfy the mathematical standards recommended by nationally and internationally recognized programs as the standards to be taught during the pre-school period, and that efforts must be made to enhance the quality of books published for such purposes. For future studies, it is recommend that researchers examine in detail how mathematical categories provided in the books are dealt with, or if and how teachers include illustrated children's books in mathematics activities performed in class.

*This study was partly presented at the 3rd International Eurasian Educational Research Congress in Muğla, 31 May – 03 June, 2016

[1] Çukurova University, TURKEY, mmyilmaz@cu.edu.tr

[2] Corresponding Author: Aksaray University, TURKEY, aysegulakinci2011@gmail.com

[3] Kırıkkale University, TURKEY, sengul.akinci@gmail.com

Introduction

The pre-school period is the phase during which the basis for mathematical thinking and mathematical concepts and skills required for real life and future academic progress are developed (Jennings, Jennings, Richey, & Dixon-Krauss, 1992). During this period, it is quite important for children's meaningful mathematic experiences to present mathematics in a certain context (NCTM, 2000). Illustrated children's books that are frequently used in pre-school education can provide the said contextual basis by relating mathematical concepts and skills to real-life situations (Shih & Giorgis, 2004).

In this age of significantly and fast change, the basis for skills that make real life easier, such as an analytical approach, versatile approaches to work and problem-solving, increasingly relies on mathematical thinking (NCTM, 2000). Therefore, it is critical to provide mathematics education in a meaningful way starting from the pre-school period (Ginsburg & Golbeck, 2004). There are two basic elements of providing education in a meaningful way: The first is to integrate the concepts into real-life problems. This is argued as a technique that stimulates motivation among both teachers and children. The second element, on the other hand, is the discussion among children about what they have learned. Thus, what has been learned is transferred to new situations, which is the basic indicator of meaningful learning (Jennings et al., 1992; NCTM; 2000). Using illustrated children's books in pre-school mathematics education is regarded as one of the strategies employed to bring these two elements together.

Given the goals of pre-school education in Turkey, it is important to teach mathematical concepts and skills that will prepare children during the pre-school period for primary school. These mathematical concepts and skills are grouped by the National Council for Teachers of Mathematics (NCTM) under the topics of numbers and operations, algebra, geometry, measurement, data analysis and probability, problem solving, reasoning and proof, communication, connection and representation (NCTM,2000). Illustrated children's books used in the process of teaching the targeted concepts and skills can introduce mathematical concepts to children and raise awareness of mathematics in them by means of story patterns based on life (Anderson, Anderson & Shapiro, 2005).

Presenting mathematical concepts under real life situations in illustrated children's books increases children's interest in mathematics and enables them to develop a deeper understanding (Barnaby, 2015). Furthermore, stimulating their interest in mathematics through stories and images makes it easier for them to create mathematical connections (Shatzer, 2008). For a child that is able to create a connection between mathematics and the story, mathematics becomes more effective and more applicable to daily life situations (Barnaby, 2015; Hong, 1996).

Considering children's future lives, the experiences they gain during the pre-school period must be so arranged as to shape their attitude toward mathematics positively (Keat&Wilburne, 2009). However, the research suggests that the traditional methods used for small children's mathematics education are not effective in developing a positive attitude toward mathematics (Baroody, Lai, & Mix, 2006; Hong, 1996).

Whereas, it would provide a basis for a positive attitude toward mathematics if teachers sought to go beyond such methods as study pages, exercise books and representation and present creative, expedient and effective methods of mathematics education (Jennings et al., 1992). Illustrated children's books are suggested as an effective means of developing a positive attitude through the integration of mathematics and children's literature (Shih & Giorgis, 2004).

A literature review suggests that the current research is mainly focused on recommending the names of quality books that may be used in mathematics education (Forrest, Schnabel & Williams, 2006; Griffiths & Clyne, 1991; Mc Duffie & Young, 2003; Radebaugh, 1981; Shatzer, 2008; Shih & Giorgis, 2004; Thatcher, 2001). On the other hand, the use of illustrated children's books in pre-school mathematics education is suggested to have a positive impact on children's understanding of mathematics (Casey, Erkut, Ceder & Mercer-Young, 2008; Hong, 1996; Van den Heuvel-Panhuizen & Elia, 2011; Van den Heuvel-Panhuizen & van den Boogaard & Doig, 2009). In addition to school activities, book reading with the participation of family has been demonstrated to be effective in children's development of mathematical concepts (Anderson, 1997; Anderson & Shapiro, 2004).

To determine of illustrated children's books in Turkey, some studies were conducted focusing on content and physical characteristics of the books (Gonen, 1993; Demircan, 2006; Isitan, 2014), gender roles (Gursimsek & Gunay, 2005), elements of violence (Daglioglu & Cakmak, 2009),values (Erdal, 2009), elements of humor (Akinci, 2015) and geographical elements (Kilic, Gulec & Genc, 2014). However, a limited number of studies have focused on the mathematical content of illustrated children's books (Ocal, Ocal, & Simsek, 2015).

The use of illustrated children's books is argued to be important in pre-school mathematics education. It appears, however, that there is little current knowledge about whether or not concepts and skills are provided in illustrated books focusing on mathematics. In this context, it is a necessity to investigate the presence of mathematical content in illustrated children's books prepared for use in pre-school mathematical education, as well as to identify any insufficiencies.

Objective

Shih and Giorgis (2004) suggest that three different types of books may be preferred to teach mathematical concepts in means of children's literature: The first is through books focusing on mathematics and written for the purpose of teaching mathematical concepts and skills. These books contain mathematics openly. Thus, children are able to create the intended content directly in their minds. The second type includes books that do not use mathematics as the basis for the story but have the story and mathematics as an inseparable whole. In these books, mathematics is embedded in the story. The third type presents activities conducted to create opportunities for mathematical thinking by using any type of illustrated children's books. In these activities, a child discovers the mathematical relationships in the book himself/herself. Accordingly, a study conducted by Ocal et al. (2005) investigated the books in the third group with regard to mathematical concepts and skills and evaluated them in view of the MoNE Pre-School Educational Program.

In contrast to the study conducted by Ocal et al. (2015), the present study focuses on investigating the mathematical content of the first and second types of books according to the classification of Shih and Giorgis (2004).Illustrated books prepared for use in mathematics education facilitate the process for children and teachers by offering the intended mathematical content in a direct and complementary manner. Given that illustrated books aimed at mathematics education would be preferred more by children and teachers from such a perspective, this study aimed to investigate the current state of such books. Identifying the mathematical concepts and skills contained and not contained in the illustrated books is believed to be important to eliminating the relevant insufficiencies.

Accordingly, the research question was determined as follows: "How do illustrated children's books for pre-school mathematics education meet the mathematical content standards set forth by NCTM?"

Method

Research Design

The present study investigating illustrated children's books for the pre-school period is a descriptive research study conducted using the survey model. Survey models are approaches aiming to reveal the current situation as is, and include researching and identifying what is intended without trying to change subjects, cases, objects, individuals or situations (Karasar, 2014). Hence, the current situation as to the mathematical content in illustrated children's books was investigated without trying to change it, and thus, the survey model was used.

Research Sample

Criterion sampling method was used to create the research study group. According to the classification by Shih and Giorgis (2004), the first and second types of books featuring mathematics as the focus of the story and integrating mathematics and the story were determined as the basic criterion. As such, 174 illustrated children's books offered for sale by bookstores in Adana and Ankara provinces and included by academicians in their private libraries were surveyed. As a result of the survey, 52 illustrated children's books that were determined to meet the criteria constituted the research study group.

Research Instruments and Procedures

A document analysis was conducted to collect research data during the study. Books included in the study group were surveyed by the researchers between March 2016 and May 2016. A "Content Control List" was created to study the illustrated children's books forming the study group in terms of mathematical content. A control list is defined as a measuring tool that enables the existence of the behavior aimed to be studied to be expressed as "x" or "yes" (Turgut & Baykul, 2013). At the stage of preparing this list, NCTM (2000) content standards were used that cover in detail the topics that are internationally recognized and constitute the basis for mathematics education. The areas of "numbers and operations, algebra, geometry, measurement, data analysis and probability"that are among these standards constitute some of the sub-dimensions of the control list created. In the study, mathematical expressions

contained in the books were taken as the data analysis unit and analyzed in parallel to the content control list.

Table 1

Content Control List

General categories	Sub categories
Numbers and Operations	Counting, and telling the number of objects counted Expressing numerical values with materials and models Recognizing numbers Ordinal numbers Object groups of equal number-division Addition Subtraction Basic fraction units Using different methods and tools in calculations
Algebra	Pattern Classification Distinction Sorting Algebra symbols and mathematical models Qualitative changes Quantitative changes
Geometry	Geometric shapes and their characteristics Spatial Relationships Symmetry and translation movements Spatial, visual reasoning and geometric modeling
Measurement	Physical measurement; length, volume, weight and area Non-Physical Measurements; time Standard units of measurement Non-standard units of measurement Measuring Tools Repeated Measures Comparison Estimation
Data Analysis and Probability	Data Collection Data Organization Data Visualization Data description Deduction Based on Data Understanding and Applying Basic Probability Concepts

An alternative way of achieving triangulation using multiple-data sources to improve the validity and reliability of qualitative research is suggested to involve the inclusion of multiple researchers in data collection and analysis processes (Merriam, 2015, p. 206). In this respect, among the books included in the study group, seven books were selected to be surveyed jointly by three researchers. Mathematical expressions that were thought to be included in these seven books were noted, and the control list was marked by a joint decision on which mathematical contents these expressions subsumed. Mathematical expressions in the remaining 45 books, on the other hand, were evaluated independently by the three researchers. Once the surveying was complete, the reliability formula suggested by Miles and Huberman (1994) (reliability= agreement/(agreement+disagreement) was used to identify the agreement coefficient among the researchers. Separate percentages were calculated for each sub-dimension comprising the Content Control List, and the agreement percentages were calculated to be 0.82 for numbers and operations; 0.86 for algebra; 0.86 for geometry; 0.78 for measurement and 0.84 for data analysis and probability. For codings on which the researchers disagreed, they worked together to reach a joint decision on which content standards the expressions met.This was intended to ensure reliability by precluding evaluation errors caused by disagreements.As for the research findings, analysis results were presented in tables and direct quotes were used in order to be faithful to the original data for the purpose of delivering accuracy in codings.

Results

This section presents in table form the frequency and percentage distribution of the results acquired with respect to mathematical content of the illustrated children's books included in the study group. First, general mathematical skills provided in the books were investigated in the study. Table 2 presents the results pertaining to general mathematical skills provided in the books.

Table 2

Frequency and Percentage of General Mathematical Content Skills in the Illustrated Children's Books

General Category	***f***	**%**
Numbers and Operations	*532*	*44.7*
Algebra	*37*	*3.1*
Geometry	*358*	*30.0*
Measurement	*256*	*21.5*
Data Analysis and Probability	*7*	*0.5*
Total	1190	100.0

When Table 2 is reviewed, the general categories of Numbers and Operations, Algebra, Geometry, Measurement, Data Analysis and Probability are provided. According to this, the illustrated children's books investigated appear to include the special sub category of Numbers and Operations (44.7%) most. This is followed by the

special sub categories of Geometry (30.0%), Measurement (21.5%), Algebra (3.1%) and Data Analysis/Probability (0.5%), respectively.

Table 3
Frequency and Percentage of Numbers and Operations Category in the Illustrated Children's Books

General Category	Special Sub Categories	*f*	%
Numbers and Operations	Object groups of equal number-division	*1*	0.2
	Expressing numerical values with materials and models	*0*	*0.0*
	Recognizing numbers	*35*	*8.3*
	Counting, and telling the number of objects counted	*252*	*60.1*
	Ordinal numbers	*56*	*13.3*
	Addition	*21*	5.0
	Subtraction	*42*	10.0
	Understanding and showing the basic fraction units	*12*	2.8
	Using different methods and tools in calculations	*0*	*0.0*
	Total	419	100.0

In Table 3, special sub categories under the general category of Numbers and Operations are provided. Accordingly, the illustrated children's books in the study group appear to include the special sub categories of Counting, Telling the Number of Objects Counted *(60.1%)* most. These are followed by Ordinal Numbers (*13.3%)*, Subtraction (10.0%), Recognizing Numbers (*8.3%)*, Object Groups of Equal Number-Addition (5.0%), Understanding and Showing the Basic Fraction Units (2.8%), Division (0.2%), Showing Numerical Values with Materials and Models (*0.0%),*and their special sub categories, respectively. Also, a sub category of Numbers and Operations skills,"Using Different Methods and Tools in Calculations" appears not to be provided in any of the illustrated children's books investigated. Failure to include this skill, a special sub category of the Numbers and Operations skills that constitute the basis for the pre-school mathematics curriculum, is considered to be an important insufficiency in these children's books. The following are examples of the Numbers and Operations Category:

"Three walnuts came from the monkey basket"
"Tonton also counted, "one bird, two birds, three birds"
"... This time you have come first"
"... He munched up half a cucumber"

"... The dog with eyes as large as a saucer is given a narrow cushion, the one with eyes as large as a plate a wide cushion, and the one with eyes as large as a pan lid the widest cushion."

Table 4

Frequency and Percentage of Algebra Category in the Illustrated Children's Books

General Category	Special Sub Categories	*f*	%
Algebra	Pattern	*2*	3.5
	Classification	*5*	*8.7*
	Distinction	*0*	*0.0*
	Sorting	*18*	*31.5*
	Algebraic symbols and mathematical models	*0*	*0.0*
	Qualitative changes	*32*	*56.1*
	Quantitative changes	0	0.0
	Total	57.0	100.0

In Table 4, special sub categories under the general category of "Algebra" are provided. Accordingly, the special sub category of "Qualitative Changes" (56.1%) appears to be provided most frequently, followed by "Sorting" (31.5%), "Classification" (8.7%) and "Pattern" (3.5%) categories, respectively. An important finding is that the categories of Algebraic Symbols and Mathematical Models, Distinction and Quantitative Changes are not included in the books at all. The followings are some examples of the Algebra category:

"... We wait in winter for summer to come to swim in the sea. Thus, there comes summer, after summer comes fall, after fall comes winter comes, after winter comes spring, and after spring comes summer again..."

"Its curved tail is getting smaller and smaller."

"... Then who ever goes down first will eat the candies."

"... Should I only put the round ones?" They are perfect for soups..."

Table 5

Frequency and Percentage of Geometry Category in the Illustrated Children's Books

General Category	Special Sub Categories	*f*	%
Geometry	Geometric Shapes and their characteristics	*135*	93.1
	Spatial relationships, Spatial/visual reasoning and geometric modeling	*8*	*5.5*
	Symmetry and translation movements	*2*	*1.3*
	Total	145	100.0

In Table 5, special sub categories under the general category of "Geometry" appear to be provided. Accordingly, the sub-category "Geometric Shapes and Their Characteristics" (393.1%) appears to be provided most frequently, followed by Spatial,

Visual Reasoning and Geometric modeling (5.5%) and Symmetry and Translation movements (1.3%) sub categories, respectively.

The followings are some examples of the Geometry Category:
"The square spoke in a fury"

"Three men put round onions and potatoes in the soup"

Table 6
Frequency and Percentage of Measurement Category in the Illustrated Children's Books

General Category	Special Sub Categories	*f*	%
Measurement	Using non-standard units of measurement	*20*	5.6
	Using standard units of measurement	*7*	*1.9*
	Measuring Tools	*0*	*0.0*
	Comparison	*91*	*25.8*
	Estimation	*4*	*1.1*
	Physical Measurements; length, area volume and weight	*162*	*46.0*
	Non-Physical Measurements; time	*68*	19.3
	Total	*352*	100.0

In Table 6, special sub categories under the general category of Algebra are provided. Physical Measurements appear to be provided most frequently, followed by Non-Physical Measurements; Comparison, Using Standard Units of Measurement, Estimation, Using Non-Standard Units of Measurement and Measuring Tools, respectively. The followings are some examples of the Measurement Category:
"She rolled a large and little heavy bolder toward Bonbon."

"...First, a lightning flashed, then the sky thundered in a fury, and the earth shook once again"

"He began to think about what it was moving inside the apple. Or...was it the fat walnut...or...the shining moon that made the apple laugh?"

"Put on the red pants we left to learn your leg length and tell us if its length is short, long, or the same"
"The little brother has the smallest feet. The middle brother's feet are bigger than the little brother's. The oldest brother's feet are bigger than the middle brother's. Three brothers ordered their shoes and left the shop."

Table 7
Frequency and Percentage of Data Analysis and Probability Category in the Illustrated Children's Books

General Category	Special Sub Categories	*f*	*%*
Data Analysis and Probability	Data Collection	*0*	*0.0*
	Data Organization	*0*	*0.0*
	Data Visualization	*0*	*0.0*
	Data description	*0*	*0.0*
	Deduction Based on Data	*6*	100.0
	Understanding and Applying Basic Probability Concepts	0	*0.0*
	Total	6	100.0

Table 7 presents the results regarding the data analysis and probability category provided in the illustrated children's books included in the study group. Per this analysis, only the special sub category of deduction based on data under the general category of data analysis and probability appears to be provided, and the special sub categories of data collection, data organization, data visualization, data description and understanding and applying basic probability concepts are not seen at all.

Discussion and Conclusion

Discussion

The findings acquired by the present research conducted to investigate the mathematical content in the illustrated children's books prepared for use in pre-school mathematics education suggest that the numbers and operations category is provided most frequently in the investigated books, followed by geometry, measurement, algebra, data analysis and probability categories, respectively. The books investigated hereunder appear to frequently provide the concepts and skills relating to the numbers and operations category that constitutes the basis for the development of mathematical concepts. The study by Öçal et al. (2015) that investigated the mathematical concepts and skills provided in the randomly selected illustrated children's books reached similar conclusions. Given that children begin using numerical terms frequently starting around age two (Avcı & Dere, 2002), it is believed to be important to provide books that use such expressions frequently and that are therefore appropriate for the development of mathematical concepts. In parallel, the skill of manipulating numbers is known to have a high percentage among the achievements and indicators provided in the MoNE 2013 Pre-School Education Program (Özen Uyar & Yılmaz Genç, 2016). In this context, the use of numbers included among the priority skills to be taught to children may be supported greatly by illustrated children's books offered during the pre-school education period. While the sub categories under the numbers and operations dimension of counting, telling the number of objects, ordinal numbers, subtraction, recognizing numbers and addition, are observed to be provided frequently in the books investigated, the missing categories of showing numerical values with materials and models and using different methods and tools in calculations give rise to the thought that the development of children in such areas is

not adequately supported by books. The findings of the study conducted by Powell and Nurnberger-Haag (2015) aiming to evaluate the numbers and the characteristics of counting show parallelism with the results of the present study. Accordingly, the study conducted by surveying books revealed that teaching of the number 0 and the numbers greater than 10 is very limited.

When the research findings regarding the measurement category are reviewed, concepts and expressions relating to physical measurement such as length, area, volume and weight and non-physical measurement such as time (Yıldırım Hacıibrahimoğlu, 2016) appear to be frequently provided. Even if children in the younger age group do not use standard units, they are able to measure using non-standard units, thus developing an understanding of the nature of the units and gaining awareness of concepts of measurability (NCTM, 2000). In this respect, it is believed that illustrated children's books that give children the opportunity to learn how to measure using standard and non-standard units with different tools throughout the course of the story should be more frequently used. However, it is believed to be as important as the provision of mathematical concepts and skills in the books that these books are integrated effectively by teachers into classroom activities. A study conducted by Jennings et al. (1992) concluded that books included in mathematics activities improved children's interest and success in mathematics, as well as increasing the number of mathematical terms they used in their free games.

When the research findings are examined, the books included in the study group appear to provide the content relating to the sub category of Geometric Shapes and their characteristics under the geometry category at a high rate. A study conducted by Casey et al. (2004), focusing on the development of children's spatial skills, concluded that the children in the group in which the blocks and illustrated story books were used together proved to be more successful in mathematics activities than the children in the group in which the story books were not used. The study conducted by Casey et al. (2008) revealed that girls' geometry skills developed more than the boys' as a result of the activities during which relevant story books were read in addition to the geometry studies. Based on the results of the present study, it may be suggested that such skills can be developed by including illustrated books featuring mathematical content in activities in areas in which children are found to lack mathematical skills.

The books investigated herein appear to provide only the examples of expressions relating to deduction, based on data from the data analysis and probability category, that is comprised of five sub categories. Considering that skills relating to this category include concepts that will form the basis for future academic life, the investigated books' providing somesuch content gives rise to the thought that this category is neglected. A study conducted by İnan (2014) found that a high number of pre-school teacher candidates believed that data collection and probability topics were not provided in pre-school mathematics education. Considering, based on such findings, that the frequency of providing data analysis and probability content in mathematics activities may be low, it is believed to be important to support such skills through the use of illustrated children's books.

Conclusion

When the research results are examined in general, the illustrated children's books included in the study group are seen not to contain the special categories related to using different methods and tools in calculations, division, algebraic symbols and mathematical models, quantitative changes, symmetry and translation movements, repeated measures, data collection, data organization, data visualization, data recognition, and understanding and applying basic probability concepts. Such results acquired hereunder suggest that the books prepared in our country fail to meet the NCTM mathematics standards.

Considering the pre-school age and developmental stage, children do not know how to read and write during that period, and therefore, illustrated children's books are frequently used in pre-school education. The studies conducted suggest that illustrated books that provide mathematical concepts and contents and offer entertaining and meaningful contents to children contribute greatly to the pre-school mathematics curriculum (Casey, Kersh, & Young, 2004; Hong, 1999; van den Heuvel-Panhuizen & Elia 2011). Rather than illustrated books containing mathematical expressions, books specifically prepared for use in pre-school mathematics activities offer mathematical concepts visually, and in this way, support children's understanding of abstract expressions (Montague-Smith & Price, 2012). It is considered important that illustrated children's books properly scripted for use in pre-school mathematics education, presenting mathematical concepts properly and containing mathematical contents, are accessible by teachers and families.

Recommendations

Only 52 of the 174 books investigated to create the study group were found to have mathematics as the focus of the story and to integrate mathematics into the story. This rate is believed to be inadequate. Also, when it is analyzed on the basis of topics, there are some categories that are not provided at all, or provided very little, in the books investigated. Particularly, it appears to be a necessity that books focusing on such areas be published and that publishing houses and authors be advised to focus on such topics.

It is recommended that teachers frequently use illustrated children's books in preschool mathematics education, making it easier to understand mathematical concepts by means of rich visual presentations and meaningful contexts. Such story books used effectively with different reading techniques and proper interventions are believed to stimulate children's motivation in mathematics activities. A list of illustrated children's books may be published to encourage the pre-school teachers to use illustrated children's books in mathematics activities. This way, teachers can have the opportunity to access such books that have been reviewed by experts and have a proper content.

This study was conducted to investigate whether and how illustrated children's books for pre-school mathematics education meet the mathematical content standards set forth by NCTM. For future studies, it is recommended that researchers examine in detail how mathematical categories provided in the books are dealt with, or whether and how teachers include illustrated children's books in mathematics activities performed in the class.

References

Akinci, S. (2015). *Resimli cocuk kitaplarında yer alan mizahi unsurların incelenmesi.* Yayinlanmamis Yuksek Lisans Tezi, Hacettepe Universitesi Egitim Bilimleri Enstitusu, Ankara.

Anderson, A. (1997). Families and mathematics: A study of parent-child interactions. *Journal for Research in Mathematics Education*, 484-511.

Anderson, A., Anderson, J., & Shapiro, J. (2004). Mathematical discourse in shared storybook reading. *Journal for Research in Mathematics Education*, 5-33.

Anderson, A., Anderson, J., & Shapiro, J. (2005). Supporting multiple literacies: Parents' and children's mathematical talk within storybook reading. *Mathematics Education Research Journal*, *16*(3), 5-26.

Avci, N., & Dere, H. (2002). Okul oncesi cocugu ve matematik. *V. Ulusal Fen Bilimleri ve Matematik Kongresi*, 16-18.

Barnaby, D. (2015). The use of Children'sLiterature to Teach Mathematics to improve Confidence and Reduce Math Anxiety. Unpublished master's dissertation, University of Toronto, Ontario, Canada.

Baroody, A. J., Lai, M. L., & Mix, K. S. (2006). The development of young children's early number and operation sense and its implications for early childhood education. *Handbook of research on the education of young children*, *2*, 187-221.

Casey, B., Kersh. J. & Young. J. (2004). Storytelling sagas: An effective medium for teaching early childhood mathematics. *Early Childhood Research Quarterly, 19*, 167–172.

Casey, B., Erkut, S., Ceder, I., & Young, J. M. (2008). Use of a storytelling context to improve girls' and boys' geometry skills in kindergarten. *Journal of Applied Developmental Psychology*, *29*(1), 29-48.

Daglioglu, H. E., & Cakmak, O. C. (2009). Okul oncesi cocuklarina yonelik yayinlanan hikâye kitaplarinin siddet ve korku ogeleri acisindan incelenmesi. *Turk Kutuphaneciligi*, *23*(3), 510-534.

Demircan, C. (2006). TUBITAK cocuk kitaplıgı dizisindeki kitaplarin dis yapisal ve iç yapisal olarak incelenmesi. *Mersin Universitesi Egitim Fakultesi Dergisi*, *2*(1).

Erdal, K. (2009). Egitim degerleri acisindan cocuk kitapları. *Akademik Bakis,* 17, 1-18.

Forrest, K., Schnabel, D., & Williams, M. E. (2006). Mathematics and literature, anyone? *Teaching Children Mathematics, 13*(4), 216-217.

Ginsburg, H. P., & Golbeck, S. L. (2004). Thoughts on the future of research on mathematics and science learning and education. *Early Childhood Research Quarterly*, *19*(1), 190-200.

Gonen, M. (1993). Anaokulu ogretmenlerinin bes ve alti yas cocuklari icin kullandiklari masal ve hikâye kitaplarinin niteliklerinin incelenmesi. *Turk Kutuphaneciligi*, *7*(2), 83-88.

Griffiths, R., & Clyne, M. (1991). The power of story: Its role in learning mathematics. *Mathematics Teaching*, *135*, 42-45.

Gursimsek, I., & Günay, D. (2010). Cocuk kitaplarında cinsiyet rollerinin islenisinde kullanilan dilsel ve dildisi gostergelerin degerlendirilmesi. *Buca Egitim Fakultesi Dergisi*, (18).

Inan, C. (2014). Okul oncesi ogretmen adaylarinin matematik dersini ogretebilme konusunda hazir bulunusluk duzeylerinin degerlendirilmesi (Diyarbakır il ornegi). *International Periodical for the Languages, Literature and History of Turkish or Turkic*, *9*(8), 537-550.

Hong, H. (1996). Effects of mathematics learning through children's literature on math achievement and dispositional outcomes. *Early Childhood Research Quarterly*, *11*(4), 477-494.

Isitan, S. (2014). Caillou cizgi film karakteri ile ilgili Turkçe resimli oyku kitaplarinin bicimsel ve icerik ozelliklerinin incelenmesi. *Ilkogretim Online*,*13*(1).

Ilter, I. (2014). Cocuk edebiyatı aracılığıyla sosyal bilgilerin matematikle butunlestirilmesi. *International Journal of Human Sciences,* 11(2), 1117-1138

Jennings, C. M., Jennings, J. E., Richey, J., & Dixon-Krauss, L. (1992). Increasing interest and achievement in mathematics through children's literature. *Early Childhood Research Quarterly*, *7*(2), 263-276.

Karasar, N. (2014). *Bilimsel arastirma yontemleri*. Ankara: Nobel.

Keat, J. B., & Wilburne, J. M. (2009). The impact of storybooks on kindergarten children's mathematical achievement and approaches to learning. *Online Submission*, *6*(7), 61- 67.

Kilic, O. O., Gulec, H., & Genc, S. Z. (2014). Okul oncesi donem resimli oyku kitaplarının cografi kavramlari icermesi yönünden incelenmesi. *Ahi Evran Universitesi Kirsehir Egitim Fakultesi Dergisi*, *15*(1).

McDuffie, A. M. R., & Young, T. A. (2003). Promoting mathematical discourse through children's literature. *Teaching Children Mathematics*, *9*(7), 385.

Merriam, S. B. (2015). *Nitel arastirma: Desen ve uygulama icin bir rehber*. (Çev. S. Turan). Istanbul: Nobel.

Miles, M.B. & Huberman, A.M. (1994). *Qualitative data analysis* (2. Baski). Thousand Oaks, CA: Sage Publications.

Montague-Smith, A. & Price, A. J. (2012). *Mathematics in the early years* (3rd ed.). London, UK: Routledge.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM Publishing.

Ocal, T., Ocal, M. F., & Simsek, M. (2015). Investigating mathematical elements in story books appropriate for pre-schoolers. *Current Research in Education*, *1*(2), 58-69.

Ozen Uyar, R. & Yilmaz Genç, M. M. (2016). MEB 2013 Okul Öncesi Eğitim Programinin Bilimsel Surec Becerileri Acisindan Analizi, 3.Uluslararasi Avrasya Egitim Arastirmalari Kongresi (Sozlu bildiri), 1-3 Haziran, Mugla.

Powell, S. R., & Nurnberger-Haag, J. (2015). Everybody counts, but usually just to 10! A systematic analysis of number representations in children's books. *Early Education and Development*, *26*(3),377- 398.

Radebaugh, M. R. (1981). Using children's literature to teach mathematics. *The Reading Teacher*, *34*(8), 902-906.

Shatzer, J. (2008). Picture book power: Connecting children's literature and mathematics. *The Reading Teacher*, *61*(8), 649-653.

Shih, J. C., & Giorgis, C. (2004). Mathematics and Literature Connection.*Teaching Children Mathematics*, 329.

Thatcher, D. H. (2001). Reading in the math class: selecting and using picture books for math investigations. *Young Children*, *56*(4), 20- 26.

Turgut, M. F., & Baykul, Y. (2013). *Egitimde olcme ve degerlendirme*. Ankara: Pegem Akademi.

van den Heuvel-Panhuizen, M., van den Boogaard, S., & Doig, B. (2009). Picture books stimulate the learning of mathematics. *Australian Journal of Early Childhood*, *34*(3), 30-39.

van den Heuvel-Panhuizen, M., & Elia, I. (2011). Kindergartners' performance in length measurement and the effect of picture book reading. *ZDM*, *43*(5), 621-635.

Yildirim Hacıibrahimoglu, B. (2016). Olcme. In B. Akman, (Ed.), *Okul oncesi donemde matematik* (5.Baskı), (pp. 148-157) Ankara: Pegem Akademi.

Özet

Resimli Çocuk Kitaplarında Yer Alan Matematiksel İçeriğin İncelenmesi

Atıf:

Yılmaz-Genç, M.M., Akıncı-Coşkun, A., & Pala, Ş. (2017). A study of mathematical content provided in ıllustrated children's books. *Eurasian Journal of Educational Research, 69*, 159-175. http://dx.doi.org/10.14689/ejer.2017.69.9

Problem Durumu: Okul öncesi dönem, gerçek yaşam koşulları ve ileriki akademik süreç için gerekli olan matematiksel düşünmenin, matematik kavram ve becerilerinin temelinin hazırlandığı evredir. Bu dönemde, çocuklara matematiğin bir bağlam çerçevesinde sunulması, anlamlı matematik deneyimleri bakımından oldukça önemlidir. Okul öncesi eğitimde sıklıkla kullanılan resimli çocuk kitapları, barındırdığı matematik kavram ve becerilerini gerçek yaşam durumları ile ilişkilendirerek, sözü edilen bağlamsal temeli sağlamaktadır. Alan yazın incelendiğinde yapılan çalışmaların çoğunlukla, matematik eğitiminde kullanılabilecek nitelikli kitap isimleri önerisine yoğunlaştığı görülmektedir. Diğer taraftan, resimli çocuk kitaplarının okul öncesi matematik eğitiminde kullanımının, çocukların matematiği anlamaları üzerinde pozitif bir etkisi olduğu belirtilmektedir. Ayrıca yalnızca okul etkinliklerinin değil, aile katılımlı kitap okuma çalışmalarının da

çocukların matematiksel kavram gelişiminde etkili olduğu sonucuna ulaşılmıştır. Resimli çocuk kitaplarının Türkiye'deki durumunu ortaya koymak üzere; içerik ve fiziksel özellikler , cinsiyet rolleri, şiddet öğeleri, değerler, mizahi ve coğrafi unsurlar alanlarında çalışmalar yapılmıştır. Ancak resimli çocuk kitaplarında yer alan matematik içeriğini incelemeye yönelik sınırlı sayıda çalışmaya rastlanmıştır. Okul öncesi dönem matematik eğitiminde resimli çocuk kitapları kullanılmasının önemli olduğu belirtilmektedir. Fakat odağı matematik olan resimli çocuk kitaplarında, kavram ve becerilerin yer almasına ilişkin mevcut durumun bilinmediği görülmektedir. Bu bağlamda, okul öncesi matematik eğitiminde kullanılmak üzere hazırlanan resimli çocuk kitaplarının, matematiksel içerik durumlarının incelenerek eksikliklerin belirlenmesine ihtiyaç duyulmaktadır.

Araştırmanın Amacı: Bu çalışma ile NCTM (National Council of Teachers of Mathematics) tarafından okul öncesi döneme yönelik belirlenen matematik öğretimi içerik standartları arasında gösterilen "Sayma ve İşlem, Cebir, Geometri, Ölçme, Veri Analizi ve Olasılık" standartlarının resimli çocuk kitaplarında yer alma durumlarının incelenmesi amaçlanmıştır.

Araştırmanın Yöntemi: Okul öncesi döneme yönelik resimli çocuk kitaplarının matematiksel içerik açısından incelendiği bu çalışma, betimsel nitelikte olup tarama modeli ile gerçekleştirilmiştir. Çalışma kapsamında, Adana ve Ankara illerindeki kitapevlerinde satışa sunulan ve akademisyenlerin özel kütüphanelerinde bulunan 174 adet resimli çocuk kitabı incelenmiştir. Bu inceleme neticesinde, ölçütleri karşıladığına karar verilen 52 adet resimli çocuk kitabı, araştırmanın çalışma grubunu oluşturmuştur. Araştırma verilerinin toplanması sürecinde doküman analizi gerçekleştirilmiştir. Çalışma grubunu oluşturan resimli çocuk kitaplarının matematiksel içerik bakımından incelenmesi amacıyla "İçerik Kontrol Listesi" oluşturulmuştur. Bu listenin oluşturulma aşamasında, uluslararası düzeyde kabul gören ve matematik eğitiminde temel olan başlıkları ayrıntılı bir şekilde kapsayan NCTM içerik standartlarından yararlanılmıştır. Bu standartlar arasında gösterilen; "sayı ve işlem, cebir, geometri, ölçme, veri analizi ve olasılık" alanları ise oluşturulan kontrol listesinin alt boyutları arasında yer almaktadır. Çalışmada veri analiz birimi olarak kitaplarda yer verilen matematiksel ifadeler alınmış ve oluşturulan içerik kontrol listesine paralel olarak analiz edilmiştir. Çalışma grubuna alınan kitaplar arasından, üç araştırmacı tarafından ortak olarak incelenmek üzere yedi kitap seçilmiştir. Bu kitaplarda yer aldığı düşünülen matematiksel ifadeler not edilmiş ve bu ifadelerin hangi matematiksel içerikleri kapsadığı konusunda ortak karar alınarak kontrol listesine işaretleme yapılmıştır. Kalan 45 kitapta yer alan matematiksel ifadeler ise üç araştırmacı tarafından bağımsız olarak değerlendirilmiştir. İncelemeler tamamlandığında, araştırmacılar arası uyuşum katsayısı hesaplanmıştır. İçerik Kontrol Listesi'ni oluşturan her alt boyut için ayrı ayrı yüzdeler tespit edilmiş ve uyum yüzdeleri; sayı ve işlem için 0.82; cebir için; 0.86; geometri için 0.86; ölçme için 0.78; veri analizi ve olasılık için ise 0.84 bulunmuştur.

Araştırmanın Bulguları: Okul öncesi matematik eğitiminde kullanılmak üzere hazırlanan resimli çocuk kitaplarındaki matematik içeriğin incelenmesi amacıyla yapılan bu araştırmadan elde edilen bulgular; incelenen kitaplarda en fazla sayı ve işlem (%44,7) kategorisine yer verildiğini göstermektedir. Bunu sırasıyla; geometri (%30,0), ölçme (%21,5), cebir (%3,1) ile veri analizi ve olasılık (%0.5) kategorilerinin takip ettiği belirlenmiştir. Kitaplarda en az ele alınan kategori olan veri analizi ve olasılık genel kategorisine ait özel alt kategorilerden yalnızca verileri temel alarak çıkarımda bulunma özel alt kategorisine yer verildiği, veri toplama, verileri düzenleme, verileri görselleştirme, verileri tanımlama ve temel olasılık kavramlarını anlama ve uygulama ait özel alt kategorilerine hiç yer verilmediği görülmüştür.

Araştırmanın Sonuç ve Önerileri: Araştırma sonuçları genel olarak incelendiğinde, çalışma grubuna alınan resimli çocuk kitaplarında; hesaplamada farklı yöntem ve araçlar kullanma, ayırma, cebir sembolleri ve matematiksel modeller, nicel değişimler, simetri ve öteleme hareketleri, tekrarlı ölçümler, veri toplama, verileri düzenleme, verileri görselleştirme, verileri tanıma ve temel olasılık kavramlarını anlama ve uygulama özel alt kategorilerine hiç yer verilmediği görülmüştür. Araştırma kapsamında elde edilen bu sonuçlar ülkemizde hazırlanan kitapların NCTM matematik standartlarını karşılamada yeterli olmadığını göstermektedir. Çalışma grubunun oluşturulmak üzere incelenen 174 kitabın yalnızca 52 tanesinde matematiğin hikâyenin odağında olduğu ve matematiğin hikâye ile bütünleştirildiği görülmüştür. Bu oranın oldukça yetersiz olduğu düşünülmektedir. Ayrıca, konu bazında bakıldığında, incelenen kitaplar arasında hiç yer verilmeyen veya çok az ele alınan kategoriler olduğu da görülmüştür. Özellikle bu alanlara yönelik kitapların yayınlanmasının gerekli olduğu düşünülmekte ve yayınevleri ve yazarlara bu konular üzerine eğilmeleri önerilmektedir. Zengin görsel sunumlar ve anlamlı bir içerik sayesinde matematiksel kavramların anlaşılmasını kolaylaştıran resimli çocuk kitaplarının okul öncesi dönem matematik etkinliklerinde öğretmenler tarafından sıklıkla kullanılması önerilmektedir. Farklı okuma teknikleri ve uygun müdahaleler ile etkili bir şekilde kullanılan bu hikâye kitaplarının, çocukların matematik etkinliklerine karşı motivasyonlarını artıracağı düşünülmektedir. Okul öncesi öğretmenlerini, matematik etkinliklerinde resimli çocuk kitaplarını kullanma konusunda teşvik etmek amacıyla önerilen resimli çocuk kitapları listesi yayınlayabilir. Bu sayede öğretmenler, uzmanlar tarafından incelenmiş ve doğru içeriğe sahip kitaplara ulaşma fırsatı bulabilir. Bu çalışma, okul öncesi matematik eğitimine yönelik resimli çocuk kitaplarının, NCTM tarafından belirlenen matematik içerik standartlarını karşılama durumunu incelemek üzere gerçekleştirilmiştir. İleri çalışmalar için; kitaplarda yer verilen matematik kategorilerinin ele alınış biçimlerinin ayrıntılı olarak incelenmesi veya öğretmenlerin sınıflarında gerçekleştirdikleri matematik etkinliklerine resimli çocuk kitaplarını dâhil etme durumlarının incelenmesi araştırmacılara öneri olarak sunulabilir.

Anahtar Sözcükler: Erken çocukluk eğitimi, okul öncesi dönem matematik, NCTM, resimli çocuk kitapları.

Eurasian Journal of Educational Research 69 (2017) 177-197

Eurasian Journal of Educational Research
www.ejer.com.tr

Development of the Academic Performance Perception Scale [1]

Recep GUR[2]

ARTICLE INFO

Article History:

Received: 14 December 2016

Received in revised form: 18 February 2017

Accepted: 08 April 2017

DOI: http://dx.doi.org/10.14689/ejer.2017.69.10

Keywords

academic effort, self-efficacy, achievement, academic proficiency, polytomous DIF

ABSTRACT

Purpose: While numerous studies about academic performance that focused on only one factor, studies aiming to measure academicians' perceptions across many factors have not been observed in the literature. The current study aims to fill this gap and become a resource for upcoming studies. The aim of this study is to develop a valid and reliable scale that measures academicians' performance perception. **Research Methods:** The first research group of the study consists of 125 academicians who have been working in or studying for post-graduate degrees at Ankara University Faculty of Educational Sciences, while the second group of researchers, who have undergone confirmatory factor analysis, consists of 147 academicians either working as professionals or acquiring post-graduate education at Erzincan University (except for one multivariate extreme value). The pre-testing form of the scale composed of 29 positive and 11 negative factors, for a total of 40 items. The expert opinions obtained about the items is evidence for content validity. **Findings:** Results indicated that the final form of the scale which was composed of 19 positive and 7 negative factors, 26 items in total, is a reliable and valid data collection tool to be used in the field of education. **Implications for Research and Practice:** Researchers may be able to use this newly developed tool to investigate the presence of a meaningful relationship between academic encouragement scores and the academic performance perceptions of academicians.

[1] This study was presented at the 3rd International Eurasian Educational Research Congress in Muğla, Turkey, 31 May-3 June 2016.

[2] Corresponding Author: Erzincan University, Faculty of Education, Department of Measurement and Evaluation, Turkey, rgur@erzincan.edu.tr

Introduction

Investment is required in society for the development of a nation, while for the development of society, investment is also required in the individual. Education is the most important factor concerning an investment in the individual. Similar to the various developments in economics, technology, and politics, there are also rapid developments in the field of education. Increasingly more scientific studies are required in order to keep up with research being executed in developed countries, to compete with them and not to lag behind.

Academicians are among the first to come to mind with regard to individuals who use their intelligence most in the sense of curiosity. The studies of the academics whose performances are at a high level play a prominent role in the development and changes in society. According to Aldakhillah and Parante (2002), performance is the efficacy and competence of the duty that the individual is obliged to do. Competence refers to the capacity that should be possessed by the individual in order to be able to effectively fulfil a task or a job (Sahin, 2004). Academic performance is the value of the academicians determined by accumulating different criteria together (Kaptanoglu and Ozok, 2006). In the determination of these values, the academic performance perceptions of the academicians are encountered. Perception is the process of interpreting or making sense of the information that was received through the sensory organs (Schunk, 2012). Furthermore, the perception of academic performance is the reflection of the responses encountered in relation to the efficacy and competence of the academician.

Physical and hardware facilities should be provided for academicians to more easily carry out qualified scientific studies. The academic performance of research assistants who are temporary assigned to realise their postgraduate education at another university is affected by numerous external factors, such as the deprivation of physical and instrumental equipment, the absence of their own room, exclusion, the intensity of lecture assistant duties apart from administrative duties that affect post-graduation success, and the anxiety provoked by their various obligations (Kahraman, 2007). Therefore, it is vital to provide the necessary environment to enable the scientists to execute academic work with ease.

The scientists who are the architects of development and changes need to have good language skills. Gastel and Day (2016) emphasised that if a scientist is not a native English speaker, he or she may be more apprehensive to publish in English. Yavuzer and Gover (2012) reached the general conclusions in their research that the aim of the foreign language exams that are being conducted in Turkey are not being fulfilled, that the scope should be expanded in a way to measure the four language skills (reading, writing, listening and speaking), and that foraging language is obligatory in the execution and in the follow-up of scientific activities. Foreign language is important for the universality in science. Therefore, it may increase the quality of the scientific studies if foreign language exams are carried out in accordance with their purposes.

Giving verbal presentations, participating in international congresses with posters, and being able to publish are the criteria of academic performance. In the research conducted by Olkun (2006), the question "Are there any issues you would like to mention in relation to candidate articles coming from developing countries?" was responded to by the editors of national and international journals in the following way:

> They noted that there was a significant increase in the number of articles coming from developing countries. However, it was also mentioned that these candidate articles were insignificant, superficial, regarding only local problems. Moreover, they were written with bad English, without complying with the writing techniques. They were weak in terms of research design and analysis, and most of the authors were not updated. It was stated that these articles were rejected because of the insufficient scientific communication, or they were rejected by requesting significant corrections. Some of the editors stated that these problems were also encountered in articles coming from developed countries. So, it turned out that this problem was not only a language problem, but it was also related to scientific writing skills (p. 45).

The results of Olkun's research are parallel with Kline's (2009) expressions emphasising the importance of scientific skills. Kline stated that the use of language spoken in everyday life, antiquated research topics, and unnecessary graphics and tables are among the most notable problems. He also stressed that, in verbal presentations, the most important measure is to decide what to say and what not to say. In verbal presentations, active presentations are needed, instead of boring the audience by reading pages full of slides filled with assorted colours and animations (Kline, 2009). Therefore, raising scientists with advanced scientific writing skills will also positively affect their academic performance.

Gender roles in society come out to be one of the factors affecting academic performance or achievement of academicians. Female academicians listed such factors as "women's multiple roles" and "prejudice against women for positions requiring higher responsibility" among the problems they face within the institution for which they worked; i.e. the causes that result in lowered academic performance (Yilmaz and Ozdemir, 2012). Likewise, a variety of studies have been conducted suggesting that traditional roles imposed on women in the home are also carried out by the majority of female research assistants (Ergol, Koc, Eroglu and Taskin, 2012); neither their level of education nor their status has succeeded in changing the traditional roles created within the home, and the roles of women in society do not tend to change in the workplace (Dikmen and Maden, 2012). Belkis (2016) found that motherhood poses an overall concern due to distractibility in academic activities, fatigue, sleeplessness, and parental issues (babysitter, school, etc.). Even though female academicians persevere at their academic career aspirations with a modern understanding of motherhood, as it creates an intense workload, the academic performance of academicians is adversely affected. Belkis highlighted the social gender inequality generated due to maternity-related stereotypes as the effective

factor on academic performance, not the motherhood itself. Naymansoy (2010) discussed the role of motherhood in his study, and further pointed out the lack of preschool institutions, like day care centres, offering mothers assistance in child care, which constitutes one of the hindrances preventing academic performance.

Although most scientists are left alone with factors that lower their academic performance, they cannot stop themselves from executing scientific works due to the motivating influence of their curiosity. The belief of self-efficacy is also an impressive factor beside the required competence and skills that should be possessed by a scientist to perform qualified work. According to Bandura (1997, p. 3), "Self-efficacy is the individual's judgement related to himself about his capacity to organize and successfully perform the activities that are necessary in order to demonstrate a certain performance". Therefore, the belief related to things that have been achieved by an individual before and can be achieved in the future again affects academic success.

Self-efficacy measures can be formulated based on criteria set for the performance. The self-efficacy belief is an affective factor that increases the performance (McCown, Driscoll, and Roop, 1996). There is not a characteristic that can compensate for lack of knowledge or skills. Therefore, academics with high self-efficacy beliefs choose high-level goals that require being stronger, spending more time, and performing better, which also increases their success (Goddard, Hoy, and Hoy, 2004). Therefore, self-efficacy measures help to explain why academics in the same field with the same level of ability demonstrate different academic performances (Hazir and Bikmaz, 2004; Lane, Hall, and Lane, 2004; Schunk and Pajares, 2005; Zimmerman and Kitsantas, 2005). Therefore, the academic performances of the academicians who believe they will succeed, even when facing difficulties, i.e. academicians with high self-efficacy, are expected to be high.

Many studies have been conducted on the self-efficacy belief, especially regarding an individual's choice of activities, the steadiness against the difficulties, the level of the effort, and its impact on the performance. The following examples can be given in this regard; career self-efficacy (Bacanli, 2006); self-efficacy beliefs in writing; academic self-efficacy beliefs (Lent, Brown and Gore, 1997), research self-efficacy (Bishop, Bieschke, and Garcia, 1993), self-efficacy beliefs related to computers (Akkoyunlu, Orhan and Umay, 2005), professional self-efficacy (Schyns, 2004). In the conducted studies, it was concluded that high academic self-efficacy had an important influence on academic life (Pajares and Graham, 1999; Schunk, 1995) and a positive effect on performance (Vrugt, Langereis and Hoogstraten, 1997).

The research self-efficacy beliefs are among the factors that affect scientific research skills. Research self-efficacy is the belief that an individual can complete a research task. This belief affects an individual's academic performance (Bard, Bieschke, Herbert and Eberz, 2000). Bailey (1999) concluded that there was a positive relationship between research self-efficacy and the motivation of the academicians, academic degrees, and scientific research experiences. Therefore, scientists with high research self-efficacy exhibit a higher academic performance.

In the globalising world, the contribution of computers to science is a fact that cannot be denied. It was concluded that academicians who show a positive attitude towards the computer are more confident during their teaching process, and display higher self-efficacy beliefs toward their academic performance because they benefit more from computers and information technology (Ipek, Tekbiyik and Ursavas, 2010). It is important for scientists to improve their ability to use technological facilities and follow technology closely to conduct scientific research.

In the literature, the research studies related to academic performance seem to focus on a single dimension, such as language competence (Ocal, 2012; Yavuzer and Gover, 2012), academic writing skills (Kline, 2009; Olkun, 2006), career self-efficacy (Bacanli, 2006), self-efficacy beliefs in writing (Parajres, Hartley, and Valiante, 2001), academic self-efficacy beliefs (Lent, Brown, and Gore, 1997), research self-efficacy (Bishop, Bieschke and Garcia, 1993), self-efficacy beliefs related to computers (Akkoyunlu, Orhan and Umay, 2005), and professional self-efficacy (Schyns, 2004). Although there are studies about the academic performance that focus only on one factor, no studies have been found in the related literature that intend to measure the academic performance perception of the academicians amongst numerous factors. The identification of the level of the academicians' academic performance perceptions is considered to be beneficial for the relevant institutions. By determine this measure, institutions may be able to detect what steps should be taken in order to enable the academicians to continue their contributions to their country scientifically and technologically, and train qualified academicians by using the existing resources of the country in the best way. A limited number of studies conducted on the academic performance are available in the literature. This study is considered to be a source for other studies in terms of eliminating deficiencies in the literature. From this point on, the aim of this study is to develop a valid and reliable scale that can measure the academic performance perceptions of academicians.

Method

Research Groups

There were two different research groups involved in this study. Principal Component Analysis (PCA) was applied for the data obtained from first research group, while Confirmatory Factor Analysis (CFA) was performed with the data obtained from the second research group. The first research group of this study consisted of 125 academicians working at Ankara University, Faculty of Educational Sciences or enrolled in postgraduate education. Of the 125 academicians, 77 were women (61.60%) and 48 were men (38.40%). The average age of the academicians was 29, and their ages varied from 22 to 51. With regard to marital status, 47 of the 125 academicians were married (37.60%), 75 were single (60.0%), and 3 were divorced (2.40%). In addition, 111 of the academicians (88.80%) had no children, 10 (8.00%) had one child, and 4 academicians had two children. Of the 125 academicians, 120 (96.00%) attended or were still attending the "Scientific Research Methods" course, while 5 of the academicians (4.00%) had not attended this course.

The second research group consisted of 147 academicians (except for one multivariate extreme value) either working as professionals or enrolled in post-graduate education at Erzincan University. Of the 147 academicians, 54 (36.80%) were Research Assistants, 13 (8.80%) were Lecturers, 62 (42.20%) were Assistant Professors, 13 (8.80%) were Associate Professors, and 5 (3.40%) were Professors. Of these academicians, 46 (61.60%) were females and 101 (38.40%) were males. The average age of the academicians was 34, ranging from 23 to 55. Of the 147 academicians, 99 (67.30%) were married, 45 (30.70%) were single, and 3 (2.00%) were divorced. Furthermore, 63 (42.90%) had no children, 26 (17.70%) had one child, 48 (32.70%) had two children, 8 (5.40%) had three children, and 2 (1.40%) had four children. A total of 131 (89.10%) had attended or were currently attending the "Scientific Research Methods" course, while the remaining 16 (10.90%) had not yet attended the "Scientific Research Methods" course.

Research Instruments and Procedures

Three academicians working at Ankara University, Faculty of Educational Sciences, were asked to write an essay describing their feelings and thoughts about the factors affecting the academic performance. As a result of the literature review and the examination of the essays written by the academicians, 13 negative and 32 positive expressions regarding the academic performance perception were created. Expressions that conveyed double negativity, uncertainty, and incoherency were avoided by considering the development steps of the Likert-type scale (Tavsancil, 2010). The prepared items were examined by two experts who conducted research on "self-efficacy", three Measurement and Evaluation experts, a specialist in Computer Teaching and Technology, an English Teacher, and a Turkish Language Expert, and were evaluated in terms of language, scope, and psychometrics.

The expressions were revised according to the opinions, and the Academic Performance Perception Scale, consisting of 40 items, was prepared for a pre-trial application. A total of 29 of these expressions were positive *("I know the concepts related to my field well enough to teach them effectively to the students", "I can perform the data analysis of my research without any help", "I closely follow the developments in technology", etc.)*, while 11 of these expressions reflected negative perceptions regarding academic performance *("I hesitate to speak at congresses held abroad", "I have difficulties in reporting statistics programs (e.g. SPSS/SAS.)", "I leave my research incomplete when facing difficulties", etc.)*. The positive and negative items were mixed in the scale.

Data Analysis

Since 29 items on the scale reflected positive perception regarding the academic performance, a scoring key graded as '1' Never, '2' Rarely, '3' Occasionally, '4' Frequently, and '5' Always was prepared, while for the 11 items reflecting negative perception, the scoring key was rated reversely. The scores of the scale were calculated according to this scoring key by collecting the scores of the academicians. Univariate outliers were not found in relation to the total scores of the academicians' academic performance perception scale because there was no z value other than the -

3< z< 3 range (Mertler and Vannatta, 2005). Histogram graphs, skewness, and kurtosis values, along with the mean, median, and mode values, were examined for the univariate normality assumption. The range between -1 and +1 was taken as the acceptable range for the skewness and kurtosis (Leech, Barrett, and Morgan, 2005).

The Pearson Product Moments Correlation Coefficient was calculated for the item-total test correlation in order to determine the item validity of the 40 items included within the test form of the scale. The item analysis method based on the difference between the sub-superior group averages was used in order to determine the item discrimination. The difference between the sub-superior groups was calculated by the t test of the independent groups. Furthermore, as proof of the item validity, the Ordinal Logistic Regression Method was utilised to detect Differential Item Functioning (DIF) in polytomous items in terms of gender (Miller and Spray, 1993). Ordinal Logistic Regression assumes a negligible/tolerable level of DIF, if ΔR^2<.13 (Level A); a medium level of DIF, if .13≤ΔR^2<.26 (Level B); and a magnitude level of DIF is present if ΔR^2>.26 (Level C) (Zumbo and Thomas, 1996). One of the most significant steps of the scale development process is to identify if there is any biased item for and against a group in a systematic manner (Camilli and Shepard, 1994; Zumbo, 1999; Allalouf, Hambleton and Sireci, 1999). A biased item in a scale has an adverse effect on the validity of the measurement results (Clauser and Mazor, 1998). Differential Item Functioning procedures should be followed as a prerequisite to identify the item bias (Zumbo, 1999). An item containing DIF in a measuring scale available for affective traits suggests that the likelihood of individuals in different subgroups (as per gender, ethnicity, etc.) with the same scale scores (same attitude/perception level) exhibiting similar responses to the relevant item will vary (Hulin, Drasgow and Parsons, 1983; cited by Dodeen and Johanson, 2003). Estimations based on expert reviews are needed to claim bias on any item that is specified to flag DIF as a result of the statistical analysis (Camilli and Shepard, 1994; Zumbo, 1999). Notwithstanding the differences in literature with regards to the sample sizes of DIF studies in polytomous items, Wood (2011) defined a small sample size to be 40 individuals, while Fidalgo, Hashimoto, Bartram, and Muñiz (2007) and Muñiz, Hambleton and Xing (2001) defined a small sample size to be 50 individuals per group.

The principal component analysis method was used to determine the construct validity of the academic performance perception scale. According to Sencan (2005), instead of the descriptive factor analysis, researchers should use the principal component analysis method if the main aim of the research is to develop a main objective scale or to detect under which dimensions the measurement items can be grouped. The KMO value and the results of the Bartlett test were examined in terms of the assumptions of the principal component analysis (Kline, 2000). In addition, the items that did not fit to the scale provided load for more than one factor, and had a factor load of less than .50 because the size of the sample was N= 125, according to Km-Yin (2004), were excluded from the scale by considering the items whose factor eigenvalue was bigger than 1 (as cited in Sencan, 2005). In order to validate the underlying factor structure derived following the implementation of Principle

Component Analysis (PCA) on the data of first research group, Confirmatory Factor Analysis (CFA) was applied for the data collected from the second research group (N=147). The criterion that required the sample size to be at least five times the observed number of variables was considered in performing CFA (Tabachnick and Fidell, 2007). Moreover, upon the examination of the z scores of academicians in relation to their CFA assumptions, coefficients for Mahalanobis distance, residual values, tolerance, VIF values, and condition index (CI), it was concluded that no univariate extreme value (±3z) was present, and there was only one multivariate extreme value ($(\chi 2 = 59.20 > \chi 2_{(26,.001)}$=54.05; p<.001). Furthermore, the data exhibited a multivariate normal distribution, test linearity was achieved, and a multicollinearity problem (tolerance<.10; VIF<10; CI<30) did not exist. The Cronbach's Alpha coefficient was calculated for the reliability of the academic performance perception scale.

Results

The descriptive statistics related to the scale scores of the academicians in the first research group are presented in Table 1.

Table 1

Descriptive Statistics Related to the Scale Scores of the Academicians

N	*Range*	*Min*	*Max*	*M*	*Med*	*Mod*	K_y	B_S	*Ss*
125	*93*	*94*	*187*	*139.23*	*140*	*140*	*-.06*	*-.05*	*17.76*

When examining Table 1, it can be stated that the group consisting of 125 academicians is heterogeneous because of the wide range and high standard deviation. The fact that the values of the skewness and kurtosis are within +-1 and that the mode, median, and arithmetic mean values are close to each other provides information regarding that the univariate normality assumption was ensured (Leech et al., 2005). In addition, the histogram graph related to the scale scores of the academicians is given in Figure 1.

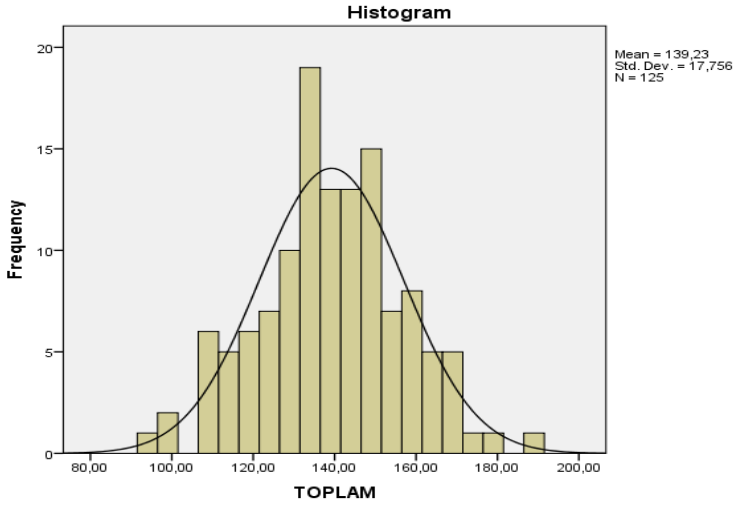


Figure 1. Histogram graph related to the scale scores of the academicians

When examining Figure 1, it is seen that the distribution of the academicians' scale scores shows a distribution which is similar to a normal distribution. The item-total test correlation was examined to determine whether there was a positive and linear relationship between the responses given to an item on the scale and the responses given to the whole scale (Erkus, 2003). The item-total test correlations were between .18 and .73, and each of the items had a significant relationship with the scale scores ($p<.05$).

When the t values were calculated for the 27% sub-superior groups to determine whether the items distinguished between those having positive perceptions regarding academic performance and those having negative perceptions, the t values were observed to vary between 1.66 and 12.77. Therefore, the item scores, except the item scores of the 17th, 24th, 32nd, and 33rd items on the academic performance perception scale, showed a significant difference according to the 27% sub-superior groups ($p<.05$). Accordingly, it can be stated that, when the non-significant items are excluded from the scale, the sub and superior groups of the scale are well-distinguished. The Ordinal Logistic Regression Method was used to explore if polytomous items that contain DIF in terms of gender exist in the scale that was developed to measure the academic performance perceptions of academicians. In line with this information, the Results of the Ordinal Logistic Regression Method for DIF Analysis is given in Table 2.

Table 2

The Results of the Ordinal Logistic Regression Method for DIF Analysis

Item No	ΔR^2	*Item No*	ΔR^2	*Item No*	ΔR^2	*Item No*	ΔR^2
1	.012	11	.001	21	.036	31	.008
2	.009	12	.003	22	.008	32	.049
3	.006	13	.008	23	.032	33	.017
4	.004	14	.009	24	.042	34	.012
5	.030	15	.045	25	.003	35	.041
6	.028	16	.034	26	.033	36	.046
7	.013	17	.001	27	.029	37	.038
8	.033	18	.019	28	.022	38	.031
9	.082*	19	.051	29	.048	39	.080
10	.064*	20	.074*	30	.023	40	.023

(*:$\chi 2_{Model3} - \chi 2_{Model1} > \chi 2_{.01}$)

When examining Table 2, it can be propounded that, even if three out of 40 items flagged DIF in terms of gender, these three items (Items no. 9, 10, and 20) have a negligible/tolerable level of DIF, i.e. (ΔR^2<.13) at Level A (Zumbo and Thomas, 1996). Hence, in addition to the item-total test correlation and the outcomes of the t test for 27% sub-superior groups, the inexistence of an item containing DIF at the levels of B and C in the scale can be claimed as evidence for item validity.

To determine the construct validity of the scale, the principal component analysis method was applied to the 36 items that were found to be significant according to the t test results for item-total test correlation and the 27% sub-superior groups. As a result of the analysis, 10 items were removed from the scale because the items provided load to more than one factor and had a factor load less than the .50 needed to be excluded.

When the results of the KMO value and the Bartlett test were examined, the KMO value was found to be .84, which showed that the data structure of the research group consisting of 125 academicians and belonging to the 26 items was good enough to perform the principal component analysis in terms of the size (Leech et al, 2005). The fact that the obtained chi-square value related to the Bartlett test result was significant ($\chi^2(406, N=125)=1861.418, p<.01$) at the p=.01 level meant that the data came from a highly variable normal distribution. For the test of the construct validity, the Varimax rotation technique was used in the analysis of the principal components, since the scale was multi-factored. The findings related to the factors are given in Table 3.

Table 3

Findings Related to Factors that Were Obtained as a Result of the Principal Component Analysis

Factor	*Eigenvalues*	*Variance Percentage*	*Total Variance Percentage*
1	7.83	27.00	27.00
2	4.49	15.47	42.47
3	2.09	7.20	49.66
4	1.46	5.02	54.68
5	1.34	4.62	59.30

When examining Table 3, it can be seen that there were five factors with an eigenvalue bigger than 1.00. The first factor described 27% of the total variance. The contribution of the factors of the total variance percentage decreased after the first factor. This situation can be seen in Figure 2.

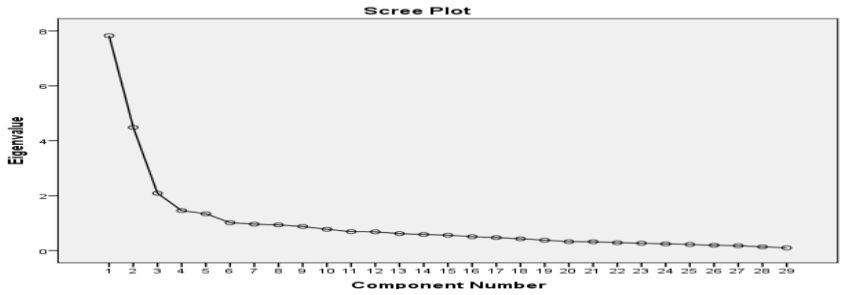


Figure 2. Scree Plot

When examining Figure 2, it can be stated that the slope experienced by high acceleration and rapid deceleration indicated a significant number of factors. After the fifth factor, it seemed that the slope started to stabilise. Therefore, it can be propounded that the eigenvalue that made the largest contribution to the total variance percentage was formed by a five-factor structure, by paying attention to the number of factors above one. The first factor consisted of 10 items, the second factor consisted of five items, the third factor of five items, and the fourth and fifth factors consisted of three items. The 26 items collected under these five factors described 59.30% of the total variance.

Table 4

Items and Factor Loads in the Dimensions as a Result of the Principal Component Analysis

Items and Factor Loads of the Foreign Language Self-Efficacy Dimension	
Item No	**Factor Load**
8	.87
10	.86
20*	.85
15	.84
21*	.79
22*	.78
7	.73
13	.72
18*	.71
23*	.65
Items and Factor Loads of the Scientific Research Self-Efficacy Dimension	
Item No	**Factor Load**
4	.82
5	.77
31*	.73
12	.63
3	.53
Items and Factor Loads of the Technology Self-Efficacy Dimension	
Item No	**Factor Load**
19	.76
9	.59
2	.58
11	.56
16	.53
Items and Factor Loads of the Effective Lecture Dimension	
Item No	**Factor Load**
26	.78
27	.72
28	.65
Items and Factor Loads of Self-efficacy versus external factors dimension	
Item No	***Factor Load***
35	.72
29	.63
36*	.62

*Items scored reversely

When examining Table 3, it can be observed that the factor loads of the items collected under the five dimensions varied between .53 (3rd item *"I can make my scientific studies fit the format of each journal"*.) and .87 (8th item *"I can engage in joint research with a foreign academician without the need for tools such as dictionaries or translation programs"*.).

As a result of the implementation of Principle Component Analysis, the Academic Performance Perception Scale consisting of 26 items with five factors was applied to the second research group. Confirmatory Factor Analysis (CFA) was conducted for the data obtained following the application to determine whether they fit the Five-Factor Model. Notwithstanding the differences in literature concerning which of the goodness of fit indexes acquired after CFA should be reported, in addition to the $\chi 2/df$ of all other indexes, Iacobbucci (2010) suggested CFI and SRMR to be reported, while Brown (2006) emphasised the reporting of RMSEA, SRMR, CFI and NNFI. However, Karagoz (2016) noted that RMSEA is highly responsive to sample sizes and, hence, RMSEA should not be reported in studies with small sample sizes (N<250). In line with this information, the Goodness of Fit Index for Academic Performance Perception Scale is provided in Table 5.

Table 5

Goodness of Fit Indexes and Values for Academic Performance Perception Scale

Goodness of fit index	*Values*
$\chi 2$	588.90
df	289
NNFI	.90
CFI	.91
SRMR	.09

Upon the analysis of Table 5, $\chi 2/df = 2.04$ index, which is below 2.50, proves to have a perfect fit (Kline, 2011); NNFI and CFI indexes, which are equal to or above .90, have an acceptable fit (Bentler, 1980; Marsh, Hau, Artelt, Baumert and Peschar, 2006); and SRMR index, which is below .10 have an acceptable fit (Hu and Bentler, 1999; Kline, 2011). Moreover standardized factor loadings of the items weren't found to be higher than 1, so it can be propounded that, the five-factor model have an acceptable fit with the data (Simsek, 2007). An overall analysis of fit indexes as a result of the Confirmatory Factor Analysis revealed the validation of five-factor structure of Academic Performance Perception Scale with 26 items. Brown (2006) stated that the examination of model stability or invariance on research groups is enabled when CFA is performed on different research groups. Therefore, the five-factor model for the Academic Performance Perception Scale can be inferred to have stable outcomes over different research groups as per CFA results.

The Cronbach's Alpha (α) coefficient, which is the internal consistency measure for the 40-item scale was .89, while it was found to be α=.88 for the final scale including 26 items, which was obtained as a result of the item-total test correlation, item analysis based on the sub-superior group averages in addition to the DIF outcomes, the analysis of the principal components. The Cronbach's alpha coefficient is quite high. This result indicates that the items forming the scale are consistent with each other, which means that the academic performance perception scale can be used reliably. After the validity and reliability analysis, the original form of the scale consists of 26 items. Nineteen items on the scale are positive, and 7 are negative. The highest score that can be obtained from the scale is 130, while the lowest score is 26.

Discussion and Conclusion

Discussion

In this study, a 26-item scale was developed to measure the academic performance perceptions of the academicians. The scale consists of five dimensions: Foreign Language Self-Efficacy, Scientific Research Self-Efficacy, Technology Self-Efficacy, Teaching Effectively, and Self-Efficacy versus External Factors. Foreign language is very important for the rapid spread of knowledge and communication within the globalising world (Ocal, 2012). In the study conducted by Yavuzer and Gover (2012), it was concluded that foreign language was obligatory in the execution and follow-up of the scientific activities. Therefore, the "Foreign Language Self-Efficacy" of the academicians is one of the factors that affects academic performance perceptions.

It is also emphasised that there is a need for educators who use computers effectively during the learning and teaching process, who are able to persevere when facing computer-related problems, i.e. who have high computer-related self-efficacy beliefs (Akkoyunlu, Orhan and Umay, 2005). In addition, there is a high level of positive correlation between the individuals' attitudes towards the computers and their self-efficacy beliefs related to academic performance (Ipek, Tekbiyik and Ursavas, 2010), which is in line with the "Technology Self-efficacy" dimension included within the academic performance perception scale. However, in the study conducted by Odaci and Celik Berber (2012), a significant negative relationship was found between problematic internet users who spend their time by browsing unnecessary pages on the Internet and academic self-efficacy; the academic procrastination behaviours of the such individuals showed a reduction in their academic self-efficacy beliefs. According to these findings, it can be stated that, if the technology is not used properly, it can negatively affect the academic performance of the academicians.

The prospect of the academicians' self-efficacy scores on their academic performances (Bishop, Bieschke and Garcia, 1993), according to the findings, demonstrate that high self-efficacy beliefs are likely to influence the willingness to teach and to be more successful in classroom management (Gibson and Dembo, 1984). In addition, individuals who display high self–efficacy do not give up while facing difficult situations (Bandura, 1997), which can be shown as evidence for the dimensions of scientific research self-efficacy, effective teaching, and the dimension of self-efficacy versus external factors.

Conclusion

In this study, it was concluded that the items were discriminating, since the item-test correlation was significant and the item scores of the academicians (except four items) showed a significant difference ($p<.05$) according to the 27% sub-superior groups, and no item containing DIF in terms of gender exists at the levels of B-C. As a result of the principal component analysis, the factor loads of the 26 items varied from .53 to .87 and described 59.30% of the total variance. Findings of CFA were

proven to demonstrate a good fit with the Five-Factor Model. This result provided evidence for construct validity. The internal consistency coefficient of the scale was .88. As a result, a reliable and valid data collection tool was developed that can be used for the determination of academicians' academic performance perceptions.

Recommendations

An application similar to the in-service training applied by the Ministry of Education for the teachers can be established by the Higher Education for the academicians by determining their academic performance perception levels. It can be investigated whether there is a significant relationship between academicians' academic incentive scores and their academic performance perceptions. Studies can be conducted on the accuracy level of the academic performance perception scale scores of the academicians working at five universities included among the best 500 universities, according to the ranking of the best universities in the world, and the academicians working at five universities that are not included within this ranking.

References

Akkoyunlu, B., Orhan, F., & Umay, A. (2005). Bilgisayar ogretmenleri icin "Bilgisayar ogretmenligi oz-yeterlik olcegi" gelistirme calismasi [Development of the self-efficacy scale for computer teachers]. *Hacettepe Universitesi Egitim Fakultesi Dergisi, 29*, 1-8.

Aldakhilallah, K. A., & Parente, D. H. (2002). Redesigning a square peg: Total quality management performance appraisals. *Total Quality Management,* 13(1), 39-51.

Allalouf, A., Hambleton, R., & Sireci, S. (1999). Identifying the causes of DIF in translated verbal items. *Journal of Educational Measurement, 36*, 185–198.

Bacanlı, F. (2006). Personality characteristics as predictors of personal indecisiveness. *Journal of Career Development, 32*(4), 320-332.

Bailey, J. G. (1999). Academics' motivation and self-efficacy for teaching and research. *Higher Education Research And Development, 18*(3), 343-359.

Bandura, A. (1997). *Self-efficacy*: *The exercise of control*. New York: Freeman.

Bard, C. C., Bieschke, K. J., Herbert, J. T., & Eberz, A. B. (2000). Predicting research interest among rehabilitation counseling students and faculty. *Rehabilitation Counseling Bulletin, 44*(1), 48-55.

Belkis, O. (2016). Anneligin akademik kariyer gelisimine etkileri uzerine nitel bir arastirma. *Egitim ve Ogretim Arastirmalari Dergisi,* 5(4), 250-263.

Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology,* 31, 419-456.

Bishop, R. M., Bieschke, K. J., & Garcia, V. L. (1993). Prediction of research self-efficacy and future research involvement (Report No. CG025398). *American Psychological Association.(ERIC Document Reproduction Service NO. ED370026)*.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: The Guilford Press.

Camili, G., & Shepard, L. A. (1994*). Methods for identifying biased test items.* London: Sage Publications.

Clauser, B. E., & Mazor, K. M. (1998).Using statistical procedures to identify differential item functioning test items. *Educational Measurement: Issues and Practice, 17*, 31–44.

Dikmen, N., & Maden, D. (2012). Kadin akademisyenlerin gorunmeyen emegi uzerine bir arastirma: Ordu Universitesi ornegi. *Istanbul Ticaret Universitesi Sosyal Bilimler Dergisi*, 11(1), 257-288.

Dodeen, H., & Johanson, G. A. (2003).An analysis of sex-related differential item functioning in attitude assessment.*Assessment & Evaluation in Higher Education, 28*(2), 129–134.

Ergol, S., Koc, G., Eroglu, K., & Taskin, L. (2012). Turkiye'de kadin arastirma gorevlilerinin ev ve is yasamlarinda karsilastiklari guçlukler. *Yuksekogretim ve Bilim Dergisi*, 2(1), 43-49.

Erkus, A. (2003). *Psikometri uzerine yazilar.* Ankara: Turk Psikologlar Dernegi.

Fidalgo, A. M., Hashimoto, K., Bartram, D., & Muñiz, J. (2007). Empirical Bayes versus standard Mantel-Haenszel statistics for detecting differential item functioning under small sample conditions. *Journal of Experimental Education, 75*, 293-314.

Gastel B., & Day, R. A. (2016). *How to write and publish a scientific paper.* California: Greenwood.

Gibson. S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology,* 76(4). 569-582.

Goddard, R. D., Hoy, W. K., & Hoy, W. A. (2004). Collective efficacy:Theoretical development, empirical evidence, and future directions. *Educational Researcher*, 33, 3-13.

Hazir Bikmaz, F. (2004). Ozyeterlik inanclari. In Kuzgun Y., & Deryakulu, D. (Ed.) *Egitimde bireysel farkliliklar* (pp. 289-314). Ankara: Nobel.

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

Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal of Consumer Psychology,* 20, 90-98.

Ipek, C., Tekbiyik, A., & Ursavas, O. F. (2010). Lisansustu ogrencilerinin arastirma oz-yeterlik inancları ve bilgisayar tutumlari. *Gaziantep Universitesi Sosyal Bilimler Dergisi, 9*(1), 127-145.

Kahraman, L. (2007). Turkiye'de ogretim uyesi yetistirme baglaminda "35. madde" uygulamasi. *Mulkiye Dergisi*, 31(256), 187–218.

Kaptanoglu D., & Ozok, A. F. (2006). Akademik performans degerlendirmesi icin bir bulanik model. *ITU Dergisi*, 5(1), 193-204.

Karagoz, Y. (2016). *SPSS ve AMOS uygulamali istatistiksel analizler*. Ankara: Nobel Yayinevi.

Kline, P. (2000). *An easy guide to factor analysis.* New York: Routledge.

Kline, R. B. (2009). *Becoming a behavioral science researcher a guide to research that matters.* New York: Guilford Press.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York: The Guilford Press.

Lane, A. M., Hall, R., & Lane, J. (2004). Self-efficacy and statistics performance among Sport Studies students. *Teaching in Higher Education, 9*(4), 435-448. doi: Doi 10.1080/1356251042000252372

Lent, R. W., Brown, S. D., & Gore, P. A. (1997). Discriminant and predictive validity of academic self-concept, academic self-efficacy, and mathematics-specific self-efficacy. *Journal of Counseling Psychology, 44*(3), 307-315. doi: Doi 10.1037//0022-0167.44.3.307.

Leech, N.L., Barrett, K.C., & Morgan, G.A. (2005). Spss *for Intermediate Statistics: Use and Interpretation*. London: Lawrence Erkbaum.

Marsh, H. W., Hau, K. T., Artelt, C., Baumert, J. & Peschar, J. L. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing,* 6(4), 311-360.

McCown, R., Driscoll, M., & Roop, P. G. (1996). *Educational psychology*. Boston: Allyn and Bacon.

Mertler, G. A., & Vannatta R. A. (2005*). Advanced and multivariate statistical methods.* USA:Pyrczak.

Miller, T. R., & Spray, J. A. (1993). Logistic discriminant function analysis for DIF identification of polytomously scored items. *Journal of Educational Measurement, 30*(2), 107–122.

Muñiz, J., Hambleton, R. K., & Xing, D. (2001). Small sample studies to detect flaws in item translations. *International Journal of Testing, 1*, 115-135.

Naymansoy, G. (2010). Turk bilim kadinlari ve bilime katkilari. *Gaziantep Universitesi Sosyal Bilimler Dergisi, 9*(1), 203-232.

Odaci, H., & Celik Berber, Ç. (2012). Universite ogrencilerinin problemli internet kullanimlarinin akademik oz-yeterlik, akademik erteleme ve yeme tutumlari ile iliskisi. *E-Journal Of New World Sciences Academy, 7*(1), 389-403.

Olkun, S. (2006). Egitim İle ilgili uluslararasi bilimsel dergilerde yayin yapma sureci: Firsatlar, sorunlar ve cozum onerileri. Karakutuk, K. (Ed.) *I. Ulusal kurultay bildirileri* (pp. 43-47). Ankara: Sosyal Bilimlerde Sureli Yayincilik.

Ocal, Ş. (2012). Yurt disi egitimi: Sorunlar ve cozum onerileri kuresellesme ve uluslararasi ogrenci harketliligi. *Egitim- Ogretim ve Bilim Arastirma Dergisi, 8*, 23-36.

Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology, 24*(2), 124-139. doi: DOI 10.1006/ceps.1998.0991

Pajares, F. Hartley, J., & Valiante, G. (2001). Response format in writing self-efficacy assessment: Greater discrimination increases prediction. *Measurement and evaluation in counseling and development; measurement and evaluation in counseling and development*.

Schunk, D. H. (1995). Self-efficacy, motivation, and performance. *Journal of Applied Sport Psychology, 7*(2), 112-137.

Schunk, D. H. (2012*). Learning theories: An educational perspective*. Boston: Pearson.

Schunk, D. H., & Pajares F. (2005). Competence perceptions and academic functioning. In A. J. Elliot and C. S. Dweck (Eds.) *Handbook of competence and motivation* (pp.85-104). New York: Guilford Press.

Schyns, B. (2004). The influence of occupational self-efficacy on the relationship of leadership behavior and preparedness for occupational change. *Journal of Career Development, 30*(4), 247-261.

Sahin, A. E. (2004). Ogretmen yeterliklerinin belirlenmesi. *Bilim ve Aklın Aydinliginda Egitim Dergisi*, 5(58).

Sencan, H. (2005). *Sosyal ve davranissal olcumlerde guvenirlik ve gecerlik.* Ankara: Seçkin.

Simsek, O. F. (2007). *Yapisal esitlik modellemesine giris temel ilkeler ve lisrel uygulamalari.* Ankara: Ekinoks Yayinlari.

Tabachnick, B. G., & Fidell L. S. (2007). *Using multivariate statistics*. USA:Pearson Education.

Tavsancil, E. (2010). *Tutumlarin olculmesi ve SPSS ile veri analizi*. Ankara: Nobel.

Vrugt, A. J., Langereis, M. P., & Hoogstraten, J. (1997). Academic self-efficacy and malleability of relevant capabilities as predictors of exam performance. *The Journal of Experimental Education*, 61-72.

Wood, W. S. (2011). Differential item functioning procedures for polytomous items when examinee sample sizes are small (Doctoral Dissertation). Graduate College of The University of Iowa, Iowa.

Yavuzer, H., & Gover, I. H. (2012). Akademik personelin yabancı dil durumu ve yabancı dil sinavlarina bakisi: Nevşehir ornegi. *NEU Sosyal Bilimler Enstitusu Dergisi, 1,* 136-158.

Yılmaz, E., & Ozdemir, G. (2012). Turkiye'de kadin akademisyen ve arastirmacilarin karsilastiklari sorunlar ve tarima bakis acilari. *Journal of Tekirdag Agricultural Faculty, 9*(2), 50-56.

Zimmermann, B. J., & Kitsantas, A. (2005). The hidden dimensionof personal competence:Self-regulated learning and practice. In A. J. Elliot and C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 509-526). New York: Guilford Press.

Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and likert-type (Ordinal) item scores.* Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.

Zumbo, B. D., & Thomas, D. R. (1996). *A measure of effect size using logistic regression procedures*. Paper presented at the National Board of Medical Examiners, Philadelphia, PA.

Akademik Performans Algı Ölçeği Geliştirme Çalışması

Atıf:

Gur, R. (2017). Development of the academic performance perception scale. *Eurasian Journal of Educational Research*, *69*, 177-197. http://dx.doi.org/10.14689/ejer.2017.69.10

Özet

Problem Durumu: Akademik performans üzerine tek boyutlu odaklanılan çalışmalar bulunmasına rağmen, akademisyenlerin birçok boyutta akademik performans algılarını belirlemeye yönelik ilgili literatürde herhangi bir çalışmaya rastlanılamamıştır. Akademisyenlerin akademik performans algılarının ne düzeyde olduğunun belirlenebilmesi, ülkelerin mevcut olanaklarını en iyi şekilde kullanarak hem nitelikli akademisyenlerin yetiştirilmesi hem de mevcut akademisyenlerin

bilimsel ve teknolojik yönden ülkelerine katkılarının devam etmesini sağlayabilmek için atılması gereken adımların neler olması konusunda ilgili kurumlara katkı sağlayacağı düşünülmektedir. Ayrıca akademik performans üzerine literatürde sınırlı sayıda araştırma yapılmıştır. Bu çalışmanın literatürdeki söz konusu eksikliği giderme açısından, diğer çalışmalara da kaynaklık edebileceği düşünülmektedir.

Araştırmanın Amacı: Bu çalışmada, akademisyenlerin akademik performans algılarını ölçebilecek geçerli ve güvenilir bir ölçek geliştirilmesi amaçlanmıştır.

Araştırmanın Yöntemi: Bu çalışmanın ilk araştırma grubunu, Ankara Üniversitesi Eğitim Bilimleri Fakültesi'nde görev yapmakta ya da lisansüstü eğitim almakta olan 125 akademisyen; doğrulayıcı faktör analizi (DFA) yapılan ikinci araştırma grubunu ise, Erzincan Üniversitesi'nde görev yapmakta ya da lisansüstü eğitim almakta olan (1 çok değişkenli uçdeğer atıldıktan sonra) 147 akademisyen oluşturmaktadır. Ölçek geliştirilirken likert tipi ölçek geliştirme adımları dikkate alınarak çift olumsuzluk taşıyan, belirsizlik yaratan, anlatım bozukluğuna sebep olan ifadelerden kaçınılmıştır. Hazırlanan maddeler, "Öz-yeterlik" üzerine çalışmalarda bulunan iki uzman; üç Ölçme ve Değerlendirme Uzmanı; bir Bilgisayar Öğretimi ve Teknoloji Uzmanı; bir İngilizce Öğretmeni ve bir Türk Dili Uzmanı tarafından incelenmiş, dil, kapsam ve psikometrik açıdan değerlendirilmiştir. Görüşler doğrultusunda ifadeler gözden geçirilmiş ve 29'u olumlu; 11'i ise olumsuz olmak üzere toplamda 40 maddeden oluşan Akademik Performans Algı Ölçeği ön deneme uygulamasına hazır duruma getirilmiştir. Yazılan maddelerin uzman görüşüne sunulması kapsam geçerliği için kanıt olarak gösterilebilir. Ölçeğin deneme formunda yer alan 40 maddenin madde geçerliğini belirlemek amacıyla madde- toplam test korelasyonları için Pearson Çarpım Momentler Korelasyon Katsayısı hesaplanmıştır. Madde ayırt ediciliklerini belirlemek için ise, alt üst grup ortalamaları farkına dayalı madde analizi yöntemi kullanılmıştır. Alt üst grup ortalamaları arasındaki fark, bağımsız gruplar t testiyle incelenmiştir. Bunların yanı sıra madde geçerliğine kanıt olarak cinsiyet değişkenine göre değişen madde fonksiyonu (DMF) gösteren çok kategorili madde bulunup bulunmadığı incelemek için sıralı lojistik regresyon yöntemi kullanılmıştır. Akademik performans algı ölçeğinin yapı geçerliğini belirlemek amacıyla, ilk araştırma grubundaki verilere temel bileşenler analizi uygulanmıştır. Temel bileşenler analizi sonucunda elde edilen faktör yapısının doğruluğunu test etmek için, ikinci araştırma grubundan toplanan veriler üzerinde doğrulayıcı faktör analizi (CFA) yöntemi uygulanmıştır. Akademik performans algı ölçeğinin güvenirliğine ilişkin olarak ise Cronbach Alfa katsayısı hesaplanmıştır.

Araştırmanın Bulguları: Bu çalışmada, madde toplam test korelasyonlarının .18 ile .73 arasında değerler aldığı ve her maddenin ölçek puanlarıyla manidar ilişki gösterdiği sonucuna ulaşılmıştır ($p < .05$). Maddelerin, akademik performans algısına olumlu yönde sahip olanlarla, olumsuz yönde sahip olanları ayırt edip etmediğini saptamak için %27'lik alt ve üst gruplar için t değerleri hesaplandığında, t değerleri 1.66 ile 12.77 arasında değişmektedir. Dolayısıyla, akademik performans algı ölçeğinde yer alan 17, 24, 32 ve 33. maddeler dışındaki madde puanları, %27'lik alt ve üst gruba göre manidar bir farklılık göstermektedir ($p<.05$). Buna göre manidar olmayan maddeler ölçekten çıkarıldığında, ölçeğin alt ve üst grubu iyi ayırt ettiği söylenebilir.

Akademisyenlerin akademik performans algılarını ölçmeye yönelik geliştirilen ölçme aracında cinsiyet değişkenine göre DMF gösteren çok kategorili maddelerin bulunup bulunmadığına ilişkin sıralı lojistik regresyon yöntemi sonuçları incelendiğinde, cinsiyet değişkenine göre, 40 maddeden üçünde DMF bulunsa da üç madde (9., 10. ve 20. Madde) için de değişen madde fonksiyonlarının, A düzeyinde (ΔR^2<.13) bir başka ifadeyle ihmal/tolerans edilebilir düzeyde olduğu bulgusuna ulaşılmıştır. Dolayısıyla madde-toplam test korelasyonu ve %27'lik alt ve üst gruplar için t testi sonuçlarının yanı sıra ölçekte B ve C düzeyinde DMF gösteren maddenin bulunmaması madde geçerliğine kanıt olarak sunulabilir. Akademik performans algı ölçeğinin yapı geçerliğini belirlemek amacıyla yapılan temel bileşenler analizi sonucunda, (KMO=.84; χ2 (406, N=125)=1861.418, p<.01) ölçeğin beş faktör toplamda 26 maddeden oluştuğu saptanmıştır. Elde edilen 26 maddenin faktör yük değerleri .53 ile .87 arasında değişmekte ve bu beş faktör toplam varyansın %59.30'nu açıklamaktadır. Temel bileşenler analizi sonucunda beş faktörden oluşan 26 maddelik Akademik Performans Algı Ölçeği ikinci araştırma grubuna uygulanmıştır. Uygulama sonucunda elde edilen verilerin, beş faktörlü model ile uyum gösterip göstermediğini tespit etmek için doğrulayıcı faktör analizi yapılmıştır. Doğrulayıcı faktör analizi sonucu elde edilen uyum indeksleri (χ2/df = 2.04; NNFI=.90; CFI=.91; SRMR=.09) genel olarak değerlendirildiğinde, 26 maddelik Akademik Performans Algı Ölçeğinin beş faktörlü yapısının doğrulandığı sonucuna ulaşılmıştır. Bir başka ifadeyle, doğrulayıcı faktör analizi sonuçlarına göre, Akademik Performans Algı Ölçeğine ilişkin beş faktörlü modelin farklı araştırma grupları üzerinde kararlı sonuçlar verdiği şeklinde yorumlanabilir. Bu sonuçlar, yapı geçerliğine kanıt sağlamaktadır. İç tutarlılık katsayısı incelendiğinde ise (α=.88) ölçeği oluşturan maddeler birbirleriyle tutarlı olduğu görülmektedir.

Araştırmanın Sonuçları ve Önerileri: 19'u olumlu, 7'si ise olumsuz toplamda 26 maddeden oluşan ölçeğin asıl formunun, akademisyenlerin akademik performans algılarını belirlemek üzere kullanılabilecek güvenilir ve geçerli bir veri toplama aracı olduğu sonucuna ulaşılmıştır. Araştırmacılar için akademisyenlerin akademik teşvik puanları ile akademik performans algıları arasında manidar düzeyde ilişki bulunup bulunmadığına yönelik bir çalışma önerilebilir. Ayrıca, en iyi dünya üniversiteleri sıralamasına göre, ilk 500'de yer alan 5 üniversitede görev yapmakta olan akademisyenler ile ilk 500'de yer almayan 5 üniversitede görev yapmakta olan akademisyenleri, akademik performans algı ölçeği puanları hangi doğruluk düzeyinde sınıflandırdığına ilişkin çalışmalar yapılabilir. Kurumlara yönelik ise, YÖK tarafından akademisyenlerin akademik performans algı düzeyleri belirlenerek akademisyenlerin ihtiyaçları doğrultusunda hizmetiçi eğitimlerin verilmesi önerilebilir. Bu eğitimler kapsamında, teknolojide yaşanan gelişmeler; yabancı dil; diksiyon ve beden dili; makale yazımında dikkat edilecek hususlar; şehirlerin ihtiyacı olan akademik çalışmalar hakkında bilgilendirme; veri analizleri; akademik çalışmalarda yararlanılabilecek istatistik ve yazılım programları hakkında seminerler verilebilir.

Anahtar Kelimeler: Akademik çaba, öz-yeterlik, başarı, akademik yeterlik.

Eurasian Journal of Educational Research 69 (2017) 199-219

Eurasian Journal of Educational Research
www.ejer.com.tr

Novice Teachers' Opinions on Students' Disruptive Behaviours: A Case Study

Senol SEZER[1]

ARTICLE INFO

Article History:

Received: 20 November 2015

Received in revised form: 06 December 2017

Accepted: 17 April 2017

DOI: http://dx.doi.org/10.14689/ejer.2017.69.11

Keywords

novice teacher, disruptive behaviour, feelings, attitude

ABSTRACT

Purpose: It is recognised worldwide that novice teachers encounter various disruptive behaviours and face many challenges that stem from problematic students. Disruptive behaviours are seen as some of the most pervasive challenges widely affecting the teaching experience of novice teachers. In this study, the aim was to determine novice teachers' opinions related to students' disruptive behaviours in the classroom environment. **Research Methods:** This study was designed in a qualitative, case study pattern. The study group consists of 24 novice teachers working at public schools in Trabzon city centre. The participants were determined by using the criterion sampling method. Data was obtained with a semi-structured interview form. **Findings** Novice teachers frequently encounter disruptive behaviours. These disruptive behaviours include chattering, distraction, engaging with mobile phones, chewing gum, abusing classmates, and disrupting the class. Novice teachers express both positive and negative feelings. The positive feelings include paying closer attention and using interesting methods, whereas negative feelings include needing support from experienced teachers, feeling anxious and stressed, and thinking about leaving the school. They exhibit punitive attitudes to cope with disruptive behaviours as well as conciliatory attitudes. **Implications for Research and Practice**: The results show that classroom management training should be imparted based on theory and practice in education faculty. Furthermore, novice teachers should be supported by experienced teachers.

[1] Corresponding Author: Senol SEZER, Ordu University, Faculty of Education, senolsezer.28@gmail.com

Introduction

Novice teachers have various backgrounds, motivations, experiences, and preparation levels in their initial teaching experiences. The initial years of service are generally considered to be the first three years of teaching, and the first year is often difficult. It is seen that novice teachers encounter the greatest challenges due to inadequate classroom management training during the beginning years of their profession.

A novice teacher is defined as someone who is teaching something new for the first time (Farrel, 2012). Novice teachers can also be defined as teachers with less than three years teaching experience, and for whom the teaching model tends to focus on survival and establishing basic classroom routines (Huberman, 1993). There is no consensus on how many years of teaching are necessary to be an experienced teacher. In some studies, a novice teacher is defined as a teacher with less than five years of teaching experience (Kim & Roth, 2011), whereas others referred to the novice teacher as a teacher with two years of teaching experience or less (Haynes, 2011). In the current study, a novice teacher was defined as a teacher with two years or less of teaching experience.

Novice teachers enter classrooms with high expectations for themselves and for their students, but they often encounter different challenges (Toprakci, 2008). In the study conducted by Ozturk & Yildirim (2013), novice teachers expressed that classroom management was the most challenging part of the profession. In another study conducted by Yalcinkaya (2002), the challenges of novice teachers were perceived to be inexperience, conflicts between pre-service training and in-service applications, pressures on new teachers, obligations to finish more tasks, fear of inspection, and adaptation to the school and environment.

In many studies, disruptive behaviours were stated to negatively affect the teaching and learning environment (Gordon & Browne, 2004; Rigby, 2003; Rogers, 2011; Seeman, 2010). Similarly, in previous studies conducted by Achinstein (2006), Amada (1994), Day, Stobart, Sammons, & Kington (2006), Mabeba & Prinsloo (2000), Ulvik, Smith & Hellege (2009) the students' disruptive behaviours were shown as the greatest challenge for novice teachers. Furthermore, many modern-day schools consist of numerous children who are psychologically unstable or disturbed and feel discouraged at school (Aitken & Harford, 2011; Karatas & Karaman, 2013; Koca, 2016; Ozturk &Yildirim, 2013; Sari & Altun, 2015). These children display a lack of self-respect and act indifferently toward activities in the classroom (Brandon, 2003). Children in low-income families, in particular, tend to feel vulnerable, and often behave overly reactive and furiously due to the strict application in many prison-like schools (Cummings, 2000; Dudley-Marling, 2007; Ozturk, 2008; Smith & Laslett, 2002; Watts & Erevelles, 2004). Additionally, various disruptive behaviours stem from insufficient teacher qualifications, including boring teaching methods, disarranged learning environment, overreacting to disruptive behaviours, punishing, displaying stereotypical attitudes, and exhibiting burnout (Gorski, 2008; Lewis, 2009; Goldstein & Brooks, 2007; Okutan, 2005). Some disruptive behaviours arise from group

dynamics, including disapproval by peer group, bullying, teasing, and displaying hostile attitudes (Henley, 2009). Consequently, these behaviours are seen as the primary challenges for novice teachers at the beginning of their teaching profession (Dunbar 2004; Goldstein & Brooks, 2007; Shelton & Brownhill, 2008; Weiner, 2003).

Novice teachers frequently fail to cope with disruptive behaviours; therefore, their job stress exacerbates, and they experience low job satisfaction and high burnout (Berliner, 1986; Espin & Yell, 1994; Walters & Frei, 2007). If novice teachers cannot sufficiently ensure discipline and manage disruptive behaviours appropriately in the classroom, the job stress affects their overall health and emotional state (Wood & McCarthy, 2002). Furthermore, novice teachers' inadequate classroom management skills negatively influence their commitment to profession (Browers & Tomic, 2000; Ingersoll & Smith, 2003). Notwithstanding that existing limited studies on the challenges encountered by novice teachers at the beginning of the teaching profession, e.g. Erdemir (2007), Erkoc (2010), Karatas & Karaman (2013), Koca (2016), Korkmaz, Saban, & Akbaslı (2004) and Sarı & Altun (2015), said studies were conducted on disruptive student behaviours. Surveys on disruptive behaviours are usually focused on management skills of teachers in the classroom environment, e.g. Bernshausen & Cunningham (2001), Emmer & Stough, (2001), Greenhalgh (2001), Ozturk (2005), Porter (2006), Romano & Gibson (2006), Rose & Gallup (2005), Wong & Wong (1998), and Zuna & McDougall (2004). An important part of the studies is generally focused on effective teacher attitudes to cope with students' disruptive behaviours, e.g. Cinkir (2004), Mathieson & Price (2003), O'Brien (2012), Oliver, Wehby, & Reschly (2011) and Safran & Oswald (2003). Consequently, it is seen that, as the main challenges for novice teachers stem from students' disruptive behaviours, the present study was conducted on this topic.

In this study, we aimed to explore in detail the opinions expressed by novice teachers related to students' disruptive behaviours in the classroom environment. For this purpose, answers were sought the following questions:

1-How often do novice teachers encounter disruptive behaviours in the classroom environment?

2-What disruptive behaviours do novice teachers encounter in the classroom environment?

3-How do novice teachers feel when encountering disruptive behaviours?

4-What methods do novice teachers use to manage disruptive behaviours?

Method

Research Design

This study was conducted in a case study design, which is a qualitative research method. Qualitative research designs are used to obtain comprehensive knowledge about a topic (Denzin & Lincoln, 2005; Marshall & Rossman, 2006; Patton, 2014; Singh, 2007). The case study allows an investigation to retain the holistic and meaningful characteristics of real-life events, and this method enables researchers to closely examine data within a specific context (Fidel, 1984; Yin, 2003; Zainal, 2007).

Research Sample

The study group consisted of 24 novice teachers working in public schools in Trabzon city centre. The participants were determined by using the criterion sampling method. The criterion sampling method is used to determine a study group with sufficient knowledge and experiences about the researched topics (Patton, 2014, 238). To have 1-2 years of seniority was taken as a prerequisite criterion, and having encountered students' disruptive behaviours in classroom was another criterion. Additionally, having worked in schools with different training programs was also selected as a criterion for the triangulation among the novice teachers' opinions. Furthermore, having felt anxiety or distress due to students' disruptive behaviours was taken as another criterion. The study group consisted of 14 female (58.3%) and 10 male (41.7%) novice teachers. Six of them work in primary schools (25%), six in secondary schools (25%), six in Anatolian high schools (25%), and six in vocational high schools (25%). The average seniority was 1.7 years and the average age was 23.5.

Research Instruments and Procedures

The study was conducted in four stages: (i) definition of the problem, (ii) preparation of the data collection instrument, (iii) data collection, and (iv) data analysis and interpretation (Mayring, 2011, 112; Yildirim & Simsek, 2013, 93-97).

(i) Definition of the problem: During interviews with novice teachers, it was understood that they frequently encounter students' disruptive behaviours and experience anxiety. Moreover, the students' disruptive behaviours reduce the morale and motivation of the novice teachers. In the first stage, the problem was defined. Then, a conceptual framework was created to be able to classify and compare the disruptive behaviours examined in the study.

(ii) Preparation of the data collection instrument: A semi-structured interview form was used to collect data. The form consisted of two parts. In the first part, there were four questions determining participants' demographic characteristics. The second part consisted of four open-ended questions: (i) How often do you encounter disruptive behaviours in the classroom environment? (ii) What disruptive behaviours do you encounter in the classroom environment? (iii) How do you feel when you encounter disruptive behaviours? (iv) What methods do you use to manage disruptive behaviours?

(iii) Data collection: Each participant was called before the interview by the researcher for an appointment, and then the semi-structured interview form was administered. Each interview lasted about 30-45 minutes, and the novice teachers expressed their views in writing in response to the open-ended questions.

(iv) Data analysis and interpretation: Data was analysed by using the descriptive analysis method. This method consists of four stages: creating a thematic framework for analysis, processing data, identifying findings, and analysis and interpretation (Yildirim & Simsek, 2013, 256). Similar responses to the same questions were evaluated in the same category. The written data were coded, and 19 sub-themes

arose, identifying students' disruptive behaviours. Two main themes were determined to identify these sub-themes: (i) indifference and (ii) disruptiveness. Nineteen sub-themes were determined to identify the feelings of the novice teachers. Two main themes were determined to identify these sub-themes. The main themes were (i) positive feelings and (ii) negative feelings. Seventeen sub-themes were determined to be related to the novice teachers' methods to manage students' disruptive behaviours, and two main themes were determined to identify these criteria. The main themes were (i) conciliatory attitudes and (ii) punitive attitudes. The views on the sub-themes were sorted according to their frequency.

Validity and Reliability

Related literature was reviewed in-depth, and a conceptual framework was created on the subject to improve the internal validity of the study. The data were coded by the researcher and another scholar who with experience in qualitative research. In the first stage, the themes were kept large enough to cover the concepts, but narrow enough to exclude unrelated concepts. The formula "Reliability= Consensus/Consensus + Dissidence × 100" was applied to determine the reliability of the coding (Miles & Huberman, 1994, 64). The agreement between the two coders was calculated as 50/ (50+6) × 100= .89. In the second phase, structural integrity was achieved by controlling the relationship between the main themes and sub-themes as well as the relationship between each sub-theme and the others. The research process was explained in detail, and all evidence was shown without comment to improve the external validity.

Results

The frequency of disruptive behaviours is shown in Figure 1.

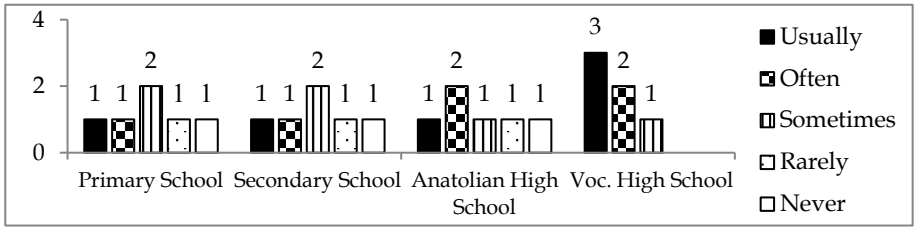


Figure 1. The frequencies of disruptive behaviours (η=24)

Disruptive Behaviours in Classroom

In this section, the novice teachers' views on disruptive behaviours are shown in two main themes: 'indifference' and 'disruptiveness'. The term participant is abbreviated as "P". It is impossible to state all the participants' opinions because of space limitations. As such, the more notable views have been included. The views related to indifference are as follows:

[P2] '*I frequently encounter students who want to go home'.*
[P7] '*I struggle to attract my students' attention to the subject'.*

[P8] 'The students do not prepare their tools and equipment (book, notebook, pencil, eraser, etc.) in my class.

[P13] '*The students talk a lot each other in my class.*

[P14] '*I frequently encounter students who complete homework during my class.*

[P16] '*Some of the students are playing computer games during my class.*

[P17] '*I encounter students eating snacks during my class.*

[P19] '*I encounter students escaping from my class to spend time at the internet café*'.

[P21] '*I encounter students being frequently late and usually sleeping during the first lesson*'.

[P23] '*I encounter students behaving in a way that is unrelated the topic and finding a pretext to participate*'.

[P24] '*I have to warn students not to engage with their mobile phones in almost all of my classes*'.

The frequency and percentages of sub-themes related to the main theme of indifference are presented in Table 1.

Table 1

Sub-Themes of Indifference

Sub-theme	*f*	*%*
Chattering	21	19.8
Distraction	19	17.9
Engaging with mobile phone	15	14.2
Tardiness	11	10.4
Bored by the lesson	10	9.4
Eating snacks	8	7.5
Not completing homework	7	6.6
Behaving in unrelated ways	6	5.7
Course material is unprepared	4	3.8
Playing computer games on tablet	3	2.8
Escaping from school	2	1.9
Total	106	100

According to Table 1, the students' indifference behaviours are chattering (f=21, 19.8%), distraction (f=19, 17.9%), engaging with mobile phone (f=15, 14.2%), tardiness (f=11, 10.4%), bored by the lesson (f=10, 9.4%), eating snacks (f=8, 7.5%), not completing homework (f=8, 6.6%), behaving in unrelated ways (f=6, 5.7%), course material is unprepared (f=4, 3.8%), playing computer game on tablet (f=3, 2.8%) and escaping from school (f=2, 1.9%). The views related to the 'disruptiveness' main theme are as follows:

[P10] '*The students use insulting language with each other*'.

[P11] '*I frequently encounter students trying to disrupt the class.*

[P12] '*I have to cope with problematic students causing distractions for classmates and hindering them from focusing on learning activities*'.

[P13] '*I encounter students mocking my teaching methods*'.
[P15] '*I struggle with students displaying rude attitudes*'.
[P17] '*I encounter students using swear words*'.
[P18] '*I encounter students attempting to cheat during my exams*'.
[P19] '*I struggle with students chewing gum*'.
[P24] '*The students verbally abuse their classmates*'.

The frequency and percentages of sub-themes related to the main theme of disruptiveness are presented in Table 2.

Table 2

Sub-themes of Disruptiveness (η=24)

Sub-theme	*f*	*%*
Chewing gum during lesson	20	23.3
Verbally abuse classmates	17	19.8
Trying to disrupt the course	15	17.4
Displaying rude attitude	11	12.8
Using insulting language	9	10.5
Attempting to cheat	7	8.1
Using swear words	5	5.8
Mocking teaching methods	2	2.3
Total	86	100

In Table 2, the students' disruptive behaviours in disruptiveness are chewing gum during lesson (f=20, 23.3%), verbally abusing classmates (f=17, 19.8%), trying to disrupt the course (f=15, 17.4%), displaying rude attitude (f=11, 12.8%), using insulting language (f=9, 10.5%), attempt to cheat (f=7, 8.1%), using swear words (f=5, 5.8%), and mocking the teaching methods (f=2, 2.3%).

The Novice Teachers' Feelings

In this section, the novice teachers' feelings on students' disruptive behaviours are evaluated in two main themes: 'positive feelings' and 'negative feelings'. The views related to the 'positive feelings' are as follows:

[P1] '*I think I should pay closer attention to my pupils*'.
[P2] '*I think that I need to use more interesting methods and ensure that the materials attract students' attention*'.
[P3] '*I think I need support from experienced teachers'.*
[P5] '*I feel that I need to improve my teaching skills'.*
[P9] '*I think that I should be in more collaboration with the parents'.*
[P10] '*I think that I need to be a role-model to my students in terms of speaking in a polite manner*'.
[P14] '*I think that I should lead them to do their homework before they come in my class'.*
[P16] '*I think that I need to motivate all my students, as well'.*
[P22] '*I feel I have to do my best to teach my pupils'.*

[P23] *'I think that the teacher candidates need to have a training based on theory and practice in education faculty'.*

The frequency and percentages of sub-themes related to the main theme of positive feelings are presented in Table 3.

Table 3

Sub-themes of Positive Feelings (η=24)

Sub-theme	*f*	*%*
Paying closer attention to the students	20	20.4
Using more interesting methods	17	17.3
Using support from experienced teachers	15	15.3
Improving teaching skills	12	12.2
Effective teacher training	11	11.2
Motivating students	10	10.2
Collaboration with the parents	7	7.2
Leading students	4	4.1
Being a role model	2	2.1
Total	98	100

According to Table 3, the positive feelings of novice teachers are paying closer attention to the students (f=20, 20.4%), using more attractive methods (f=17, 17.3%), utilising support from experienced teachers (f=15, 15.3%), improving teaching skills (f=12, 12.2%), effective teacher training (f=11, 11.2%), motivating students (f=10, 10.2%), collaboration with the parents (f=7, 7.2%), leading students (f=4, 4.1%), and being a role model (f=2, 2.1%). The views related to the 'negative feelings' main theme are as follows:

[P7] '*I feel that I am insufficient*'.

[P11] *'I feel stressed'.*

[P12] *'I think that I need to leave this school to continue my profession'.*

[P13] *'When the students sarcastically criticise my teaching methods, I feel anxious'.*

[P15] *'When I unsuccessfully cope with disruptive behaviours, I think that I must resign'.*

[P17] *'When I unsuccessfully cope with the students swearing, I feel helpless'.*

[P18] *'I think that I could not continue long term in the teaching profession'.*

[P19] *'Because of the students disregarding my warnings, I feel as if I have lost my self-respect'.*

[P20] *'When the students attempt to disrupt my concentration and the course flow, I feel unsuccessful'.*

[P24] *'If the students are in an especially disregardful mood, I feel angry'.*

The frequency and percentages of sub-themes related to the main theme of negative feelings are presented in Table 4.

Table 4

Sub-Themes of Negative Feelings (η=24)

Sub-theme	*f*	*%*
Feeling anxious	22	21.8
Feeling stressed	18	17.8
Thinking about leaving the school	15	14.9
Feeling angry	13	12.9
Feeling unsuccessful	11	10.9
Feeling helpless	8	7.8
Feeling insufficient	6	5.9
Regretting teaching profession	4	4.0
Thinking about resignation	3	3.0
Losing self-respect	1	1.0
Total	101	100

According to Table 4, the negative feelings of novice teachers are feeling anxious (f=22, 21.8%), feeling stressed (f=18, 17.8%), thinking about leaving the school/profession (f=15, 14.9%), feeling angry (f=13, 12.9%), feeling unsuccessful (f=11, 10.9%), feeling helpless (f=8, 7.8%), feeling insufficient (f=6, 5.9%), regretting teaching profession (f=4, 4%), thinking about resignation (f=3, 3%), and losing self-respect (f=1, 1%).

Disruptive Behaviour Management

In this section, the novice teachers' disruptive behaviour management methods are evaluated under two main themes: conciliatory attitudes and punitive attitudes. The views related to the conciliatory attitudes are as follows:

[P1] *'I pay close attention to student exhibiting disruptive behaviours'.*

[P3*] 'I try to benefit from the experienced teachers' suggestions'.*

[P5] *'I act in a tolerant manner'.*

[P7] *'I talk to the problematic student'.*

[P8] *'I clarify the student's disruptive behaviours'.*

[P9] *'I consult with parents'.*

[P13] *'I give information to the school counsellor about the student's disruptive behaviours'.*

[P16] *'I utilise support from an experienced teacher'.*

[P20] *'I give extra responsibilities'.*

[P22] *'I explain the classroom rules and principles at the beginning of the academic year, and I ensure behaviour compliance with them'.*

The frequency and percentages of sub-themes related to the main theme of conciliatory attitude are presented in Table 5.

Table 5

Sub-themes of Conciliatory Attitudes (η=24)

Sub-theme	*f*	*%*
Acting in a tolerant manner	21	21.9
Trying to understand the problem	15	15.6
Setting classroom rules	13	13.5
Using support from experienced teachers	11	11.5
Talking with problematic student	10	10.4
Ignoring student's disruptive behaviour	9	9.4
Paying close attention	7	7.3
Consulting parents	5	5.2
Informing school counsellor	4	4.2
Giving extra responsibilities	1	1.0
Total	96	100

In Table 5, the conciliatory attitudes of novice teachers are acting in a tolerant manner (f=21, 21.9%), trying to understand the problem (f=15, 15.6%), setting classroom rules (f=13, 13.5%), using support from experienced teachers (f=11, 11.5%), talking to problematic student (f=10, 10.4%), ignoring student's disruptive behaviour (f=9, 9.4%), paying close attention (f=7, 7.3%), consulting parents (f=5, 5.2%), informing school counsellor (f=4, 4.2%), and giving extra responsibilities (f=1, 1%). The views related to the punitive attitudes main theme are as follows:

[P11] *'I complain to the parents about the student's disruptive behaviours'.*

[P14] *'I give information to the school administrator'.*

[P15] *'I report them to the school disciplinary board'.*

[P17] *'I discreetly handle the students who persist with disruptive behaviours'.*

[P18] *'I act more decisively'.*

[P19] *'I impose punishments'.*

[P21] *'I warn sharply'.*

[P23] *'I refer to the disciplinary board'.*

[P24] *'I deliberately give extra homework to the students disrupting the course'.*

The frequency and percentages of sub-themes related to the main theme of punitive attitude are presented in Table 6.

Table 6

Sub-themes of Punitive Attitude (η=24)

Sub-theme	*f*	*%*
Giving information to the school administrator	21	22.8
Reporting to disciplinary board	18	19.6
Complaining to parents	15	16.3
Impose punishments	13	14.1
Giving extra homework	10	10.9
Warning sharply	8	8.7
Acting decisively	5	5.4
Handling discreetly	2	2.2
Total	92	100

In Table 6, the punitive attitudes of novice teachers are giving information to the school administrator (f=21, 22.8%), reporting to disciplinary board (f=18, 19.6%), complaining to parents (f=15, 16.3%), impose punishments (f=13, 14.1%), giving extra homework (f=10, 10.9%), warning sharply (f=8, 8.7%), acting decisively (f=5, 5.4%), and handling discreetly (f=2, 2.2%).

Discussion and Conclusions

The present study aimed to explore in detail the opinions of novice teachers related to students' disruptive behaviours. The findings show that novice teachers frequently encounter disruptive behaviours in the classroom environment. Similarly, in studies conducted by Guclu (2004) and Ogel, Tari, & Yilmazcetin-Eke (2006), the beginner teachers working at high schools frequently encountered students' disruptive behaviours. Additionally, Goldstein & Brooks (2007) stated that adolescents are in troubled mood during their later stages of the educational process, which can increase the ratio of the disruptive student behaviours.

It is commonly stated by the novice teachers that the students behave indifferently in the classroom. Novice teachers encounter disruptive behaviours such as chattering, distraction, engaging with mobile phones, tardiness, being bored by the lesson, eating snacks, not completing homework, behaving in an unrelated way, and playing computer games. Comparable results are seen in previous studies conducted by Akpinar & Ozdas (2013), Gordon & Browne (2004), and Siyez (2009).

In addition, novice teachers encounter the students behaving in a disruptive context. These disruptive behaviours are chewing gum during course, abusing classmates verbally, trying to disrupt the class, displaying a rude attitude, using insulting language, attempt to cheat, and using swear words. Similarly, in studies conducted by Melnick & Meister (2008), Nguyen (2013), and O'Brien & Christie

(2005), teachers with three years or less of seniority encountered disruptive behaviours more than twice as likely as experienced teachers. Likewise, in the study conducted by Ulvik, Smith, & Helleve (2009), beginning teachers encountered more disruptive behaviours than experienced teachers.

Novice teachers have positive and negative feelings on students' disruptive behaviours. Positive feelings include paying closer attention to students, using more interesting methods, utilising support from experienced teachers, improving teaching skills, effective teacher training, motivating students, collaborating with the parents, leading students, and being a role model. These feelings are corroborated in previous studies conducted by Daly, Witt, Martens, & Dool (1997), Dudley-Marling (2007), and Seeman (2010). Similar results were attained in previous studies conducted by Bomer, Dworin, May, & Semingson (2008), Marais & Meier (2010), and Scherff (2007).

Negative feelings include feeling anxious, feeling stressed, thinking about leaving the school, feeling angry, feeling unsuccessful, feeling helpless, feeling insufficient, regretting teaching profession, thinking about resigning, and losing self-respect. Similar findings are seen in previous studies conducted by Eren (2012), Erturk & Kececioglu (2012), Yilmaz-Toplu (2012), and Yuksel & Yuksel (2014).

The results show that the novice teachers display conciliatory attitudes to manage students' disruptive behaviours. Novice teachers' conciliatory attitudes include acting tolerantly, trying to understand the problem, setting classroom rules, utilising support from an experienced teacher, talking to problematic student, ignoring student's disruptive behaviour, paying close attention, consulting with parents, informing school counsellor, and giving extra responsibilities. Similar findings are seen in previous studies conducted by Forbes (2004) and Little (2005).

Novice teachers also display punitive attitudes, including giving information to the school administrator, reporting to the school disciplinary board, complaining to the parent, imposing punishments, giving extra homework, warning sharply, acting decisively, and handling discreetly. Similar results are seen in previous studies conducted by Aydin (2014) and Oliver & Reschly (2007). In the mentioned studies, novice teachers expressed being overwhelmed due to the bullying and disrespectful behaviours of the students and that they prefer acting more reactively and using punitive methods.

As a result, novice teachers are often forced to cope with the students' disruptive behaviours. The findings show that novice teachers struggle with disruptive behaviours in the classroom, and that the novice teachers are adversely affected due to the students' disruptive behaviours. The novice teachers should be supported by experienced teachers. Additionally, peer coaching can be effective to improve the classroom management skills of novice teachers. Providing efficient peer coaching in accordance with their needs can reduce the problems that novice teachers experience when coping with students' disruptive behaviours in the classroom environment.

Based on the findings, it can be suggested that classroom management training should be given based on theory and practice. In order to efficiently prepare candidate teachers for the profession, alternative family structures should be analysed and student characteristics should be considered by teacher training institutions. School administrators should exhibit conciliatory attitudes toward novice teachers and, thus, said attitudes should be displayed toward fewer problematic students. Effective school-parent collaboration can be used to prevent students' disruptive behaviours. The reasons behind students' disruptive behaviours can be examined as another topic. Further research can be conducted on the ideal school and school happiness.

References

Achinstein, B. (2006). New teacher and mentor political literacy: Reading, navigating and transforming induction contexts. *Teachers and Teaching: Theory and Practice, 12*(2), 123-138. DOI: 10.1080/13450600500467290.

Aitken, R., & Harford, J. (2011). Induction needs of a group of teachers at different career stages in a school in the Republic of Ireland: Challenges and expectations. *Teaching and Teacher Education, 27*(2), 350-356. DOI: 10.1016/j.tate.2010.09.003.

Akpinar, B., & Ozdas, F. (2013). Lise disiplin sorunlarinin sinif degiskeni acisindan incelenmesi. [Investigation the high school discipline problems in terms of gender variable]. *Kastamonu Egitim Dergisi, 21*(4) (Ozel Sayi), 1365-1376.

Amada, G. (1994). *Coping with the disruptive college student: A practical model.* Asheville, NC: College Administration Publications.

Aydin, A. (2014). *Sinif yonetimi*. [Classroom management]. Ankara: PegemA Yayincilik.

Berliner, D. C. (1986). In pursuit of expert pedagogue. *Educational Researcher. 15*(7), 5-13.

Bernshausen, D., & Cunningham, C. (2001). The role of resiliency in teacher preparation and retention. *53rd Annual Meeting of the American Association of Colleges for Teacher Education,* Dallas, Texas.

Bomer, R., Dworin, J., May, L., & Semingson, P. (2008). Miseducating teachers about the poor: A critical analysis of Ruby Payne's claims about poverty. *Teachers College Record, 110*(12), 2497-2531.

Brandon, W.W. (2003). Toward a white teachers' guide to playing fair: Exploring the cultural politics of multicultural teaching. *Qualitative Studies in Education, 16*(1), 31-50.

Browers, A. & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education, 16*(2), 139-253.

Cummings, C. B. (2000). *Winning strategies for classroom management*. Virginia: ASCD Press.

Cinkir, S. (2004). Okulda etkili ogrenci-ogretmen iliskisinin yonetimi. [Effective management of student-teacher relationship in school]. *Milli Egitim Dergisi,* 161.

Daly, E. J., Witt, J. C., Martens, B. K., & Dool, E. J. (1997). A model for conducting a functional analysis of academic performance problems. *Social Psychology Review, 26*(4), 554-574.

Day, C., Stobart, G., Sammons, P., & Kington, A. (2006). Variations in the work and lives of teachers: relative and relational effectiveness. *Teacher and Teaching: Theory and Practice, 12*(2), 169-192. DOI: 10.1080/13450600500467381.

Denzin, N. K., & Lincoln, Y. S. (2005). *The sage handbook of qualitative research.* California: Sage Publications.

Dudley-Marling, C. (2007). Return of the deficit. *Journal of Educational Controversy, 2*(1), 1-12.

Dunbar, C. (2004). *Best practices in classroom management.* USA: Michigan State University Press.

Emmer, E. T., & Stough, L. M. (2001). Classroom management: A critical part of educational psychology with implications for teacher education. *Educational Psychologist, 36*(2), 103-112.

Erdemir, N. (2007). Meslege yeni baslayan fen bilgisi ogretmenlerinin karsilastiklari sorunlar ve sikayetleri. [The problems and complaints faced by science teachers who are new in their profession]. *Elektronik Sosyal Bilimler Dergisi, 6*(22), 135-149. Retrieved March 17, 2017, from http://www.esosder.com/?sayfa=dergilist&sayi=22

Eren, A. (2012). Ogretmen adaylarinin mesleki yönelimi, kariyer gelistirme arzulari ve kariyer secim memnuniyeti. [Prospective teachers' professional engagement, career development aspirations, and career choice satisfaction]. *Kastamonu Egitim Dergisi, 20*(3), 807-826.

Erkoc, A. (2010). *Aday ogretmenlerin goreve basladiklarinda karsilastiklari sorunlar. [The problems candidate teachers faced when they start their profession].* (Unpublished master thesis), Sutcu Imam Universitesi, Sosyal Bilimler Enstitüsü, Kahramanmaras.

Erturk, E., & Kececioglu, T. (2012). Calisanların is doyumlari ile mesleki tükenmislik duzeyleri arasindaki iliskiler: Ogretmenler üzerine ornek bir uygulama. [Relations between the levels of employees job satisfaction and burnout: A sample application on the teachers]. *Ege Akademik Bakis, 12*(1), 39-52.

Espin, C. A., & Yell, M. L. (1994). Critical indicators of effective teaching for pre-service teachers: Relationships between teaching behaviours and ratings of effectiveness. *Teacher Education and Special Education, 17,* 154-169.

Farrell, T. S. C. (2012). Novice-service language teacher development: Bridging the gap between preservice and in-service education and development. *TESOL Quarterly, 46*(3), 435-449.

Fidel, R. (1984). *The case study method: A case study.* Retrieved April, 14, 2017, from http://s3.amazonaws.com/academia.edu.documents/33415874/study_case.pdf?AWSAccessKeyId

Forbes, C.T. (2004), Peer mentoring in the development of beginning secondary science teachers: Three case studies. *Mentoring & Tutoring: Partnership in Learning, 12*(2), 219-239.

Goldstein, S., & Brooks, R. B. (2007). *Understanding and managing children's classroom behaviour.* New Jersey: John Wiley & Sons.

Gordon, A. M., & Browne, K. W. (2004). *Beginnings and beyond: Foundations in early childhood education.* Australia: Delmar Learning.

Gorski, P. C. (2008). Peddling poverty for profit: Elements of oppression in Ruby Payne's framework. *Equity & Excellence in Education, 41*(1), 130-148.

Greenhalgh, P. (2001). Ingredients of effective practice with pupils who have emotional and behavioural difficulties. In J. Wisser., H. Daniels & T. Cole (Eds.). *Emotional and behavioural difficulties in mainstream schools.* (pp.13-28). Oxford: Elsevier.

Guclu, M. (2004). Ortaogretim kurumlarinda disiplin cezasi alan ogrencilerin sosyo-ekonomik yonden incelenmesi (Kayseri Ornegi). [Examination of the students' in socio-economic aspects who receive disciplinary punishment in secondary schools (Kayseri Sample)] (Unpublished master thesis). Erciyes Universitesi, Kayseri.

Haynes, L. (2011). Novice teachers' perceptions of their mentoring experiences. Retrieved March 10, 2017, from http://search.proquest.com/docview/922678446?accountid=13014

Henley, M. (2009). *Classroom management: A proactive approach.* Cambridge: Pearson Publishing.

Huberman, M. (1993). *The lives of teachers*. New York: Teachers College Press.

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

Ingersoll, R. M., & Smith, T. M. (2003). The wrong solution to the teacher shortage. *Educational Leadership, 60*(8), 30-33.

Karatas, P., & Karaman, A. C. (2013). Challenges faced by novice language teachers: Support, identity, and pedagogy in the initial years of teaching. *The International Journal of Research in Teacher Education, 4*(3), 10-23.

Kim, K., & Roth, G. (2011). Novice teachers and their acquisition of work-related information. *Current Issues in Education, 14*(1) Retrieved February, 11, 2017, from https://cie.asu.edu/ojs/index.php/cieatasu/article/view/268/152

Koca, A. (2016). Problems of novice teachers: Challenges vs. support. *Journal of Education in Black Sea Region, 1*(2), 92-101.

Korkmaz, İ., Saban, A., & Akbasli, S. (2004). Goreve yeni baslayan sinif ogretmenlerinin karsilastiklari guclukler. [Professional challenges encountered by beginning classroom teachers]. *Kuram ve Uygulamada Egitim Yonetimi, 10*(2), 266-277.

Lewis, R. (2009). *Understanding pupil behaviour: Classroom management techniques for teachers.* London: Taylor & Francis Group.

Little, P. F. B. (2005). Peer coaching as a support to collaborative teaching. *Mentoring & Tutoring: Partnership in Learning, 13*(1), 83-94.

Mabeba, M. Z., & Prinsloo, E. (2000). Perceptions of discipline and ensuing discipline problems in secondary education. *South African Journal of Education, 20,* 34-41.

Marais, P., & Meier, C. (2010). Disruptive behaviour in the foundation phase of schooling. *South African Journal of Education, 30,* 41-57.

Marshall, C., & Rosman, B. G. (2006). *Designing qualitative research.* California: Sage Publications.

Mathieson, K., & Price, M. (2003). *Better behaviour in classrooms.* London: Taylor & Francis Group.

Mayring, P. (2011). *Nitel sosyal arastirmaya giris.* [Introduction to qualitative social research]. (Trans. A. Gumus, & M. S. Durgun). Ankara: Bilgesu.

Melnick, S., & Meister, D. (2008). A comparison of beginning and experienced teachers' concerns. *Educational Research Quarterly, 31*(3), 39-56.

Nguyen, H. T. (2013). Peer Mentoring: A Way Forward for Supporting Preservice EFL Teachers Psychosocially During the Practicum. *Australian Journal of Teacher Education, 38*(7), 30-44.

O'Brien, J., & Christie, F. (2005). Characteristics of support for beginning teachers: evidence from the new teacher induction scheme in Scotland. *Mentoring & Tutoring: Partnership in Learning, 13*(2), 189-203.

O'Brien, T. (2012). *Promoting positive behaviour.* New York: Routledge.

Okutan, M. (2005). Siniftaki olaylari ve ogrenci davranıslarini dogru anlamak. [To understand the events and student behaviors correctly]. In Emin Karip (Ed.). *Sinif yönetimi*. (pp.155-159). Ankara: PegemA Yayincilik.

Oliver, R. M., & Reschly, D. J. (2007). *Effective classroom management.* Washingthon, DC: National Comprehensive Center for Teacher Quality.

Oliver, R. M., Wehby, J. H., & Reschly, D. J. (2011). *Teacher classroom management practices: Effects on disruptive or aggressive student behaviour.* Oslo: The Campbell Collaboration.

Ogel, K., Tari, I., & Yilmazcetin-Eke, C. (2006). *Okullarda suc ve siddeti onleme.* [Prevention of the crime and violence in schools]. Istanbul: Yeniden Yayinlari.

Ozturk, B. (2005). Sinifta istenmeyen davranislarin onlenmesi ve giderilmesi. [Prevention and elimination of disruptive behaviors in classroom]. In Emin Karip (Ed.). *Sınıf yönetimi*. (pp.5-13). Ankara: PegemA Yayincilik.

Ozturk, M. (2008). *Induction into teaching: Adaptation challenges of novice teachers.* (Unpublished master thesis). Middle East Technical University, Ankara.

Ozturk, M., & Yildirim, A. (2013). Adaptation challenges of novice teachers. *Hacettepe Universitesi Egitim Fakultesi Dergisi, 28*(1), 294-307.

Patton, M. Q. (2014). *Nitel arastirma ve degerlendirme yontemleri*. [Qualitative research and evaluation methods]. (Trans. Eds. M. Butun & S. B. Demir). Ankara: Pegem Akademi.

Porter, L. (2006). *Behaviour in schools: Theory and practice for teachers*. England: Open University Press.

Rigby, K. (2003). *Stop the bullying: A handbook for schools*. Melbourne: ACER Press.

Rogers, B. (2011). *You know the fair rule*. Melbourne: ACER Press.

Romano, M., & Gibson, P. (2006). Beginning teacher successes and struggles: An elementary teacher's reflections on the first year of teaching. *The Professional Educator, 28*(1), 1-26.

Rose, L. C., & Gallup, A. M. (2005). The 37th annual Phi Delta Kappa/Gallup poll of the public's attitudes toward the public schools. *Phi Delta Kappan, 87*(1), 41-57.

Safran, S. P., & Oswald, K. (2003). Positive behaviour supports: Can schools reshape disciplinary practices? *Exceptional Children, 69*(3), 361-373.

Sari, M. H., & Altun, Y. (2015). Goreve yeni baslayan sinif ogretmenlerinin karsilastiklari sorunlar. [Problems faced by beginning primary education teachers]. *Hacettepe Universitesi Egitim Fakultesi Dergisi, 30*(1), 213-226.

Scherff, L. (2008). Disavowed: The stories of two novice teachers. *Teaching and Teacher Education, 24*(5) 1317-1332.

Seeman, H. (2010). *Preventing disruptive behavior in colleges*. Maryland: Rowman & Littlefield Publishers.

Shelton, F., & Brownhill. S. (2008). *Effective behaviour management in the primary classroom.* London: Open University Press.

Singh, K. (2007). *Quantitative social research methods.* New Delhi: Sage Publications.

Siyez, D. M. (2009). Liselerde gorev yapan ogretmenlerin istenmeyen ogrenci davranislarina yonelik algilari ve tepkileri. [High school teachers' perceptions of and reactions towards the unwanted student behaviors]. *Pamukkale Universitesi Egitim Fakultesi Dergisi, 25*(1), 67-80.

Smith, C. J., & Laslett, R. (2002). *Effective classroom management*. London: Taylor & Francis Group.

Toprakci, E. (2008). *Sinifa dayali yonetim*. [Classroom-based management]. Ankara: PegemA Yayincilik.

Ulvik, M., Smith, K., & Helleve, I. (2009). Novice in secondary school-the coin has two sides. *Teaching and Teacher Education, 25*(6), 835–842.

Walters, J., & Frei, S. (2007). *Managing classroom behavior and discipline*. M. E. Kessler (Ed.). CA: Shell Education.

Watts, I. E., & Erevelles, N. (2004). These deadly times: Reconceptualising school violence by using critical race theory and disability studies. *American Educational Research Journal, 20*(41), 271-299.

Weiner, L. (2003). Why is classroom management so vexing to urban teachers? *Theory into Practice, 42*(4), 305-312.

Wong, H. K., & Wong, R. T. (1998). *How to be an effective teacher: The first days of school.* CA: Harry K. Wong Publications.

Wood, T., & McCarthy, C. (2002). *Understanding and preventing teacher burnout.* Washington, DC: ERIC Clearinghouse on Teaching and Teacher Education.

Yalcinkaya, M. (2002). Yeni ogretmen ve teftis. [New teacher and inspection]. *Milli Egitim Dergisi, 150*, 153-154. Retrieved January, 8, 2017, from http://yayim.meb.gov.tr/dergiler/153-154/yalcinkaya.htm

Yildirim, A., & Simsek, H. (2013). *Sosyal bilimlerde nitel arastırma yontemleri.* [Qualitative research methods in the social sciences]. Ankara: Seckin Yayinevi.

Yilmaz-Toplu, N. (2012). *Okul oncesi ve ilkogretim ogretmenlerinin tukenmislik düzeyleri. [Burnout levels of preschool and primary school teachers].* (Unpublished master thesis). Adnan Menderes Universitesi. Sosyal Bilimler Enstitüsü, Egitim Bilimleri Anabilim Dali, Aydin.

Yin, R. K. (2003). *Case study research: Design and methods.* (Second Edition). London: Sage Publications.

Yuksel, H., & Yuksel, M. (2014). Isdoyumu ile isten ayrilma duzeyi arasindaki iliski: Ilkogretim ogretmenleri uzerinde bir uygulama. [The correlation between job satisfaction and job turnover: A case study on the secondary school teachers]. *Uluslararasi Sosyal Arastırmalar Dergisi, 7*(32), 559-572. Retrieved from http://www.sosyalarastirmalar.com/cilt7/sayi32_pdf/7egitim/yuksel_hasan_meldayuksel.pdf

Zainal, Z. (2007). *Case study as a research method*. Retrieved April, 14, 2017, from www.researchgate.net/publication/41822817

Zuna, N., & McDougall, D. (2004). Using positive behavioural support to manage avoidance of academic tasks. *Teaching Exceptional Children, 37*(1). 18-24.

Yeni Başlayan Öğretmenlerin Olumsuz Öğrenci Davranışlarına İlişkin Görüşleri: Bir Durum Çalışması

Atıf:

Sezer, S. (2017). Novice teachers' opinions on students' disruptive behaviours: A case study. *Eurasian Journal of Educational Research*, *69*, 199-219. http://dx.doi.org/10.14689/ejer.2017.69.11

Özet

Problem Durumu: Yeni başlayan öğretmen, ilk kez yeni şeyler öğreten kişi olarak tanımlanır. Yeni başlayan öğretmen, aynı zamanda üç yıldan az tecrübeye sahip olan ve öğretim modeli günü kurtarmaya ve temel sınıf rutinlerini oluşturmaya dayalı öğretmen olarak tanımlanabilir. Bu öğretmenler kendileri ve öğrencileri için yüksek beklentilerle sınıflara girer, ancak çoğu kez farklı zorluklarla karşılaşır.

Birçok çalışmada olumsuz davranışların öğrenme ve öğretme ortamını olumsuz etkilediği belirtilmektedir. Bu nedenle sınıf ortamındaki olumsuz öğrenci davranışları, yeni başlayan öğretmenler açısından en büyük zorluk olarak görülmektedir. Bu öğretmenler olumsuz davranışlarla sıklıkla başa çıkamazlar, iş stresi artar ve bu nedenle düşük iş doyumu ve yüksek tükenmişlik yaşarlar. Göreve yeni başlayan öğretmenlerin yaşadığı zorluklara ilişkin yapılan çalışmalar sınırlı olmakla birlikte çalışmaların çoğu sınıf ortamında karşılaşılan olumsuz öğrenci davranışlarını konu almaktadır. Öte yandan yapılan araştırmalar büyük oranda sınıf içi etkili öğretmen davranışlarına odaklanmaktadır.

Araştırmanın Amacı: Bu çalışma, sınıf ortamında öğrencilerin yıkıcı davranışlarıyla ilgili acemi öğretmenlerin görüşlerini ayrıntılı olarak araştırmayı amaçlamaktadır. Bu amaçla aşağıdaki sorulara yanıt aranmaktadır:

1. Yeni başlayan öğretmenler sınıf ortamında olumsuz davranışlarla hangi sıklıkta karşılaşmaktadır?
2. Yeni başlayan öğretmenlerin karşılaştıkları olumsuz davranışlar nelerdir?
3. Acemi öğretmenler olumsuz davranışlarla karşılaştıklarında nasıl hissetmektedir?
4. Yeni başlayan öğretmenler olumsuz davranışları yönetmek için hangi yöntemleri kullanmaktadır?

Araştırmanın Yöntemi: Nitel araştırma yöntemlerinden durum çalışması deseninde yürütülen araştırmanın çalışma grubunu, Trabzon il merkezinde yer alan devlet okullarda görev yapan ve 1-2 yıl meslek deneyimine sahip 24 öğretmen oluşturmaktadır. Araştırmanın çalışma grubu, ölçüt örneklem yöntemiyle belirlenmiştir.

Veri Analizi: Araştırmanın verileri araştırmacı tarafından geliştirilen yarı yapılandırılmış görüşme formu ile elde edilmiştir. Katılımcılar, görüşlerini form üzerine yazılı olarak ifade etmiştir. Veriler, betimsel analiz yöntemi ile analiz edilmiştir. Katılımcıların aynı soruya verdiği benzer yanıtlar, aynı kategoride değerlendirilmiştir. Yazılı veriler kodlanmış ve öğrencilerin olumsuz davranışlarını betimlemeyen 19 alt tema oluşturulmuştur. Bu alt temaları tanımlamak için iki ana tema belirlenmiştir: Ana temalar, (i) kayıtsızlık ve (ii) bozuculuk şeklindedir. Yeni başlayan öğretmenlerin duygularını betimleyen 19 alt tema tespit edilmiştir. Bu alt temaları tanımlamak için iki ana tema belirlenmiştir. Ana temalar (i) olumlu duygular ve (ii) olumsuz duygular şeklindedir. Yeni başlayan öğretmenlerin olumsuz öğrenci davranışlarını yönetmede başvurduğu yöntemleri betimleyen on sekiz alt tema ve bu alt temalara ilişkin iki ana tema belirlenmiştir. Ana temalar (i) uzlaşmacı tutum (ii) cezalandırıcı tutum şeklindedir. Alt temalarda yer alan görüşler, tablo üzerinde frekans ve yüzde değerlerine göre sıralanmıştır.

Araştırmanın Bulguları: Araştırmadan elde edilen bulgular, yeni başlayan öğretmenlerin sınıf ortamında olumsuz öğrenci davranışları ile sıklıkla karşılaştıklarını göstermektedir. Bu öğretmenler, sınıf ortamında gevezelik, dikkat dağınıklığı, cep telefonu ile uğraşma, ders esnasında sakız çiğneme, sınıf arkadaşını istismar etme, dersi kesintiye uğratmaya çalışma şeklindeki olumsuz öğrenci davranışları ile karşılaşmaktadır. Yeni başlayan öğretmenler, olumsuz öğrenci davranışlarını yönetmede öğrencilerle daha yakından ilgilenmekte, daha ilginç yöntemler kullanmakta ve deneyimli öğretmenlerden destek almaktadır. Olumsuz davranışlar, yeni başlayan öğretmenlerin kaygılar yaşamasına, stresli hissetmesine ve okuldan ayrılmayı düşünmesine neden olmaktadır. Yeni başlayan öğretmenler, olumsuz öğrenci davranışları ile başa çıkmak için cezalandırıcı tutum ve uzlaşmacı tavır sergilemektedir.

Araştırmanın Sonuçları ve Öneriler: Göreve yeni başlayan öğretmenler, sınıf ortamında olumsuz öğrenci davranışları ile sıklıkla karşılaşmaktadır. Öğrenciler derslere ilgisiz davranmakta, saldırgan ve müdahaleci bir tutum sergilemekte, öğretim yöntemini sıkıcı bulmakta ve teknolojik araçları ders ortamında amaç dışı kullanmaktadır. Bu bağlamda öğrencilerin etkili öğrenme ve olumlu davranış geliştirmeleri için

izlenmesi gereken yeni eğitim politikaları araştırılabilir. Yeni öğretmenler, kendilerini yetersiz, stresli, kızgın ve gergin hissetmektedir. Bu şekilde mesleği uzun süre devam ettiremeyeceklerini ve istifa etmeyi düşünmektedir. Öğretmen yetiştiren kurumlar, değişen öğrenci yapısı ve veli beklentilerini karşılayacak nitelikte teori ve uygulamaya dayalı bir eğitim modeli oluşturmalı ve donanımlı öğretmenler yetiştirmelidir. Göreve yeni başlayan öğretmenler, görece cezalandırıcı bir tutum sergileyerek sınıf ortamında karşılaştıkları olumsuz davranışları yönetme eğilimindedir. Bu öğretmenlere deneyimli ve yetenekli öğretmenler tarafından akran koçluğu desteği verilmeli, okul-veli işbirliği artırılarak sorunlu öğrencilerin olumsuz davranışlarını azaltmada öğrenci velilerinden daha fazla destek alınmalıdır. Aday öğretmenlerin meslek için etkin bir şekilde hazırlanabilmesi için değişen aile yapıları analiz edilmeli ve öğretmen yetiştirme kurumları tarafından öğrenci karakteristikleri dikkate alınmalıdır.

Araştırmadan elde edilen bulgulara dayalı olarak, sınıf yönetimi eğitiminin teori ve pratiğe dayalı olarak verilmesi önerilebilir. Okul yöneticileri acemi öğretmenlere kolaylaştırıcı bir tutum sergileyerek nispeten daha az sorunlu öğrencilerin bulunduğu sınıfların sorumluluğunu verebilir. Sonuçlara dayanarak, olumsuz öğrenci davranışların önlemede okul-aile işbirliğinden yararlanılması önerilebilir. Öğrencilerin okul ortamında rahatsızlık duyduğu olaylar ve nedenleri, farklı araştırma konuları olarak ele alınabilir. İdeal okul ve okul mutluluğu konusunda araştırmalar yürütülebilir.

Anahtar Sözcükler: Yeni öğretmen, olumsuz davranış, duygular, tutum.

Eurasian Journal of Educational Research 69 (2017) 221-240

Eurasian Journal of Educational Research
www.ejer.com.tr

Predictive and Explanatory Relationship Model between Procrastination, Motivation, Anxiety and Academic Achievement

Ugur AKPUR[1]

ARTICLE INFO

Article History:

Received: 10 August 2015

Received in revised form: 03 December 2016

Accepted: 25 February 2017

DOI: http://dx.doi.org/10.14689/ejer.2017.69.12

Keywords
affective factors,
language learning,
structural equation modelling

ABSTRACT

Purpose: The purpose of this study is to determine the predictive and explanatory relationship model between procrastination, motivation, anxiety and academic achievement of university students. **Research Methods:** In this study, a causal research design was used. The study group consisted of 211 participants. In order to determine their motivation levels, Academic Motivation Scale (AMS); their anxiety levels, Foreign Language Classroom Anxiety Scale (FLCAS) and their procrastination levels, Aitken Procrastination Inventory (API) were all applied. Students' grades during the term were considered as the criteria for academic achievement. Data were obtained through Structural Equation Model (SEM). **Findings:** Within the theoretical background, the proposed model was tested, and after path analysis it was modified and verified by testing through fitness indexes. The values confirmed that the model is compatible and the goodness-of-fitness values are within the limits. Findings reveal that, there is not a significant relationship between anxiety and academic achievement. However, both the relationship between academic procrastination and academic achievement and the relationship between motivation and academic achievement are significant. According to findings, the relationship between academic procrastination and motivation is significant and negative. Considering the findings, the predictive and explanatory relationship pattern between procrastination, motivation anxiety and academic achievement was suggested as a model. **Implications for Research and Practice:** Procrastination and motivation are significant variables in predicting academic achievement. Future studies could focus on other affective variables that are thought to have relationship to academic achievement.

[1] Yildiz Teknik Universitesi, uakpur@yahoo.com

Introduction

Many responsibilities that the students have to take on during their educational lives influence the efficiency of their education as well as their success to some extent. However, it is frequently observed that students generally procrastinate in their duties. The term procrastination derives from the Latin word "*procrastinat*-deferred till the morning" (Procrastination, 2015). According to Solomon and Rothblum, (1984, 503), procrastination is "the act of needlessly delaying tasks to the point of experiencing subjective discomfort." Ackerman and Gross (2005, 5) define the term "as the delay of a task or assignment that is under one's control." Furthermore, Steel (2007), along with Gustavson and Miyake (2017), while defining the term, mentions the notion of voluntariness. According to this, procrastination comprises intentional choice of one action over the other choices. Similarly, Shraw, Wadkins and Olafson (2007, p. 12) describe the term as "intentionally delaying or deferring work that must be completed". In other words, it is a "self-report tendency to nearly always or always put off academic tasks and to nearly always or always experience problematic levels of anxiety associated with procrastination" (Rothblum, Beswick & Mann, 1984, as cited in Rothblum, Solomon & Murakami, 1986, 387). Solomon and Rothblum (1984) argue that procrastination involves something more than time spent on studying or attitudes towards a subject. Rather, it encompasses anxiety, being indecisive, rebellion against control, and so on.

Procrastination is closely associated with academic performance, and in the literature there are many studies aiming to explain the notion's frequency and consequences. It is estimated that nearly 95% of college students procrastinate on academic assignments (Ellis & Knaus, 1977, as cited in Onwuegbuzie & Jiao, 2000). Solomon and Rothblum (1984) asserted that undergraduate students procrastinate on academic tasks such as term papers, preparing for exams and reading assignments within the range of 27 to 46 percent. Onwuegbuzie and Jiao (2000) offer that 60 percent of graduate level students procrastinate on academic tasks. Similarly, Onwuegbuzie (2004) in his study reports that 41.7% of graduate students nearly always or always procrastinate on writing their term papers, 39.3% of students procrastinate on preparing for their exams, and finally 60.0% of students procrastinate on doing their weekly reading assignments. Klassen and Kuzucu (2009) assert that 83% of adolescents procrastinate at least one hour per day on writing tasks. Ebadi and Shakoorzade (2015) in their study argue that more than half of students almost always procrastinate or always procrastinate. Steel (2007, 80), in his meta-analysis, found that, across 41 studies, there are consistently negative relationships between academic performance and procrastination with the average correlation of -19. That is, procrastination although sometimes "harmless," is generally detrimental; however "never helpful." Likewise, Kim and Seo (2015) conducted a meta-analysis of 33 studies which involved 38,529 participants and their research shows that procrastination is negatively correlated with academic performance. Similarly, Klassen et al. (2010) and You (2015) found that procrastination has a negative influence upon academic performance. As observed,

the great majority of studies assert that there is a negative correlation between procrastination and academic performance.

There have been many attempts to define the reasons why individuals procrastinate or keep procrastinating despite knowing its consequences. While Lay and Silverman (1995) argue that there is not a significant relationship between anxiety and procrastination, Rothblum et al. (1986) claim that the notion of procrastination contains cognitive and affective constituents and has a significant relationship to anxiety. They also asserted the idea that more than 40 percent of the participants in their study claimed a considerable amount of stress. Another study conducted by Senecal, Koestner and Vallerand (1995) reveals that individuals with high intrinsic motivation procrastinate less, and those with high extrinsic motivation procrastinate more. This supports the claim that procrastination is a motivational matter. Likewise, Lee (2005) asserts that intrinsic motivation has important effects on procrastination. Thus, while considering the reasons for procrastination, one has to take motivational factors into consideration.

In the conceptual framework, it is clear that motivational, affective and cognitive aspects should be taken into account to apprehend procrastination (Muszynski & Akamatsu, 1991; Senecal et al., Koestner & Vallerand, 1995). As Klassen, Krawchuk, Lynch and Rajani (2008, 137) assert, while motivation expresses something to do with struggle, determination and endeavor to a special purpose, procrastination, then, might be considered a kind of "anti-motivation," evasion and postponement. Thus, procrastination suggests lack of motivation, and this might be a disadvantage to academic success (Dunn, Rakes & Rakes, 2014) because it limits or even hinders the individual's potential to fulfill certain tasks.

Like procrastination, the notion of motivation has also gained much attention among researchers. Despite its popularity, the definition of the term has not been specifically stated (Oxford & Shearin, 1994), and Kleinginna and Kleinginna (1981) assert that reaching a consensus among the definitions of the term is a major problem as there are 102 different categories of explanations of the term. For instance, Ryan and Deci (2000, 54) describe it as "to be moved to do something," while Cheng and Dörnyei (2007, 153) frame the term as the "initial engine to generate learning". Notwithstanding the different explanations, the term itself generally connotes an impulse-like feeling.

Although it is mostly treated as a unitary notion by classical and modern theories, Self-Determination Theory (SDT) handles motivation from a different viewpoint. From this perspective, the types of motivation are far more important than the total amount of motivation (Deci & Ryan, 2008). That is, rather than the amounts, kinds of motivation are stressed (Ryan & Deci, 2000). In SDT, a basic distinction is made clear between intrinsic and extrinsic motivation (Ryan & Deci, 2000; Lee, 2011; Dornyei, 2003; Pelletier, 2002). Intrinsic motivation refers to inner satisfaction, interest or joy. On the other hand, extrinsic motivation refers to a reward, praise, wish or order from the outer world (Ryan & Deci, 2000; Deci & Ryan,

1985; Vallerand & Ratelle, 2002). In short, despite the general tendency, SDT focuses on the separation of the motivation types.

There are many views and studies asserting that motivation does affect academic achievement. For instance, Gardner (1985) states that there is a close relationship between students' motivation, their aptitude in a foreign language and their academic performance. Mallik (2017) mentions the crucial role of motivation in acquiring a foreign language. Goodman et al. (2011) in their study, which aims to determine the relationship between university students' motivation and academic performance, have found that there are significant relationships between intrinsic motivation, extrinsic motivation and academic performance. Further, it was revealed that their intrinsic and extrinsic motivation influenced the level of efforts they made to fulfill the targeted outcome. Similarly, a study conducted by Bidin et al. (2009) revealed the fact that motivation is an important variable in the language learning process, and a high extrinsic motivation level especially enhances academic achievement. Examining the relationship between procrastination and motivation, Klassen, Krawchuk and Rajani (2008) hold that procrastination suggests lower levels of motivation and mirrors lack of motivation. They maintain the idea that motivation has a negative correlation with procrastination which influences students' academic performance unsatisfactorily.

Being generally associated with poor academic performance, anxiety (Hussain, 2011; Kitano, 2001; Matsuda & Gobel, 2013; Rassaei, 2015; Tuncer & Dogan, 2015; Bensalem, 2017; Kuscu, 2017) is another variable of procrastination (Solomon & Rothblum, 1984). Akbay and Gizir (2010) put forward the idea that even though a momentary feeling of relaxation emerges just after the academic procrastination behavior, in the long term, this feeling changes into a kind of anxiety that has negative effects on academic performance. In their study, Solomon and Rothblum (1984) reveal that there is a positive relationship between academic procrastination and particular types of anxiety, like test anxiety and social anxiety, pointing to the same opinion. Similarly, Scher and Osterman (2002) argue that anxiety is a prevalent reason for procrastination. Likewise, Ferrari, O'Callaghan and Newbegin (2005) together with Haycock, McCarthy and Skay (1998), assert that procrastination is linked with inadequate academic performance and higher anxiety levels. Onwuegbuzie (2004), in his study examining the prevalence of procrastination among graduate students, reached the conclusion that academic procrastination is related significantly to test and class anxiety which influences academic performance in a negative way. The notion of anxiety also plays an important role in the language learning process (Onwuegbuzie, Bailey & Daley, 2000; Horwitz, 2010; MacIntyre & Gardner, 1991; Young, 1991; Cakici, 2016) while Gregersen and Horwitz (2002, 566) found that anxious learners expressed "avoidance and procrastination in their language learning," whereas not even a single non-anxious learner mentioned procrastination or work avoidance.

As an overall conclusion, the findings of the aforementioned studies reveal that procrastination displays a negative influence on academic performance (Dunn et al., 2014; Steel, 2007; Kim & Seo, 2015); the students who procrastinate have lower

motivation to fulfill a certain task (Klassen et al., 2008; Steel 2007) and as the procrastination level increases, the level of anxiety increases, as well (Solomon & Rothblum, 1984; Gregersen & Horwitz, 2002; Ferrari et al., 2005; Onwuegbuzie 2004). As a consequence, in light of the literature review, it is argued that academic procrastination adversely affects academic performance as well as motivation. On the other hand, there seems a positive and significant relationship between academic procrastination and anxiety. In this framework, the purpose of this study emerged as follows:

What is the predictive and explanatory relationship model between procrastination, motivation, anxiety and academic achievement?

After reviewing the theoretical background and empirical research, the proposed model was shaped as follows:

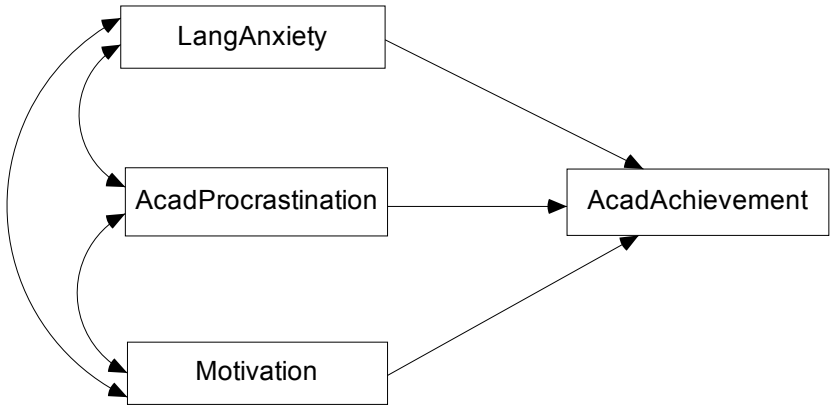


Figure 1. Proposed model

In Figure 1, the proposed model was constituted in accordance with the theoretical context of the independent variables (procrastination, motivation, anxiety) and dependent variable (academic achievement).

Method

Research Design

The study was conducted in causal research design. The cause and effect relationship between variables was analyzed through Structural Equation Modeling (SEM).

Research Sample

The population of this study includes university students attending Yıldız Technical University preparatory classes due the fall term of the Academic Year 2014-2015. The study group consisted of 229 students. Eighteen questionnaire sheets were ignored due to the poor feedback. In the end, 211 students, who were chosen randomly, formed the study group. The data gained from 211 students, 87 (41.2%) being female and 124 (58.8%) being male, were assessed.

Research Instruments and Procedure

In order to determine the students' procrastination levels, the Aitken Procrastination Inventory (API) was applied. Developed by Aitken (1982), the inventory was adapted into Turkish by Balkis (2006). Consisting of 16 items, the inventory is a five-point Likert scale ranging from false (1) to true (5). For each of the items, the participants are supposed to choose the item which is more or less convenient for them. High scores display the participants' high level of procrastination while the low scores indicate just the opposite. The inventory's internal consistency coefficient was calculated .89 Cronbach's Alpha, and test-retest reliability coefficient was found to be .87 (p<.001) (Balkis, 2006).

The Foreign Language Classroom Anxiety Scale (FLCAS), which was originally developed by Horwitz, Horwitz and Cope (1986), aims to determine levels of anxiety among students in foreign language classes. Composed of 33 items, the scale was adapted into Turkish by Aydin (2001). Being a five-point Likert scale, FLCAS was conducted on 300 university students who were studying in the foreign language department, and factor analysis indicated that internal consistency coefficient was .93 Cronbach's Alpha. Test-retest process was conducted for eight weeks and test-retest reliability coefficient was found to be .83 (p = .001) (Aydin, 2001).

Students' motivation levels were assessed through the Academic Motivation Scale (AMS), which was developed by Vallerand and Ratelle (1992) and adapted into Turkish by Karatas and Erden (2012). The scale is made up of 27 items and its internal consistency coefficient was found to be .97 Cronbach's Alpha (Karatas & Erden, 2012). In this study, four items (5, 12, 19, 26 items) that belong to the Amotivation dimension were excluded. Consequently, the inventory consisting of 23 items was applied in the study. The coefficient reliability of the scale in this form was found to be .89 Cronbach's Alpha.

The students' academic achivement was assessed through their grade point average for the fall term of the 2014-2015 academic year. The assessment criteria were as folows:

Table 1

The Assessment of Academic Achievement

2 Midterm Exams	40%
3 Pop-Quizzes	20%
2 Reading Exams	10%
Portfolio Work	10%
Presentation and Oral Exam	15%
Class Participation	5%
Total	**100**

Data Analysis

The data gained from the study were analyzed through SEM and statistically evaluated by means of AMOS software. SEM, which may shortly be depicted as a bunch of statistical methods, allowed us to comprehend "the relationship between one or more than one independent variables and one or more than one dependent variables" (Ullman & Bentler, 2013, 661). Further, it offers a broad and flexible evaluation between the observed and latent variables (Hoyle & Smith, 1994). Additionally, it can also be used to test, analyze and comprehend the multidimensional structure of a model. In this way, determining and removing the weaknesses of a hypothesized model and displaying multifaceted interactions can be clearly accomplished (Weston & Gore, 2006; Kline, 1998; Anderson & Gerbing, 1998).

Although there is not a consensus on the appropriate sample size for SEM (Hoe, 2008; Raoprasert & Islam, 2010), Hoe (2008) reports that a sample size of more than 200 provides adequate statistical value for an analysis. Likewise, Kline (2005) asserts that a sample size of less than 100 is considered a small sample, a size between 100 and 200 is a medium sample, and a size more than 200 is a large sample. Hoelter (1983) also holds 200 as the critical sample size.

Results

In the model to be tested, the relationship pattern between the variables of procrastination, motivation, anxiety and academic achievement was analyzed through path analysis.

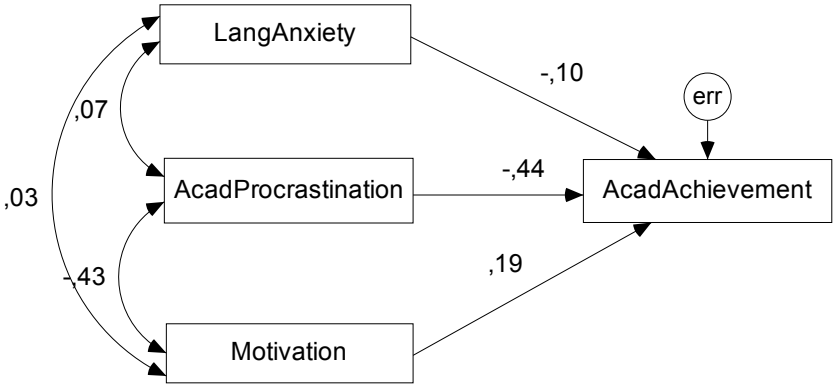


***Figure* 2.** Values of the proposed model

In Figure 2, the values of the proposed model along with the relationship pattern between variables are displayed.

In order to test the model, the maximum likelihood process was applied in the AMOS program. Among the ways of testing a model, determining the values of some goodness-of-indexes and comparing them with the acceptable values can be regarded as a reliable method (Schermelleh-Engel, Moosbrugger & Muller, 2003).

The values of good fit and acceptable fit along with the values of the proposed model displayed in Table 2.

Table 2

Recommendation for Model Evaluation

Fit Measure	Good Fit	Acceptable Fit	Proposed Model
χ2/df	.0 ≤ χ2/df ≤ 2	2 ≤ χ2/df ≤ 3	.0
RMSEA	0 ≤ RMSEA ≤ .05	0 ≤ RMSEA ≤ .08	.30
NFI	.95 ≤ NFI ≤ 1.00	.90 ≤ NFI ≤ .95	.1
CFI	.97 ≤ CFI ≤ 1.00	.95 ≤ CFI ≤ .97	.1
GFI	.95 ≤ GFI ≤ 1.00	.90 ≤ AGFI ≤ .95	.1
AGFI	.90 ≤ AGFI ≤ 1.00	.85 ≤ AGFI ≤ .90	.94

RMSEA = Root Mean Square Error of Approximation, NFI = Normed Fit Index, CFI = Comparative Fit Index, GFI = Goodness-of-Fit Index, AGFI = Adjusted Goodness-of-Fit-Index (Schermelleh-Engel et al., 2003).

In the proposed model, the value of chi-square is "0", should be less than three when divided by the degree of freedom. This shows that the model has a suitable index value regarding the value of chi-square.

The results of the research also demonstrated that the goodness-of-fit indexes of the proposed model were as follows: NFI = .1(.95 ≤ NFI ≤ 1.00); CFI = .1(.97 ≤ CFI ≤ 1.00); GFI = .1(.95 ≤ GFI ≤ 1.00); AGFI = .94 (.90 ≤ AGFI ≤ 1.00). These figures demonstrate that the model's fitness was acceptable. Nevertheless, RMSEA value was found to be .25, which is not within the limits of the recommended value (0 ≤ RMSEA ≤ .05). Thus, after the necessary path analysis, the model was reviewed again and modified.

To obtain the suitability of the model as a whole, the two-headed row between language anxiety and motivation was omitted and after this adjustment, the model was re-evaluated as in Figure 3.

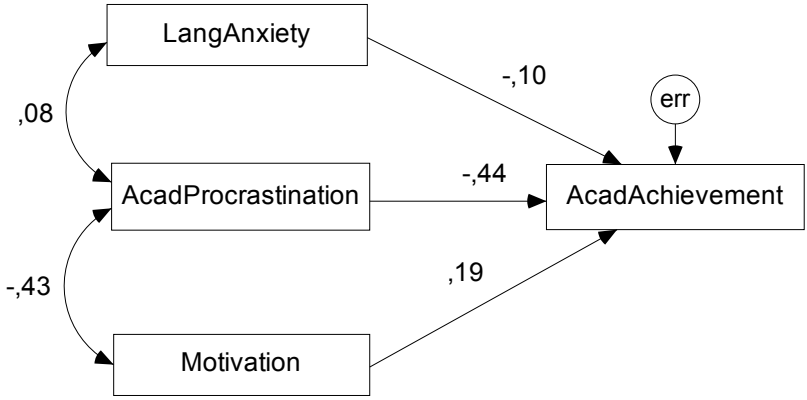


Figure 3. Values of the last model

In Figure 3, the proposed model was modified and after the necessary adjustments it was evaluated again.

Table 3

The Values of the Last Model

Fit Measure	Good Fit	Acceptable Fit	The Last Model
χ2/df	.0 ≤ χ2/df ≤ 2	2 ≤ χ2/df ≤ 3	.20
RMSEA	0 ≤ RMSEA ≤ .05	0 ≤ RMSEA ≤ .08	.01
NFI	.95 ≤ NFI ≤ 1.00	.90 ≤ NFI ≤ .95	.99
CFI	.97 ≤ CFI ≤ 1.00	.95 ≤ CFI ≤ .97	.98
GFI	.95 ≤ GFI ≤ 1.00	.90 ≤ AGFI ≤ .95	.97
AGFI	.90 ≤ AGFI ≤ 1.00	.85 ≤ AGFI ≤ .90	.99

The figures displayed in Table 3 indicate that, when the two-headed row between Language Anxiety and Motivation is omitted, the model is compatible with the goodness-of-fit indexes. The value of chi-square when divided by the degree of freedom (df: 1), was found to be .20, which can be considered a good fit.

Similarly, the values of NFI = .99 (.95 ≤ NFI ≤ 1.00); CFI = .98 (.97 ≤ CFI ≤ 1.00); GFI = .97 (.95 ≤ GFI ≤ 1.00); AGFI = .99 (.90 ≤ AGFI ≤ 1.00) provided verification that the model is compatible and the goodness-of-fitness values of it are within the limits. Contrary to the initial model, the value of RMSEA was found to be .01, which is within the limits of the recommended value (0 ≤ RMSEA ≤ .05).

Table 4

Regression Weights, Standard Errors, Critical Ratios and 'p' Values of the Variables of the Last Model

Variable			Estimate	St. Err.	Critical Ratio	p
Lang. Anxiety	→	Acad. Achiev.	-.134	.075	-1.786	.07*
Acad. Procr.	→	Acad. Achiev.	-.579	.084	-6.898	.00**
Motivation	→	Acad. Achiev.	.121	.041	2.946	.00**

Total Effect Value: .88.36 **p<.05, **p<.01.

Table 4 shows that the predictive power of language anxiety to predict academic achievement is -.134; the power of academic procrastination to predict academic achievement is -.579; and the power of motivation upon academic performance is .121. The total effect value of anxiety, procrastination and motivation is 88.36.

Table 4 also shows that the relationship between language anxiety and academic achievement is not significant (Critical Ratio-CR = -1.786; p<.05). On the other hand, the relationship between academic procrastination and academic achievement is significant (CR = -6.898; p<.01). Similarly, there is a significant relationship between motivation and academic achievement, as well (CR = 2.946; p<.01).

In Table 5, correlations, standard errors, critical ratios and 'p' values of the variables of the last model are itemized.

Table 5

Correlations, Standard Errors, Critical Ratios and 'p' Values of the Variables of the Last Model

Variable			Estimate	St. Err.	Critical Ratio	p
Lang. Anxiety	↔	Acad.Achiev.	11.45	9.09	1.260	.20*
Acad. Procr.	↔	Motivation	-125.60	21.93	-5.726	.00**

*p<.05, **p.01

Table 6 shows that there is not a significant relationship between language anxiety and academic achievement (CR = 1.260; p<.05). Nevertheless, the relationship between academic procrastination and motivation is significant in a negative way (CR = -5.726; p<.01).

Discussion and Conclusions

The current study's aim was to determine and propose a model analyzing the relationship pattern between academic procrastination, motivation, anxiety and academic achievement. In the current study, it was found that the relationship between academic procrastination and academic achievement is significant in a negative way. This confirms the view that procrastination although sometimes "harmless," is generally detrimental however; it is "never helpful" (Steel, 2007, 80). The results obtained in the study are compatible with most of the research performed in various countries (Steel, 2007; Kim & Seo, 2015; Klassen et al. 2010; Onwuegbuzie, 2004; Dunn et al., 2014). Therefore, being aware of the consequences of the notion could increase the quality and efficiency of education.

As for the notion of motivation, in the present study, it was discovered that motivation was a significant predictor of academic achievement. As anticipated, the findings displayed the same results. The results are also consistent with various research from different sorces and samples. (Bidin et al. 2009; Goodman et al. 2011; Mo, 2011; Cheng, Lin & Su, 2011; Nishitani & Matsuda, 2011). Thus, it is clear that, motivating students in learning environments will bring about enhanced academic achievement. In other words, once students are motivated to do something, they will perform their responsibilities and duties simply by virtue of the wish and resulting contentment.

With regard to anxiety, it was found out that foreign language anxiety is not a significant predictor of language achievement, and there is not a significant relationship between anxiety and language achievement. This supports the idea that facilitating anxiety may play an important role in academic performance (Scovel, 1978; Skehan, 1990; Eysenck, 1979). On the other hand, there are various studies suggesting that anxiety is generally associated with poor academic performance (Gardner, 2010; MacIntyre, Noels & Clément, 1997; Hussain, 2011; Kitano, 2001; Matsuda & Gobel, 2013). This controversy may arise from the fact that, as Horwitz (2010, 154) claimed, the notion of anxiety is "multi-faceted and psychologists have differentiated a number of types of anxiety including trait anxiety, state anxiety, achievement anxiety, and facilitative-debilitative anxiety".

Commenting on the findings, some limitations should be considered. The results are limited by the size of the sample and the findings should be evaluated in this context. Apart from university students, further research could be carried out with bigger samples from different schools and grades. Further, it would be advisable to study other affective variables that are thought to have relationship with academic achievement. What is more, conducting the study at the end of the academic year may have affected the results, and the participants may respond to the items differently, whereby the reliability and the validity of the model could vary.

Outside of its limitations, one of the important aspects of the study is that it proves a unique model for analyzing the relationships between academic procrastination, motivation, language anxiety and academic achievement. In accordance with the analysis of the literature, it was acknowledged that the

aforementioned variables have close relationships with academic achievement, and the model tested offers a thorough description of their interactions. Furthermore, the findings, which generally bear resemblance to other studies, provide illumination to decision-makers in planning, applying and evaluating the educational programs.

As an overall conclusion, the findings of the study, which are thought to aid educators comprehending the relationships between the aforesaid variables and the roles they play in an educational context, demonstrate that foreign language anxiety and academic procrastination have negative effects on academic achievement. On the other hand, it has been revealed that motivation has positive effects on academic achievement. As for the correlations between the variables, although there is not a significant relationship between language anxiety and academic achievement, it has been found that the relationship between academic procrastination and motivation is significant. Thus, it is apparent that procrastination, motivation and anxiety can be noted as important affective variables that affect academic achievement, and they should therefore be treated with utmost attention.

References

Ackerman, S. D., & Gross B. L. (2005). My instructor made me do it: Task characteristics of procrastination. *Journal of Marketing Education,* 27, 5-13.

Akbay, S. E., & Gizir, C. A. (2010). Cinsiyete gore universite ogrencilerinde akademik erteleme davranisi: Akademik gudulenme, akademik ozyeterlik ve akademik yukleme stillerinin rolu [Academic procrastination among university students according to gender: The role of academic motivation, academic self-efficacy and academic attributional style] *Mersin Universitesi Egitim Fakultesi Dergisi,* 6, 60-78.

Anderson, J. C., & Gerbing, D. W. (1998). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin,* 103, 411-423.

Aydın, B. (2001). *Konusma ve Yazma Derslerinde Yabancı Dil Ogrenimindeki Kaygi Nedenleri* [*A Study of Sources of Foreign Language Classroom Anxiety in Speaking and Writing Classes*]. (Unpublished Doctoral Dissertation). Anadolu Universitesi, Eskisehir.

Balkıs, M. (2006). *Ogretmen Adaylarinin Davranislarindaki Erteleme Egilimlerinin Dusunme ve Karar Verme Tarzlari ile Iliskisinin Incelenmesi* [*The Relationships between Student Teachers' Procrastination Behaviors and Thinking Styles and Decision Making Styles*]. (Unpublished Doctoral Dissertation). Dokuz Eylul University, Izmir.

Bensalem, E. (2017). Foreign Language Learning Anxiety: The Case of Trilinguals. *Arab World English Journal,* 8 (1), 234-249. DOI: https://dx.doi.org/10.24093/awej/vol8no1.17

Bidin, S., Jusoff, K., Abdul-Aziz, N., Mohamad, M., Salleh, M. M., & Tajudin, T. (2009). Motivation and Attitude in Learning English among UITM Students in the Northern Region of Malaysia. *English Language Teaching,* 2, 16-20.

Çakıcı, D. (2016). The correlation among EFL learners' test anxiety, foreign language anxiety and language achievement. *English Language Teaching,* 9(8), 190-203.

Cheng, H. F., & Dornyei, Z. (2007). The use of motivational strategies in language instruction: The case of EFL teaching in Taiwan. *Innovation in Language Learning and Teaching,* 1, 153-174.

Cheng, P. Y., Lin, M., & Su, C. K. (2011). Attitudes and motivations of students taking professional certificate examinations. *Social Behavior and Personality,* 39(10), 1303-1314.

Deci, E., & Ryan, R. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum.

Deci, E., & Ryan, R. (2008). Self-Determination Theory: A macrotheory of human motivation, development, and health. *Canadian Psychology,* 49, 182-185.

Dornyei, Z. (2003). Attitudes, orientations, and motivations in language learning: Advances in theory, research and applications. *Language Learning,* 53, 3-32.

Dunn, K. E., Rakes, G. C., & Rakes, T. A. (2014). *The Influence of Perfectionism and Achievement Goal Orientation on Procrastination in Online Graduate Students.* Paper presented at the American Educational Research Association Conference, Philadelphia.

Ebadi, S., Shakoorzadeh, R. (2015). Investigation of academic procrastination prevalence and its relationship with academic self-regulation and achievement motivation among high-school students in Tehran City. *International Education Studies,* 8 (10), 193-199.

Eysenck, W. M. (1979). Anxiety, learning and memory: A reconceptualization. *Journal of Research in Personality,* 13, 363-385.

Ferrari, J. R., O'Callaghan, J., & Newbegin, I. (2005). Prevalence of procrastination in the United States, United Kingdom, and Australia: Arousal and avoidance delays among adults. *North American Journal of Psychology,* 7, 1 -6.

Gardner, C. R. (1985). *Social psychology and language learning: The role of attitudes and motivation*. London: Edward Arnold.

Gardner, C. R. (2010). *Motivation and second language acquisition*. New York: Peter Lang Publishing.

Goodman, S., Jaffer, T., Keresztesi, M., Mamdani, F., Mokgatle, D., Musariri, M., Pires, J., & Schlechter, A. (2011). An investigation of the relationship between students' motivation and academic performance as mediated by effort. *South African Journal of Psychology,* 41, 373-385.

Gregersen, T., & Horwitz, E. K. (2002). Language learning and perfectionism: Anxious and non-anxious language learners' reactions to their own oral performance. *The Modern Language Journal,* 86, 562-570.

Gustavson, D. E., Miyake, A. (2017). Academic procrastination and goal accomplishment: A combined experimental and individual differences investigation. *Learning and Individual Differences,* 54, 160–172.

Haycock, L. A., McCarthy, P., & Skay, C. L. (1998). Procrastination in college students: The role of self-efficacy and anxiety. *Journal of Counseling & Development,* 76, 317-324.

Hoe, L. S. (2008). Issues and procedures in adopting structural equation modeling technique. *Journal of Applied Quantitative Methods,* 3(1), 76-83.

Hoelter, D. R. (1983). The analysis of covariance structures: Goodness-of-fit indices. *Sociological Methods and Research,* 11, 325-344.

Horwitz, K. E. (2010). Foreign and second language anxiety. *Language Teaching,* 43, 154-167.

Hoyle, R. H., & Smith, G. T. (1994). Formulating clinical research hypotheses as structural equation models: A conceptual overview. *Journal of Consulting and Clinical Psychology,* 62, 429-440.

Hussain, M. A. (2011). *Relationship of Classroom Environment with Anxiety and Attitude of Secondary School Students towards the Learning of English* (Unpublished Doctoral Dissertation). International Islamic University, Islamabad.

Karatas, H., & Erden, M. (2012). Akademik motivasyon olceginin dilsel esdegerlik, gecerlik ve guvenirlik calismasi [Bilinqual equivalance, validity and reliability of academic motivation scale]. *e-Journal of World Sciences Academy,* 7, 983-1003.

Kim, R. K., & Seo, H. E. (2015). The relationship between procrastination and academic performance: A meta-analysis. *Personality and Individual Differences,* 82, 26–33

Kitano, K. (2001). Anxiety in the college Japanese language classroom. *The Modern Language Journal,* 85, 549-566.

Klassen, R. M., Krawchuk, L. L., Lynch, S. L., & Rajani, S. (2008). Procrastination and motivation of undergraduates with learning disabilities: A mixed-methods inquiry. *Learning Disabilities Research & Practice,* 23, 137-147.

Klassen, R. M., Krawchuk, L. L., & Rajani, S. (2008). Academic procrastination of undergraduates: Low self-efficacy to self-regulate predicts higher levels of procrastination. *Contemporary Educational Psychology,* 33, 915-931.

Klassen, M. R., & Kuzucu, E. (2009). Academic procrastination and motivation of adolescents in Turkey. *Educational Psychology,* 29, 69-81.

Klassen, M. R., Ang, R. P., Chong, W. H., Krawchuk, L. L., Huan, V. S. , Wong, I. Y. F., & Yeo, L. S. (2010). Academic procrastination in two settings: Motivation correlates, behavioral patterns, and negative impact of procrastination in Canada and Singapore. *Applied Psychology: An International Review,* 59, 361-379.

Kleinginna, R. P., Kleinginna, A. M. (1981). A categorized list of motivation definitions with a suggestion for a consensual definition. *Motivation and Emotion,* 5, 263-291.

Kline, R. (1998). Software review: Software programs for structural equation modeling: Amos, EQS, and Lisrel. *Journal of Psychoeducational Assessment,* 16, 343-364.

Kline, R. (2005). *Principles and practice of structural equation modeling (2nd ed.).* New York: Guilford.

Kuscu E. (2017). Teaching the anxiety of learning a foreign language that influences high school students in learning French as a second foreign language The case of Denizli. *Journal of Language and Linguistic Studies,* 13(1), 88-102.

Lay, C., & Silverman, S. (1995). Trait procrastination, anxiety and dilatory behavior. *Personal Individual Differences,* 21, 61-67.

Lee, E. (2005). The relationship of motivation and flow experience to academic procrastination in university students. *The Journal of Genetic Psychology: Research and Theory on Human Development,* 166(1), 5-15.

Lee, W. (2011). *Neural Substrates of Intrinsic Motivation* (Unpublished Doctoral Dissertation). The University of Iowa, Iowa.

MacIntyre, D. P., & Gardner, C. R. (1991). Language anxiety: Its relation to other anxieties and to processing in native and second languages. *Language Learning,* 41, 513-534.

MacIntyre, D. P., Noels, K., & Clément, R. (1997). Biases in self-ratings of second language proficiency: The role of language anxiety. *Language Learning,* 42, 265-287.

Mallik, S. (2017). Motivation as a promoting determinant in second and foreign language classroom: a review. *European Journal of Foreign Language Teaching*, 2 (1), 90-102.

Matsuda, S., & Gobel, P. (2013). Anxiety and predictors of performance in the foreign language classroom. *System,* 32, 21-36.

Mo, S. (2011). An exploratory study of intrinsic & extrinsic motivators and student performance in an auditing course. *American Journal of Business Education,* 4(2), 19-25.

Muszynski, S. Y., & Akamatsu, T. J. (1991). Delay in completion of doctoral dissertations in clinical psychology. *Professional Psychology: Research and Practice,* 22, 119-123.

Nishitani, M., & Matsuda, T. (2011). The relationship between language anxiety, interpretation of anxiety, intrinsic motivation and the use of learning strategies. *US-China Education Review,* B 3, 438-446.

Onwuegbuzie, A. J. (2004). Academic procrastination and statistics anxiety. *Assessment & Evaluation in Higher Education,* 29, 3-19.

Onwuegbuzie, J. A., & Jiao, Q. G. (2000). I'll go to the library later: The relationship between academic procrastination and library anxiety. *College and Research Libraries,* 61, 45-54.

Onwuegbuzie, J. A., Bailey, P., & Daley, E. C. (2000). The validation of three scales measuring anxiety at different stages of the foreign language learning process: The input anxiety scale, the processing anxiety scale and the output anxiety scale. *Language Learning,* 50, 87-118.

Oxford, R., & Shearin, J. (1994). Language learning motivation: Expanding the theoretical framework. *The Modern Language Journal,* 78, 12-28.

Pelletier, G. L. (2002). A motivational analysis of self determination for pro-environmental behaviors. In R. Ryan, & E. Deci (Eds.): *Handbook of self determination research* (pp. 205-232). Rochester: University of Rochester Press.

Procrastination. (2015). In oxforddictionaries.com Retrieved from https://en.oxforddictionaries.com/definition/procrastinate

Raoprasert. T., & Islam, S. M. (2010). *Designing an efficient management system.* New York: Springer.

Rassaei, E. (2015). Oral corrective feedback, foreign language anxiety and L2 development. *System*, 49, 98 - 109

Rothblum, E. D., Solomon, L. J., & Murakami, J. (1986). Affective, cognitive and behavioral differences between high and low procrastinators. *Journal of Counseling Psychology,* 33, 387-394.

Ryan, R., & Deci, E. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology,* 25, 54–67.

Senecal, C., Koester, R., & Vallerand, R. (1995). Self-regulation and academic procrastination. The *Journal of Social Psychology,* 135, 607-619.

Scher, J. S., & Osterman, N. M. (2002). Procrastination, conscientiousness, anxiety, and goals: Exploring the measurement and correlates of procrastination among school-aged children. *Psychology in the Schools,* 39, 385-398.

Schermelleh-Engel, K., Moosbrugger, H., & Muller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online,* 8, 23-74.

Scovel, T. (1978). The effect of affect on foreign language learning: A review of the anxiety research. *Language Learning,* 28, 129-142.

Shraw, G., Wadkins, T., & Olafson, L. (2007). Doing the things we do: A grounded theory of academic procrastination. *Journal of Educational Psychology,* 99 (1), 12-25.

Skehan, P. (1990). *Individual differences in second language learning*. London: Edward Arnold Publishing.

Solomon, J. L., & Rothblum, E. D. (1984). Academic procrastination: Frequency and cognitive-behavioral correlates. *Journal of Counseling Psychology,* 31, 503-509.

Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin,* 133, 65-94.

Tuncer, M., & Dogan, Y. (2015). Effect of foreign language classroom anxiety on Turkish university students' academic achievement in foreign language learning. *Journal of Education and Training Studies,* 3 (6), 14-19.

Ullman, B. J., & Bentler, M. (2013). Structural equation modeling. In: B. Weiner, (Ed.), *Handbook of psychology* (pp. 661-690). California: John Wiley & Sons Inc.

Vallerand, J. R., & Ratelle, C. F. (2002). Intrinsic and extrinsic motivation: A hierarchical model. In: R. Ryan & E. Deci (Eds.): *Handbook of self determination research* (pp. 37-66). Rochester: University of Rochester Press.

Weston, R., & Gore, P. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist,* 34, 719-751.

You, J. W. (2015). Examining the effect of academic procrastination on achievement using LMS data in e-learning. *Educational Technology & Society,* 18 (3), 64–74.

Young, D. J. (1991). Creating a low-anxiety classroom environment: What does language anxiety research suggest? *The Modern Language Journal,* 75, 426-439.

Erteleme, Motivasyon, Kaygı ve Akademik Başarı Arasındaki Yordayıcı ve Açıklayıcı İlişkiler Modeli

Atıf:

Akpur, U. (2017). Predictive and explanatory relationship model between procrastination, motivation, anxiety and academic achievement. *Eurasian Journal of Educational Research, 69*, 221-240. http://dx.doi.org/10.14689/ejer.2017.69.12

Özet

Problem Durumu: Öğrencilerin sorumlulukları arasında, kendilerine verilen ödevleri ya da projeleri zamanında teslim etmek ve çalışma programlarını sekteye uğratmadan eğitimlerine devam etmek büyük önem taşımaktadır. Bu durum, alınan eğitimin etkinliğini ve başarısını etkileyen en önemli unsurlardan biridir. Akademik görevlerin bilinçli olarak ileri bir tarihe ertelenmesi akademik başarıyı azaltmakta ve eğitimin niteliğini olumsuz yönde etkilemektedir. Genel olarak, bitirilmesi gereken görevleri bireyin kendi isteği doğrultusunda ertelemesi ya da ötelemesi şeklinde tanımlanan erteleme davranışı, zaman yönetiminde karşılaşılan problemlerden çok, bünyesinde kaygı, motivasyon eksikliği, kararsız olma durumu ve otoriteye karşı koyma gibi unsurları barındırmaktadır. Gerçekleştirilen araştırmalarda öğrencilerin büyük bir bölümünün akademik görevlerde erteleme davranışı sergilediği; bu oranın yazma ödevlerinin yanı sıra sınavlara hazırlanma ve özellikle okuma ödevlerinde büyük bir artış gösterdiği ifade edilmektedir. Erteleme davranışlarının arkasında yatan nedenlerin belirlenmesi amacıyla çeşitli araştırmalar gerçekleştirilmiş ve bu konuda farklı yaklaşımlar ortaya konmuştur. Bu görüşlerden biri, erteleme davranışlarının motivasyonel bir durum olduğu şeklindedir. Yapılan araştırmalarda Öz-Belirleme Teorisine göre bir motivasyon çeşidi olan içsel motivasyonun, erteleme davranışları üzerinde etkili olduğunu öne sürülmektedir. Bir diğer deyişle, erteleme davranışı sergileyen bireylerin herhangi bir eylemi gerçekleştirmede daha düşük bir motivasyon düzeyine sahip olduğu, bu durumun da beraberin de akademik başarıyı olumsuz etkilediği ifade edilmektedir. Diğer bir görüş ise, erteleme davranışlarını tetikleyen unsurun, bireylerin sahip olduğu yüksek kaygı düzeyi olduğuna dikkat çekmektedir. Buna göre, bireylerin kaygı düzeyleri ile sergiledikleri erteleme davranışları arasında anlamlı bir ilişki bulunmakta ve her iki değişken de akademik başarıyı olumsuz yönde etkilemektedir. Gerçekleştirilen bazı araştırmalarda ise erteleme davranışının temel sebebinin kaygı durumu olduğu özellikle ifade edilmektedir. Sonuç olarak, alan yazın incelendiğinde akademik erteleme davranışlarının genellikle akademik başarı ve motivasyon üzerinde olumsuz etkilerinin olduğu; öte yandan akademik erteleme davranışları ile kaygı düzeyleri arasında ise olumlu bir ilişki olduğu görülmektedir.

Araştırmanın Amacı: Bu çalışmada akademik erteleme, motivasyon, kaygı ve akademik başarı arasındaki yordayıcı ve açıklayıcı modelin belirlenmesi ve söz konusu değişkenler arasındaki ilişkiler örüntüsünün saptanması amaçlanmıştır.

Araştırmanın Yöntemi: Bu araştırmada değişkenler arasında neden-sonuç ilişkisi olacağı düşünüldüğü için nedensel araştırma deseni kullanılmıştır. Çalışmanın evrenini, 2014-2015 akademik yılı güz yarıyılında Yıldız Teknik Üniversitesi, Yabancı Diller Yüksekokulu, Temel İngilizce (Hazırlık) Bölümüne devam eden öğrenciler oluşturmuştur. Çalışma grubunda ise, evrenden tesadüfi olarak seçilen ve araştırmaya gönüllü olarak katılan 229 öğrenci bulunmaktadır. Gerçekleştirilen uygulama sonucunda 18 ölçek eksik doldurulduğundan dolayı değerlendirilmeye alınmamış ve sonuç olarak 221 öğrenci çalışmaya dâhil edilmiştir. Öğrencilerin akademik erteleme davranışlarını ölçek amacıyla, Aitken (1982) tarafından geliştirilen ve Balkıs (2006) tarafından Türkçeye uyarlanan "Aitken Erteleme Eğilimi Ölçeği"; motivasyon düzeylerini belirlemek amacıyla, Vallerand ve diğ. (1992) tarafından geliştirilen ve Türkçeye Karataş ve Erden (2012) tarafından uyarlanan "Akademik Motivasyon Ölçeği" ile kaygı düzeylerini ölçmek için Horwitz, Horwitz ve Cope (1986) tarafından geliştirilen ve Türkçeye Aydın (2001) tarafından uyarlanan "Yabancı Dil Sınıf Kaygısı Ölçeği" kullanılmıştır. Öğrencilerin 2014-2015 akademik yılı güz yarıyılı boyunca aldıkları dönem içi notları da akademik başarı ölçütü olarak değerlendirilmiştir. Çalışma sonucunda elde edilen veriler SPSS 22.0 ve AMOS programları yardımıyla, yapısal eşitlik modeli kullanılarak analiz edilmiştir.

Araştırmanın Bulguları: Alan yazın incelemesine paralel olarak önerilen model test edilmiş ve path analizi sonucunda gerekli değişiklikler yapılarak tekrar test edilip doğrulanmıştır. Yapılandırılmış modelin uyum indeksleri incelendiğinde, modelin verilerle uyumlu olduğu görülmüştür (NFI = .99; CFI = .98; GFI = .97; AGFI = .99; RMSEA = .01). Elde edilen sonuçlar kaygı ile akademik başarı arasında anlamlı bir ilişkinin olmadığını, öte yandan erteleme davranışları ile akademik başarı arasında ve motivasyon ile akademik başarı arasında anlamlı bir ilişki olduğunu ortaya koymuştur. Ayrıca, erteleme davranışları ile motivasyon arasında olumsuz bir ilişki olduğu da saptanmıştır. Bulgulara paralel olarak erteleme davranışları, motivasyon ve kaygı ile akademik başarı arasındaki açıklayıcı ve yordayıcı ilişkiler örüntüsü bir model olarak verilmiştir.

Araştırmanın Sonuçları ve Öneriler: Elde edilen bulgular ışığında, eğitim-öğretim etkinliklerinde, erteleme davranışları, motivasyon, kaygı gibi değişkenlerin belirlenmesi ve olumsuz sonuçlarla karşılaşılmaması amacıyla gerekli önlemlerin alınması başarıyı arttıran unsurlar olarak değerlendirilmektedir. Öğrencilerin akademik erteleme davranışlarının azaltılması amacıyla kendilerine verilen görevleri önem sırasına koyarak bir takım planlamalarda bulunmaları ve bunlar gerçekleştirmek için de zaman aralıkları belirlemeleri önerilmektedir. Buna ek olarak, erteleme davranışlarının, motivasyon ve kaygı kavramlarıyla yakından ilişkili olmasından dolayı, öğrencilerin motivasyonlarının arttırılması gerektiği düşünülmektedir. Bu amaçla, öğrenme ortamlarının öğrencilerin kendi amaç ve ilgileri doğrultusunda düzenlenmesinin motivasyonlarını arttıracağı ifade edilmektedir. Bunun yanı sıra, eğitim-öğretim aktivitelerinde merak uyandıran

etkinliklerin planlanması, öğrencilere kendi yetenek ve kabiliyetleri doğrultusunda gerçekçi ve ulaşılabilir amaçlar belirlenmesi, öğrencilerin kendilerini rahat hissedebilecekleri bir öğrenme ortamının sağlanması, öğrencilere kendilerini ifade edebilmeleri için uygun fırsatların sunulması ve onların bu konuda teşvik edilmesi gibi etkinliklerin motivasyon düzeylerini arttıracağı ve buna paralel olarak da kaygı düzeylerini azaltacağı düşünülmektedir. Bu tür uygulamaların, aynı zamanda, akademik başarıyı Bu araştırmada test edilen model, yabancı dil eğitimi alan öğrencilerden elde edilen verilerle şekillendirilmiştir. Söz konusu model farklı program ve sınıflara devam eden öğrenciler üzerinde test edilebilir. Böylece değişkenler arasındaki açıklayıcı ve yordayıcı ilişkiler farklı perspektiflerle değerlendirilebilir.

Anahtar Kelimeler: Duyuşsal faktörler, dil öğrenimi, yapısal eşitlik modeli.

Eurasian Journal of Educational Research 69 (2017) 241-252

Eurasian Journal of Educational Research
www.ejer.com.tr

Development of Speaking Skills through Activity Based Learning at the Elementary Level

Zahoor-UL-HAQ[1] Bushra Ahmed KHURRAM[2] Arshad Khan BANGASH[3]

ARTICLEINFO

Article History:

Received: 10 January 2016

Received in revised form: 25 April 2016

Accepted: 12 March 2017

DOI: http://dx.doi.org/10.14689/ejer.2017.69.13

Keywords

Speaking skill, activity based learning, second language

ABSTRACT

Purpose: This paper discusses an effective instructional method called 'activity based learning' that can be used to develop the speaking skills of students in the elementary school level. The present study was conducted to determine the effect of activity based learning on the development of the speaking skills of low and high achievers in a 6th grade class at the elementary school level. **Research Methods:** The research was based on pre-test post-test equivalent group design. A total of 50 male students served as participants in this research. The participants were selected randomly from two sections and then a pre-test was administered. Based on pre-test scores, the participants were divided equally into experimental and control groups. High and low score achievers in each group were also identified. The experimental group received instruction based on the activity based learning method, and the control group was instructed through the traditional language teaching method. At the end of the experiment, a post-test was administered to measure the development of speaking skills in students. The independent sample t-test was used to test the significance of difference between the mean scores of groups at the 0.05 level. **Implications for Research and Practice**: The findings of the study suggested that activity based learning was an effective way to enhance students' speaking skills since the experimental group post test score was significantly different from the control group post test score. Based on the result of the study, it is recommended that the activity based learning method be used in class to develop and enhance the speaking skills of students. It is also recommended that teachers be provided training to implement the activity based method in language lessons.

[1]Corresponding Author: Zahoor UL-HAQ, Department of Education, Bacha khan University Charsadda, zahoorulhaqnsr@gmail.com

[2]Department of English, University of Karachi, Pakistan, bushrakhurram@gmail.com

[3]Department of Sociology & Gender Studies, Bacha Khan University Charsadda, arshad.dwhh@gmail.com

Introduction

Activity Based Learning (henceforth, ABL) is generally defined as 'any instructional method that engages students in the learning process' in the classroom (Prince, 2004, 223). Harfield et al. (2007) indicates that during ABL students are not passive recipients of knowledge; rather, they actively participate in learning experiences. This is so, since ABL is based on the constructivist theory of learning that indicates that 'humans cannot be given information, which they immediately understand and use; instead, humans must construct their own knowledge' based on their previous experiences and usually in collaboration with others (Powell & Kalina 2009, 242).

Research shows that activity based teaching facilitates learning. For instance, Harfield et al.'s (2007) study indicates that activity based teaching results in marked improvement in student engagement in the classroom and in grades in relation to those from the previous class. Besides that, Churchill's (2003) study indicates that ABL could also facilitate development of higher order thinking skills in students.

Some researchers, such as Kolb (1984), point out that demonstrative activity based teaching as compared to conventional ways of teaching, is more suitable for facilitating learning. Domin (2007) also states that teachers could provide successful learning experiences to learners through engaging them in activities.

The study of Zahoor-ul-Haq et al., (2015) explored that students actively participated in the language classroom who were taught through activity based learning method (ABL) because activity based instruction provided students many opportunities to develop their listening skill. Students taught through activity based teaching method outscored students who were taught through traditional language teaching methods in the listening skill on post-test. The low achievers of experimental group showed a significant jump over the low achievers of control group on post-test in listening. The results also proved that high achievers who were taught through activity based teaching method showed better performance in listening than those high achievers who were instructed through conventional/traditional way of language teaching.

Speaking

Speaking is the 'ability to carry out a conversation in the language' (Numan, 1991, 12). Speaking is regarded as a vital language skill (Grainger 2000), since the most important function of a language is 'facilitating communication with others' (Littlewood, 1992, 9). Indeed, the art of speaking is considered 'the single most important' (Numan, 1991, 39) and 'most rewarding' (Haley & Austin, 2004, 20) aspect of language learning.

Learning to speak, whether in a first or other language, necessitates that students develop linguistic and sociolinguistic aptitudes (Mahbub-ul-Alam & Khan, 2014). To elaborate, learning to speak a language requires learners to use accurate grammar, pronunciation, and vocabulary. Moreover, it involves developing knowledge about when and how to communicate (Burns & Seidlhofer, 2002). Since learning to speak is

challenging, teachers need to carefully select a task to give students speaking practice and should guide the students when and how to perform the task (Mercer, 1998). Moreover, the teacher should not only be encouraging, but also keep in mind that learners learn while making mistakes during an assigned task (Baker & Westrup, 2000).

Speaking activities used during the study:

In this study, a few warm up activities suggested by Holmes (2003) were used with the experimental group to develop and enhance students' speaking skills. Warm up activities help the students get to know each other in a comfortable atmosphere. Warm up activities can be divided into two types (Holmes, 2003). The first type gives students an opportunity to interview each student, without too much teacher supervision. Conducted in pairs, students often find these activities non-threatening since they have only one listener (Klippel, 1984). During the second type of speaking activity the students play games that are fun and rather interesting.

Activity 1: Getting to Know You Interview

In this warm up activity, students are grouped in pairs and each receives a guideline to interview each other in English. The two students in each pair then interview each other in order to fill in the blanks on the guideline. Next, each student introduces his/her partner to the class, in no more than two to three minutes, using the guideline as a memory aid. A guideline might include: Name, date of birth, place of birth, hobbies, education, like/dislikes, favorite book, favorite food, prizes/awards, travel experience,

Activity 2: Simon Says…

This fun game is based on carrying out the actions of simple sentences. In this activity students stand in a circle and one of the nominated students gives commands for the other students to follow. Examples of commands include: Simon says, 'Touch your eyes', Simon says, 'Open your mouth', Simon says, 'Show me your hands'. What students must remember during this activity is that if the command begins with 'Simon says', they must follow the order. If a command does not begin with 'Simon says' students must not carry out the order, or else they will get disqualified and drop out of the circle. Although this game is usually played by 10-16 year olds, it can be played with more mature students, making the commands more demanding, such as Simon says, 'Tell us about your father-in-law's profession'.

Activity 3: It's in the Bag

For this activity, a durable plastic bag that has enough capacity to contain about twenty items is bought. In this bag the teacher puts items such as plastic fruit, balls, a cell phone, toffee, a battery cell, an eraser, a pen, or any other item that is not scary or dangerous. Students are asked to put their hand in the bag, blindly select one item, and then try to describe it to other students by feeling it. They can describe the shape, texture, weight, size, and material of what they have in their hand. The student can keep describing the selected item until the other students guess what it is.

Activity 4: The One-Minute Game

In this game the class is divided into two teams that compete against each another. The game starts when a member of one team is given an impromptu topic to speak on for one full minute. This member is supposed to speak on the given topic without making any grammatical mistakes, stopping or hesitating, mispronouncing the words, or using inappropriate vocabulary. While the speaker speaks, the members of the opposing team listen carefully and disqualify the speaker if he/she makes any of the previously mentioned mistakes. The student of the opposing team who points out the error is then given a topic to speak on for one minute without making any mistakes. Any speaker who fluently and accurately speaks for one minute scores one point for his/her team.

Holmes (2003) suggests that context appropriate topics could be more engaging and fun. A few suitable topics for a Pakistani classroom could be:

How can the electric power shortage crisis in Pakistan be solved?

If you were elected as the prime minister of Pakistan, what would be the first problem you would solve?

How can tourism in Pakistan be enhanced?

Activity 5: Detective

In this very interesting game, three students who claim to have had the same unusual experience stand at the front of the room. Out of these three students, only one student had the real experience while the other two are imposters. Each member of the class acts as a detective and asks one question from the three students about their experience to catch the two imposters. At the end of the questioning, the class votes on which student they think had the unusual experience. Then the student who had the unusual experience steps forward.

Purpose of the Study

This study investigates the effect of activity based learning on the development of speaking skills of low and high achievers in a 6th grade class at the elementary school level. More specifically, the study tests the following null hypotheses:

Ho1: There is no significant difference between the mean scores of experimental and control groups with respect to achievement in speaking.

Ho2: There is no significant difference between mean scores of high achievers and low achievers of experimental and control groups with respect to achievement in speaking skills.

In a Pakistani context, the development of speaking skills through activity based learning is a new initiative since most language teachers in Pakistani public sector elementary schools use conventional teaching methodologies in their classrooms (Khan, 2011).

Method

Research Design

In line with the purpose of the study a quantitative approach was employed and pretest-posttest equivalent group experimental design was used. The experimental research in language learning is usually 'conducted within a language classroom, which can be viewed as a real-life laboratory' and aims to understand aspects of language learners' learning in a controlled environment (Phakiti, 2014, 2).

Research Sample

A sample of 50 male students was selected randomly from the two sections of class 6 of Government High School Tarkha district Nowshera..For this purpose a teacher-made pre-test was served to the sample. On the basis of students achievements in pre-test scores two equal groups i.e. experimental and control groups were formed. Further, low achieving and high achieving students were also identified in both groups. Those students who achieved who were above the mean scores were named as high achieving students and those who were below the mean were named low achieving students in both experimental and control groups.

Research Instrument

An achievement test developed by the researchers was used to measure to what extent the participants developed their speaking skills. This test was administered on the participants twice, as a pre-test and a post-test. The pre-test was administered to distribute the participants into experimental and control groups. The post-test was administered to the participants at the completion of the study.

A table of specifications was prepared for the purposes of test development. Based on the specifications table, 10 test items related to speaking skills were written for the selected lessons. The qualitative data was converted into quantitative data by giving specific score to the specification table. The draft test was consulted for content validity by an expert panel of academicians who specialized in English language teaching, including the English language and other language experts. Some test items were revised based on their feedback. The reliability of the test was measured by using the split-half (odd-even) technique. For this purpose, the test items were divided into halves, ensuring that each half was matched in terms of item difficulty and content. Each half was marked separately. The reliability was calculated by using the Spearman-Brown formula: Reliability = $2r/1+r$, where r = the actual correlation between the halves of the instrument. The alpha reliability coefficient of the test was estimated to be 0.88.

Research Procedure

Two teachers from G.H.S. Tarkha, district Nowshera, who had masters in English from the University of Peshawar, were hired for the study. Both had relatively equal teaching experience and teaching competencies. The teacher who volunteered to teach the experimental group was already trained by the DCTE and KPK in teaching English through activity based learning.

Lesson plans were prepared in accordance with the activity based learning method for implementation in the experiment group. The four p's(preparation, presentation, practice, and production) lesson plan format was used for designing the lessons. The lesson plans were consulted with the supervisor and other language experts who specialized in activity based learning. Based on their feedback, some of the lesson plans were revised.

To make the lessons relevant and interesting, activities were selected from the British Council's E.T.T.E (English for Teaching, Teaching for English) Project. These activities included Simon Says, Chinese Whispers, Name Revision Ball Game, Action Song (heads, shoulder…), Words, Role Play, and Dialogues. The activities provided in the textbook were also used in the lessons. Since English is taught as a second/foreign language in Pakistan, in the first two lessons are only warm up activities and reconducted to motivate and engage students. Later, other activities are introduced during the lessons. The treatment was done for seven weeks between 1 October 2014 and 20 November 2014. The duration of each lesson was forty minutes.

Students in the control group were taught using the conventional teaching method for seven weeks. For teaching purposes, the teacher used the activities given in the assigned textbook.

The first researcher and other language experts observed the teaching of both the teachers. The views of students on their learning experience were also sought during the lessons.

Data Analysis

Relevant data was analyzed to test the hypothesis. Mean, standard deviation, and difference of means were computed for each group. To measure the significance of the difference between the means of the two groups, a t-test of independent sample was applied. Significance of difference between the mean scores of both the experimental and control groups on the variable of pre-test and post-test scores was tested at a 0.05 level.

Collected data were fed into the ''statistical package for social sciences'' (SPSS) program. Data were analyzed by applying the t-test for independent samples.

Results

The significance of difference between the mean scores of the experimental and control groups were found on the pre-test and post-test by applying the t-test. Obtained results, along with analysis and interpretation, are presented below.

Table 1

Significance of Difference Between the Mean Scores of the Experimental and Control Groups on Pre-Test with Respect to Achievement in Speaking

Group	***N***	***Mean***	***SD***	***t-value***	
				Table value	***Calculated value***
Experimental	*25*	*4.32*	*3.15*	*1.68*	*0.30**
Control	*25*	*4.04*	*3.45*		

**Not Significant d.f.=48 Significance level = 0.05*

Table 1 indicates that the calculated result of t was 0.30 and the table value of t was 1.68. Results were tested at 0.05 (level of significance) and the degree of freedom was 48.Hence, the table value of t (1.68) was greater than the t (0.30) obtained value. Thus, Ho1 was accepted because no significant difference between the mean scores was found. In this way, the experimental and control groups were similar with respect to previous knowledge of speaking skills on the pre-test.

Table 2

Significance of Difference Between the Mean Scores of Low Achievers of the Experimental and Control Groups on Pre-Test with Respect to Achievement in Speaking

Group	***N***	***Mean***	***SD***	***t-value***	
				Table value	***Calculated value***
Low achievers of the experimental group	*14*	*1.78*	*0.97*	*1.703*	*0.949**
Low achievers of the control group	*15*	*1.5*	*0.83*		

*Not Significant d.f. = 27 Significance level = 0.05

Table 2 reflects that the obtained result of t was 0.949 and the table value of t was 1.703. Results were tested at 0.05 (level of significance) and degree of freedom was 27. Hence, the table value of t (1.703) was greater than the t (0.949) obtained value. Thus, Ho1 was approved because no significant difference between the mean scores was found. Hence, the low achievers of the experimental and control groups were the same with respect to prior knowledge of speaking skills on the pre-test.

Table 3

Significance of Difference Between the Mean Scores of High Achievers of The Experimental and Control Groups on Pre-Test with Respect to Achievement in Speaking

Group	N	Mean	SD	t-value	
				Table value	Calculated value
High achievers of the experimental group	11	7.55	1.44	1.729	-0.5*
High achievers of the control group	10	7.9	1.79		

*Not Significant d.f. =19 Significance level = 0.05

Table 3 indicates that the obtained result of t was -0.5 and the table value of t was 1.729. Results were tested at 0.05 (level of significance) and degree of freedom was 19.Hence, the table value of t (1.729) was greater than the t (-0.5) calculated value. This is why Ho1 was approved: because no significant difference between the mean score was found. In this way, the high achievers of the experimental and control groups were identical with respect to achievement in speaking skill on pre-test.

Table 4

Significance of Difference Between the Mean Scores of the Experimental and Control Groups On Post-Test with Respect to Achievement in Speaking

Group	N	Mean	SD	t-value	
				Table value	Calculated value
Experimental	25	21.72	4.05	1.68	8.319*
Control	25	10.48	5.40		

* Significant d.f.=48 Significance level = 0.05

Table 4 shows that the obtained result of t was 8.319 and the table value of t was 1.68. Results were tested at 0.05 (level of significance) and the degree of freedom was 48.Hence, the table value of t (1.68) was less than the t (8.319) obtained value. This is why Ho2 was discarded: because a significant difference between the mean scores of the experimental and control groups was found. The group taught through activity based learning showed dominance over the control group in the speaking skills on the post-test.

Table 5

Significance of Difference Between the Mean Scores of The Low Achievers of the Experimental and Control Groups on Post-Test with Respect to Achievement in Speaking

Group	N	Mean	SD	t-value	
				Table value	Calculated value
Low achievers of the experimental group	14	17.92	5.28	1.703	7.992*
Low achievers of the control group	15	6.66	1.34		

* Significant d.f. = 27 Significance level = 0.05

Table 5 indicates that the obtained result of t was 7.992 and the table value of t was 1.703. Results were tested at 0.05 (level of significance) and the degree of freedom was 27.Hence, the table value of t (1.703) was less than the t (7.992) obtained value. This is why Ho2 was discarded: because a significant difference was found between the mean scores of the low achievers of the experimental and control groups. In this way, the low achievers who were taught through activity based learning showed superiority over the low achievers of the control group with respect to achievement in speaking skills on the post-test.

Table 6

Significance of Difference Between the Mean Scores of High Achiever of the Experimental and Control Groups on Post-Test with Respect to Achievement in Speaking

Group	N	Mean	SD	t-value	
				Table value	Calculated value
High achievers of the experimental group	11	24.72	3.46	1.729	5.383*
High achievers of the control group	10	16.20	3.79		

*Significant d.f. = 19 Significance level = 0.05

Table 6 shows that the obtained result of t was 5.383 and the table value of t was 1.729. Results were tested at 0.05 (level of significance), while degree of freedom was 19.Hence, the table value of t (1.729) was less than the t (5.383) obtained value. This is why Ho2 was discarded: because a significant difference between the mean scores of high achievers of the experimental and control groups was found. In this way, the low achievers who were taught through activity based learning outscored the low achievers of the control group with respect to achievement in speaking skills on the post-test.

Discussion and Conclusion

According to the results of the research, there were no significant differences in the pre-test scores of speaking skills between the experiment and control groups. However, the experimental group performed significantly better than the control group on the post-test with respect to achievement in speaking. The difference between the post-test mean scores of both groups was significant at the (0.05) level. Similarly, low achievers and high achievers of the experimental group outscored the control group with respect to achievement in speaking skills on the post-test. Therefore, the null hypothesis was abandoned. It was concluded that activity based learning significantly increased the level of student achievement in speaking skills. The results supported the findings of Bailey (2005) and Songsiri (2007), who stated that speaking ability and self-belief in speaking might be enhanced if a suitable program of study, teaching methods, adequate activities, and resources could be provided to students. The literature also suggests that teachers should conduct a variety of speaking activities in the classroom to enhance their speaking abilities (Zhang, 2009). This study confirms the views of Zahoor-ul-Haq et al. (2015) who were of the opinion that low achievers who had learned through activity based learning outscored control group in the language skills acquistion.

Recommendations

This study concluded that activity based learning was effective in enhancing student speaking skills. Based on the results of the present study, it is recommended that in order to enhance the speaking skills of students, teachers should use the activity based learning method in the classroom. It is also recommended that teachers should be provided training to implement the activity based learning method in language lessons.

References

Abdelhamid, T. S. (2003). Evaluation of teacher-student learning style disparity in construction management education. *Journal of Construction Education*, *8*(3), 124-145.

Bailey, K. M. (2005). *Practical English language teaching: Speaking*. New York: McGraw-Hill.

Baker, J., & Westrup, H. (2000). *The English language teacher's handbook*: London: Bloomsbury Group: Continuum.

Brown, G., & Yule, G. (1999). *Teaching the spoken language*. Cambridge, UK: Cambridge University Press.

Brown, H. D. (1994). *Teaching by principles: an interactive approach to language pedagogy*. Englewood Cliffs, NJ: Prentice Hall Regents.

Burns, A., & Seidlhofer, B. (2002). Speaking and pronunciation. In N, Schmitt (ed). *An introduction to applied linguistics* (pp. 211-232). Routledge.

Burns, A., & Joyce, H. (1997). *Focus on speaking*. Sydney: National Center for English Language Teaching and Research.

Churchill. D. (2003). *Effective design principles for activity-based learning: the crucial role of 'learning objects' in Science and engineering education.* Retrieved of October 10, 2011 from http://www.learnerstogether.net/PDF/Effective-Design-Principles. pdf

Domin, D. S. (2007). Students' perceptions of when conceptual development occurs during laboratory instruction. *Chemistry Education Research and Practice*, *8*(2), 140-152.

Goh, C. (2007). *Teaching speaking in the language classroom*. Singapore: SEAMEO Regional Language Centre.

Haley, M. H., & Austin, T. Y. (2013). *Content-based second language teaching and learning: An interactive approach.* Boston: Pearson Education.

Harfield, T., Davies, K., Hede, J., Panko, M., & Kenley, R. (2007). Activity-based teaching for Unitec New Zealand construction students. *Emirates Journal for Engineering Research,* 12(1), 57-63.

Hein, G. (1991). Constructivist learning theory. Retrieved from on 19 November 2011.

http://www.exploratorium.edu/IFI/resources /constructivistlearning.html.

Holmes, D. (2003). *Speaking activities for the classroom.* Bangkok.

Hull, D. (1999). *Teaching science contextually*. Retrieved on December 04, 2011 from: http://www.cord.org/uploadedfiles/ Teaching_ Science_ Contextually.pdf

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Englebrook, NJ: Prentice Hall.

Littlewood, W. (1992). *Teaching oral communication: A methodological framework*. England: Blackwell Publishers.

Mahbub-ul-Alam, A., & Khan, M. E. I. (2014). Speaking in a second language in Bangladesh. *NU Journal of Humanities, Social Sciences & Business Studies, 1*(1), 135-142.

Mercer, N. (1998). English as a classroom language. In N. Mercer & J. Swann (Eds.), *Learning English development and diversity* (pp. 119-147). New York: Routledge

Murray, P., Donohoe, S., & Goodhew, S. (2004). Flexible learning in construction education: A building pathology case study. *Structural Survey*, *22*(5), 242-250.

Numan, D. (1991). *Language teaching methodology: A textbook for teachers.* New York: Prentice Hall.

Panko, M., Kenley, R., Davies, K., Piggot-Irvine, E., Allen, B., Hede, J. & Harfield, T. (2005). *Learning styles of those in the building and construction sector.* Report for Building Research New Zealand, Inc. Unitec New Zealand, Auckland.

Patil, Z. N. (2008). Rethinking the objectives of teaching English in Asia. *Asian EFL Journal*, *10*(4), 227-240.

Phakiti, A. (2014). *Experimental research methods in language learning.* New York: Bloomsbury publishing.

Powell, K. C. and Kalina, C. J. (2009). Cognitive and social constructivism: Developing tools for an effective classroom. *Education,* 130(2), 241-250.

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education,* 93(3), 223-231.

Songsiri, M. (2007). *An action research study of promoting students' confidence in speaking English*. (Dissertation of Doctor of Education Degree), School of Arts, Education and Human Development, Victoria University, Australia. Retrieved from eprints.vu.edu.au/1492/1/Songsiri.pdf

Suydam, M. N., & Higgins, J. L. (1997). *Activity-based learning in elementary school mathematics: Recommendations from research*. Columbus, Ohio: ERIC/SMEE.

Wyse, D., & Jones, R. (2001). *Teaching English language and literacy.* London: Taylor & Francis Group.

Zahoor-ul-Haq; Khan, A., & Tabassum, R. (2015). Effect of ABL method on students' performance in listening skill at Grade-VI. *Journal of Humanities & Social Sciences (Pakistan), 23*(3), 95-108.

Zhang, Y. (2009). Reading to speak: Integrating oral communication skills. *English Teaching Forum*, *47*(1), 32-34.

Submission Checklist

Makale Sunumu Kontrol Çizelgesi

Indicate that this submission is ready to be considered by this journal by checking off the following.

Aday makalenin değerlendirilmeye hazır olduğunu aşağıdakilerin her biri ile karşılaştırarak kontrol ediniz.

Manuscript Agreement:

I / We hereby accept that, the article after being accepted for publication in the Eurasian Journal of Educational Research (EJER), the author (s) as, all rights related to the article has been transferred to the ANI PUBLISHING under the laws of the, "copyright transfer".

Makale Eurasian Journal of Educational Research (EJER) dergisinde basıma kabul edildikten sonra, yazar(lar) olarak; makale ile ilgili tüm hakları, "Telif Hakkı Devir" yasaları uyarınca, ANI YAYINCILIĞA devrettiğimizi kabul ediyoruz.

Manuscript Submission:

I have prepared my manuscript based on the criteria listed below and I accept all submission conditions.

Makalemi aşağıda belirtilen kriterlere göre hazırladım ve makale gönderme koşullarının tamamını kabul ediyorum.

Indicate that this submission is ready to be considered by this journal by checking off the following.

Aday makalenin değerlendirilmeye hazır olduğunu aşağıdakilerin her biri ile karşılaştırarak kontrol ediniz.

1	☒	*The manuscript is a report of original educational research or a discussion article on research topics in education.* Aday makale, eğitim alanıyla doğrudan ilgili bir araştırma ya da tartışma yazısıdır.
2	☒	*The submission has not been previously published, nor sent to another journal for consideration.* Sunulan çalışma daha önce herhangi bir yerde yayınlanmamıştır, başka bir derginin incelemesinde değildir.

3	☒	*Within a year, I have not submitted a manuscript to EJER as an author or co-author for review other than this manuscript.* Son bir yıl içerisinde, yazarı olduğum ya da yazarları arasında bulunduğum başka bir çalışma değerlendirilmek üzere EJER'e sunulmamıştır.
4	☒	*All authors of this manuscript are subscribers of the EJER and accept to pay **400 EURO** for the cost of proofreading and 10 loose copies of their articles if the manuscript is accepted.* Aday makalenin kabul edilip basılması halinde makalelerinin İngilizce son okuma (proofreading) ve 10 serbest kopya gönderimi nedeniyle oluşan **1275 TL** maliyeti Anı Yayıncılık'a ödemeyi kabul ederler.
5	☒	*The entire manuscript is written in English.* Aday makalenin bütününün yazım dili İngilizce'dir.
6	☒	*The original manuscript is typed on A4 paper. The margins are 2.5 cm.* Aday makale kenar boşlukları 2.5 cm olan A4 kağıda yazılmıştır.
7	☒	*Each paragraph is longer than two sentences.* Her bir paragraf en az üç cümle içermektedir.
8	☒	*The entire manuscript - including quotations, references, author note, content footnotes, figure captions, and all parts of tables – is double-spaced.* Aday makalenin tamamı, alıntılar, kaynakça, şekil ve tablo başlıkları da dâhil olmak üzere çift aralıklı yazılmıştır.
9	☒	*The submission file is in Microsoft Word document file format. 12-point Times New Roman font is used in entire manuscript.* Aday makale, tamamında 12 punto Times New Roman yazı tipi kullanılarak hazırlanmış bir Microsoft Word dokümanıdır.
10	☒	*The text has had the authors' names removed. If an author is cited, "Author" and year are used in the bibliography and footnotes, instead of author's name, paper title, etc. The author's name has also been removed from the attached document.* **Aday makale, yazar adları çıkarılarak sunulmuştur**. Eğer yazar kendisine atıfta bulunduysa yazarın adına ve çalışma başlığına yer verilmeyecek, sadece "Author" yazılarak çalışmanın yılı belirtilecektir. Eklenen dosyada yazar adı belirtilmeyecektir.
11	☒	*The title is 10 to 12 words.* Aday makalenin başlığı 10-12 sözcük uzunluğundadır.

12	☒	*The maximum length of the manuscript-including structured abstract in English, tables, and references is 6000 words. This limitation does not include Turkish extended abstract (750-1000 words) which is placed after the references section.* Aday makale, İngilizce abstract, tablolar ve kaynakça vb. tüm ögeler dâhil olmak üzere en fazla 6000 sözcüktür. Kaynakça'nın ardından yer verilen uzun Türkçe özet (750-1000 sözcük) bu sayıya dâhil değildir.
13	☒	*The article is preceded by* ***English Structured Abstract of not more than 250 words*** *and not less than 200 using five required headings:* **Purpose***: State the problem in field. Then explain the purpose of the study.* **Method:** *Specify the research design, sample, and research instrument and data analysis in brief.* **Findings***: Highlight the significant, interesting or surprising results.* **Implications for Research and Practice** . (These headings may need some adaptation in the case of discussion papers: *Background, Purpose of Study, Sources of Evidence, Main Argument, and Conclusions)*. More information available from (http://www.tandf.co.uk/journals/authors/rereabstracts.asp) Yapılandırılmış İngilizce öz 200-250 sözcük uzunluğunda olup, aday makalenin başında yer almakta ve **Purpose** *(İlk önce alanda karşılaşılan sorunu belirtelim. Daha sonra araştırmanın amacını bir cümle ile veriniz),* **Method** *(Araştırma deseni, örneklem, very taoplama aracı ve verilerin analizinini kısaca açıklayınız),* **Findings** *(En önemli ve çarpıcı araştırma bulgularını verelim)***Implications for Research and Practice, (***Uygulama ve ileriye dönük araştırmalar için olası çıkarımlarınız)* başlıklarını içermektedir. Bu başlıklar tartışma yazıları için: *Çalışmanın Temeli, Çalışmanın Amacı, Kanıt Kaynakları, Ana Tartışma ve Sonuçlar* şeklinde olabilir. Daha fazla bilgi için http://www.tandf.co.uk/journals/authors/rereabstracts.asp adresine başvurunuz.
14	☒	*Following the structured abstract in English four to six keywords are included. They should* ***represent*** *the content of your manuscript and* ***be specific*** *to your field or sub-field. Avoid using keywords form the title of the paper.* Yapılandırılmış İngilizce özden sonra 4-6 anahtar sözcüğe yer verilmiştir. Anahtar kelimeler çalışmanızı temsil etmeli ve kendi alanınıza ya da alt alanlara özgü olmalıdır. **Makale adındaki kavramları anahtar kelime olarak seçmekten kaçınınız**.
15	☒	*An extended (750-1000 words) Turkish structured abstract is placed following the "References" section using five required headings: Problem Statement, Purpose of* *Study, Methods, Findings and Results, and Conclusions and Recommendations*. (These headings may need some adaptation in the case of discussion papers: *Background, Purpose of Study, Sources of Evidence, Main Argument, and Conclusions)*. More information

		available from http://www.tandf.co.uk/journals/authors/rereabstracts.asp Kaynakça'dan sonra 750-1000 sözcükten oluşan Türkçe yapılandırılmış öze yer verilmiştir. Türkçe yapılandırılmış öz *Problem Durumu, Araştırmanın Amacı, Araştırmanın Yöntemi, Araştırmanın Bulguları, Araştırmanın Sonuçları ve Önerileri* başlıklarını içermektedir. Bu başlıklar tartışma yazıları için: *Çalışmanın Temeli, Çalışmanın Amacı, Kanıt Kaynakları, Ana Tartışma ve Sonuçlar* şeklinde olabilir. Daha fazla bilgi için; http://www.tandf.co.uk/journals/authors/rereabstracts.asp
16	☒	*Following the Turkish structured abstract, four to six keywords are included.* Uzun Türkçe özetten sonra 4-6 anahtar sözcüğe yer verilmelidir.
17	☒	*References are not cited in the structured abstracts in English and in Turkish.* İngilizce abstract ve Türkçe öz içerisinde atıfta bulunulmamıştır.
18	☒	*The format of headings, tables, figures, citations, references, and other details follow the **APA 6 style** as described in the Publication Manual of the American Psychological Association, 6th edition, available from http://www.apa.org* Aday makalenin başlıkları, tabloları, şekilleri, atıfları, kaynakçası ve diğer özellikleri tamamen APA altıncı baskıda belirtildiği şekildedir.
19	☒	*All illustrations, figures, and tables are placed within the text at the appropriate points, rather than at the end.* Aday makalenin şekilleri ve tabloları metin içerisinde bulunmaları gereken uygun yerlere yerleştirilmiştir. Makale sonunda sunulmamıştır.
20	☒	Citations in the text of the document include the author's surname, the year of publication, and, when there is a specific quote from a source used, a page number where the quote is located in the text. Example: Nothing seemed so certain as the results of the early studies (Tatt, 2001, p. 445). It was precisely this level of apparent certainty, however, which led to a number of subsequent challenges to the techniques used to process the data (Jones & Wayne, 2002, p. 879). There were a number of fairly obvious flaws in the data: consistencies and regularities that seemed most irregular, upon close scrutiny (Aarns, 2003; West, 2003, p. 457). With studies by two authors, always include both author names:

		(Anderson & Bjorn, 2003) As Anderson and Bjorn (2003) illustrated in their recent study As recently as 2003, a prominent study (Anderson & Bjorn) illustrated When a study has 3, 4, or 5 authors, include the names of all the authors the first time the work is cited: (Anderson, Myers, Wilkes, & Matthews, 2003) For all subsequent citations of this work, use "et al.": (Anderson et al., 2003) When a work has 6 or more authors, use et al.: (Bell et al., 2003) For unsigned works, include the title, enclosed in parentheses. Put quotation marks for short work titles, and italicize the titles of reports, books, and other significant works: ("Recent Developments," 2004) (Dictionary of Tetrathalocigistic Diseases, 2004) Metin içindeki atıfları üstte verilen örneklere uygundur.
21	☒	Three levels of headings are used: Level 1, Level 3 and Level 4. The headings are formatted as follows: Centered Uppercase and Lowercase Heading (Level 1) *Flush Left, Italicized, Uppercase and Lowercase Side Heading* (Level 3) *Indented, italicized, lowercase paragraph heading ending with a period.* Start writing after the period (Level 4). Aday makale içerisinde üç farklı düzey başlık kullanılmıştır. Düzey 1, Düzey 2, Düzey 3. Başlıklar bu düzeylere uygun olarak aşağıdaki şekilde biçimlendirilmiştir: Ortalı ve Her Sözcüğün İlk Harfi Büyük Yazılmış Başlık (Düzey 1) *Tam Sola Dayalı, İtalik ve Her Sözcüğün İlk Harfi Büyük Yazılmış Başlık* (Düzey 3) *İçeriden, italik, tamamı küçük harflerle yazılmış ve nokta ile bitten başlık.* Noktadan sonra normal metin yazımına devam edilmeli (Düzey 4).
22	☒	References are listed in alphabetical order. Each listed reference is cited in text, and each text citation is listed in the References. Basic formats are as follows: Haag, L., & Stern, E. (2003). In search of the benefits of learning Latin. *Journal of Educational Psychology, 95,* 174–178.

		Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley. Johnson, D. W., & Johnson, R. T. (1990). Cooperative learning and achievement. In S. Sharan (Ed.), *Cooperative learning: Theory and research* (pp. 173–202). New York: Praeger. **Turkish References Only:** Çınkır, Ş., & Çetin, S. K. (2010). Öğretmenlerin okullarda mesleki çalışma ilişkileri hakkındaki görüşleri [Teachers' opinions about the professional working relationships in schools]. *Kuram ve Uygulamada Eğitim Yönetimi, 1 6*(3), 353-371. **Article in an Internet-only journal/Periodical, database** Fredrickson, B. L. (2000, March 7). Cultivating positive emotions to optimize health and well being. *Prevention & Treatment, 3,* Article 0001a. Retrieved November 20, 2000, from http://journals.apa.org/prevention/volume3/pre0030001a.html More information is available from: http://citationonline.net/CitationHelp/csg04-manuscripts-apa.htm#references Kaynakçanın yazımı üstte verilen örneklere uygundur.
23	☒	*Order of the main parts in the manuscript is as follows:* *Main title in English (max. 12 words)* *Structured abstract (min. 200- max.250 words length)* *Keywords (in English, min. four-max. six)* *Main text* *References* *Main Title in Turkish (max. 12 words* *Extended structured abstract (min.750-max.1000 words length in Turkish)* *Keywords (in Turkish, min. four-max. six)* Aday makaleyi oluşturan ana öğeler aşağıdaki düzendedir: İngilizce Ana Başlık (En fazla 12 sözcük) Yapılandırılmış İngilizce Abstract (En az 200, en fazla 250 sözcük) Anahtar Sözcükler (İngilizce, en az dört, en fazla altı) Ana Metin Kaynakça

		Türkçe Ana Başlık (En fazla 12 sözcük) Yapılandırılmış Türkçe Öz (En az 750, en fazla 1000 sözcük) Anahtar Sözcükler (Türkçe, en az dört, en fazla altı)
24	☒	**Structure of the Manuscript should be as follows:** Manuscript Title **English Abstract** (200-250 words) Sub-headings: *Purpose:…* *Purpose of Study:…* *Method:….* *Findings:….* *Implications for Research and Parctice:…* *Keywords:..* **Introduction Method** (sub-headings): *Research Design* *Research Sample* ♦*Research Instrument and Procedure* *Data Analysis* **Results, Discussion and Conclusion, References** **Extended Turkish Abstract (750-1000 words) Sub-headings:** *Problem Durumu:….* *Araştırmanın Amacı:….* *Araştırmanın Yöntemi:…* *Araştırmanın Bulguları:…* *Araştırmanın Sonuçları ve Önerileri:…..*
25	☒	***EJER editor reserves the right to reject manuscripts and not to publish it that do not comply fully with the EJER Manuscript Submission Guidelines.***

♦ Reliability and the validity of the research instrument used or adapted in the work must be provided, and explained in detail.

SUBSCRIPTION/ABONELİK

Annually / Yıllık

***İndividual** / Kişi Aboneliği ..... : 80 TL*

***İnstitution** / Kurum Aboneliği .....: 250 TL*

***Abroad** / Yurtdışı Aboneliği: .....: 250 $*

***Student** / Öğrenci Aboneliği: .....: 70 TL*

For single issue / Tek sayılar için

İndividual / *Kişiler .....: 25 TL*

***İnstitution** / Kurumlar .....: 75 TL*

Payment / Ödeme

İş Bankası Kızılay Şubesi, Anı Yayıncılık 4214-0962561

Account number/numaralı hesap