EURASIAN JOURNAL OF EDUCATIONAL RESEARCH

Bimontly Peer-Reviewed Journal, Year: 18 Issue: 75 / 2018
İki Ayda Bir Yayımlanan Hakemli Dergi, Yıl: 18 Sayı: 75 / 2018

EURASIAN JOURNAL OF EDUCATIONAL RESEARCH

Bimontly Peer-Reviewed Journal, Year: 18 Issue: 75 / 2018
İki Ayda Bir Yayımlanan Hakemli Dergi, Yıl: 18 Sayı: 75 / 2018

PUBLISHING MANAGER / Sahibi ve Yazı İşleri Müdürü
Anı Yayıncılık Eğitim ve Danışmanlık Reklam Kırtasiye Sanayi Ticaret Ltd. Şti. adına
Anı Publishing Education and Consultancy Advertisement Stationary Industry Trade Co. Ltd. in the name of
Özer DAŞCAN

EDITORIAL PRODUCTION ADMINISTRATOR / Yayın Yönetim
Dilek ERTUĞRUL

HEADQUARTER / Yönetim Merkezi
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: 20.05.2018
Printing Address / Matbaa Adresi: Vadi Grafik
İvedik OSB 770. Sokak No: 101 Yenimahalle-Ankara
Tel: +90 312 395 85 71
Broadcast Type / Yayın Türü: Local Broadcast / Yerel Süreli Yayın
Cover Design / Kapak Tasarımı: Anı Publishing / Anı Yayıncılık
Composition / Dizgi: Göksel ÇAKIR

Eurasian Journal of Educational Research (Print ISSN 1302-597X, e-ISSN: 2528-8911) is a bimonthly (6 issues per year) peer-reviewed journal published by Anı Yayıncılık.

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 75th ISSUE
75. Sayı Hakemleri

Begüm Aytemur

Burcu Pehlivan Tunç

Deha Doğan

Duygu Anıl

Emrah Gül

Fazilet Taşdemir

Funda Nayır

Gülşen Taşdelen Teker

Hasan Çakır

Kıvanç Aycan

Leman Tarhan

Mehmet Tekerek

Mesut Türk

Murat Özdemir

Münevver Başman

Nilgün Tosun

Saadet Kuru Çetin

Sema Buz

Şenel Poyrazlı

Tuncay Öğretmen

Tuncay Sarıtaş

Vesile Alkan

Eurasian Journal of Educational Research 75 (2018) 1-18

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

A Review on Effectiveness of Cello Etudes Created for Longas and Syrtos

Burcu AVCI AKBEL [1]

ARTICLE INFO

Article History:
Received: 17 Jan. 2018
Received in revised form: 22 Mar. 2018
Accepted: 20 Apr. 2018
DOI: 10.14689/ejer.2018.75.1

Keywords
Turkish music, instrumental music, cello education, preparatory exercises.

ABSTRACT

Purpose: This study has been performed with a view to find a solution to challenges experienced in cello performances of pieces in the forms of Longas and Syrtos within the category of instrumental Turkish Music – due to their inherent agility, technical difficulties, etc., to ensure that the pieces can be more precisely and easily played in technical and musical terms. **Methods:** Both descriptive and experimental methods are used in this study. As for the preparatory exercises, the etude writing model adapted from the Taba-Tyler model is used. Longas and Syrtos, which will be examined in the scope of this study, were determined through interviews held with cello teachers at Turkish music conservatories in Turkey, whereas an experimental method was used for the second part of the study. In this study, a pretest-posttest control group model (N=8) was used. The data obtained upon the performance evaluation were compared by applying the Wilcoxon signed-rank test and the Mann–Whitney U test through the SPSS program. **Findings:** According to the statistical data obtained, significant differences were found in the arithmetic means of the experimental group, whereupon it was concluded that the preparatory exercises created according to this result are highly effective. **Implications for Research and Practice:** Within the scope of this study, etudes were created to facilitate the performance of the pieces, and a significant difference was observed in the performances of the students. Many other etudes can be written to overcome challenges in other genres and observe their effect on student success.

[1] Yıldırım Beyazıt University, TURKEY, e-mail:burcuavci812002@yahoo.com,
ORCİD: https://orcid.org/0000-0002-3128-9295

Introduction

Turkish Music is grouped into two forms: instrumental music and vocal music, and these are further divided into various categories. Although vocal music is at the forefront of Turkish Music, there are also certain forms that are specific to instruments, such as *Pesrev*, *Taksim*, *Medhal*, *Saz Semâi, Longa*, *Syrto*, and Traditional Dance Music.

Longas and Syrtos are forms of instrumental music. İsmail Hakki Ozkan (2000) defines the term Longa as follows: "Longa refers to swift and lively traditional dance music, mainly measured at a rhythmic pattern of 2/4 as *Nîm Sofyan usûl*, typically in 2, 3 or 4 hanes (parts of a musical piece), with or without teslim (the part returned to after every hâne)". Ozkan (2000) defines the term Syrto as follows: "Syrto refers to musical pieces composed for stringed instruments, with a relatively free structure compared to Longas, with or without hânes, allowing *usûl geckisi* (rhythm transitions) in final hânes, and having a very swift minor *usûl*, basically composed to accompany a dance called Syrto. Syrtos have their special rhythmic patterns" (p.81). As these works, which have their origins in the Balkans, were often composed with maqams in Turkish Music, they went through a transformation within Turkish Music and thus inherited the character of Turkish Music. Although Longas and Syrtos are considered Traditional Dance Music, they have often been used as energetic music for stringed instruments but not performed as a *raks* (dance) element in Turkish Music. All of these musical pieces involve agility, instrumental control, high level of technical knowledge, experience, and practice.

A performer must possess the ability to overcome technical difficulties arising from the structure of the related instrument. Today's performers being trained in Turkey are expected to be capable of using all technical possibilities of the instrument in the western sense while at the same time remaining familiar to their own culture and music. In this regard, it has been concluded that Longas and Syrtos, all of which require agility, instrumental control, high technical knowledge, experience and practice, should be supported with etudes as well as preliminary preparatory studies aimed at facilitating their cello performance.

Etudes are the shortest and the most systematic form of progress, both technically and musically, in instrument education. "Etudes refer to exercises which are performed to enhance performance in technical terms and which generally form the content of instrumental methods" (Türkel & Sen, 2016, p.914). "Etudes are used as a practice to build up knowledge and skills on style, phrasing, technique, tone, intonation and harmony" (Lancaster, 1994, p.16). As for instrument education, etudes fulfill various functions including, for example, helping in the acquisition of technical and sight-reading skills, improving speed, and developing musical sense (Ercan, 2008, p.96). As is indicated, the currently available etudes are used to teach and strengthen overall technical and musical knowledge, and the mentioned etudes should be utilized as well. The exercises prepared within the frame of this study, on the other hand, were especially written to ensure that the performer can overcome the difficulties in the selected works. The etudes function as a preparation for musical pieces, supporting

effective time management, easier and faster acquisition of targeted behaviors, and eliminating difficulties in pieces.

The preparatory etudes created in this study are aimed at fulfilling particular requirements and targets. At this point, the basic observations of the researcher are of considerable importance. For example, an instructor who observes that students make similar mistakes on the same subjects when performing a particular musical piece can create small exercises or melodic patterns in order to overcome related technical challenges (Yalcinkaya, 2010, p.27). The researcher has years of experience as a cello instructor and performer, which ensured her ability to make observations at formal and informal musical education institutions for this study. Moreover, the need for such research has been verified by interviews held with persons who are experienced and specialized in the field.

Longas and Syrtos are different from other pieces of Turkish Music both in the sense that they are included in Turkish Music forms and that they contain musical elements of Western Music, which are musical dynamics, chromatic passages, double sounds, swiftly playing passages, arpeggios, large intervals, and chords. This study is based upon the assumption that the preparatory exercises created for cello with the help of the data obtained upon analysis on Longas and Syrtos will facilitate and improve performance if such exercises are practiced in advance.

Method

Research Design

The study employed a mixed model consisting of descriptive and experimental methods. There are comparisons and checks in the experimental part of the research. Therefore, the research pattern is a pre-test-posttest control group model (N=8). Eight participants were randomly assigned to the control and experimental groups, with four students in each.

The pre-test-posttest control group model includes two groups formed through objective assignment. One of them is employed as the experimental group while the other is employed as the control group. Pre-test and posttest measurements are performed in both groups. The availability of pre-tests in the model helps recognition of the similarity levels of groups prior to the tests and correction of the results of posttests accordingly (Karasar, 2007, p.97).

As for the descriptive part of the research, the etude writing model adapted from the Taba-Tyler model was used while creating the preparatory exercises. According to this model, the following basic steps of programming must be followed when designing an etude: "a) Identifying the requirements for writing an etude, b) Identifying the distant goal, general and special objectives, c) Selecting and arranging the content special for the designated objectives, d) Selecting and arranging the learning experiences" (Yalcinkaya, 2012, p.23–24).

Below are the programming steps for this study: a) According to the data obtained from the researcher's observations and interviews held with experts, the challenges encountered in the performance of pieces in the Longa and Syrto forms by cello create the need for exercises aimed at facilitating the performance of such pieces. b) The distant goal ascertained is to introduce the forms of Longa and Syrto in Turkish Music, and also pieces and preparatory exercises in the mentioned forms to both Turkish Music performers and cellists interested in musical genres other than Turkish Music by preparing such pieces and preparatory exercises in a key that is suitable for cello and in bass clef. This research aims at serving vocational education, whereas the special objective of the research is to find solutions to the problems experienced in the performance of pieces in the Longa and Syrto forms that are included in the musical category for stringed instruments in Turkish Music – due to reasons such as agility, technical difficulties, etc. c) Opinions of experts were received while selecting the content of preparatory exercises. The experts stated that there are usually challenges experienced in terms of "coordination between the right hand and the bow, detaché bow technique, transitions between strings". Thus, the mentioned challenges provided the basis for the content of preparatory exercises. d) The preparatory exercises were created on the basis of the 'known to unknown' principle in this research. In order to ensure that these exercises would be more efficient and practiced more consciously, written explanations were provided under each exercise, describing which behaviors are meant to be improved through such exercises, and which points are to be taken into consideration.

Research Sample

The study group in the scope of this research was determined by a monographic sampling method, which is among the non-probability sampling methods. In monographic sampling, the researcher works on a cluster of single or multiple subjects, which the researcher thinks can represent the universe based on information and predictions about the universe (Ural, Kilic, 2011, p.44). The basic idea in monographic sampling is to select as a sample any sub-group of the universe, which must be described on the basis of the information available according to the assumption that the sub-group will represent the universe (Arslanturk, 1997, p.109). In light of these data, the study group was selected among senior undergraduate students and graduate students at Marmara University and Gazi University. This selection was made on the basis of directives made by their cello teachers: eight students with the capability to play the etudes in the scope of the study who are at similar levels in technical terms were selected. These students were confirmed by experts to be at technically similar levels.

Research Instrument and Procedure

The interview technique was used to determine the basis of the research, to select the musical pieces, and to reveal the challenges in the pieces. Interviews were held with three academics. Moreover, a rubric consisting of five options was prepared as another data collection tool. This rubric was designed to find out whether the basic

behaviors that the performer is expected to achieve in cello performance were achieved. The opinions of the experts were received during the preparation of the rubric; the evaluation criteria were finalized in line with their opinions. The etudes were prepared with the aim to facilitate the performance of pieces such as Longa and Syrto, and they were submitted to the experts for their opinions by means of the rubric created. The experts were asked to evaluate the results of the pre-test and posttest obtained by the study group through the rubric. The results of the pre-test and posttest obtained by the study group were examined by three different experts in the field of cello education (1 TRT Artist, 2 academics), and the rubric was filled out according to these data. Analyses were performed by using the data obtained from the rubrics and non-parametric tests; the results obtained were compared and interpreted. The data are presented in the findings section.

The first part of this research is a qualitative study of descriptive type. This part consists of two stages: During the first step, the general screening model was used and a literature review was performed to collect information about the subject with a view to determine the basis and specify the subject ascertained for the research. Scores of all Longas and Syrtos were accessed in the sheet music archive of TRT Ankara Radio. In order to determine the basis for the research, select the Longas and Syrtos to be reviewed in scope of the research, and identify the challenges encountered in the pieces, one-to-one interviews were held with three persons who are cello teachers at classical Turkish music conservatories in Turkey. According to Buyukozturk and his friends , "Semi-structured interview technique is used as it combines both fixed alternative questions and the possibility of an in-depth analysis in the relevant field". During the interview, answers were sought to a number of questions regarding which pieces of Longa and Syrto are taught, which of them pose challenges in performance, and what kind of challenges are encountered, if there is any requirement for writing preparatory exercises for such challenging pieces, which points must be taken into consideration when writing such exercises, and what kind of opinions can be further included. Through interviews held with experts, the pieces were selected and the topics were identified that needed to be focused on in the preparatory writing exercises on the basis of the topics posing challenges and the suggestions of the experts.

During the second stage, the Longas and Syrtos selected during interviews held with experts were first transposed into bass clef and then transposed into a key suitable for cello. Later, the scores of these pieces were analyzed and the parts that could be used in the cello education were selected. The content of the analysis involves identification of technical and musical elements in these works. Preparatory exercises were written to support the performance of the selected pieces according to the results of analyses on such works, and basic techniques and appropriate bow signs suitable for the works were added. Preparative exercises were created in conformity with the structure of the analyzed Longas and Syrtos, the sound field of the cello, easy playability on cello, the aim to achieve a good tone, the data obtained from interviews held with experts, and their suggestions. These exercises were created with regard to the requirements that the exercises must be aimed at overcoming the challenges

encountered in performance of Longas and Syrtos, they must be easily decipherable and universal, and must follow the 'known to unknown' and 'simple to complex', etc. principles. These exercises were finalized after technical and musical revisions were made to them, including legatos, finger numbers, ornaments, nuance signs, etc. Bass clef was used when writing the cello exercises so that this study would address all persons interested in Turkish Music or other musical genres.

An experimental method is used for the second part of the study. During this stage, the sample group was divided into two groups of four persons through unbiased assignment, including one as the experimental group and the other as the control group. During the pre-test stage, the students both in the experimental and control groups were provided with the musical pieces, and they were asked to decipher the pieces after a review of five minutes. The researcher took part in these sessions solely as the observer. During the next two weeks, the students were given training for one hour daily for five days a week. The experimental group was asked to practice the etudes while the control group was asked to practice the musical pieces, as they are instructed per the teaching method applied in the educational system at conservatories. During the posttest stage of the experiment, the musical pieces were performed both by the experimental group and the control group. Videos of the pre-test, posttest, and training stages were recorded. All performance records taken after the end of the experiment stage were evaluated by three experts in the field of cello education through the evaluation scale prepared for this research. The results were statistically interpreted.

Data Analysis

During this stage, the mean of the numerical data obtained as a result of the evaluation performed by the experts was taken and the mean scores were subjected to two statistical analyses via the SPSS program. Non-parametric statistics were used in the analysis as the sizes of the experimental and control groups were fewer than 15 individuals. The Wilcoxon signed-rank test was applied to compare the pre-test and posttest scores of the study group according to the pieces and persons. The Mann–Whitney U test was employed to find the difference between the evaluations made by the experts.

Results

Findings on Selection of the Musical Pieces

In order to select the Longa and Syrto pieces to be included in the research, it was examined whether the musical pieces of this genre are taught at Turkish music conservatories. For this purpose, three academicians lecturing at different universities were interviewed as experts, and they were asked to indicate the pieces that they used in their classes, and the pieces in Longa and Sirto forms that are challenging for students. Out of the musical pieces in Longa and Syrto form that are commonly used

in education, *Sehnaz Longa* composed by Santuri Ethem Efendi, *Nihavend Longa* composed by Kemani Kevser Hanim, *Kurdilihicazkar Longa* composed by Kemani Sebuh Efendi, and *Sehnaz Hicaz Syrto* composed by Sultan Abdulaziz are taught by all of the academics who were interviewed. Furthermore, all of the academics interviewed stated that their students had difficulties performing these pieces. Therefore, the scope of this research covers the mentioned four pieces. After the pieces were selected, the difficulty levels of these works were identified, and the etudes were prepared accordingly. The results of Kolmogorov-Smirnov test are presented in Table 1.

Findings to Be Used on Determining the Tests

Table 1

Kolmogorov-Smirnov Test

		a	b	c	d	e	f	g	TMV
PreTest	Number of observations (n)	96	96	96	96	96	96	96	96
	Test statistic	0,194	0,217	0,255	0,233	0,263	0,266	0,235	0,150
	Probability value (p)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Post Test	Number of observations (n)	96	96	96	96	96	96	96	96
	Test statistic	0,202	0,202	0,244	0,223	0,249	0,219	0,240	0,141
	Probability value (p)	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000

Note: a, b, c, d, e, f, g, express evaluation criteria. a: 'ability to play precise and clean', b: 'ability to play with rhythmic accuracy', c: 'ability to apply right-hand techniques', d: 'ability to apply left-hand techniques', e: 'ability to play with musicality', f: 'ability to play with fluidity', g: 'ability to play at the designated tempo'. H0: All variables are normally distributed. H1: All variables are not normally distributed.
TMV: Total mean values

Since the probability value (p) $<\alpha = 0,05$, the H0 hypothesis is rejected and the H1 hypothesis is accepted. The variables are not normally distributed. For this reason nonparametric hypothesis tests should be used. Therefore, the Wilcoxon signed-rank test and Mann–Whitney U test are used for the study. The scores given by each of the experts to both groups and an evaluation of the results are presented Table 2.

Findings on the Evaluation Scores Given by Experts

Table 2

Individual Evaluation Scores of Control and Experimental Groups for Experts

		Pretest								Posttest							
		g	f	e	d	c	b	a	TMV-pre	g	f	e	d	c	b	a	TMV-post
Control Group	Number of Observations	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
	Mean	3,5	4,13	4,06	3,75	3,31	3,88	3,94	26,56	3,75	4,19	4,25	4,06	3,75	4,25	4,31	28,56
Experimental Group	Number of Observations	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
	Mean	2,56	2,88	3,5	2,81	2,19	2,44	2,69	19,06	3,06	3,25	3,63	3,13	2,5	3	3,13	21,69
Total	**Number of Observations**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**
	Mean	**3,03**	**3,5**	**3,78**	**3,28**	**2,75**	**3,16**	**3,31**	**22,81**	**3,41**	**3,72**	**3,94**	**3,59**	**3,13**	**3,63**	**3,72**	**25,13**
	Mann–Whitney U (p)	**0,012**	**0,000**	**0,058**	**0,003**	**0,004**	**0,003**	**0,002**	**0,002**	**0,119**	**0,042**	**0,036**	**0,018**	**0,004**	**0,084**	**0,005**	**0,056**
Control Group	Number of Observations	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
	Mean	2,75	2,69	2,81	2,94	2,5	2	2,25	17,94	2,81	3,13	2,94	3	2,94	2,31	2,69	19,81
Experimental Group	Number of Observations	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
	Mean	2,5	1,87	1,81	2,13	1,75	1,56	1,13	12,75	4,5	4,75	4,5	4,69	4,56	4,63	4,69	32,31
Total	**Number of Observations**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**
	Mean	**2,63**	**2,28**	**2,31**	**2,53**	**2,13**	**1,78**	**1,69**	**15,34**	**3,66**	**3,94**	**3,72**	**3,84**	**3,75**	**3,47**	**3,69**	**26,06**
	Mann–Whitney U (p)	**0,762**	**0,013**	**0,001**	**0,023**	**0,012**	**0,327**	**0,000**	**0,013**	**0,000**	**0,000**	**0,000**	**0,000**	**0,000**	**0,000**	**0,000**	**0,000**
Control Group	Number of Observations	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
	Mean	1,56	1,56	2	2	1,69	1,62	1,75	12,19	2,25	2,31	2,25	2,38	2,38	2,06	2,31	15,94
Experimental Group	Number of Observations	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
	Mean	1,31	1,31	1,94	1,62	1,25	1,38	1,31	10,13	3,25	3,31	3,62	3,38	3,25	3,38	3,5	23,69
Total	**Number of Observations**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**	**32**
	Mean	**1,44**	**1,44**	**1,97**	**1,81**	**1,47**	**1,5**	**1,53**	**11,16**	**2,75**	**2,81**	**2,94**	**2,87**	**2,81**	**2,72**	**2,91**	**19,81**
	Mann–Whitney U (p)	**0,245**	**0,184**	**0,771**	**0,051**	**0,056**	**0,245**	**0,030**	**0,036**	**0,005**	**0,001**	**0,000**	**0,002**	**0,009**	**0,001**	**0,001**	**0,000**

Note: a, b, c, d, e, f, g, expresses evaluation criteria. a: 'ability to play precise and clean', b: 'ability to play with rhythmic accuracy', c: 'ability to apply right-hand techniques', d: 'ability to apply left-hand techniques', e: 'ability to play with musicality', f: 'ability to play with fluidity', g: 'ability to play at the designated tempo'.

If probability value (p)<α=0,05, the test is statistically significant.

TMV-pre: Total mean values of pre test, TMV-post: Total mean values of post test

When the results of both groups were examined for Expert 1, it was observed that the evaluation scores of the control group were higher than the evaluation scores of the experimental group in both the pre-test and posttest. The data show that the etudes applied to the experimental group between the pre-test and the posttest were not found to be successful by Expert 1.

When the pre-test characteristics of both groups were examined by Expert 2 and Expert 3, it was observed that the evaluation scores of the control group were higher than the evaluation scores of the experimental group. On the other hand, when the posttest characteristics of the groups were examined, it was observed that the evaluation scores of the experimental group were higher than the evaluation scores of the control group. The data show that the etudes applied to the experimental group between the pre-test and the posttest were found to be successful by Expert 2 and Expert 3. The scores given to the experimental group and the control group for all musical pieces, and an evaluation of the results are presented Table 3.

Findings on the Evaluation Scores Given According to Musical Pieces

Table 3

Individual Evaluation Scores of Experimental and Control Groups for All Musical Pieces

		Pre-test								Posttest							
All Musical Pieces		a	b	c	d	e	f	g	TMV-pre	a	b	c	d	e	f	g	TMV-post
Control Group	Number of Observations	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48
	Mean	2,60	2,79	2,96	2,90	2,50	2,50	2,65	18,90	2,94	3,21	3,15	3,15	3,02	2,88	3,10	21,44
Experimental Group	Number of Observations	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48
	Mean	2,12	2,02	2,42	2,19	1,73	1,79	1,71	13,98	3,60	3,77	3,92	3,73	3,44	3,67	3,77	25,90
Total	**Number of Observations**	**96**	**96**	**96**	**96**	**96**	**96**	**96**	**96**	**96**	**96**	**96**	**96**	**96**	**96**	**96**	**96**
	Mean	**2,36**	**2,41**	**2,69**	**2,54**	**2,11**	**2,15**	**2,18**	**16,44**	**3,27**	**3,49**	**3,53**	**3,44**	**3,23**	**3,27**	**3,44**	**23,67**
	Mann–Whitney U (p)	**0,050**	**0,002**	**0,012**	**0,001**	**0,000**	**0,006**	**0,000**	**0,001**	**0,005**	**0,015**	**0,001**	**0,010**	**0,049**	**0,005**	**0,007**	**0,004**

Note: a, b, c, d, e, f, g, expresses evaluation criteria. a: 'ability to play precise and clean', b: 'ability to play with rhythmic accuracy', c: 'ability to apply right-hand techniques', d: 'ability to apply left-hand techniques', e: 'ability to play with musicality', f: 'ability to play with fluidity', g: 'ability to play at the designated tempo'. If probability value (p)<α=0,05, the test is statistically significant.
TMV-pre: Total mean values of pre test, TMV-post: Total mean values of post test

When the pre-test characteristics of the experimental group and the control group were examined for all pieces, it was observed that the evaluation scores of the control

group were higher than the evaluation scores of the experimental group. On the other hand, when the posttest characteristics of the groups were examined, it was observed that the evaluation scores of the experimental group were higher than the evaluation scores of the control group. The data show that the etudes applied to the experimental group between the pre-test and the posttest were successful in all the pieces. The scores given to the experimental group and the control group for each of the individuals forming the study group, and an evaluation of the results are presented Table 4.

Findings on the Evaluation Scores Given According to Individuals

Table 4

Comparison of the Mean Scores of the Pre-test and Posttest According to the IndMV-preividuals in the Experimental Group and the Control Group

Persons	pre	a-post	a-pre	b-post	b-pre	c-post	c-pre	d-post	d-pre	e-post	e-pre	f-post	f-pre	g-post	g-pre	TMV	TMV	TMV
N	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
C.G. Person 1	2,00	2,42	2,50	2,92	2,58	2,67	2,42	2,83	1,92	2,50	2,25	2,50	2,25	2,75	15,92	18,58	2,66	
N	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	
C.G. Person 2	2,50	2,83	2,50	3,00	3,08	3,17	2,75	3,08	2,17	2,83	2,17	2,67	2,50	2,92	17,67	20,50	2,83	
N	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	
C.G. Person 3	2,75	2,92	2,92	3,33	2,92	2,92	3,08	2,92	2,92	3,17	2,75	3,00	2,92	3,08	20,25	21,33	1,08	
N	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	
C.G. Person 4	1,92	4,33	2,25	4,50	2,75	4,25	2,42	4,42	3,00	3,58	2,83	3,33	2,92	3,67	21,75	25,33	3,58	
Kruskal-Wallis Test (p)	0,125	0,057	0,369	0,455	0,464	0,076	0,142	0,145	0,024	0,071	0,430	0,477	0,386	0,295	0,104	0,054		
N	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	
E.G. Person 1	2,75	4,25	2,58	4,25	2,67	4,42	2,83	4,25	2,58	4,08	2,75	4,33	2,50	4,33	18,67	29,92	11,25	
N	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	
E.G. Person 2	2,25	3,33	1,75	3,25	2,42	3,92	2,17	3,50	1,75	3,25	1,67	3,42	1,67	3,75	13,67	24,42	10,75	
N	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	
E.G. Person 3	1,58	2,50	1,50	3,08	1,83	3,08	1,33	2,75	1,08	2,67	1,00	2,58	1,08	2,83	9,42	19,50	10,08	
N	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	
E.G. Person 4	3,17	3,58	3,25	3,58	3,25	3,83	3,33	3,75	1,50	3,75	1,75	4,33	1,58	4,17	14,17	29,75	15,58	
N	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	
Kruskal-Wallis Test (p)	0,024	0,001	0,061	0,003	0,184	0,000	0,000	0,001	0,000	0,062	0,000	0,003	0,004	0,026	0,001	0,002		

Note: a, b, c, d, e, f, g, expresses evaluation criteria. a: 'ability to play precise and clean', b: 'ability to play with rhythmic accuracy', c: 'ability to apply right-hand techniques', d: 'ability to apply left-hand techniques', e: 'ability to play with musicality', f: 'ability to play with fluidity', g: 'ability to play at the designated tempo'.
TMV: Total mean values, TMV-Dif: TMV-Difference between Pre-test and Post-test
The cells highlighted in yellow are pre-test data. N: Number of observations, C.G.: Control Group, E.G.: Experimental Group.
If probability value (p)<α=0,05, the test is statistically significant.

The posttest mean scores of both the control group and the experimental group are higher than the pre-test mean scores. In other words, the persons performing the piece made progress later (in the posttest) compared to the beginning (pre-test), and the evaluation scores increased as a result of the posttest evaluation. However, increases in the mean score of the students in the control group are much higher than the increases in the mean score of the students in the experimental group. This can be explained by the fact that the etudes created in preparation for the musical pieces are more effective than traditional educational methods. The scores given by each of the experts to the experimental group and the control group for each musical piece, and an evaluation of the results are presented Table 5.

Findings on the Evaluation Scores Given by Each of the Experts for Each Musical Piece

Table 5

Comparison of The Mean Scores of Pre-test and Post-test According to Experts

Group	Expert		**Piece1**		**Piece2**		**Piece 3**		**Piece 4**		**Total**	
			TMV pretest	TMV posttest	TMV pretest	TMV posttest	TMV pretest	TMV posttest	TMV pretest	TMV posttest	TMV pretest	TMV posttest
Control Group	Expert 1	N	4	4	4	4	4	4	4	4	**16**	**16**
		Ort.	29,75	31,25	25,50	29,75	25,25	26,00	25,75	27,25	**26,56**	**28,56**
	Expert 2	N	4	4	4	4	4	4	4	4	**16**	**16**
		Ort.	23,50	24,25	18,75	22,75	15,00	16,00	14,50	16,25	**17,94**	**19,81**
	Expert 3	N	4	4	4	4	4	4	4	4	**16**	**16**
		Ort.	12,00	17,00	13,00	17,00	13,25	15,75	10,50	14,00	**12,19**	**15,94**
	Total	**N**	**12**	**12**	**12**	**12**	**12**	**12**	**12**	**12**	**48**	**48**
		Ort.	**21,75**	**24,17**	**19,08**	**23,17**	**17,83**	**19,25**	**16,92**	**19,17**	**18,90**	**21,44**
	Kruskal-Wallis Test (p)		**0,034**	**0,157**	**0,011**	**0,016**	**0,025**	**0,024**	**0,011**	**0,017**		
Experimental Group	Expert 1	N	4	4	4	4	4	4	4	4	**16**	**16**
		Ort.	22,25	24,00	20,75	23,50	16,25	19,25	17,00	20,00	**19,06**	**21,69**
	Expert 2	N	4	4	4	4	4	4	4	4	**16**	**16**
		Ort.	12,75	32,50	13,75	32,75	12,75	32,00	11,75	32,00	**12,75**	**32,31**
	Expert 3	N	4	4	4	4	4	4	4	4	**16**	**16**
		Ort.	11,00	23,75	9,50	25,25	10,50	23,00	9,50	22,75	**10,13**	**23,69**
	Topl.	**N**	**12**	**12**	**12**	**12**	**12**	**12**	**12**	**12**	**48**	**48**
		Ort.	**15,33**	**26,75**	**14,67**	**27,17**	**13,17**	**24,75**	**12,75**	**24,92**	**13,98**	**25,90**
	Kruskal-Wallis Test (p)		**0,078**	**0,099**	**0,024**	**0,069**	**0,205**	**0,061**	**0,087**	**0,069**		

Note: If probability value (p)<α=0,05, the test is statistically significant.
TMV-pre: Total mean values of pre test, TMV-post: Total mean values of post test

The pre-test and posttest mean evaluation scores of the seven characteristics of the experimental group and the control group — individually for all the pieces — were examined, and it was checked whether there is any difference between the experimental group and the control group according to the etude exercises given. It was found out that Expert 1 gave higher scores to the control group in the pre-test and posttest evaluations on average for all pieces. These results reveal that the etudes

applied to the experimental group did not yield successful results according to Expert 1. When the scores given by Expert 2 and Expert 3 to the experimental group and the control group for all musical pieces were examined, it was observed that the mean scores of the control group were higher than the scores of the experimental group. However, it was seen in the posttest evaluations made after the etudes that the mean scores of the experimental group were higher than the mean scores of the control group. This shows that the etudes applied between the pre-test and the posttest yielded successful results according to Expert 2 and Expert 3.

Discussion and Conclusion

Discussion

The study by Parasiz (2009) is a review, in terms of functionality and effectiveness, of preparatory exercises aimed at improving performance in the practice of Modern Turkish Music pieces being used in violin education. For this reason, the study comprises a test model and pre-test-posttest model with a control group within that model. As a result of this study, it has been concluded that the preparatory exercises are effective in performances of Modern Turkish Music pieces that are used in violin education.

The study by Kaya and Gokbudak (2011) analyzed the suitability of maqam etudes and exercises for use in the process of cello education in branches of music education at departments of fine arts education at faculties of education, and their possible contributions to cello education. An experimental method is used in the study. It has been concluded in the frame of the study that maqam etudes and exercises are effective and can be used in parallel with the tonal etudes and exercises used to ensure that the performer acquires basic techniques in the process of cello education.

For inscription of practicing patterns and etudes, several parts were chosen for notation among Necdet Yasar's *taksim*s under the study by Bilgin (2011); later, exercises and etudes were formed by using such parts. The observation method was employed in the study, manifesting the style characteristics in Necdet Yasar's performance.

As for the results of the research conducted, it has been observed that the etudes formed for the instruments used in Turkish Music are aimed at ensuring easier comprehension of the styles of certain composers or facilitating the performance of certain musical pieces. However, no study has been identified that includes preparatory etudes for cello so as to facilitate the performance of Turkish Music pieces. This study was conducted in an effort to make a contribution to the field upon realization of the lack of sources available on the subject.

Conclusion

When the findings on the evaluation scores given by the experts were examined, it was observed that the evaluation scores by Expert 1 of the control group were higher than the evaluation scores of the experimental group both in the pre-test and posttest. When the pre-test characteristics of the experimental group and the control group were examined for Expert 2 and Expert 3, it was observed that the evaluation scores of the control group were higher than the evaluation scores of the experimental group. On the other hand, when the posttest characteristics of the groups were examined, it was observed that the evaluation scores of the experimental group were higher when compared to the evaluation scores of the control group. The data show that the etudes applied to the experimental group between the pre-test and the posttest were not found successful by Expert 1, while they were found successful by Experts 2 and 3.

When the findings on the evaluation scores given according to the musical pieces were examined, it was seen that the same result was achieved for all the pieces. When the pre-test characteristics of the experimental group and the control group were examined, it was observed that the evaluation scores of the control group were higher than the evaluation scores of the experimental group. On the other hand, when the posttest characteristics of the groups were examined, it was observed that the evaluation scores of the experimental group were higher when compared to the evaluation scores of the control group. The data show that the etudes applied to the experimental group between the pre-test and the posttest were successful for all of the four pieces.

When the findings on the evaluation scores according to the individuals were examined, it was seen that the posttest mean scores of the persons both in the control group and the experimental group were higher than their pre-test mean scores. In other words, the persons performing the piece made progress later (in the posttest) compared to the beginning (pre-test), and the evaluation scores increased as a result of the posttest evaluation. However, increases in the mean score of the students in the control group are much higher than the increases in the mean score of the students in the experimental group. This can be explained by the fact that the etudes created in preparation for the musical pieces are more effective than traditional educational methods.

Finally, the pre-test and posttest mean evaluation scores of the seven characteristics of the experimental group and the control group — individually for all the pieces — were examined and it was checked whether there is any significant level of difference between the experimental group and the control group. It was found out that Expert 1 gave higher scores to the control group in the pre-test and posttest evaluations on average for all pieces. It is seen that the etudes applied to the experimental group did not yield successful results according to Expert 1. When the scores given by Expert 2 and Expert 3 to the experimental group and the control group for all musical pieces were examined, it was observed that the mean scores of the control group were higher than the scores of the experimental group. However, it was seen in the posttest

evaluations made after the etudes that the mean scores of the experimental group were higher than the mean scores of the control group. This shows that the etudes applied between the pre-test and the posttest yielded successful results according to Expert 2 and Expert 3.

In conclusion, it was discovered that the etudes written with the aim to improve and facilitate the performance of musical pieces in the scope of this study created a significant difference in students' skills, which include the 'ability to play precise and clean', 'ability to play with rhythmic accuracy', 'ability to apply right-hand techniques', 'ability to apply left-hand techniques', 'ability to play with musicality', 'ability to play with fluidity', and 'ability to play at the designated tempo'.

Recommendations

Within the scope of this study, etudes were created to facilitate the performance of the pieces, and a significant difference was observed in the performances of the students. Similarly, instructors have been provided with the opportunity to create small but functional etudes, as well as various course materials, for their students when applying the etude writing system on the students by using a Turkish Music piece. There are a large number of methods and etudes to teach the cello technique. Therefore, as in this study, any etudes that are written should aim at enabling the student to perform musical pieces by helping them overcome difficulties unique to the piece in terms of technique, musicality, and style.

References

Arslanturk, Z. (1997). *Sosyal bilimler icin arastirma metod ve teknikleri* [Research methods and techniques for social sciences] (3rd ed.). İstanbul: Marmara Universitesi İlahiyat Fakultesi Yay.

Bilgin, K. (2011). *Necdet Yasar'in taksimlerinden hareketle tanbur etudleri* [Tanbur etudes about Necdet Yasar's taksims] (Master's thesis). Yildiz Teknik Universitesi Sosyal Bilimler Enstitusu, İstanbul. Retrieved March 19, 2018, from https://tez.yok.gov.tr/UlusalTezMerkezi/

Buyukozturk S., Buyukozturk, S., Kilic-Cakmak, E., Akgun, O. E., Karadeniz, S., & Demirel, F. (2016). *Bilimsel arastirma yontemleri* [Scientific Research Methods] (21st ed.). Ankara: Pegem Akademi.

Ercan, N. (2008). *Piyano egitiminde ilke ve yontemler* [Principles and Methods in Piano Education]. Ankara: Sozkesen Matbaasi.

Karasar, N. (2007). *Bilimsel arastirma yontemi* [Scientific Research Method] (17th ed.). Ankara: Nobel Yayin Dagitim LTD. STİ.

Kaya, E.E. & Gokbudak, Z.S. (2011). Viyolonsel egitiminde makamsal etut ve egzersizlerin kullanim durumuna iliskin bir arastirma [A study on the use of

makam etudes and exercises in violoncello education]. *e-Journal of New World Sciences Academy*, 6(4), 532-559.

Lancaster, L. K. (1994). *Analyzing flute pedagogy: a discussion with selected pedagogues* (Doctoral dissertation). University of Missouri D.M.A. Conservatory of Music, Kansas City.

Ozkan, İ.H. (2000). *Turk Musikisi nazariyati ve usulleri kudum velveleleri* [Turkish Music Theory and Principles] (6th ed.). İstanbul: Otuken Nesriyat A.S.

Parasiz, G. (2007). *Keman ogretiminde kullanilan Cagdas Turk Muzigi eserlerinin seslendirilmesine yonelik olarak olusturulan hazirlayici alistirmalarin etkililik ve islevsellik yonunden incelenmesi* [Investigation of the preparatory exercises prepared for the performed of Contemporary Turkish Music pieces used in violin teaching in terms of effectiveness and functionality] (Doctoral dissertation). Gazi Universitesi Egitim Bilimleri Enstitusu, Ankara. Retrieved March 19, 2018, from https://tez.yok.gov.tr/UlusalTezMerkezi/

Turkel, L., & Sen, Y. (2016). Cagdas Turk Muzigi flut eserleri icrasinda makamsal flut egzersizlerinin kullanilabilirligi ve ogrenci gorusleri (Ataturk Universitesi ornegi), [The availability of maqamic flute exercises in performs of Contemporary Turkish Music flute pieces and students views (example of Ataturk University)]. *İdil Dergisi*, *5*(23), 909-930, Doi: 10.7816.

Ural, A., & Kilic, İ. (2011). *Bilimsel arastirma sureci ve SPSS ile veri analizi* [Scientific Research Process and Data Analysis with SPSS]. Ankara: Detay Yay.

Yalcinkaya, B. (2010). *Flut egitiminde Geleneksel Turk Muzigi eserlerinin seslendirilmesine yonelik olusturulan etut yazma modeli ve bu modelle bestelenen ozgun etutlerin ogrenci basarisi uzerine etkileri* [Etud writing model for performing Traditional Turkish Music pieces in flute education and the effects of the original studies composed with this model on student success] (Yayimlanmamis Doktora Tezi). Gazi Universitesi Egitim Bilimleri Enstitusu, Ankara. Retrieved March 1, 2018, from: https://tez.yok.gov.tr/UlusalTezMerkezi/

Yalcinkaya, B. (2012). Calgi egitiminde etut yazma modeli, [Etude writing model in instrument education]. *e-Journal of New World Sciences Academy (NWSA)*, 7(1), 21-34.

Longa ve Sirtoların Viyolonsel ile Seslendirilmesine Yönelik Olarak Oluşturulan Etütlerin Etkililik Yönünden İncelenmesi

Atıf:

Avci Akbel, B. (2018). A review on effectiveness of cello etudes created for longas and syrtos. *Eurasian Journal of Educational Research*, 75, 1-18, DOI: 10.14689/ejer.2018.75.1

Özet

Problem Durumu: Türk Müziği türlerinden olan Longa ve Sirtoların seslendirilmesinde, öğrenciler tarafından aynı noktalarda benzer yoğunlukta hatalar yapıldığı tespit edilmiştir. Hataların yapıldığı kısımlardaki teknik ve müzikal sorunların giderilmesine yönelik olarak etütler yazılmasının, bu sorunları giderebileceği varsayımıyla hareket edilerek bu araştırmaya başlanmıştır. Literatürde bu konuda yapılan çalışmalar incelendiğinde, bazı bestecilerin üsluplarını anlamaya yönelik olarak veya bazı eserlerin seslendirilmesini kolaylaştırmak adına Türk Musikisinde kullanılan enstrümanlar için etüt yazma yoluna gidilmiştir. Fakat Türk Musikisi eserlerinin seslendirilmesini kolaylaştırmak adına viyolonsel için hazırlayıcı alıştırmaların yazıldığı bir çalışmaya rastlanmamıştır. Bu çalışma, bu boşluğun fark edilmesi üzerine bu alana katkı sağlayabilmek amacıyla yapılmış bir çalışmadır.

Araştırmanın Amacı: Bu araştırma, Türk Müziğinde çalgı müziği içerisinde yer alan Longa ve Sirto formundaki eserlerin -içerdikleri ajilite, teknik zorluklar, vb. sebeplerle- viyolonsel ile icra edilmelerinde yaşanan zorluklara çözüm bulunabilmesi ve bu yolla bu eserlerin teknik ve müzikal olarak daha doğru ve daha kolay çalınabilir hale getirilmesi amacıyla yapılmıştır. Bu alanda yapılan araştırmaların sonuçlarına bakıldığında, bazı bestecilerin üsluplarını anlamaya yönelik olarak veya bazı eserlerin seslendirilmesini kolaylaştırmak adına Türk Musikisinde kullanılan enstrümanlar için etüt yazma yoluna gidilmiştir. Fakat Türk Musikisi eserlerinin seslendirilmesini kolaylaştırmak adına viyolonsel için hazırlayıcı alıştırmaların yazıldığı bir çalışmaya rastlanmamıştır. Bu çalışma, bu boşluğun fark edilmesi üzerine bu alana katkı sağlayabilmek amacıyla yapılmış bir çalışmadır.

Araştırmanın Yöntemi: Araştırmada karma model kullanılmıştır. Araştırmanın deneysel kısmında karşılaştırma ve kontrol etme durumları söz konusudur. Bu sebeple araştırma; ön-test son-test kontrol gruplu model olarak desenlenmiştir. Ön-test son-test kontrol gruplu modelde, deney grubu ve control grubu olmak üzere yansız atama ile oluşturulmuş iki grup bulunur. Bu iki grupta da deney öncesi ve sonrası ölçmeler yapılır. Modelde ön testlerin bulunması, grupların deney öncesi benzerlik derecelerinin bilinmesine ve son test sonuçların buna göre düzeltilmesine

yardım eder (Karasar, 2007, 97). Hazırlayıcı alıştırmaların oluşturulmasında ise Taba-Tyler Modelinden uyarlanan etüt yazma modeli kullanılmıştır.

Bu araştırmada çalışma grubu, monografik örneklem seçme yöntemiyle belirlenmiştir. Monografik örneklem, olasılık dışı örneklem oluşturma yöntemlerinden biridir. Bu örneklem türünde araştırmacı, evreni temsil edebileceğini düşündüğü bir küme ya da bir-birkaç denek üzerinde çalışır (Ural, Kılıç, 2011, 44). Arslantürk (1997)'e göre monografik örneklemede örneklem, mevcut bilgiler yardımıyla betimlenmek istenen ana kütlenin herhangi bir alt grubunun ana kütleyi temsil edeceği varsayılarak seçilmektedir. Bu bilgilere göre araştırmanın çalışma grubu, Marmara Üniversitesinde ve Gazi Üniversitesinde lisans dört ve lisansüstü düzeylerde öğrenim gören öğrenciler arasından seçilmiştir. Bu seçim yapılırken viyolonsel öğretmenlerinin yönlendirmeleri doğrultusunda hareket edilmiş, araştırma kapsamındaki etütleri çalabilecek ve teknik açıdan benzer seviyede olan sekiz öğrenci belirlenmiştir. Bu öğrencilerin teknik olarak benzer düzeyde oldukları uzman kişiler tarafından onaylanmıştır.

Birinci aşamada, genel tarama modelinden yararlanılmıştır. Bu araştırma için belirlenen konunun doğru bir temele oturtulması ve bu konuda yapılan çalışmaların içeriğinin belirlenmesi için literatür taraması yapılarak konuyla alakalı veriler toplanmıştır. TRT Ankara Radyosu nota arşivinde bulunan tüm Longa ve Sirtoların notalarına ulaşılmıştır. Araştırmada incelenecek olan Longa ve Sirtolar Türkiye'deki Klasik Türk Müziği Konservatuarlarında viyolonsel öğretimi veren kişiler ile görüşme yapılarak belirlenmiştir. Belirlenen Longa ve Sirtolar öncelikle fa anahtarına dönüştürülmüş, viyolonsele uygun olan bir perdeye transpoze edilmiştir. Belirlenen eserlerin analiz edilmesi ile ortaya çıkan sonuçlara göre bu eserlerin çalınmasını destekleyecek hazırlayıcı alıştırmalar yazılmış ve bu eserlere uygun yay işaretleri ile temel teknikler eklenmiştir.

Araştırmanın ikinci bölümünde ise deneysel yöntem kullanılmıştır. Bu bölümde Longa ve Sirtoların viyolonsel öğretiminde icra edilmesini kolaylaştırmaya yönelik olarak hazırlanan alıştırmaların öğrenci performansı üzerinde ne derece etkili olduğunu ölçmek amacıyla, ön-test son-test kontrol guruplu modelden yararlanılmıştır. Çalışma grubundaki tüm öğrencilerin kayıtları üç uzman kişi tarafından –belirlenen yedi müzikal unsur açısından- değerlendirilmiştir. Çalışmanın sonunda, uzman kişiler öğrencilerin performanslarını değerlendirmişlerdir. Bu değerlendirmeler yoluyla elde edilen bilgiler, SPSS programı yardımıyla Wilcoxon işaretli sıralar testi ve Mann-whitney u testi uygulanarak karşılaştırılmıştır.

Araştırmanın Bulguları: Elde edilen istatistiksel verilere göre, deney grubunun aritmetik ortalamalarında anlamlı farklılıklar bulunmaktadır. Buradan hareketle, oluşturulan alıştırmaların öğrenci performansında başarıyı önemli ölçüde artırdığı belirlenmiştir. Bu araştırmada eserlerin icralarını iyileştirmek ve kolaylaştırmak amacıyla yazılan etütlerin, öğrencilerin 'temiz ve doğru çalabilme', 'doğru ritm ile çalabilme', 'sağ el teknik davranışlarını gerçekleştirebilme', 'sol el teknik davranışları gerçekleştirebilme', 'müzikal çalabilme', 'akıcı çalabilme' ve 'belirtilen tempoda çalabilme' becerilerinde anlamlı farklılık oluşturduğu bulunmuştur.

Araştırmanın Sonuçları ve Önerileri: Bu çalışma kapsamında eserlerin icrasını kolaylaştırmaya yönelik olarak etütler oluşturulmuş ve öğrencilerin icralarında kayda değer bir farklılık tespit edilmiştir. Aynı şekilde öğretim hayatında Türk Musikisi eserlerinden hareketle etüt yazma yöntemini öğrencisine uygulayan öğretmen, farklı ders kaynaklarına ek olarak öğrencisine yönelik küçük ama işlevsel etütler üretebilecektir. Viyolonsel tekniğini öğretmeye yönelik metotlar ve etütler çok fazla sayıdadır. Bu sebeple bundan sonra yazılacak etütler, bu çalışmada da olduğu gibi eserlere özgü teknik, müzikal, üslup vb. açılardan yaşanan çalım zorluklarını kolaylaştırmaya dönük olmalıdır. Bunun yanı sıra, viyolonsel ve diğer enstrümanlarda yaşanan çalım zorlukların giderilmesine yönelik olarak farklı Türk Müziği türlerindeki eserler üzerine de çalışmalar yapılması önerilir.

Anahtar Kavramlar: Türk Müziği, enstrümantal müzik, çello eğitimi, hazırlık alıştırmaları.

Eurasian Journal of Educational Research 75 (2018) 19-36

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

Examining Mental Health Professionals' Social Justice Attitudes in Turkey

Dilek Yelda KAGNICI[1] Serkan DENIZLI[2]

ARTICLE INFO

Article History:

Received: 13 Oct. 2017

Received in revised form: 18 Jan. 2018

Accepted: 28 Apr. 2018

DOI: 10.14689/ejer.2018.75.2

Keywords
Social justice
mental health professionals
belief in a just world
moral identity
open-mindedness
social dominance orientation

ABSTRACT

Purpose: Mental health professionals' personal values have a critical role in providing efficient services to culturally diverse clients, being social advocates, and promoting social justice. In the present study, mental health professionals in Turkey were assumed to be grouped under three categories: the ones who are sensitive, who are rigid, and who are apathetic based on their Belief in a Just World, moral identity, open-mindedness, and social dominance orientation scores. The study aimed to examine the differences among these three groups of mental health professionals in Turkey in relation to social justice.

Methods: A total of 232 mental health professionals participated in the study. The Social Justice Scale, Moral Identity Scale, Open-Mindedness Scale, Social Dominance Scale, and Belief in a Just World Scale were used to collect data, and cluster analysis was performed to analyze the data. **Findings**: The results indicated that there were statistically significant differences between three categories regarding social justice scores. Results pointed out that sensitive Turkish mental health professionals were more to prone to social justice than the other groups. **Implications for Research and Practice:** According to the results, there is a potential risk that there might be cultural groups outside of the boundaries of these mental health professionals' scope of justice. In order to prevent these risks, some precautions should be taken during mental health professionals' training. Mental health professionals should be equipped with current competencies to provide service to diverse groups.

[1] Corresponding Author: Ege University, TURKEY, e-mail: yelda.kagnici@ege.edu.tr, ORCID: https://orcid.org/ 0000-0002-5201-4784

[2] Ege University, TURKEY, e-mail: serkan.denizli@ege.edu.tr, ORCID: https://orcid.org/ 0000-0002-8623-841X

Introduction

Counselors and psychologists have been providing mental health services to a more culturally diverse population than ever. Providing sufficient services to diverse groups requires multicultural competencies. As Vera and Speight (2003) pointed out, commitment to multiculturalism also requires commitment to social justice.

Multiculturalism, social justice, and advocacy are highly empathized concepts of both counseling and psychology fields in the United States of America. Multicultural competencies (Sue, Arredondo, & Davis, 1992), social justice competencies, and advocacy competencies (American Counseling Association [ACA], 2003) have been formulated; there are divisions (e.g., American Psychological Association [APA] Division 45) and associations (e.g., Association for Multicultural Counseling and Development) working on these issues. However, the case is different in Turkey. Although Turkey is a multicultural country, multiculturalism is still a very hot topic and viewed in the context of religion or ethnicity (Kagnici, 2015). Studies regarding multiculturalism in psychology and psychiatry in the Turkish context are very limited (Goregenli, 2010; Keklik, 2010) due to academicians' positivistic approach (Cimilli, 2002; Goregenli, 2010; Kaygusuz, 2009) and the risks associated with talking about diversity. In counseling, although it has become an issue of interest (e.g., Erdur Baker, 2007; Bektas; 2006, Kagnici 2011; 2013; 2014; Karairmak, 2008; Koc, 2003), the number of studies conducted are very limited. Advocacy and social justice are also new research areas in Turkey. Although there are a few studies regarding social justice and education (Aslan & Gulacti; Gezer, 2017; Polat, 2007; Ozdemir, 2017; Tomul, 2009), in counseling literature only a few studies (Kagnici, 2015; 2017; Karairmak, 2015; Keklik, 2010) have mentioned advocacy and social justice.

Although social justice is a new research area in counseling literature in Turkey, social justice is not a new concept. As defined by Lee (2007, p. 1), "social justice involves promoting access and equity to ensure full participation in the life of a society, particularly for those who have been systematically excluded on the basis of race/ethnicity, gender, age, physical or mental disability, education, sexual orientation, socioeconomic status, or other characteristics of background or group membership". Besides, advocacy is accepted as a professional role for counselors and psychologists in order to make environmental changes for clients (Lewis & Bradley, 2000, 11). As Motulsky, Gere, Saleem, and Trantham (2014) underlined, sensitivity to injustice is only a start; counselors and psychologists should also be agents of change. In order to be agents of change, counselors and psychologists first need to be culturally aware, in other words, work hard on their values, personal characteristics, and potential cultural biases.

Cultural awareness is the mental health professionals' sensitivity to their personal values and biases and awareness of how these may influence perceptions of the client, the client's problem, and the counseling relationship (Sue, Arredondo, & McDavis, 1992). This basic competence indicates that mental health professionals' personal values have a critical role in providing efficient services to culturally diverse clients,

being social advocates, and promoting social justice. Therefore examining personal variables related with social justice seems to be critical.

Moral identity is one of the individual personal characteristics considered to be related to social justice attitude. According to Aquino and Reed (2002), moral identity is a self-schema that is organized around specific moral trait associations that are closely linked in memory, which has both a public and private aspect. As Damon and Hart (1992, cited in Aquino & Reed, 2002) proposed, people who formulate their self-concept according to their moral beliefs are likely to translate those beliefs into action. In other words, when justice and care are deeply rooted in one's self-concept, they guide people's attitudes and actions towards out-group members (Sam, Hardy, Bhattacharjee, Reed, & Aquino, 2010). As Hoffman (2000, cited in Eisenberg & Morris, 2001) mentioned, activation in moral cognitions might result in affective motivation to rectify violations of justice or simply to improve another's situation.

Belief in a Just World (BJW) was introduced by Lerner in 1965. According to BJW, good things happen to good people and bad things happen to bad people (Furham, 2003). The BJW is associated with a tendency to not defend the interests of victims and needy others, which results in victim blaming (Furham & Proctor, 1992; Montada, 1998; Overcash, Calhoun, Cann, & Tedeschi, 1996). Since the main point in BJW is that people live in a world where they generally get what they deserve, in cases of any inequality, inequality is observed as "being deserved". Having a lower BJW is more likely to have higher and more positive attitudes toward social justice (Parihk, Post, & Flowers, 2011; Van Soest, 1996).

Personality is also considered as a related variable. As one of the five multicultural personality characteristics, open-mindedness refers to "an open and unprejudiced attitude towards different groups and towards different cultural norms and values" (Van der Zee & Van Oudenhoven, 2000, 294). As Garmon (2010) mentioned, individuals who are open-minded tend to learn more than individuals who are closed-minded and also less likely to distance themselves from ethnically different individuals (Hello, Scheepers, & Sleegers, 2006). Open-minded people are likely to try different activities, have intellectual curiosity, and be ready to examine various ideologies (Ward, Leong, & Low, 2004). Therefore, certain personality variables such as empathy and openness affect people in terms of prejudice against subordinates (Akrami & Ekehammar, 2006; Altemeyer, 1998; Pratto, Sidanus, Satallworth, & Malle, 1994; Stephan & Finlay, 1999).

Social dominance orientation is the other variable that is considered to be related to social justice attitude. Social dominance theory argues that members of dominant groups have good things in life such as good health and members of subordinate groups receive poor housing and poor health in life (Sidanius & Pratto, 2011). In other words some groups in a society have more advantages than the others, and this is how it should be due to group-based social hierarchies. According to Sidanius (1993; cited in Sidanius & Pratto, 2011), the individuals with low social dominance orientation are more likely to adopt equality-based ideologies and own roles that decrease inequalities in society.

The current state of Turkish mental health professionals' social justice level is unknown. As Pederson (2002) pointed out, psychology has been moving towards a more multicultural basis. During this transition some mental health professionals are expected to be under the influence of cultural encapsulation. These people are insensitive to cultural variations among individuals and assume their own view is the only right one (Wrenn, 1962 as cited in Pederson, 2002). Since this transition is painful, some other mental health professionals are expected to be uninterested in diversity issues. However, a majority of mental health professionals are expected to be eager to become culturally competent. Moreover, based on their attitudes, people make either positive, negative, or neutral evaluations of objects (Haddock & Maio, 2005). Similarly, based on the previous literature it was expected that mental health professionals would be grouped into three groups based on their social justice attitudes, those with positive, negative, and neutral evaluations of social justice issues. Accordingly, in the present study mental health professionals in Turkey were assumed to be grouped under three categories: those who are sensitive, rigid, and apathetic based on their Belief in a Just World, moral identity, open-mindedness, and social dominance orientation scores. The aim was to examine the differences among these three groups of mental health professionals in Turkey in relation to their attitudes to social justice.

Method

Research Design

In the present study, a descriptive correlational research design was utilized. Correlational research aims to accurately describe a situation or phenomenon without manipulation variables in the study (Christensen, Johnson, & Turner, 2015). Specifically, a classification strategy using cluster analysis (Heppner, Kivlighan, & Wampold, 2008) was used to classify counselors into subgroups, which were hypothesized to present differences on their social justice attitudes.

Research Sample

A total of 283 mental health professionals participated in the study. The sample included 170 counselors, 89 psychologists, and 24 psychiatrists. In addition, psychiatrists were excluded from participating due to the small sample size. Univariate and multivariate outliers were excluded from the analysis. Consequently the final data included 232 mental health professionals including 149 counselors and 83 psychologists. Ages of the participants ranged from 22 to 56, with a mean of 30.57 (sd = 6.76); and 72.4 % of the participants were females and 27.6 % of the participants were males. Regarding experience levels of the participants, 8.6 % had 1 year or less experience in the profession, 40.09 % had 1 to 5 years of experience, 26.3 % had 6 to 10 years of experience, 12.1 % had 11 to 15 years of experience, and 12.1 % had more than 15 years of professional experience.

Research Instruments and Procedures

In the present study, it was aimed to reach mental health professionals from all over Turkey. For this purpose, a personal information form and all instruments were uploaded to an online platform (i.e., Google forms). An online survey method has the advantage of reaching a broad range of participants from a large population (Heppner, Kivlighan, & Wampold, 2008). Later invitations for participating in the study were sent to mental health professionals via e-mail groups and social media. Mental health professionals that were aimed to be reached were counselors, psychologists, and psychiatrists.

Social Justice Scale. The Social Justice Scale (SJS) was developed by Torres-Harding, Siers, and Olsın (2012). The SJS aims to measure attitudes regarding social justice, values, perceived behavioral control, and intentions. The scale has 24 items and four factors: a) social justice attitudes, b) perceived behavioral control, c) subjective norms, and d) behavioral intentions. Torres-Harding et al. demonstrated that confirmatory factor analysis (CFA) fit indices indicated an acceptable fit for the factor structure of the SJS. Cronbach's alpha coefficients were between .82 to .95. The scale was adapted to Turkish by Cirik (2015). The fit indices of the second-order CFA demonstrated a good fit for the model (χ^2 = 671.15, *df = 246,* $\chi 2/d$ = 2.72, GFI = .90 AGFI = .88, RMSEA = .05, NFI = .96, NNFI = .97, CFI = .97, SRMR = .04). The internal consistency coefficients of the Turkish form were .92 for total scale and ranged from .84 to.92 for the subscales. In the present study, Cronbach's alpha was calculated as .92 for the total scale, and the coefficient alpha ranged from .88 to .93 for the subscales.

Moral Identity Scale. The Moral Identity Scale (MIS) was developed by Aquino and Reed (2002). The original scale has two subscales measuring symbolization (five items) and internalization (five items). The MIS has acceptable internal consistency reliabilities of .73 and .82 for internalization and symbolization factors, respectively. The Turkish adaptation study of the MIS was conducted by Yilmaz and Yilmaz (2015). Yilmaz and Yilmaz confirmed the factor structure of the Turkish form of the MIS with CFA, and fit indices indicated good fit for the measurement model of the Turkish version of the MIS (CFI = .96, NNFI = .95, IFI = .96, GFI = .95, AGFI = .92, RMSEA = .052, NFI = .91, SRMR = .057, PNFI = .067, PGFI = .57). The test–retest reliability coefficient was .71 for the internalization subscale and .74 for the symbolization sub-scale in the Turkish form of the scale. Cronbach's alpha coefficients were .77 and .76 for the internalization subscale and symbolization subscale, respectively. The MIS is a seven-point scale for answers ranging from "1- totally disagree" to "7-totally agree". In the present study Cronbach's alpha was found to be .65 for the internalization subscale, .77 for symbolization subscale, and .76 for the total scale.

Open-Mindedness Scale. The Open-Mindedness Scale is the subscale of Multicultural Personality Questionnaire (MPQ) developed by Van der Zee and Van Oudenhoven (2000) to measure multicultural effectiveness. The Open-Mindedness Scale measures open-mindedness, in other words being in a state of not having prejudice towards individuals from different cultural groups and with different cultural norms and values. The Open-Mindedness Scale has 16 items that respondents evaluate on a five-

point scale ranging from "1-not applicable" to "5-totally applicable" to rate their open mindedness levels. Exploratory factor analysis indicated that the Turkish form of the MPQ has five factors including open-mindedness, and Cronbach's alpha for open-mindedness scale was .84 (Kagnici, 2011). In the present study, Cronbach's alpha was .86 for the Open-Mindedness Scale.

Social Dominance Scale. The Social Dominance Orientation Scale (SDOS) (Sidanius & Pratto, 1999) was used to measure the social dominance orientations of the participants. SDOS is a seven-point scale ranging from "1-Totally True" to "7-Totally Wrong" and has 16 items. The scale was adapted to Turkish by Karacanta (2002), and the factorial structure was further investigated by Hasta and Karacanta (2013). The Turkish form of SDOS presented a two-factor solution as Jost and Thompson (2000) found. One factor is called "opposition to equality," and the other is "group-based dominance". Cronbach's alpha for the subscales for the Turkish form was .85 for the total scale and .87 and .70 for opposition to equality and group-based dominance subscales, respectively. In the present study, Cronbach's alpha coefficient was. 86 for the opposition to equality subscale, .78 for the group-based dominance subscale, and .84 for the total scale.

Belief in a Just World Scale. The Belief in a Just World Scale (BJS) was developed by Dalbert (1987, 1999; cited in Cetinalp-Sahin, 2014) and adapted to Turkish by Goregenli (2003; cited in Cetinalp-Sahin, 2014). The BJS has six items and is a five-point Likert-type scale. This scale measures how an individual perceives the world as a just place. Cronbach's alpha for the original form was .78 (Dalbert, 1999) and was .69 for the Turkish form. Cronbach's alpha was calculated as .86 for the present study.

Data Analysis

Prior to data analysis, assumptions of the ANOVA and normal distribution were tested. Normal distribution was assured by calculating skewness and kurtosis values, and there were no values exceeding the normal distribution. Later, k-means cluster analysis was performed on 232 mental health professionals to explore whether the mental health professionals constitute subgroups based on their (i) social dominance orientation inventory scores, (ii) moral identity inventory scores, (iii) belief in a just world inventory scores, and (iv) open-mindedness inventory scores. Cluster analysis yielded three subgroups in the sample, and one-way analysis of variance was used to see whether the three subgroups of professionals differ in their social justice attitude scores.

Results

The cluster analysis presented three interpretable clusters (Cluster 1, n = 70; Cluster 2, n = 76; and Cluster 3, n = 86). Cluster 1, named "Rigid Professionals", was constituted by mental health professionals with the highest BJW levels (z = 1.13) and social dominance orientations (z = .48) and lower levels of open-mindedness (z = .05) and moral identity (z = .03) as expected. Cluster 2, named as "Sensitive Professionals", had the highest open-mindedness (z = .75) and moral identity scores (z =.69) and the lowest BJW (z = -.52) and social dominance orientations (z = -.62) levels among mental

health professionals. Cluster 3, labeled as "Apathetic Professionals", included mental health professionals with a low social dominance orientation (z = .94), and their scores were concurrently low on levels of open-mindedness (z = -.62), BJW (z = -.45) and moral identity (z = -.81). The results of the cluster analysis and the z-scores of the clusters are presented in Table 1.

Table 1

Final Cluster Centers and z-Scores of the Clusters on Social Dominance, Open-Mindedness, Moral Identity, and Belief in a Just World Scale

	Cluster		
	Rigid Professionals (n = 70)	Sensitive Professionals (n = 76)	Apathetic Professionals (n = 86)
Social dominance	.48	-.62	.09
Open-mindedness	.056	.75	-.62
Moral identity	.34	.69	-.81
Belief in a Just World	1.13	-.52	-.45

A one-way analysis of variance (ANOVA) was conducted to examine whether the social justice levels differ among the Turkish mental health professionals who were grouped into the three clusters (rigid, sensitive, and apathetic). All scores of the subscales of Social Justice Attitudes Scale, which were attitudes towards social justice, perceived behavioral control, subjective norms subscale, and behavioral intention subscale, were included in the analysis. Standard deviations of the subscales of the Social Justice Scale are presented in Table 2.

Table 2

Descriptive Statistics Related to the Social Justice Subscale Scores of Mental Health Professionals

Subscale	Group	*n*	*M*	*SD*	*SE*
Attitude	Rigid	70	70.84	6.14	.73
	Sensitive	76	75.21	3.52	.40
	Apathetic	86	69.52	7.84	.84
	Total	232	71.78	6.63	.43
Behavior	Rigid	70	29.65	3.09	.37
	Sensitive	76	31.32	3.28	.37
	Apathetic	86	27.26	4.03	.43
	Total	232	29.31	3.90	.25
Norms	Rigid	70	20.29	4.30	.51
	Sensitive	76	20.00	6.09	.69
	Apathetic	86	19.97	5.36	.57
	Total	232	20.07	5.30	.34
Intention	Rigid	70	24.47	2.73	.32
	Sensitive	76	26.11	2.52	.28
	Apathetic	86	23.30	3.77	.40
	Total	232	24.57	3.30	.21

The results of the ANOVA indicated that there were statistically significant differences, F (2,229) = 18.19, $p < .01$ between the three clusters regarding attitudes towards social justice (ATSJ) subscale scores. Levene's test was significant for the ATSJ scores. Thus in order to find the differences among the groups, we chose Dunnet's C test for the follow-up analysis which does not assume homogeneity of variances. The results of the Dunnett's C test presented that rigid professionals had lower levels of ATSJ when compared to sensitive professionals and sensitive professionals had a higher level of ATSJ scores when compared to apathetic professionals. There were no significant differences between rigid professionals and apathetic professionals regarding ATSJ scores. Means and standard deviations for attitudes towards the social justice subscale are available in Table 2.

Regarding the perceived behavioral control (PBC) subscale scores, the results of the ANOVA indicated that there were significant differences among all clusters [F (2, 229) = 27.17, $p < .01$]. Levene's test was significant for the PBC scores, thus Dunnett's C follow-up test, which does not assume homogeneity of variances for the follow-up analysis, was chosen. Sensitive professionals had significantly higher scores on PBC when compared to rigid professionals. Moreover, apathetic professionals had significantly lower scores on PBC scores than rigid professionals and sensitive professionals. Means and standard deviations for perceived behavioral control (PBC) are presented in Table 2.

There were no significant differences among clusters regarding the subjective norms subscale [F (2, 229) = 0.08, $p > .05$). However, there were significant differences among clusters in the behavioral intention (BI) subscale scores [F (2, 229) = 16.47, $p < .01$]. Levene's test was significant for PBC scores, Dunnet's C follow-up test, which does not assume homogeneity of variances as the follow-up analysis, was again chosen. Sensitive professionals had higher scores on BI when compared with both rigid professionals and apathetic professionals. There were no significant differences in BI rigid professionals and apathetic professionals. Means and standard deviations for the subjective norms subscale are available in Table 2.

Discussion and Conclusion

The present study aimed to examine the differences among three groups of mental health professionals in Turkey in relation to social justice. In order to examine social justice levels in terms of BJW, moral identity, open-mindedness, and social dominance orientation, cluster analysis was performed. Results pointed out that the *sensitive* Turkish mental health professionals were more to prone to social justice than the other groups, namely r*igid and apathetic* professionals. These mental health professionals had the highest open-mindedness and moral identity scores and the lowest BJW and social dominance orientation levels among the participating mental health professionals. In other words, these mental health professionals are able to be considered to be open to various experiences and sensitive, to have high moral development and behave according to universal ethical principles, to not observe inequality as "being deserved" and to not believe in group-based social hierarchies. The finding that these mental

health professionals' social justice levels are higher than the other groups' is not surprising since all these qualifications seem essential for being advocates of disadvantaged groups. Thus, this finding is parallel with the literature that the people who are more open to cultural differences, who have high moral identity, whose social dominance orientation and BJW are lower have positive social justice attitudes (Parihk, Post, & Flowers, 2011; Yildirim & Akgun, 2013; Van Soest, 1996).

According to the findings of the study, there is a group of Turkish mental health professionals with high levels of social dominance orientation and low open-mindedness and moral identity levels whose social justice levels are also low. This specific finding is notable because it may be assumed that these mental health professionals might face difficulties while working with culturally diverse groups, and might even provide insufficient services to these groups. According to Aguiar, Vala, Correia, and Pereira (2008), when the suffering of victims falling outside the boundaries of an individual's scope of justice, that individual's justice concerns might not be activated. Turkey is a multicultural country, and under these circumstances, there is a potential risk that there might be cultural groups outside the boundaries of this group of mental health professionals' scope of justice.

In order to prevent these risks, it is believed that some precautions should be taken during mental health professionals' training. Multicultural competence has emerged as an important topic in human service professions and their teaching (Ridley & Kleiner, 2003). Societies are more diverse than ever, and mental health professionals are providing mental health services to diverse groups. Since culture plays an important role in the counseling process (Vontress, 2000), mental health professionals should have the cultural competence to include cultural dimensions into counseling process. Moreoever, social justice awareness and advocacy skills should be integrated in counseling and psychology training. It seems critical that mental health professionals should be challenged in terms of multiculturalism, social justice, and advocacy during their training. There are study findings that training might have significant positive effects on participants' social justice attitudes, belief in a just world, moral identity, and open-mindedness levels (Hurtado, Engberg, & Ponjuan, 2003 cited in Broido & Reason, 2005; Kagnici, 2011; Miranda, Radliff, Cooper, & Eschenbrenner, 2014; Nagda, Gurin, & Lopez, 2003).

In Turkey mainly graduates of counseling and psychology undergraduate programs work as mental health providers, and licensing or certification is not mandated. In order to work as a counselor and psychologist in Turkey, it is sufficient to have an undergraduate degree. Therefore, providing such training in bachelor's degree programs is crucial in Turkey. As D'Andrea and Daniels (1991) proposed, counselor education programs have two levels: the cultural encapsulation level and the conscientious levels. In the cultural encapsulation level, multicultural counseling training is almost nonexistent; in the conscientious level, students systematically receive multicultural counseling training. Currently, it is hard to claim that counseling and psychology programs in Turkey are at the conscientious level. As mentioned, multicultural counseling and social justice are new concepts, and currently the focus of the available studies is mainly on introducing the concepts and competencies.

Although integrating multiculturalism, social justice, and advocacy issues into the counseling and psychology curriculum is crucial at the moment, it is also a fact that integration of social justice in counseling psychology programs is limited and not easy (Motulsky, Gere, Saleem, & Trantham, 2014). As underlined by Miranda and colleagues (2014), social justice needs to be integrated into the program description, course content, activities, and field-based training in counseling and psychology bachelor's programs in Turkey. Fostering a supportive atmosphere for learning experiences; engaging in difficult dialogues that challenge students to critically examine and reflect upon their worldviews, their privilege, and their biases; and providing opportunities for students to engage in service-oriented collaborative projects with surrounding communities as emphasized by Bemak, Chung, Talleyrand, Jones, and Daquin (2010) seem to be the key points in this integration.

Multicultural competence, social justice, and advocacy studies are mainly conducted by collecting data about perceptions of the participants. In future studies, participants' actual actions related to these concepts needs to be examined. Also, further studies need to examine not only the personal factors related to these concepts, but also the cultural factors in order have a more comprehensive understanding of social justice. Also the variables of belief in a just world, moral identity, and social dominance orientation need to examined through qualitative studies.

The present study has some limitations. One of the limitations is related to its data collection procedure. The data was collected through an online survey. The online survey method for collecting data is criticized because of its limitations in assuring the representativeness of the sample (Heppner, Kivlighan, & Wampold, 2008). Moreover, precautions that could be taken during face-to-face data collection were not able to be taken.

References

ACA Advocacy Competencies (2003). *Advocacy competencies.* http://www.counseling.org/

Aguiar, P., Vala, J., Correia, E. I., & Pereira, E. C. (2008). Justice in our world and in that of others: Belief in a just world and reactions to victims. *Social Justice Research*, *21*, 50–68.DOI 10.1007/s11211-007-0059-3

Akrami, N., & Ekehammar, B. (2006). Right-wing authoritarianism and social dominance orientation. Their roots in big-five personality factors and facets. *Journal of Individual Differences, 27,* 1-10.

Altemeyer, B. (1998). The other "authoritarian personality." In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 47–92). Orlando, FL: Academic Press.

Aquino, K., & Reed, A., (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology, 83*, 1423–1440.

Aslan, S., & Gulacti, F. (2013). Üniversite öğrencilerinin sosyal adalete ilişkin görüşleri [The Views of university students regarding social justice]. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi, 15, 203-214.*

Bektas, D. Y. (2006). Kültüre duyarlı psikolojik danışma yeterlikleri ve psikolojik danışman eğitimindeki yeri [Multicultural counseling competences and the place of multicultural competences in counseling education]. *Ege Eğitim Dergisi*, 1, 43-59.

Bemak, F., Chung, R. C., Talleyrand, R. M., Jones, H., & Daquin, J. (2011). Implementing multicultural social justice strategies in counselor education training programs. *Journal for Social Action in Counseling and Psychology, 3,* 29-43.

Broido, E. M., & Reason, R. D. (2005). The development of social justice attitudes and actions: An overview of current understandings. *New Directions for Student Services, 110,* 17-28.

CD'Andrea, M., & Daniels, J. (1991). Exploring the different levels of multicultural counseling training in counselor education. *Journal of Counseling & Development, 70*, 78-85.

Christensen, L. B., Johnson, R. B., & Turner, L. A. (2015). *Research methods, design and analysis.* (12th ed.). Pearson Education Limited: Harlow-Essex.

Cimilli, C. (2002). Kültür ve depresyon. Türkiye'den bir bakış [Culture and depression. A view from Turkey]. In K. Sayar (Ed.), *Kültür ve ruh sağlığı [In culture and mental health]* (p. 163-184). Ankara: Metis Yayıncılık

Cirik, İ. (2015). Psychometric characteristics of the social justice scale's Turkish form and a structural equation modeling. *Eurasian Journal of Educational Research, 61,* 23-44. http://dx.doi.org/10.14689/ejer.2015.61.2

Cetinalp Sahin, P. (2014). *Polis okulu öğrencilerinin adil dünya inancı, otoriterlik,sistemimeşrulaştırma, vatanseverlik ve insan hakları tutumlarının incelenmesi [Investigating Police School Students' BJW, authoritarianism, system legitimation, patriotism and human rights attitudes]* (Yayımlanmamış yüksek lisans tezi) [Unpublished masters thesis]. Ege Üniversitesi, İzmir.

Eisenberg, N., & Morris, A:S. (2001). The origins and social significance of empathy-related responding. a review of empathy and moral development: Implications for caring and justice by M. L. Hoffman. S*ocial Justice Research, 14*, 95-120.

Erdur-Baker, O. (2007). Psikolojik danışma ve kültürel faktörler [Counseling and cultural factors]. *Türk Psikolojik Danışma ve Rehberlik Dergisi [Turkish Psychological Counseling and Guidance Journal], 27*, 109-122

Furhnam, A. (2003). Belief in a just world: research progress over the past decade. *Personality and Individual Differences, 34,* 795–817.

Garmon, M. A. (2010). Six key factors for changing preservice teachers' attitudes/beliefs about diversity. *Educational Studies, 38,* 275-286.

Gezer, M. (2017). Öğretimde Sosyal Adalet İnançları Ölçeği'nin Türkçe uyarlaması [The Turkish adaptation of Learning to Teach for Social Justice–Beliefs Scale]. *Sakarya University Journal of Education, 7*(2), 295-309.

Goregenli, M. (2010). Psikolojinin Kürt sorunu ile imtihanı [Psychology's examination with Kurdish question]. *Eleştirisel Psikoloji Bülteni [Critical Psychology Bulletin], 3-4,* 1-11.

Haddock, G., & Maio, G. R. (2005) *Contemporary perspectives on the psychology of attitudes.* Psychology Press: Hove and Newyork.

Hasta, D. & Karacanta, H. (2013). *Sağ, merkez ve sol görüşlü öğrencilerde yetkecilik, sosyal baskınlık yönelimi ve benlik saygısı [authoritarianism, social donimance orientaiton and self-respect in right, center and left wing students]*. (Yayınlanmamış Araştırma Raporu) [Unpublished research report].

Hello, E., Scheepers, P., & Sleegers, P. (2006). Why the more educated are less inclined to keep ethnic distance: An empirical test of four explanations. *Ethnic and Racial Studies, 29 (5),* 959–985.

Heppner, P.P., Kivlihgan, D. M. & Wampold, B. E. (2008). *Research design in counseling. (*3rd ed.). Thomson Brooks/Cole: Belmont.

Kagnici, D. Y. (2011). Teaching multicultural counseling: An example from a Turkish counseling undergraduate program. *Eurasian Journal of Educational Research*, 44, 118-128.

Kagnici, D. Y. (2013). Çok kültürlü psikolojik danışman eğitiminin rehberlik ve psikolojik danışmanlık lisans programlarına yerleştirilmesi. [Accommodating multicultural counseling training in the guidance and counseling undergraduate programs]. *Türk Psikoloji Danışma ve Rehberlik Dergisi [Turkish Psychological Counseling and Guidance Journal], 5(40),* 222-231.

Kagnici, D. Y. (2014). Reflections of a multicultural counseling course: A qualitative study with counseling students and counselors. *Educational Sciences: Theory & Practice.* DOI: 10.12738/estp.2014.1.1965.

Kagnici, D. Y. (2015). Psikolojik danışman eğitiminde cinsel yönelim olgusunun irdelenmesi: Benim Çocuğum Filmi [Examination of sexual orientation in counseling education: "My Child film"]. *Türk Psikoloji Danışma ve Rehberlik Dergisi [Turkish Psychological Counseling and Guidance Journal], 44,* 83-95.

Kagnici, D. Y. (2017). Suriyeli mülteci çocukların kültürel uyum sürecinde okul psikolojik danışmanlarına düşen rol ve sorumluluklar [School counselors' roles and responsibilities in cultural adaptation process of Syrian refugee children] *Elementary Education Online, 16*(4), 1768-1776. doi: 10.17051/ilkonline.2017.342990

Karacanta, H. (2002). *Üniversite öğrencilerinin sosyal baskınlık yönelimi ve başka bazı değişkenler açısından karşılaştırılması [Comparing university students' social dominance orientation in terms of various variables]*. (Yayımlanmamış Yüksek Lisans Tezi) [Unpublished Masters Thesis]. Ankara Üniversitesi, Ankara.

Karairmak, Ö. (2008). Çok kültürlülük, kültürel duyarlılık ve psikolojik danışma [Multiculturalism, cultural sensitivity and counseling]. *Türk Psikolojik Danışma ve Rehberlik Dergisi [Turkish Psychological Counseling and Guidance Journal], 532*, 115-126.

Karairmak, Ö. (2015). Kültüre duyarlı ve hak savunucusu psikolojik danışman eğitimi [Changing paradigm in counseling field: Culturally sensitive and advocate counselor training] In. F. Kurter & J.S. Mattis (Ed.), *Türkiye'de uygulamacıların gözüyle psikolojik danışmada kültüre duyarlı olma [Being culturally sensitive in counseling from practioners point of view in Turkey]* (pp. 362-385) İstanbul: Bahçeşehir University Press

Kaygusuz, C. (2009). İnsan bilimleri, olgu değer sorunu ve akademik bilimlere yansıması [Human sciences, phenomonen value problem and reflections to academic sciences] *Eleştirisel Psikoloji Bülteni [Critical Psychology Bulletin], 2,* 48-60.

Keklik, İ. (2010). Psikolojik danışma alanının hak savunuculuğu bağlamında birey ötesi sorumlulukları [Advocacy: Responsibilities of the field of counseling beyond the individual client]. *Türk Psikolojik Danışma ve Rehberlik Dergisi [Turkish Psychological Counseling and Guidance Journal], 4 (33)*, 89-99.

Koc, Z. (2003). Kültüre duyarlı psikolojik danışma yaklaşımı: Kuramsal bir açıklama [Multicultural counseling approach: A theoretical study]. *Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Dergisi [Gazi University Journal of Industrial Arts Education Faculty], 11,* 1-17.

Lee, C. C. (2007). *Social justice: A moral imperative for counselors. ACA Professional Counseling Digest.* Retrived November 2014, from www.counseling.org/ docs/defaultsource/library-archives/professional-counselordigest/ acapcd-07.pdf?sfvrsn=4

Lee, C. C. (2007). *Social justice: A moral imperative for counselors. ACA Professional Counseling Digest.* Retrived November 2014, from www.counseling.org/ docs/defaultsource/library-archives/professional-counselordigest/ acapcd-07.pdf?sfvrsn=4

Lewis, J., & Bradley, L. J. (2000). *Advocacy in counseling* (Ed.). Greensboro: Caps Publication.

Parihk, S.B., Post, P., & Flowers, C. (2011). Relationship between a belief in a just world and social justice advocacy attitudes of school counselors. *Counseling and Values,* 56, 56-72.

Polat, S. (2007). *Eğitim politikalarının sosyal adalet açısından sonuçları konusunda yönetici ve öğretmen görüşleri [E*ducational policies social justice in terms of results about view of administrator and teachers]. (Yayınlanmamış Doktora Tezi) [Unpublished Phd thesis]. Ankara Üniversitesi, Ankara.

Pratto, F., Sidanius, J, Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology 67,* 741-763.

Miranda, A. H., Radliff, K. M., Cooper, J.M., & Eschenbrenner, C. R. (2014). Graduate student perceptions of the impact of trainıng for social justice: Development of a training model. *Psychology in the Schools, 51*(4), 348-365.

Motulsky, S.L., Gere, S. H., Saleem, R. & Trantham, S. M. (2014). Teaching social justice in counseling psychology. *The Counseling Psychologist, 42*(8), 1058–1083.

Nagda, B. A. , Gurin, P. & Lopez, G. E. (2003) Transformative pedagogy for democracy and social justice. *Race Ethnicity and Education, 6,* 165-191, DOI: 10.1080/13613320308199 65-191, DOI: 10.1080/13613320308199.

Overcash, W. S., Calhoun, L. G., Cann, A., & Tedeschi, R. G. (1996). Coping with crises: An examination of the impact of traumatic events on religious beliefs. *Journal of Genetic Psychology, 157*, 455-464.

Ozdemir, M. (2017). Examining the relations among social justice leadership, attitudes towards school and school engagement. *Education and Science, 42,* 267-281. DOI: 10.15390/EB.2017.6281

Pedersen, P. B. (2002). Ethics, competence and Professional issues in cross cultural counseling. In P.B. Petersen, W.J. Lonner, J.E. Tribmble & J.G.Dragun (Ed.), *Counseling across cultures* (5th Ed., pp. 3-28). Sage Publications.

Ridley, C. R., & Kleiner, A. J. (2003). Multicultural counseling competence: History, themes and issues. In D. B. Pope Davis, H. L. K. Coleman, W. M. Liu & R. L. Toporek (Ed.), *Handbook of multicultural competencies in counseling and psychology*. California: Sage Publications

Sam A., Hardy, S.A., Bhattacharjee, A., Reed, A., & Aquino, A. (2010). Moral identity and psychological distance: The case of adolescent parental socialization. *Journal of Adolescence 33,* 111–123.

Sidanius, J. & Pratto, F. (2011). Social dominance theory. In Paul A. M. Van Lange, Arie W. Kruglanski, E. Tory Higgins (Ed.), *Handbook of theories of social psychology* (Volume two). Sage Publication.

Stephan, W. G., & Finlay, K. (1999). The role of empathy in improving intergroup relations. *Social Issues, 55,* 729-743.

Sue, D. W., Arredondo, R., & McDavis, R. J. (1992). Multicultural counseling competencies and standards: A call to the profession. *Journal of Multicultural Counseling and Development, 20,* 64-88.

Tomul, E. (2009). İlköğretim okullarındaki sosyal adalet uygulamalarına ilişkin yönetici görüşleri [Opinions of administrators on social justice practices in elementary schools]. *Eğitim ve Bilim [Education and Science], 34(152*), 126-137.

Torres-Harding, S.R., Siers, B., & Olson, B. (2012). Development and psychometric evaluation of the Social Justice Scale (SJS). *American Journal of Community Psychology*, 59 (1-2), 77-88. DOI: 10.1007/s10464-0111-9478-2

Van Soest, D. (1996). Impact of social work education on student attitudes and behavior concernıng oppression. *Journal of Social Work Education, 32*, 191-202.

Vera, E. M., & Speight, S. L. (2003). Multicultural competence, social justice and counseling psychology: Expanding our roles. *The Counseling Psychologist, 31*, 253-272.

Van der Zee, K. I., & Van Oudenhoven, J. P. (2000). Psychometric qualities of the multicultural personality questionnaire: a multidimensional instrument of multicultural effectiveness. *European Journal of Personality, 14*, 291–309.

Vontress, C. E. (2002). Culture and counseling. In W. J. Lonner, D. L. Dinnel, S. A. Hayes, & D. N. Sattler (Ed.), *Online readings in psychology and culture* (Unit 10, Chapter 1). Retrieved from http://www.wwu.edu/~culture.

Ward, C., Leong, C., & Low, M. (2004). Personality and sojourner adjustment. An exploration of big five and cultural fit proposition. *Journal of Cross-Cultural Psychology, 35*, 137–151.

Yildirim, N., & Akgun, S. (2013). Sivil toplum kuruluşu gönüllülerinin sosyal sistemin meşruiyetine ilişkin algıları, adil dünya inançları ve sosyal baskınlık yönelimleri [Perceived legitimacy of the social system, belief in a just world and social dominance orientation of volunteers in non-governmental organizations]. *Toplum ve Sosyal Hizmet [Community and Social Service],* 22, 115-128.

Yilmaz, F., & Yilmaz, F. (2015). Ahlâkî Kimlik Ölçeği Türkçe formunun geçerlik ve güvenirlik çalışması [The validity and reliability study of Moral Identity Scale's Turkish version]. *Celal Bayar Üniversitesi Sosyal Bilimler Dergisi [Celal Bayar University Social Sciences Journal] 13*,111-133.

Türkiye'deki Ruh Sağlığı Uzmanlarının Sosyal Adalet Tutumlarının İncelenmesi

Atıf:

Kagnici, D. Y., & Denizli, S. (2018). Examining mental health professionals' social justice attitudes in Turkey. *Eurasian Journal of Educational Research*, 75, 19-36, DOI: 10.14689/ejer.2018.75.2

Özet

Problem Durumu: Ruh sağlığı uzmanları her geçen gün daha fazla oranda farklı kültürel özelliklere sahip olan bireylerle çalışmaktadır. Bu durum ruh sağlığı uzmanlarının çok kültürlülük, sosyal adalet ve hak savunuculuğu gibi güncel bir takım yeterliklere sahip olmalarını zorunlu kılmaktadır. Dezavantajlı gruplar sunulan hizmetlerden ve var olan olanaklardan toplumdaki diğer bireyler kadar yararlanamamaktadır. Hak savunuculuğu eşit oranda hizmetlere erişimi için dezavantajlı gruplar adına mücadele etmektir. Bu rol ve sorumluluk ruh sağlığı uzmanlarının güncel sorumluluklarının başında yer almaktadır. Türkiye'de ruh sağlığı alanında hak savunuculuğu ve sosyal adalet ile ilgili araştırma oldukça sınırlıdır. Temel bir sorumluluk ve rol olarak tanımlanan hak savunuculuğu sosyal adalet tutumu ile doğrudan ilişkilidir. Bu bağlamda Türkiye'deki ruh sağlığı uzmanlarının sosyal adalet tutumlarının irdelenmesi önemli görülmektedir.

Araştırmanın Amacı: Bu araştırmanın amacı Türkiye'deki ruh sağlığı uzmanlarının sosyal adalet tutumlarının incelenmesidir. Çok kültürlü psikolojik danışma yaklaşımının tarihsel sürecinde bazı uzmanlar çok kültürlülük konusunda duyarlılığa sahipken, bazı uzmanların çeşitlilik konularına mesafeli oldukları ve hatta çalışmalarında bu konulara yer vermedikleri görülmüştür. Ruh sağlığı uzmanının farkındalığı çok kültürlü psikolojik danışma yeterliliklerinin temelini oluşturmaktadır. Farklı kültürel özelliklere sahip bireylere ruh sağlığı hizmeti sunacak ruh sağlığı uzmanlarının kendi tutum, davranış ve değerlerinin farkında olması ön koşuldur. Bu bağlamda bu araştırmada ruh sağlığı uzmanlarının sosyal adalet tutumları kişisel değişkenler olarak ele alınan adil dünya inancı, açık görüşlülük, sosyal baskınlık yönelimi ve ahlâkî kimlik bağlamında incelenmiştir.

Araştırmanın Yöntemi: Araştırmaya Türkiye genelinden 232 ruh sağlığı (149 psikolojik danışman ve 83 psikolog) uzmanı katılmıştır. Katılımcı psikiyatrist sayısının az olması nedeniyle (n = 24), psikiyatristlerin verileri analizlere dahil edilmemiştir. Araştırmaya katılan ruh sağlığı uzmanlarının % 72.4'ü kadın, % 27.6'sı erkektir. Araştırma çevrimiçi olarak sosyal medya ve derneklerin iletişim grupları aracılığıyla yürütülmüştür. Araştırmanın verisi Sosyal Adalet Ölçeği, Açık Görüşlülük Ölçeği, Sosyal Baskınlık Yönelimi Ölçeği, Adil Dünya İnancı Ölçeği ve Ahlâkî Kimlik Ölçeği ile toplanmıştır. Araştırmada küme analizi kullanılmıştır.

Araştırmanın Bulguları: Küme analizi sonucunda üç temel kategori ortaya çıkmıştır. Sosyal baskınlık yönelimi, adil dünya inancı puanı düşük, ahlâkî kimlik ve açık

görüşlülük puanı yüksek olan ruh sağlığı uzmanları "duyarlı"; sosyal baskınlık yönelimi ve adil dünya inancı puanı yüksek, ahlâkî kimlik ve açık görüşlülük puanı düşük olan ruh sağlığı uzmanları "katı"; sosyal baskınlık yönelimi, adil dünya inancı, ahlâkî kimlik ve açık görüşlülük puanı düşük olan ruh sağlı uzmanları ise "ilgisiz" olarak tanımlanmıştır. ANOVA sonuçlarına göre sosyal adalet alt ölçeğinde bu üç kategori arasında istatistik açıdan anlamlı farklılıklar bulunmuştur (F (2,229) = 18.19, p < .01). "Duyarlı" ruh sağlığı uzmanlarının sosyal adalet tutum puanlarının diğer ruh sağlığı uzmanlarından yüksek olduğu görülmüştür. Benzer şekilde algılanan davranışsal kontrol alt ölçeğinde de bu üç kategori arasında istatistik açıdan anlamlı farklılıklar bulunmuştur (F (2, 229) = 27.17, p < .01). "Duyarlı" ruh sağlığı uzmanlarının algılanan davranışsal kontrol alt ölçeği puanlarının diğer ruh sağlığı uzmanlarından yüksek olduğu sonucuna ulaşılmıştır. Davranışsal niyet alt ölçeğinde de bu üç kategori arasında istatistik açıdan anlamlı farklılıklar bulunmuştur (F (2, 229) = 16.47, p < .01). "Duyarlı" ruh sağlığı uzmanlarının niyet alt ölçeği puanlarının diğer ruh sağlığı uzmanlarından yüksek olduğu sonucuna ulaşılmıştır. Öznel normlar alt ölçeğinde ise anlamlı bir farklılık bulunmamıştır.

Araştırmanın Sonuçları ve Önerileri: Araştırmanın bulgularına göre "duyarlı" olan ruh sağlığı uzmanlarının sosyal adalet tutum puanları "katı" ve "ilgisiz" ruh sağlığı uzmanlarından daha yüksek bulunmuştur. Sonuçlar incelendiğinde Türkiye'deki ruh sağlığı uzmanlarının bir kısmının kültürel konularda "katı" bir kısmının ise "ilgisiz" olduğu görülmüştür. Katı olarak nitelendiren grubun sosyal baskınlık yöneliminin ve adil dünya inancının yüksek olması, yani toplumda yaşayan bazı grupların diğerlerinden daha üstün olduğunu ve herkesin hak ettiği şekilde yaşadığını düşünmeleri oldukça düşündürücüdür. Bu grubun aynı zamanda açık görüşlülük ve ahlâkî kimlik puanları da düşük çıkmıştır. Tüm bu bulgular birlikte değerlendirildiğinde toplumdaki tüm gruplara ruh sağlığı hizmeti sunma görevi üstlenmiş olan bir grup ruh sağlığı uzmanının bazı kültürel grupları daha üstün görebileceği; dezavantajlı grupların karşılaştıkları ayrımcı tutumları bir şekilde hak ettiklerini düşünebileceği; farklılıklar karşısında açık görüşlü olamayabileceği ve ahlâkî açıdan olması beklenen düzeyde gelişemediği gibi bir olasılık ortaya çıkmaktadır. Yine bulgular incelendiğinde bir grup ruh sağlığı uzmanının ise "ilgisiz" olduğu görülmüştür. Her iki grubun da dezavantajlı gruplarla çalışma konusunda yetersiz kalma olasılıklarının oldukça yüksek olduğu açıktır.

Toplumdaki tüm bireylerin etkili bir şekilde ruh sağlığı hizmetlerinden faydalanmaları gerekmektedir. Bunun içinde ruh sağlığı hizmeti sunan uzmanların farklılık konusunda duyarlı, bilgili ve farklılıklara sahip bireylerle çalışabilecek yeterliğe sahip olması gerekmektedir. Bu konuda ruh sağlığı eğitimcilerine ciddi sorumluluklar düşmektedir. Özellikle gerek psikoloji gerekse psikolojik danışma eğitiminde çok kültürlülük ve sosyal adalet konularına yer verilmeli ve ruh sağlığı uzmanları farklı kültürel özelliklere sahip danışanlara ve gruplara ruh sağlığı hizmeti sunabilecek yeterliklerle donatılmalıdır.

*Anahtar Kelimeler***:** Sosyal adalet, ruh sağlığı uzmanları, adil dünya inancı, sosyal baskınlık yönelimi, açık görüşlülük, ahlâkî kimlik.

Eurasian Journal of Educational Research 75 (2018) 37-58

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

As a Potential Source of Error, Measuring the Tendency of University Students to Copy the Answers: A Scale Development Study

Ergul DEMIR [1]

ARTICLE INFO

Article History:
Received: 07 Jan. 2018
Received in revised form: 22 Mar. 2018
Accepted: 28 Apr. 2018
DOI: 10.14689/ejer.2018.75.3

Keywords
answer-copying
dishonesty
scale development
copy-detection.

ABSTRACT

Purpose: The answer-copying tendency has the potential to detect suspicious answer patterns for prior distributions of statistical detection techniques. The aim of this study is to develop a valid and reliable measurement tool as a scale in order to observe the tendency of university students' copying of answers. Also, it is aimed to provide evidence with more comprehensive validity and reliability studies than the previously available researches. **Research Methods**: This is a scale development study. The "Answer-Copying Tendency Scale in University Students (ACTS)" was developed under the "Classical Test Theory". Other theories were also considered, especially "Item Response Theory". After preliminary studies and item writing, a trial application with 711 students and main applications with 909 students was conducted. Structural validity, item and test descriptive statistics, item discriminations, inconsistency and test–retest reliability, classification accuracy, and item bias with differential item functioning were examined. **Findings**: The ACTS composed of 2 factor and 20 items. Total scores and item scores distributions are normal. Item discriminations are very high and over 0.40. α inconsistency coefficients are over 0.88 and test–retest reliability coefficient is 0.804. It provides highly correct classifications according to the students' answer-copying positions. There is no significant and serious DIF on items. **Implications for Research and Practice**: Unlike similar examples, it was studied on the large groups and used more comprehensive techniques to obtain evidences. Results show that the validity and reliability levels of the ACTS are very high. The ACTS can be used to understand the nature of the answer-copying. Also, and more importantly, it is thought that the ACTS can be used to detect suspicious answer patterns for prior distributions.

[1] Ankara University, Educational Science Faculty, TURKEY, e-mail:erguldemir@ankara.edu.tr, ORCID: https://orcid.org/0000-0002-3708-8013

Introduction

Academic integrity, dishonesty, cheating, or, in a more limited use, answer-copying has become an increasingly important problem in schools. Academic integrity is a more comprehensive definition including answer-copying, having proxies, plagiarism, academic misconduct, falsifying, etc. (Mullens, 2000). Some research shows that answer-copying has been observed at very serious proportions (Burke, 1997; Genereux & McLeod, 1995; May & Loyd, 1993; McCabe, 1993). Dwyer and Hecht (1994) stated that such dishonest behaviors by students have been increasing with increasing class sizes and reduced instructional resources. Similarly, McCabe and Trevino (1996) reported that the percentage of cheating behaviors among college students rose slightly. The inability to take control of this tendency leads to the injured reputation of institutions and hurts honest students, their families, and all of society (Aaron & Georgia, 1994).

Statistically, answer-copying is a source of systematic error and bias on items or tests. It may lead to unfair results. Unfortunately, it is very hard to detect answer-copying because initial response patterns are not easy to determine accurately for the statistical models. Improving the statistical techniques to detect answer-copying has been studied since the 1970s (Angoff, 1974; Frary, Tideman, & Watts, 1977; Holland, 1996; Maeda & Zhang, 2016; van der Linden & Sotaridona, 2006; Wesolowsky, 2000; Wollack, 1997). According to Dwyer and Hact (1994), some probabilistic techniques to detect cheaters have been used in American higher education since the 1920s. Generally, these techniques depend on matching responses between the copier and the source with a complex standardization process. Thus, the copiers and the sources have to be defined accurately. Also, there are some other challenges to using these techniques. As stated by van der Linden & Sotaridona (2006), these common techniques are limited by the null distribution of the set of items on which the statistics is defined. It is based on population-based statistics, and it is possible to arrive at unfair results. Alternatively, performances of IRT-based techniques also depend on the accurate estimation of copier ability, and it is obvious that the responses of copiers may be contaminated. For these techniques, especially, suspicious answer patterns should be defined accurately for prior and posterior distributions. It is understood that although there are some statistical techniques to detect answer-copying, these are limited or show doubt, a potential source of error, or just a probability.

Another way to understand the nature of copying is to consider related factors (Gerdeman, 2000; Hughes & McCabe, 2006). Some research reported that dishonest behaviors, mostly including answer-copying, were related to individual characteristics like GPA, age, gender, etc. (Crown & Spiller, 1998; McCabe & Trevino, 1997; Selçuk, 1995; Whitley, 1998). Students with lower GPAs, younger students, and males are most likely to cheat. Also, these groups show more tolerance for cheating behaviors.

Although they are less likely to engage in dishonesty or copying, dishonest behaviors seem to be related to particular educators and institutional policies (Aaron, 1992; Genereux & McLeod, 1995; McCabe & Trevino, 1996). Other important factors are attitudes, perceptions, and tendencies toward copying (Genereux & McLeod, 1995; Hughes & McCabe, 2006; McCabe & Trevino, 1997; Öztürk & Yeşilyaprak, 1997). The students with high work ethic, self-esteem, and lower test anxiety are less likely to cheat. On the other hand, the prevalence of cheating and the perception of cheating as acceptable increases cheating behaviors.

In many studies, questionnaires or self-reporting have been used to observe students' perceptions or tendencies on dishonest behavior such as copying (Bolin, 2004; McCabe & Travino, 1997; Selçuk, 1995). There are a few measurement tools available developed in these contexts (Eminoğlu & Nartgün, 2009; Gardner & Miller, 1988; LaGrange, 1992). Unfortunately, these tools are very limited in use and mostly unavailable. Generally, these tools were developed on a small and limited group and have comprehensive context. Most of them provide less proof about validity and reliability. Most of them were developed a long time ago. So, although cheating has been studied for a long time, it is obvious that valid and reliable tools are still needed.

Research Objectives

The aim of this study is to develop a valid and reliable measurement tool as a scale in order to observe the tendency of university students to copy answers. Also, it is aimed to provide evidence with comprehensive validity and reliability studies. With this aim, psychometric studies have been carried out for the "Answer-Copying Tendency Scale in University Students (ACTS)": (1) Structural validity, (2) item and test descriptive statistics, (3) item discriminations, (4) inconsistency and test–retest reliability, (5) classification accuracy, and (6) item bias with differential item functioning.

Statistically, the detection of answer-copying is a challenge. On the other hand, research shows that attitudes, perceptions, and tendencies are related to answer-copying behavior. In order to understand the nature of answer-copying and then detect it, we need a valid, reliable, and, most importantly, available tool to be used in this context. Also, the answer-copying tendency has the potential to make suspicious answer patterns detectable for prior/initial distributions of statistical detection techniques.

Method

This is a scale development study. The "Answer-Copying Tendency Scale in University Students (ACTS)" was developed under the "True Score Theory" or "Classical Test Theory". In this process, other theories were also considered, especially

"Item Response Theory". The scale development steps were followed by considering DeVellis (2003).

Preliminary Studies and Item Writing

Developing the ACTS began with the preliminary studies. First, the trait and the aim were defined. The aim of the scale was defined as to measure the tendency of university students to copy answers. After that, the observable behaviors of students' copying tendencies was tried to be linked with the related literature. Simultaneously, around 80 university students were asked to write an essay on their perceptions and views about answer-copying. These documents were analyzed, and a total of 123 draft expressions were prepared for review by experts. The prepared draft form was sent to six specialists and academics working in the field of educational sciences. According to their opinions, 35 items were cancelled, 14 items were reorganized, and 79 items were accepted as they were. Of the 93 items, there were 34 items with negative direction. At the same time, the experts were asked about which type of scale would be used most effectively. Some experts recommended the Likert-type scale with five categories, while others recommended more than five categories. In response, and also by considering the related literature, it was decided that two differently numbered scoring categories would be used in the trial application. The first would be the classic Likert-type scale with five categories, and the second would be to score between 0 and 10 points. At the end of the preliminary studies, the trial application form was organized with 93 items.

Trial and Main Application

Trial application was executed with 711 undergraduate students from 16 universities and 18 faculties in May and June of 2017 in Turkey as a paper-pencil test. The main application was executed with 909 undergraduate students from 29 universities and 30 faculties in November and December of 2017 in Turkey as a web based application. Some demographic characteristics of the students whom participated in trial and main applications are given at Table 1.

Data Analysis

Defining the factor structure, exploratory and confirmatory studies were executed. "Principal component analysis" and "multidimensional scaling with ALSCAL method" was used with the trial application data. "Confirmatory factor analysis" was used with the main application data. As the test and item statistics, it was analyzed the distribution of the scores and calculated the descriptive statistics. As the item discrimination, item-total score correlations were calculated with Pearson' Product Moments correlation coefficient. For reliability, α inconsistency coefficient and test-retest reliability were calculated. For classification accuracy, discriminant function

analysis was used with four classification model. Finally, item bias was examined with the "lordif" techniques based on Item Response Theory. Before all analyses, main assumptions were checked in detail. The softwares SPSS 24.0 and Lisrel 8.7 and R 3.4.4 were used for these analyses.

Table 1

Students' Demographic Characteristics for Trial and Main Application Samples

Trial Application		*n*	%	*Main Application*		*n*	%
Gender	Female	525	73.8	Gender	Female	674	74.1
	Male	158	22.2		Male	235	25.9
	Missing	28	3.9		Total	909	100.0
	Total	711	100.0	University	Ankara University	325	35.8
University	Ankara University	447	62.9		Trakya University	115	12.7
	Hacettepe University	161	22.6		Gazi University	77	8.5
	TED University	29	4.1		Hacettepe University	69	7.6
	Konya Selçuk University	23	3.2		Other	321	35.3
	Other	22	2.9		Missing	2	0.2
	Missing	29	4.1		Total	909	100.0
	Total	711	100.0	Faculty	Education	271	29.8
Faculty	Educational Sciences	263	37.0		Educational Sciences	204	22.4
	Education	173	24.3		Science & Literature	145	16.0
	Theology	55	7.7		Language & History-Geography	88	9.7
	Science	47	6.6		Others	201	22.1
	Other	145	20.2		Total	909	100.0
	Missing	28	3.9	Class	Preparation and 1	212	23.3
	Total	711	100.0		2	235	25.9
Class	1	152	21.4		3	194	21.3
	2	206	29.0		4	218	24.0
	3	132	18.6		5 and Graduate	50	5.6
	4	158	22.2		Total	909	100.0
	Graduate	34	4.8				
	Missing	29	4.1				
	Total	711	100.0				

Results

Structural Validity: Exploratory and Confirmatory Studies

Exploratory studies for the ACTS were executed on the trial application data, and confirmatory analysis was executed on the main application data. As an exploratory analysis, "principal component analysis (PCA)" and "multidimensional scaling with ALSCAL model" were used respectively. PCA was executed with both the classic five categories of Likert-type scores and the scores between 0 and 10 separately. Before analysis, some items were recoded in order to equalize the way of all items. Missing values were checked, and there was no serious missing data problem. Nor were there any outliers (max. Mahalanobis distance < χ^2_{kritik}=25).

For the five-category Likert-type scale data, sampling is adequate (KMO=0.974), and multiple correlations among variables are statistically significant (for Bartlett's test of sphericity, χ^2=8895.179, df=190 and p<0.001). After the data reduction, 20 items with one factor could be identified as a structure. Communalities of each item are between 0.412 and 0.725. The eigenvalue of one factor is 11.463, and the total variance explained is 57%. Factor loadings are between 0.649 and 0.852.

Similarly, for scoring 0 to 10, sampling is adequate (KMO=0.959), and multiple correlations among variables are statistically significant (for Bartlett's test of sphericity, χ^2=6585.353, df=190 and p<0.001). After the data reduction with varimax and oblique rotation, 20 items with two factors could be identified as a structure. The communalities of each item are between 0.485 and 0.744. The eigenvalue of the first factor is 9.684 and second is 2.659. Total variance explained for the first factor is 48.42%, the second is 13.29%, and the total is 61.71%. The correlation between factors is statistically significant an, shows a negative and moderate relationship (r=-0.49 and p<0.05). These items and factor loadings are shown in the Table 2.

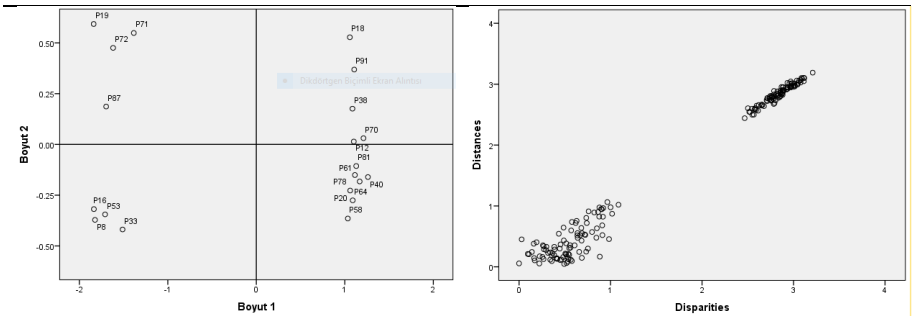
As seen at Table 2, 12 of 20 items are in the first factor and eight items are in the second factor. These eight items are the items with negative directions. The first factor was named "negative perception of exam and grade (NEGALGI)", and the second factor was named "ethical value (ETIK)". These two factors have negative correlations to each other (r=-0.49 and p<0.01). According to these, after standardization, students who have high ethical values also have lower levels of negative perception about exams and grades.

Table 2

Factor Loadings for ACTS Items

No.	Item	F1	F2
12	I'll copy answers when I get the chance.	0.791	
18	Having friends I know that get higher scores by copying answers, I also have the ambition to copy answers.	0.789	
20	I think about copying answers for the examinations that I have not prepared enough for.	0.836	
38	Copying answers can be enough to pass the exam.	0.801	
40	I feel compelled to copy answers myself in some lessons.	0.817	
58	It makes sense for me to copy answers in lessons that I will forget and not use in the future.	0.785	
61	Even if I do not do it, it gives me confidence to know that I can copy answers at the exam.	0.792	
64	It makes sense for me to copy answers to the questions I do not know.	0.821	
70	I will copy answers if I know I will not be punished.	0.808	
78	It makes more sense to copy answers when I memorize so much information that will not work in my own life.	0.752	
81	Anxiety about earning high grades pushes me to copy answers.	0.823	
91	If I am not afraid of getting caught, I'll copy answers.	0.811	
8	Copying answers makes a fool of a person.		-0.763
16	I see copying answers as an unfair advantage.		-0.761
19	I'm absolutely against copying answers.		-0.755
33	Copying answers is disrespectful to the teacher's endeavors.		-0.751
53	Copying answers is not my achievement but my deceit.		-0.744
71	Defending copying answers is completely nonsense.		-0.741
72	I would like to take real deterrent measures to prevent students from copying answers.		-0.694
87	I believe that the copying answers is immoral behavior.		-0.674
	Eigenvalue	9.684	2.659
	Total Variance Explained (%)	48.42	13.26

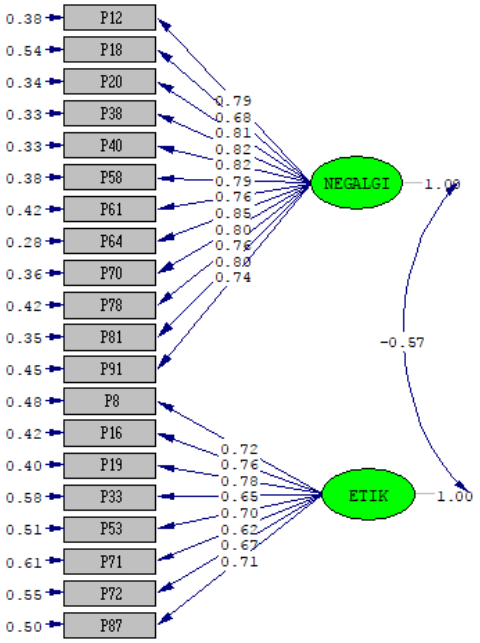
After PCA, the multidimensional scaling ALSCAL model provides visual and supportive evidence of the structure. According to the Euclidean distances, item locations are shown in Graphic 1.



Graphic 1. Item Distribution According to Multidimensional Scaling (ALSCAL) Euclidean Distances

As seen in Graphic 1, items are clearly separated by two factors. In the right-hand graphic, eight items are on one factor and twelve items are on the other factor. Furthermore, in the left-hand graphics, these two factors show a linear relationship. Also, outputs of MDA-ALSCAL show that the model provides good-fit with high variance accounted for (Stress=0.08623 and RSQ=0.98503).

After the exploratory studies, "confirmatory factor analysis (CFA)" was conducted on the main application data. Graphic 2 shows the standardized values for each path.



Graphic 2. Standardized Path Coefficients According to the Results of the CFA

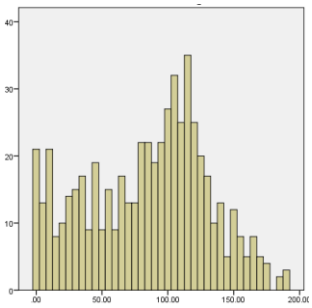
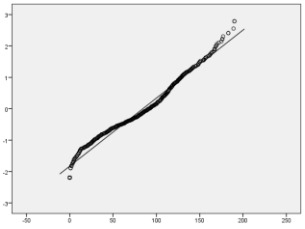
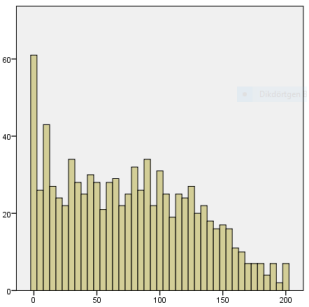
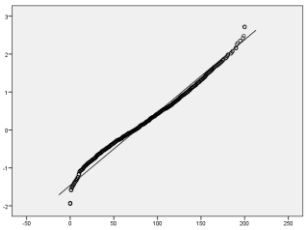
In Graphic 2, all paths are statistically significant (t_i>1.96 and p_i<0.05). Standardized solutions of the errors are under 0.90. Similar to the exploratory studies' results, there are moderate and negative correlations between the factors. Also, goodness of fit statistics show at least acceptable model-data fit (χ^2/sd=2.79, RMSEA=0.056, SRMR=0.036, GFI=0.92, NFI=0.98, CFI=0.99, ECVI Free model =41.45 and Saturated model=0.75, AIC Free model=2344.37 and Saturated Model=420.0).

Total Score Distributions and Descriptive Statistics

The ACTS has 20 items, which are each scored between 0 and 10. The available total scores are between 0 and 200. In Table 3, the total score distributions and descriptive statistics are given for both trial and main applications.

Table 3

ACTS Total Score Distributions and Descriptive Statistics

	Statistics	Value	Histogram	Normal Q-Q Plot
Trial Application	Mean	84.29		
	5% Trim.Mean	84.12		
	SE (Mean)	1.942		
	Median	91		
	Std.Dev	46.127		
	Minimum	0		
	Maximum	190		
	Skewness	-0.146		
	Kurtosis	-0.806		
Main Application	Mean	76.13		
	5% Trim.Mean	74.48		
	SE (Mean)	1.727		
	Median	74		
	Std.Dev.	52.054		
	Minimum	0		
	Maximum	200		
	Skewness	0.288		
	Kurtosis	-0.877		

As seen in Table 3, total scores distribute normally. The mean, trimmed mean, and median are close to each other. Skewness and kurtosis are between (-1, +1). Also, the graphics supporting the normality.

Item Statistics and Item-Total Score Correlations

The item statistics and item-total score correlations of the ACTS items were calculated for both trial and main application data. The results are given in Table 4.

Table 4

Item Statistics and Item-Total Score Correlations of the ACTS Items

Item	Trial Application						Main Application					
	Mean	Med.	Std.D.	Skew.	Kurt.	r	Mean	Med.	Std.D.	Skew.	Kurt.	r
12	4.49	5	3.380	0.144	-1.246	.736**	3.83	3	3.412	0.549	-1.021	.803**
18	4.08	4	3.447	0.320	-1.206	.613**	3.49	2	3.586	0.590	-1.110	.617**
20	4.69	5	3.422	-0.029	-1.329	.752**	4.13	4	3.491	0.315	-1.276	.779**
38	4.26	5	3.215	0.154	-1.129	.766**	3.82	3	3.501	0.440	-1.155	.816**
40	4.66	5	3.649	0.038	-1.435	.771**	3.78	3	3.666	0.488	-1.240	.760**
58	5.06	5	3.464	-0.076	-1.280	.760**	4.65	5	3.865	0.091	-1.559	.796**
61	4.76	5	3.563	0.013	-1.360	.721**	3.94	3	3.774	0.390	-1.378	.714**
64	4.90	5	3.405	-0.094	-1.253	.812**	3.98	4	3.543	0.368	-1.234	.844**
70	4.80	5	3.618	0.047	-1.374	.769**	4.50	4	3.816	0.216	-1.471	.812**
78	4.56	5	3.532	0.119	-1.319	.747**	3.67	3	3.607	0.517	-1.173	.779**
81	4.83	5	3.557	0.008	-1.347	.764**	4.52	4	3.759	0.170	-1.493	.806**
91	4.27	5	3.498	0.203	-1.287	.696**	3.87	3	3.604	0.466	-1.202	.819**
8*	6.69	7	3.210	-0.551	-0.951	.596**	7.27	8	3.093	-0.986	-0.113	.621**
16*	6.79	8	3.104	-0.621	-0.751	.608**	7.41	9	3.105	-1.028	-0.110	.682**
19*	5.32	5	3.545	-0.061	-1.328	.658**	5.77	6	3.538	-0.231	-1.306	.749**
33*	6.64	7	2.880	-0.551	-0.545	.529**	6.43	7	3.461	-0.559	-1.022	.569**
53*	6.58	7	3.120	-0.570	-0.749	.571**	6.80	8	3.395	-0.769	-0.714	.676**
71*	5.60	5	3.310	-0.156	-1.124	.488**	6.05	6	3.516	-0.355	-1.238	.646**
72*	5.77	6	3.303	-0.264	-1.085	.586**	5.83	6	3.626	-0.275	-1.332	.697**
87*	5.88	6	3.398	-0.245	-1.217	.576**	6.50	7	3.459	-0.578	-1.015	.678**

* Second factor items were recoded before the estimations.
**p<0.01

As seen in Table 4, trial and main application results are similar. Item means and medians are close to each other. Skewness values are between (-1, +1). Kurtosis values are mostly under -1, but between (-1.5, +1.5). Tabachnick and Fidel (2013, p.80) stated that non-strong skewness and kurtosis violations cannot lead the difference for the statistics in large samples. As a result, item scores show reasonable normality. Also, all item-total score correlations are significant at the level of 0.01. This is strong evidence of the discriminative validity of the items and the test.

Reliability Studies

For the reliability of the ACTS, first α inconsistency coefficients were calculated for both trial and main applications. If there is no violation of normality, α is an available estimation. The results are given in Table 5.

Table 5

a Coefficients for the ACTS

Factors	k	Trial Application		Main Application	
		n	α	n	α
Negative perception of exams and grades (NEGALGI)	12	711	0.950	909	0.955
Ethical value (ETIK)	8	711	0.884	909	0.907
Total	20	711	0.942	909	0.955

As seen in Table 5, the inconsistency of the ACTS is very high for both the factor level and the total test. Besides the inconsistency estimates, the test–retest reliability was considered. Test–retest application was conducted with 95 students in December of 2017 after two weeks from the main application. There was no normality violation, and the Pearson product-moment correlation coefficient was used to estimate. Results show that the ACTS has a high level of test–retest reliability ($r=0.804$ ve $p<0.001$).

Classification Accuracy

In the trial and main applications of the ACTS, participants were asked whether they copied answers, gave answers for others to copy, and/or witnessed answer-copying. Similar results were obtained from both the trial and main applications. So, just the main application data was considered for further analysis.

Table 6

Taking, Giving, and Witnessing the Answer-Copying among Undergraduates

	I did		I gave		I witnessed	
	n	%	n	%	n	%
Yes	444	48.8	360	39.6	126	13.9
No	465	51.2	549	60.4	783	86.1
Total	909	100.0	909	100.0	907	99.8

As seen in Table 6, most of the students stated that they did not copy answers (48.8%) or did not gave answers for others to copy (39.6%). On the other hand, the vast majority stated that they witnessed answer-copying (86.1%). The proportion of copying answers (60.4%) was more than giving answers for others to copy (51.2%). As validity evidence, it is expected that the ACTS can classify students according to their answer-copying positions. For this purpose, "discriminant function analysis (DFA)" was used. A total of four discriminant models were identified depending on the

answer-copying positions and ACTS factors. Before the analyses, the main assumptions were checked. As mentioned before, there are no violations about missignes, outliers, and normalities. Furthermore, the homogeneity of the variance-covariance matrices was checked by using Box's M statistics. These values are not significant at the level of 0.001, and there is no violation (Hair, Black, Babin, & Anderson, 2014, p.250). Descriptive statistics for each model are given in Table 7. And the test result for group differences and canonical discriminant functions are given in Table 8.

Table 7

Descriptive Statistics for Classification Models

Model	Predictor(s)	Grouping Variable		n	Mean	Std.Dev.
1[a]	ACTS_Total	Answer-Copying	Yes	444	52.80	48.132
			No	465	98.40	45.529
			Total	909	76.13	52.054
2[b]	ACTS_NEGALGI	Answer-Copying	Yes	444	31.28	32.178
			No	465	64.31	31.156
			Total	909	48.18	35.694
	ACTS _ETIK	Answer-Copying	Yes	444	58.48	20.291
			No	465	45.91	20.173
			Total	909	52.05	21.175
3[c]	ACTS _Total	Giving answer-copy	Yes	360	54.43	48.529
			No	549	90.35	49.345
			Total	909	76.13	52.054
4[d]	ACTS _NEGALGI	Giving answer-copy	Yes	360	33.09	33.052
			No	549	58.07	33.880
			Total	909	48.18	35.694
	ACTS _ETIK	Giving answer-copy	Yes	360	58.66	19.800
			No	549	47.72	20.941
			Total	909	52.05	21.175

[a]Box' M=1.402, F=1.400, df_1=1, df_2=2464420.891 and p=0.237
[b]Box' M=8.786, F=2.922, df_1=3, df_2=158752599.4 and p=0.033
[c]Box' M=0.120, F=0.120, df_1=1, df_2=2194294.214 and p=0.729
[d]Box' M=7.901, F=2.627, df_1=1, df_2=30459579.07 and p=0.049

Table 8

Test Results for Group Difference and Canonical Discriminant Functions

Model	Predictors	Equality of group means					Canonical Discriminant Functions					
		Wilks ′λ	F	df_1	df_2	p	Wilks ′λ	χ^2	df	p	Eigen value	Cann. Corr.
1	ACTS_Total	0.808	215.5	1	907	0.000	0.808	193.2	1	0.000	0.238	0.438
2	ACTS_NEGALGI	0.786	247.2	1	907	0.000	0.786	218.4	2	0.000	0.273	0.463
	ACTS_ETIK	0.912	87.7	1	907	0.000						
3	ACTS _Total	0.886	116.7	1	907	0.000	0.886	109.8	1	0.000	0.129	0.338
4	ACTS _NEGALGI	0.883	120.5	1	907	0.000	0.881	114.5	2	0.000	0.135	0.345
	ACTS _ETIK	0.936	62.0	1	907	0.000						

As seen in Table 8, each model shows significant group differences. Also, each model is statistically significant. After that, the maximum chance criteria were calculated according to the recommendation of Tabachnick and Fidel (2012, p.406) for unequal group sizes. The proportions of the correct classifications and related criteria are given in Table 9.

Table 9

Proportions of the Correct Classifications and Maximum Chance Criteria

Model	Correct Classification (%)	Maximum Chance (%)
1	69.9	50.0
2	71.9	50.0
3	68.2	52.2
4	68.4	52.2

As seen in Table 9, each model provides correct classifications beyond chance. For 1st and 2nd models, this ratio is higher. It is expected that the correct classification values should be higher, at least over 10% of maximum chance values (Hair et al.., 2014, p.261). According to this criteria, all models have high power for classifications. As a result, the ACTS can predict significantly both answer-copying and giving the answers to copy.

Differential Item Functioning Studies

DIF studies for the ACTS were conducted on the main application data. Gender sub-groups and faculties are considered the group variable. For gender, females were defined as the reference group (n=674, 74.1%), and males were defined as the focal group (n=235, 25.9%). For faculties, education faculties were defined as the reference group (n=440, 48.4%), and other faculties were defined as the focal group (n=469, 51.6%). "Logistic Ordinal Regression Differential Item Functioning using IRT (lordif)" (Choi, Gibbons, & Crane, 2011) was used for the estimations. This technique was developed for polythomous items and based on IRT with GPCM or GRM models. Both uniform and nonuniform DIF can be detected. DIF results for the ACTS are given in Table 10 and Table 11.

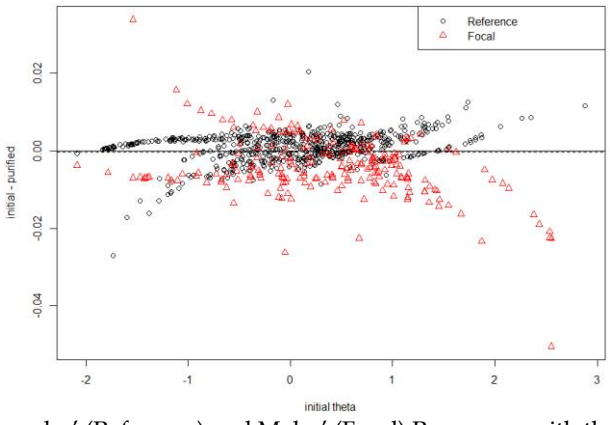
Table 10

*DIF Results for Gender Sub-Groups Obtained with lordif**

	Probability			Cox & Snell			Nagelkerke			McFadden			β_{12}
	χ^2_{12}	χ^2_{13}	χ^2_{23}	R^2_{12}	R^2_{13}	R^2_{23}	R^2_{12}	R^2_{13}	R^2_{23}	R^2_{12}	R^2_{13}	R^2_{23}	
1	0.8367	0.9478	0.7992	0.0000	0.0001	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0e+00	0.0013
2	0.0162	0.0239	0.1940	0.0017	0.0022	0.0005	0.0017	0.0022	0.0005	0.0014	0.0018	4e-04	0.0036
3	0.9611	0.9694	0.8069	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0e+00	0.0003
4	0.9632	0.1598	0.0556	0.0000	0.0023	0.0023	0.0000	0.0024	0.0024	0.0000	0.0010	1e-03	0.0003
5	0.0106	0.0351	0.6859	0.0032	0.0033	0.0001	0.0033	0.0033	0.0001	0.0016	0.0017	0e+00	0.0201
6	0.0248	0.0800	0.9156	0.0017	0.0017	0.0000	0.0017	0.0017	0.0000	0.0012	0.0012	0e+00	0.0054
7	0.3088	0.4712	0.4933	0.0008	0.0011	0.0004	0.0008	0.0012	0.0004	0.0003	0.0004	1e-04	0.0071
8	0.6653	0.5713	0.3342	0.0001	0.0003	0.0003	0.0001	0.0003	0.0003	0.0000	0.0003	2e-04	0.0013
9	0.5654	0.7148	0.5593	0.0001	0.0003	0.0001	0.0001	0.0003	0.0001	0.0001	0.0002	1e-04	0.0024
10	0.0263	0.0225	0.1037	0.0030	0.0046	0.0016	0.0031	0.0047	0.0016	0.0013	0.0020	7e-04	0.0201
11	0.2684	0.3812	0.4016	0.0004	0.0007	0.0002	0.0004	0.0007	0.0002	0.0003	0.0005	2e-04	0.0033
12	0.6771	0.8728	0.7534	0.0001	0.0001	0.0000	0.0001	0.0001	0.0000	0.0000	0.0001	0e+00	0.0023
13	0.1974	0.3201	0.4324	0.0004	0.0005	0.0001	0.0004	0.0005	0.0001	0.0004	0.0006	2e-04	0.0065
14	0.1736	0.1418	0.1517	0.0006	0.0012	0.0006	0.0006	0.0012	0.0006	0.0005	0.0010	5e-04	0.0076
15	**0.0044**	**0.0067**	**0.1691**	**0.0052**	**0.0064**	**0.0012**	**0.0053**	**0.0065**	**0.0012**	**0.0020**	**0.0025**	**5e-04**	**0.0246**
16	0.6195	0.8237	0.7070	0.0002	0.0002	0.0001	0.0002	0.0002	0.0001	0.0001	0.0001	0e+00	0.0030
17	0.2253	0.4451	0.7001	0.0005	0.0006	0.0001	0.0006	0.0006	0.0001	0.0004	0.0004	0e+00	0.0048
18	0.5429	0.4023	0.2284	0.0001	0.0005	0.0004	0.0001	0.0005	0.0004	0.0001	0.0004	4e-04	0.0021
19	0.3798	0.5318	0.4832	0.0005	0.0008	0.0003	0.0005	0.0008	0.0003	0.0002	0.0003	1e-04	0.0065
20	0.0664	0.1539	0.5412	0.0009	0.0010	0.0001	0.0009	0.0010	0.0001	0.0008	0.0009	1e-04	0.0096

*Replication=100, α=0.01, ΔR^2 =0.02, Δβ=0.1

As seen in Table 10, except the 15th item, there is no significant DIF. The 15th item shows DIF with all χ^2 , R^2 , and β values. These differences are observed between the 1st-2nd and 1st-3rd models. So, it is possible that DIF should be uniform. The 15th item of the ACTS is "defending copying answers is completely nonsense". Distributions of females' and males' responses with their tendency levels (θ) are given in Graphic 3.



Graphic 3. Females' (Reference) and Males' (Focal) Responses with their Tendency Levels (θ) in the 15th Item of the ACTS

As seen in Graphic 3, at each θ level, females are located in the middle, whereas males have different locations. At the lower θ levels, males show more admittance to the 15th item. At the higher θ levels, the opposite is the case. This can show real differences between the gender sub-groups. Indeed, some research reported that males are more prone to cheating and copying (Crown & Spiller, 1998; Genereux & McLeod, 1995; Selçuk, 1995; Whitley, 1998). If there is a real difference in the 15th item, it can be stated that the ACTS has no DIF between gender sub-groups.

Table 11

*DIF Results for Faculty Sub-Groups Obtained with lordif**

	Probability			Cox & Snell			Nagelkerke			McFadden			β_{12}
	χ^2_{12}	χ^2_{13}	χ^2_{23}	R^2_{12}	R^2_{13}	R^2_{23}	R^2_{12}	R^2_{13}	R^2_{23}	R^2_{12}	R^2_{13}	R^2_{23}	
1	0.0998	0.2560	0.8968	0.0018	0.0019	0.0000	0.0019	0.0019	0.0000	7e-04	0.0008	0.0000	0.0029
2	0.6035	0.3700	0.1898	0.0001	0.0006	0.0005	0.0001	0.0006	0.0005	1e-04	0.0005	0.0004	0.0003
3	0.9914	0.4724	0.2207	0.0000	0.0009	0.0009	0.0000	0.0009	0.0009	0e+00	0.0004	0.0004	0.0000
4	0.9297	0.9890	0.9044	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0e+00	0.0000	0.0000	0.0001
5	**0.5920**	**0.0083**	**0.0023**	**0.0001**	**0.0047**	**0.0046**	**0.0001**	**0.0048**	**0.0046**	**1e-04**	**0.0024**	**0.0023**	**0.0003**
6	0.7174	0.8093	0.5888	0.0000	0.0001	0.0001	0.0000	0.0001	0.0001	0e+00	0.0001	0.0001	0.0003
7	0.6390	0.2473	0.1086	0.0002	0.0021	0.0020	0.0002	0.0021	0.0020	1e-04	0.0007	0.0007	0.0003
8	0.8153	0.6698	0.3874	0.0000	0.0002	0.0002	0.0000	0.0002	0.0002	0e+00	0.0002	0.0002	0.0001
9	0.0867	0.1553	0.3741	0.0012	0.0015	0.0003	0.0012	0.0015	0.0003	8e-04	0.0010	0.0002	0.0027
10	0.1785	0.3856	0.7576	0.0011	0.0012	0.0001	0.0011	0.0012	0.0001	5e-04	0.0005	0.0000	0.0028
11	0.2993	0.4566	0.4837	0.0004	0.0005	0.0002	0.0004	0.0005	0.0002	3e-04	0.0004	0.0001	0.0009
12	0.2539	0.1787	0.1433	0.0006	0.0016	0.0010	0.0006	0.0017	0.0010	3e-04	0.0009	0.0006	0.0006
13	0.3633	0.5833	0.6160	0.0002	0.0002	0.0001	0.0002	0.0002	0.0001	2e-04	0.0003	0.0001	0.0006
14	0.7429	0.9080	0.7700	0.0000	0.0001	0.0000	0.0000	0.0001	0.0000	0e+00	0.0000	0.0000	0.0001
15	0.6603	0.9051	0.9373	0.0001	0.0001	0.0000	0.0001	0.0001	0.0000	0e+00	0.0000	0.0000	0.0000
16	0.6691	0.8854	0.8053	0.0001	0.0001	0.0000	0.0001	0.0002	0.0000	0e+00	0.0001	0.0000	0.0001
17	0.8401	0.3451	0.1485	0.0000	0.0008	0.0008	0.0000	0.0008	0.0008	0e+00	0.0005	0.0005	0.0000
18	0.7289	0.9405	0.9600	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0e+00	0.0000	0.0000	0.0002
19	0.1360	0.0623	0.0680	0.0014	0.0034	0.0021	0.0014	0.0035	0.0021	6e-04	0.0014	0.0009	0.0022
20	0.5212	0.3237	0.1744	0.0001	0.0006	0.0005	0.0001	0.0006	0.0005	1e-04	0.0006	0.0005	0.0004

*Replication=100, α=0.01, ΔR^2 =0.02, Δβ=0.1

As seen in Table 11, there is no DIF on any items except the 5th item. The 5th item shows DIF according to the χ^2 values. It is known that χ^2 is affected by sample size (Tabachnick & Fidel, 2013; Hair et al., 2014). On the other hand, the R^2 and β values are very small and close to 0. This shows that there are no significant DIF on the items of the ACTS.

Discussion and Conclusions

In this study, it was thought that the answer-copying tendency would be one of the indicators of the answer-copying behaviors. It is highly possible that the students with higher levels of tendency will be answer-copying. It is obvious that the statistical detection techniques couldn't provide exact solutions. These known techniques need initial response patterns for defining the suspicious focal group of answer-copying.

Therefore, we need such indirect solutions at least in order to understand the nature of answer-copying.

The "Answer-Copying Tendency Scale in University Students (ACTS)" was developed within this context. Unlike similar examples, it was studied on large groups and used more comprehensive techniques to obtain psychometric evidence. Results show that the validity and reliability levels of the ACTS are very high. The ACTS can be used to understand the nature of answer copying. Also, and more importantly, it is thought that the ACTS can be used to define suspicious answer patterns for prior distributions.

References

Aaron, R.M. (1992). Student academic dishonesty: Are collegiate institutions addressing the issue? *NASPA Journal, 29(2)*, 107-113.

Aaron, R.M., & Georgia, R.T. (1994). Administrator perceptions of student academic dishonesty in collegiate institutions. *NASPA Journal, 31(2),* 83-91.

Angoff, W.H. (1974). The development of statistical indices for detecting cheaters. *Journal of the American Statistical Association, 69,* 44–49.

Bolin, A.U. (2004). Self-control, perceived opportunity, and attitudes as predictors of academic dishonesty. *The Journal of Psychology, 138(2),* 101-114.

Burke, J. L. (1997). *Faculty perceptions of and attitudes toward academic dishonesty at a two year college.* (Unpublished doctoral dissertation). University of Georgia.

Choi, S.W., Gibbons, L.E., & Crane, P.K. (2011). Lordif: An R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations. *Journal of Statistical Software, 39*(8), 1-30.

Crown, D.F., & Spiller, M.S. (1998). Learning from the literature on collegiate cheating: A review of empirical research. *Journal of Business Ethics, 17*, 683-700.

DeVellis, R.F. (2003). *Scale development, theory and applications (Second edition).* Sage Publication, Inc.

Dwyer, D.J., & Hecht, J.B. (1994). *Cheating detection: Statistical, legal and policy implications*. Illinois State University.

Eminoglu, E., & Nartgun, Z. (2009). Universite ogrencilerinin akademik sahtekarlık egilimlerinin olculmesine yonelik bir olcek gelistirme calısması [A scale development study to measure academic dishonesty tendency of university students]. *Uluslararası Insan Bilimleri Dergisi, 6(1),* 215-240.

Frary, R.B., Tideman, T.N., &Watts, T.M. (1977). Indices of cheating on multiple-choice tests. *Journal of Educational Statistics, 6,* 152–165.

Gardner, W.M. & Kenneth, B.M. (1988). A scale for measuring attitude toward cheating. *Bulletin of the Psychonomic Society, 26(5),* 429-432.

Genereux, R.L., & McLeod, B.A. (1995). Circumstances surrounding cheating: A questionnaire study of college students. *Research in Higher Education, 36(6),* 687-704.

Gerdeman, R.D. (2000). *Academic dishonesty and the community college.* ERIC Digest, ED447840.

Hair Jr., J.F., Black, W.C., Babin, B.J., & Anderson, R.E. (2014). *Multivariate data analysis (seventh edition)*. Edinburg Gate, England: Pearson Education Limited.

Holland, P.W. (1996). *Assessing unusual agreement between the incorrect answers of two examinees using the K-index: Statistical theory and empirical support (Research Report RR 96-7).* Princeton, NJ: Educational Testing Service.

Hughes, J.M.C., & McCabe, D.L. (2006). Understanding Academic Misconduct. *Canadian Journal of Higher Education, 36(1)*, 49 – 63.

Maeda, H., & Zhang, B. (2016). An algorithm to improve test answer copying detection using the omega statistic. *International Journal of Testing, 17(1)*, 55-73.

May, K.M., & Loyd, B.H. (1993). Academic dishonesty: The honor system and students' attitudes. *Journal of College Student Development, 34*, 125-129.

McCabe, D.L. (1993). Academic integrity: What the latest research shows. *Synthesis: Law and Policy in Higher Education, 5(1),* 340-343.

McCabe, D.L., & Trevino, L.K. (1996). What we know about cheating in college: Longitudinal trends and recent developments. *Change, 28(1),* 28-33.

McCabe, D.L., & Trevino, L.K. (1997). Individual and contextual influences on academic dishonesty: A multi-campus investigation. *Research in Higher Education, 38*(3), 379-396.

Mullens, A. (2000). Cheating to win. *University Affairs, 41*(10), 22-28.

Ozturk, B. & Yesilyaprak, B. (1997). Universite ogrencilerinin kopya cekme olayına iliskin tutumları [Attitudes of the university students toward copying]. *Egitim Yonetimi, 3*(2), 227-240.

Selcuk, Z. (1995). Bir egitim ve rehberlik sorunu: Okullarda kopya cekme [An education and consultansy problem: Copying in schools]. *Kuram ve Uygulamada Egitim Yonetimi, 1*(3), 397-418.

Tabachnick, B.G., & Fidell, L.S. (2013). *Using multivariate statistics (sixth edition)*. New Jersey, USA: Pearson Education.

van der Linden, W.J., & Sotaridona, L.S. (2006). Detecting answer copying when the regular response process follows a known response model. *Journal of Educational and Behavioral Statistics, 31(3)*, 283–304.

Wesolowsky, G. (2000). Detecting excessive similarity in answers on multiple choice exams. *Journal of Applied Statistics, 27*(7), 909-921.

Whitley Jr, B.E. (1998). Factors associated with cheating among college students: A review. *Research in Higher Education, 39*(3), 235-274.

Wollack, J.A. (1997). A nominal response model approach to detect answer copying. *Applied Psychological Measurement, 21,* 307–320.

Appendix A.

Kopya Çekme Eğilimleri Ölçeği (KÇE)

[Answer-Copying Tendency Scale in University Students (ACTS)]

Aşağıdaki ifadelere katılma düzeylerinize 0 (hiç katılmıyorum) ile 10 (tamamen katılıyorum) arasında puan veriniz. [Point your participation levels below between 0 (I do not agree) and 10 (I fully agree).]

		Katılma Düzeyiniz [Participation Level]
1*	Kopya çekmek, insanın kendisini kandırmasıdır. [Copying answers makes a fool of a person.]	
2	Fırsatını yakaladığım durumlarda kopya çekerim. [I'll copy answers when I get the chance.]	
3*	Kopya çekmeyi haksız bir kazanç olarak görüyorum. [I see copying answers as an unfair advantage.]	
4	Kopya çekerek yüksek puanlar aldığını bildiğim arkadaşların olması bende de kopya çekme hırsı uyandırıyor. [Having friends I know that get higher scores by copying answers, I also have the ambition to copy answers.]	
5*	Kopya çekilmesine kesinlikle karşıyım. [I'm absolutely against copying answers.]	
6	Yeterince hazırlanamadığım sınavlarda kopya çekmeyi düşünürüm. [I think about copying answers for the examinations that I have not prepared enough for.]	
7*	Kopya çekmek, öğretmenin emeğine saygısızlıktır. [Copying answers is disrespectful to the teacher's endeavors.]	
8	Dersi geçmeye yetecek kadar kopya çekilebilir. [Copying answers can be enough to pass the exam.]	
9	Bazı derslerde kendimi kopya çekmeye mecbur hissediyorum. [I feel compelled to copy answers myself in some lessons.]	

10*	Kopya çekmem benim başarımı değil hilekârlığımı gösterir. [Copying answers is not my achievement but my deceit.]	
11	Zaten unutacağım ve ileride kullanmayacağım konuları içeren derslerde kopya çekmek bana mantıklı geliyor. [It makes sense for me to copy answers in lessons that I will forget and not use in the future.]	
12	Çekmesem bile sınavda kopya çekebilecek olduğumu bilmek bana güven verir. [Even if I do not do it, it gives me confidence to know that I can copy answers at the exam.]	
13	Bilmediğim konulardan gelen sorularda kopya çekmek bana makul geliyor. [It makes sense for me to copy answers to the questions I do not know.]	
14	Herhangi bir ceza almayacağımı bilsem kopya çekerim. [I will copy answers if I know I will not be punished.]	
15*	Kopya çekmenin savunulması tamamen saçmalıktır. [Defending copying answers is completely nonsense.]	
16*	Kopya çekilmesini önlemeye yönelik gerçekten caydırıcı önlemler alınmasını istiyorum. [I would like to take real deterrent measures to prevent students from copying answers.]	
17	İşime yaramayacak onca bilgiyi ezberleyeceğime kopya çekmek daha mantıklı geliyor. [It makes more sense to copy answers when I memorize so much information that will not work in my own life.]	
18	Not kaygısı beni kopya çekmeye itiyor. [Anxiety about earning high grades pushes me to copy answers.]	
19*	Kopya çekmenin ahlak dışı bir davranış olduğuna inanıyorum. [I believe that the copying answers is immoral behavior.]	
20	Yakalanma korkum olmasa kopya çekerim. [If I am not afraid of getting caught, I'll copy answers.]	

*İkinci faktöre (ETIK) ait maddelerdir. Ölçek toplam puanı için ters kodlanması gerekir. [Belongs to the second factor (ETIK). Should be recode for total score.]

Potansiyel Bir Hata Kaynağı Olarak Üniversite Öğrencilerinde Kopya Çekme Eğilimlerinin Ölçülmesi: Bir Ölçek Geliştirme Çalışması

Atıf:

Demir, E. (2018). As a potential source of error, measuring the tendency of university students to copy the answers: A scale development study. *Eurasian Journal of Educational Research, 75 (2018),* 37-58, DOI: 10.14689/ejer.2018.75.3

Özet

Problem Durumu: Kopya çekmeyi de içeren akademik sahtekarlık okullarda giderek daha önemli bir sorun haline gelmektedir. Bu sorun, sınıf büyüklüklerinin artırılması ve öğretim kaynaklarının azaltılması ile artmaktadır. Kopya çekmenin tespit edilmesi

için bazı istatistiksel teknikler olmasına rağmen, bunlar sınırlıdır ve ancak bir hata, olası bir hata kaynağı veya bir olasılık gösterebilmektedir. Kopya çekmenin yapısını ve doğasını anlamak için bir başka yol, tutumlar ya da eğilimler gibi ilişkili diğer faktörleri dikkate almaktır. Çalışma ahlakı yüksek olan, sınav kaygısı düşük olan, öz-benlik saygısı yüksek olan vb. öğrencilerin daha az kopya çektiği gözlenmektedir. Diğer taraftan kopya çekmenin yaygınlaşması, normal bir davranış olarak görülmesi ve kabul edilmesi durumlarında, kopya çekme davranışları daha sık görülmektedir. Pek çok çalışmada öğrencilerin sahtekarlık ya da kopya çekmeye yönelik görüş ve önerileri, ancak anketler aracılığıyla gözlenmiştir. Ölçek düzeyinde olan yani toplam puan alınabilen çok az sayıda araç vardır. Bu araçların kullanımı ve erişimi oldukça sınırlıdır. Ayrıca bu araçların genellikle, küçük ve sınırlı gruplarda geliştirildiği görülmektedir. Kapsamları geniştir. Çoğunluğu, geçerlik ve güvenirlik düzeylerine yönelik az sayıda kanıt içermektedir. Çoğunlukla bu araçlar geliştireli çok zaman geçmiştir. Anlaşıldığı üzere, uzun süredir konu üzerinde çalışılıyor olmasına ragmen, geçerli ve güvenilir araçlara olan ihtiyaç devam etmektedir.

Araştırmanın Amacı: Bu çalışmanın amacı, üniversite öğrencilerinin kopya çekme eğilimlerini gözlemlemek için ölçek olarak geçerli ve güvenilir bir ölçme aracı geliştirmektir. Ayrıca, daha kapsamlı geçerlilik ve güvenilirlik çalışmaları ile kanıt sağlamayı amaçlamaktadır.

Araştırmanın Yöntemi: Bu bir ölçek geliştirme çalışmasıdır. "Klasik Test Teorisi" altında "Üniversite Öğrencilerinde Kopya Çekme Eğilimi Ölçeği (KÇE)" geliştirilmiştir. Ayrıca diğer teoriler, özellikle "Madde Tepki Kuramı" da dikkate alınmıştır. Ön çalışmalar ve madde yazıma çalışmalarında, 80 civarı öğrenciden kopya çekmeye yönelik görüşlerini kompozisyon biçiminde yazmaları istenmiş, bu dokumanlar analiz edilerek 123 maddeden oluşan bir taslak form oluşturulmuştur. Bu form eğitim bilimleri alanında çalışan 6 akademisyenin görüşlerine sunulmuştur. Görüşler doğrultusunda gerekli düzenlemeler yapılarak 93 maddelik deneme formu oluşturulmuştur. Bu çalışmalardan sonra 711 öğrenci ile deneme uygulaması ve 909 öğrenci ile ana uygulamalar yapılmıştır. Elde edilen veriler kullanılarak, yapı geçerliği, madde ve test betimsel istatistikleri, madde ayırıcılıkları, içtutarlılık ve test-tekrar test güvenirliği, sınıflandırma doğruluğu ve değişen madde fonksiyonu ile madde yanlılığı incelenmiştir.

Araştırmanın Bulguları: Deneme ve esas uygulama verileri üzerinde ayrı ayrı yürütülen "Temel Bileşenler Analizi (TBA)" sonuçları birbirini destekler niteliktedir. Bu sonuçlara göre KÇE, 2 faktör ve 20 maddeden oluşmaktadır. İlk faktör 12 maddeden oluşmaktadır ve "negatif sınav ve not algısı (NEGALGI)" olarak tanımlanmıştır. İkinci faktör 8 maddeden oluşmaktadır ve "etik değerler (ETIK)" olarak tanımlanmıştır. Açıklanan varyans yüzdeleri, birinci faktör için %48.42, ikinci faktör için %13.29 ve toplamda %61.71'dir. Bu faktörler arasında negatif yönlü ve orta düzey bir ilişki

bulunmaktadır (r=-0.49 ve p<0.01). TBA'nın yanı sıra hem destekleyici kanıt sağlamak hem görsel sunum sağlamak için ALSCAL tekniği ile Öklit uzaklıklarına dayalı çok boyutlu ölçekleme çalışması da yapılmıştır. Elde edilen saçılma grafikleri de yapının belirgin bir şekilde 2 boyuttan oluştuğunu göstermiştir. KÇE'de her bir madde 0 ile 10 arasında puanlanmaktadır. Toplam puanlar 0 ile 200 arasında değişmektedir. Toplam puanlar ve madde puanlarına yönelik dağılımlar, normaldir. Ayrıca madde-toplam puan korelasyonlarına yönelik olarak madde ayırıcılıkları, çok yüksek ve 0,40'ın üzerinde gözlenmiştir. Güvenirlik çalışmaları olarak iki ayrı yöntem kullanımıştır. Öncelikle hem deneme uygulaması hem esas uygulama verileri üzerinde, hem test geneli hem faktörler düzeyinde α içtutarlılık katsayıları hesaplanmıştır. Bu değerler; deneme uygulaması sonuçlarına göre birinci faktör için 0.950, ikinci faktör için 0.884 ve ölçek geneli için 0.942; esas uygulama sonuçlarına göre ise sırasıyla 0.955, 0.907 ve 0.955'tir. Bu değerler çok yüksek güvenirlik düzeylerine işaret etmektedir. Ayrıca esas uygulamaya katılan 95 öğrenciye tekrar test uygulaması yapılarak test-tekrar test güvenirliği hesaplanmıştır. Bu değer de 0.804'tür ve yüksek düzeyde güvenirliğe işaret etmektedir. Bir diğer inceleme olarak KÇE'nin öğrencileri kopya çekme ve kopya verme durumlarına göre ne düzeyde doğru sınıflandırabildiği "Diskriminant Fonksiyon Analizi" ile analiz edilmiştir. Bu kapsamda test edilen dört ayrı modelde de yüksek sınıflama doğrlukları elde edilmiştir. Madde yanlılığı çalışmaları her bir madde düzeyinde değişen madde fonksiyonlaşması incelenerek yapılmıştır. Bu amaçla, çok kategorili maddelere yönelik olarak MTK'ya dayalı bir şekilde geliştirilmiş "lordif" tekniği kullanılmıştır. Analiz sonuçlarına göre maddeler üzerinde manidar ve ciddi düzeyde bir DIF gözlenmemiştir.

Araştırmanın Sonuçları ve Önerileri: Bu çalışmada ana hipotez, kopya çekme eğiliminin kopya çekme davranışlarının bir göstergesi olacağı yönündeydi. Daha yüksek düzeyde bir eğilime sahip olan öğrencilerin, kopya çekme davranışı göstermeleri olasılığı yüksektir. İstatistiksel tespit teknikleri kesin çözümler sağlayamadığından, kopya çekmenin doğasını anlamak için en azından bu tür dolaylı çözümlere ihtiyacımız var. Bu kapsamda bu çalışmada "Üniversite Öğrencilerinde Kopya Çekme Eğilim Ölçeği (KÇE)" geliştirilmiştir. Benzer örneklerden farklı olarak, bu çalışmada büyük gruplar üzerinde çalışılmış ve kanıt elde etmek için daha kapsamlı teknikler kullanılmıştır. Sonuçlar, KÇE'nin geçerlilik ve güvenilirlik düzeylerinin çok yüksek olduğunu göstermektedir. KÇE, kopya çekmenin doğasını anlamak için kullanılabilir. Ayrıca ve daha önemlisi, KÇE'nin kopya belirleme tekniklerinin öncül dağılımları için şüpheli cevap örüntülerini tespit etmek amacıyla da kullanılabileceği düşünülmektedir.

Anahtar Sözcükler: kopya çekme eğilimi, sahtekarlık, ölçek geliştirme, kopya-belirleme

Eurasian Journal of Educational Research 75 (2018) 59-80

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

The Pedagogical Beliefs and Instructional Design Practices: Pre-Service IT Teachers' Case*

Emine SENDURUR [1]

ARTICLE INFO

Article History:
Received: 02 Dec. 2017
Received in revised form: 17 Mar. 2018
Accepted: 28 Apr. 2018
DOI: 10.14689/ejer.2018.75.4

Keywords
Novice teachers, instructional designer, teacher beliefs, teaching practice.

ABSTRACT

Purpose: Although studies have separately examined the instructional design processes of novice designers and the pedagogical beliefs of novice teachers, the manner through which they are linked is still unclear. In this study, the aim was to describe how pre-service teachers' pedagogical beliefs were embodied through instructional design practices within a real context. Therefore, having a closer look at the espoused and enacted beliefs in comparison to teaching performance from the instructional design perspective may contribute to facilitating and understanding novice teachers' transfer of knowledge to practice. **Research Methods**: This was an instrumental case study with 20 pre-service teachers. The data were collected through observation, portfolios, and interviews.
Findings: The results indicated that the participants tended towards blended beliefs, but they failed to transfer them into practice. Their teaching practices mostly included patterns of teacher-centered approach. Finally, the match between espoused and enacted beliefs was not very common among the observed participants. **Implications for Research and Practice**: Since pre-service teachers attend school practice during the last year of university, they may feel isolated and unprepared for the application of student-centered approaches. That may be the reason why they put too much emphasis on control of the classroom. If they had been practicing these approaches from the beginning of their university studies, they would have left such anxieties behind and be more capable of transferring their espoused beliefs into enacted ones.

*A preliminary version of this paper was presented at ECER 2014, Porto, 1-5 Sept. 2014

[1] Ondokuz Mayıs University, TURKEY, e-mail: eminesendurur@gmail.com, ORCID: https://orcid.org/0000-0002-0340-6378

Introduction

In teacher education practices, the transfer of pedagogical knowledge may not always be smooth. Since fields such as computer education are somewhat contingent upon tools, differences among school settings may hinder the transferring process. In addition to real settings, being a novice can lead to hesitations to utilize one's knowledge. Therefore, having a closer look at the espoused and enacted beliefs regarding teaching performance from an instructional design perspective may contribute to facilitating and understanding novice teachers' transfer of knowledge to practice.

Two main objectives construct almost every type of teacher education institution: providing students with a theoretical background and enabling students to transfer theoretical knowledge into practice. Unlike theoretical parts, it is not always easy to provide appropriate experiences that are meaningful for novice teachers. Project-based learning experiences can be considered valuable, due to their affordances with regards to transfer and practice of theoretical knowledge. In teacher education, there is a gradual transition from being students to being pre-service teachers; therefore, these individuals must be ready to teach in real settings. "School experience" and "teaching practice" are two final year courses. In the former, the candidate makes observations within real settings, whereas the latter requires real teaching. The imperfect settings of schools may contribute to the transfer of knowledge as well as challenges with implementation.

Similarly to other education programs, equipping Information Technology (IT) teachers with detailed pedagogical content is vital. Their primary tools are computers, which can serve as an enabler for the transfer of constructivist pedagogical beliefs (Becker, 2000; Tondeur, Braak, Ertmer, & Ottenbreit-Leftwich, 2017). Although the nature of IT teaching is quite appropriate for this line of beliefs, IT teaching can become more aim-oriented rather than tool-oriented within other pedagogical beliefs. The way teachers integrate or use the technology can reflect their beliefs. Teacher beliefs can be defined as "implicit assumptions about students, learning, classroom, and the subject matter to be taught" (Kagan, 1992a, pp. 65-66). Beliefs are different from knowledge and are embodied in different forms. For example, a teacher may not believe in the value of technology in the classroom, whereas another teacher may plan each lesson with the inclusion of a variety of technological tools. Although Kagan (1992a) named the term *teacher belief,* the term has been recently referred to as *pedagogical belief.* According to Pajares (1992), people "have beliefs about everything" (p. 315); although beliefs are connected within the system, there is a distinction between specific and general beliefs of teachers. In this study, based on Pajares' (1992) belief system, I used the term similarly to Ertmer (2005), in that learning and teaching specific beliefs of teachers are the focus.

Despite being very complex in nature, pedagogical beliefs can be roughly divided into three types: traditional, constructivist, and blended. Teachers with traditional beliefs value the final outcomes, rather than the process; thus, they value the retention of facts. The teacher controls the classroom and gives lectures (Ertmer,

Gopalakrishnan, & Ross, 2001). Teachers with constructivist pedagogical beliefs provide students with a wide range of learning experiences, resources, tools, and materials within authentic and rich learning environments through meaningful activities (Windschitl, 2002). Teachers give opportunities, collaborate with the students' knowledge construction processes, and both monitor and allow students to monitor their own learning via diverse assessment approaches (Ertmer et al., 2001). In short, unlike traditional beliefs, constructivist beliefs are quite flexible and open to the discovery of students' potentials. Learning outcomes can diversify, depending on the way the students learn and experience. Teachers with blended beliefs can show patterns of either constructivist or traditional beliefs. For example, a teacher can value being the only authority in the class, yet benefit from group work. If we consider the two aforementioned belief types as opposite ends, then blended beliefs fall into the middle.

Like general beliefs, professional beliefs can be affected by many factors. Similarly, a teacher's pedagogical belief can affect his/her own way of teaching and learning (Ertmer, Gopalakrishnan, & Ross, 2001; Tondeur et al., 2017), acting (Pajares, 1992), and using/integrating technology (Ertmer, Ottenbreit-Leftwich, Sadik, Sendurur, & Sendurur, 2012). Researchers generally study teacher beliefs within two dimensions: enacted (practiced) vs. espoused (expressed) beliefs. Although a match between both dimensions is expected, this does not always happen due to various barriers (Chen, 2008). Nevertheless, there are rare cases showing a match between both dimensions, regardless of the barriers (Ertmer et al., 2012). There might be variety of factors contributing to the evolution of teaching beliefs, but teachers themselves are at the top of these factors (Al-Amoush, Markic, Usak, Erdogan, & Eilks, 2014); therefore, teaching beliefs can change or develop over time. The way teachers are trained as well as their culture can also play an important role in shaping their beliefs (Tang, Lee, & Chun, 2012; Tekindal, Roehrig, Jakiel, Arrastia, Rawls, & Izci, 2017). Therefore, the interventions of belief systems may be sophisticated due to the complex connections among cognitive structures.

Anything in a teacher's life may affect his/her overall belief system. Therefore, the components included in teacher education programs can be an important issue. For example, IT teachers are offered an obligatory course named "instructional design" (ID), which is not included in other subject areas. Since preparing for teaching is a part of the ID process, the pre-service IT teacher can be considered as a double novice (either as a teacher or an instructional designer). ID can be defined as "a construct that refers to the principles and procedures by which instructional materials, lessons, and whole systems can be developed in a consistent and reliable fashion" (Molenda, Reigeluth, & Nelson, 2003, p. 574). Therefore, ID is a more comprehensive process than lesson planning, as the latter can be considered as one of the products of ID. Despite having different names and approaches (e.g., linear, iterative), there are numerous ID models with similar elements (Dick & Carey, 1996; Kemp, Morrison, & Ross, 1994; Seels & Glasgow, 1997; Willis & Wright, 2000). They all include the elements of the ADDIE framework—the abbreviation of Analyze, Design, Develop, Implement, and Evaluate. Nevertheless, expert instructional

designers (IDers) may not follow these models and prefer an eclectic approach instead. Novice IDers may have difficulties until they find their own style.

Being a novice in any field can be distinguished from being an expert, regardless of context. In the ID field of expertise, individuals' approaches to cases were observed as different (Hardre, Ge, & Thomas, 2005; Perez & Emery, 1995; Stefaniak, 2017) due to ill-structured nature of ID problems (Ertmer & Cennamo, 1995). Like in teaching practice, transferring ID models into practice may become distracting in real settings (Ertmer, York, & Gedik, 2009; Yanchar, South, Williams, Allen, & Wilson, 2010). Similarly to the instructional design field, school settings are considerably complex for novice teachers. In terms of classroom management skills, which are among the main concerns of novice teachers (Chesley & Jordan, 2012), the lack of expertise can bring about ineffective practices. A recent eye-tracking study clearly exemplifies this phenomenon. In that study, a group of pre-service teachers were observed as they taught, and the results showed the limited attention that was distributed among only a few students (Stürmer, Seidel, Müller, Haüsler, & Cortina, 2017). Before field experiences, pre-service teachers start constructing their own schema about the teaching context to assume unreal conditions with respect to students' attitudes, classroom facilities, etc. However, real situations might be very different than expected ones (Kagan, 1992b), which may even result in ineffective learning as novice teachers engage in real settings (Amador, 2016; Wolff, Bogert, Jarodzka, & Boshuizen, 2015). Therefore, observation within authentic practices that do not match with their espoused beliefs would be beneficial for pre-service teachers, as they would be aware of the potential of other pedagogical beliefs with regards to practice (Ertmer, 2005). As a result, the practitioners may become more eclectic.

In the literature, studies have shown how novice IDers design, as well as the pedagogical approaches of novice teachers. However, they are all separate studies. In the computer education and instructional technologies department, students are expected to gain expertise of both; therefore, it is important for pre-service IT teachers to practice both. This study aims to explore the transfer process of teaching practice with the integration of ID practice. Moreover, during this transfer, the shape and transformation of their pedagogical beliefs were the focus, as it is assumed that IT teachers should not have problems integrating technology into the classroom. However, there are limited studies investigating whether IT teachers can successfully integrate technology. Therefore, this study can provide insights into the integration process of pre-service IT teachers while relating to their pedagogical beliefs. In short, throughout the study, I tried to explore the following research questions:

1. How do pre-service IT teachers' enacted pedagogical beliefs match with espoused ones in terms of technology integration, classroom management, teaching methods, and evaluation?
2. How are pre-service IT teachers' instructional design practices transformed into real context in combination with pedagogical beliefs?

Method

Research Design

In this study, I used case study methodology. Since I am interested in the issue of pedagogical beliefs in instructional design practices of pre-service IT teachers, this can be considered as an instrumental case study. Researchers conducting instrumental case studies try to understand "something more than just a particular case" (Fraenkel & Wallen, 2006, p.429). The reported findings are all limited to the participants. Pre-service IT teachers, who registered in the same course and practiced in the same elementary school, constructed the case itself. In order to find the answers of research questions and insights of the participants, I used more than one data source.

Research Sample

Twenty pre-service IT teachers (7 men; 13 women) were assigned to practice in a public elementary school. The aim of the inclusion of these pre-service teachers was to observe their way of practicing as both teachers and IDers, which is not a regular case for other teacher education programs. They were all final year students who had completed the ID course in the 2nd semester and the "School Experience" course in the 7th semester. None of the participants have out-of-school experience as a teacher or an IDer. In addition to the pre-service teachers, there were two mentor teachers per 10 pre-service teachers, and one supervisor, who is the author of this paper.

Research Instruments and Procedures

In the beginning of the semester, the supervisor, mentor, and the administrator of the public school met all the pre-service teachers to inform them about expectations and procedures. They were assigned to attend the regular IT courses in that school in addition to providing assistance with hardware-related issues. The pre-service teachers both observed the mentor and contributed to the teaching. As they practiced, they were responsible for creating a portfolio including weekly reflections, lesson plans, activities, evaluations, articles, and any additional components related to their professional development. After each week's observation, they wrote reflections about the lesson, mentor, students, materials, etc. The schedule for each pre-service IT teacher's teaching practice was decided with the mentor. They prepared a 40-minute lesson from the current IT curriculum. The pre-service teachers were all responsible for the lesson plan, materials, and evaluation. They had the opportunity to practice or observe the entire stages of instructional design as well as the teaching itself. In addition, they were active during their entire practice time, i.e., they helped the IT teacher in variety of activities, including fixing an overhead projector, designing network cables, dealing with students' problems, etc.

Data Analysis and Trustworthiness

Data were collected through observations, field notes, portfolios, and unstructured interviews with students and the mentors. Portfolios of the pre-service teachers consisted of printed reflection papers and any resource used during the semester. Worksheets, curriculum documents, observation notes, and administrative documents are a few examples of other printed sources in the portfolios. During the teaching practice, both the mentor and the supervisor filled the observation sheet with the following:

1. Content Knowledge: The student-teacher knows basic concepts, links them appropriately, uses suitable visuals, and makes interdisciplinary connections.
2. Pedagogical Knowledge: The student-teacher knows special instructional methods, benefits from the literature, answers students' questions, and ensures a secure learning context.
3. Planning: The student-teacher prepares clear lesson plans, states objectives and goals, selects appropriate methods, tools, and materials, decides on appropriate evaluation, and relates the content to other fields.
4. Instructional Process: The student-teacher uses time, methods, and techniques efficiently, is sensitive to individual differences, promotes student participation, uses tools and materials, provides feedback, and evaluates.
5. Classroom Management: The student-teacher prepares an effective introduction, strikes attention, provides a democratic learning environment, motivates, summarizes whenever needed, gives appropriate assignments, and takes care of students.
6. Communication: The student-teacher communicates effectively, provides clear and simple instructions, asks meaningful questions, uses appropriate intonation, listens to the students carefully, and uses effective gestures.
7. Evaluation and Recording: The student-teacher prepares appropriate evaluation materials, provides suitable feedback, grades student products, and keeps records of students' grades.

A colleague and I investigated all qualitative data separately for topics of relevance. Unstructured interviews, reflection papers, and observation notes from both supervisors and mentors were used as the main sources of data. I was also the supervisor of the pre-service teachers and conducted unstructured interviews with each participant and mentor separately. The overall issues mentioned during the interviews consisted of classroom management, teaching methods and materials, the use of technology, and overall instructional practices. As a result, these issues guided the formation of main themes. Content analysis was utilized to analyze all data sources. In the light of the main themes, the reflection papers, observation papers, and transcribed interviews were coded, then grouped in either categories and

subcategories, and then the overall data were reexamined to ensure the agreement between researchers as well as the coherence of relations.

To ensure the trustworthiness and credibility of the study, data were triangulated using different forms of data. The open-ended questions in the unstructured interviews aimed to foster deeper insights of participants; this provided detailed data, which is crucial to ensuring credibility. Moreover, confirmability was increased with the inclusion of two researchers during content analysis.

Results

Espoused Pedagogical Beliefs

Reflection papers in students' portfolios and interviews with students clearly revealed espoused pedagogical beliefs. In this study, I classify pedagogical beliefs as student-centered, teacher-centered, and blended. The findings from the qualitative data suggest that none of the participants have student-centered pedagogical beliefs. The majority of them (*N*=18) can be denominated with blended pedagogical beliefs. Only two pre-service IT teachers expressed prominent patterns of teacher-centered pedagogical beliefs.

Students with *teacher-centered* beliefs specifically focused on classroom management issues. They mentioned how hard and important it was to maintain control of the class and gain attention. They specifically emphasized the assurance of silence in the class. The following excerpts clearly indicate these concerns:

> "Since teaching includes gaining students' attention, making them willing to learn, transferring one's own knowledge, and reinforcing everything, I do not think I was able to achieve it all. During the introduction part of the lesson, I had trouble gaining the students' attention." (TC-1)

> "I was confident with my topic, but I was worried about controlling the students. I decided to stand on a strategic place that can help me dominate the class. I tried to make eye contact with students to keep them silent. I got around the class and explain the topic. Whenever I detected a movement, I immediately went near that student. In this way, I kept control of the class." (TC-2)

From the excerpts, the value placed on the authority of the teacher is clear. They believe that effective lessons are in the hands of the teacher who strikes attention all the time. Moreover, the pre-service IT teachers seemed to consider the teacher as the only source of the content. They believed that their instruction would fail if students talk, move, or giggle.

The themes extracted from the reflection papers and interviews showed that the majority of the participants' espoused beliefs are *blended*. In other words, their beliefs are both teacher- and student-centered from a pedagogical perspective. The pre-service IT teachers emphasized the importance of communication, student

participation, higher-order thinking, practice, and the physical structure of the laboratory. The following excerpts are good examples of blended beliefs:

> "I was confident and prepared, which facilitated the way I communicated. The exact and simple answers that I gave students made them eager to listen. Maintaining eye contact, listening to them, and valuing their answers increased their level of attention." (B1)

> "Active participation of students is important while teaching. Once it is provided, students are not interested in irrelevant things and then teacher does not get disturbed." (B2)

> "It is always a big deal for me to gain attention. I tried to overcome this with popular topics. For example, I gave examples of Facebook and Twitter database examples while explaining database subjects. I realized that students paid more attention and learned better." (B4)

> "I observed that students are too active, talkative, and sarcastic. They hardly pay attention to the lesson and they usually prefer to play computer games. That is why, like X teacher, I tried to strike their attention first. I preferred starting lessons with visuals, stories, and other interesting activities." (B6)

> "Because of their age, they can easily get bored. That's why I tried to integrate educational games into the lesson, so that they can enjoy it." (B10)

As the excerpts suggest, the pre-service IT teachers searched for ways to integrate technology or interesting tools to gain attention. They seemed to be sensitive to the needs of students. In addition, they considered the active participation of the students. However, it can be inferred that the source of the content is still the teacher.

The overall espoused beliefs can be summarized under four themes: *classroom management skills, method, challenges,* and *technology integration*. First, in terms of classroom management skills, participants are not very different from each other. They believe in the importance of eye contact, gaining attention, intonation, knowing students' names, and giving rewards, immediate feedback, and reinforcement. Unlike participants with teacher-centered beliefs, the participants with blended beliefs do not prefer to apply the ignorance strategy. Moreover, they value students' needs and consider the importance of participation. Second, in terms of methodology, participants with different beliefs varied except for a few issues. For example, they both emphasized linking the content of the lesson to real life. In addition, they do not prefer to stick to only one method/strategy/technique during a lesson. Instead, they are flexible to shift across them.

Participants with blended beliefs focused on the importance of brainstorming, providing examples, activating higher-order thinking, and establishing interdisciplinary links, unlike the participants with teacher-centered beliefs. They also valued untraditional methods, such as games or authentic practices. Third, in terms of challenges, both groups emphasized the challenge of keeping control of the students as well as being prepared for the lesson. They also agreed regarding the

challenge of the physical structure of the lab. Participants with teacher-centered beliefs find earning students' respect challenging. While they associated the quality of the lesson with the fluent flow, others associated it with fun. Fourth, in terms of technology integration, both groups mentioned the importance of practice, motivation, and rewards. However, unlike others, participants with blended beliefs relate their concerns about technology integration with providing opportunities for authentic experiences, exploration, and supplementary materials. Table 1 summarizes the findings.

Table 1
Themes & Categories for Espoused Beliefs

Espoused Beliefs	$f_{Teacher\text{-}Centered\text{-}Beliefs}$	$f_{Blended\text{-}Beliefs}$
Classroom Management Skills		
Communication		
Eye contact	2	2
Ignorance	2	-
Gaining attention	2	3
Intonation	2	4
Value		
Students' needs	-	5
Student participation	-	4
Knowing students' names	1	1
Giving rewards/immediate feedback/reinforcemer	2	2
Method		
Cognitive concerns		
Linkage to real life	2	2
Brainstorming	-	3
Examples	-	4
Activating higher-order thinking	-	1
Flexibility		
Shift across methods/strategies/techniques	1	1
Interdisciplinary links	-	1
Untraditional methods		
Games	-	1
Providing authentic practices	-	1
Challenges		
Classroom Management		
Controlling students	2	2
Preparedness	1	1
Earning students' respect	2	-
Quality of lesson		
Fluent	2	-
Funny	-	1
Technical problems		
Physical structure of lab	2	2
Technology Integration		
Practice	2	4
Motivation	2	1
Exploration	-	5
Authentic experiences	-	1
Supplementary materials	-	1
Rewards	2	1

Enacted Pedagogical Beliefs

The data from both the observation notes during teaching practice and the interviews demonstrated the pre-service teachers' enacted beliefs. Like the espoused beliefs, the enacted beliefs of the participants did not include student-centered pedagogical beliefs. However, contrary to the espoused ones, the majority of the participants (*N*=17) expressed teacher-centered pedagogical beliefs. Only three pre-service IT teachers expressed prominent patterns of blended pedagogical beliefs.

Participants with *teacher-centered* espoused beliefs (*N*=2) valued the direct instruction method. They believed in the functionality of providing reinforcement in time. Since silence is a must during their instruction, they tried to eliminate noise whenever possible. Hence, their enacted and espoused beliefs matched well. The following statements show how they tried to maintain control during the lesson.

> "I tried more up-to-date and interesting ways for students to gain attention." (TC-1)

> "One of the most frequent errors was that I couldn't answer all the questions because different questions were asked by more than one student at the same time. I think this could have caused students to dislike the lesson." (TC-2)

There are inconsistencies across espoused and enacted beliefs for the majority of the participants. Although they expressed their pedagogical beliefs closer to the blended point, this was absent in performance, except for three pre-service IT teachers. They gave importance to discipline, including authority and control of the class, but they were also aware of the value of student-centered methods within a meaningful context. Since real settings are full of surprises, some novice teachers could not transfer their blended beliefs into reality. Only three students presented a successful practice of the blended approach, which was generally close to the constructivist end of the spectrum. The following are some examples of them:

> "…I think I taught interactively, consistently, and practically." (B7)

> "I realized that the students' thinking style is not abstract enough. After realizing this fact, I provided more appropriate ways of teaching." (B13)

Table 2 summarizes the observed themes. Pre-service IT teachers were observed spending too much time trying to gain attention due to the desire to maintain the sole control of the lesson. Although some of them value their students' needs, they spent a considerable amount of time keeping the silence. In their espoused beliefs, they explained the importance of untraditional methods, but most of them failed to turn this into practice. They mostly benefitted from brainstorming and advanced organizers as an introduction. Some of them were capable of shifting across methods/strategies/techniques. The participants' anxiety and concerns about gestures/intonation were clearly observed in their practice but were not mentioned in their espoused beliefs. The rest of the challenges were similar to their espoused beliefs.

Table 2

Frequency of Observed Themes during Teaching Practice

Enacted Beliefs	*f*
Classroom Management Skills	
Communication	
Eye contact	7
Ignorance	4
Gaining attention	10
Intonation	7
Value	
Silence of the class	9
Students' needs	7
Student participation	9
Democratic participation	1
Giving rewards/immediate feedback/reinforcement	8
Transfer of knowledge	1
Method	
Cognitive concerns	
Statement of objectives	1
Brainstorming	10
Examples/Analogies	7
Activating higher-order thinking	1
Use of advanced organizers	11
Flexibility	
Shift across methods/strategies/techniques	4
Use of multiple resources	11
Untraditional methods	
Games	5
Discussion	2
Problem solving	1
Cooperative learning	1
Drama	2
Direct instruction	5
Challenges	
Classroom Management	
Controlling students	4
Anxiety/fear	12
Gestures/intonation	5
Time management	1
Quality of lesson	
Fluent	5
Funny	4
Technical problems	
Physical structure of lab	3

Espoused vs. Enacted Beliefs

Comparing the espoused beliefs with the enacted ones, an interesting but explainable pattern showed up (Table 3). Participants with teacher-centered beliefs showed the same patterns during practice and, thus, were consistent. However, participants with blended beliefs mostly failed to transfer what they believed into action. Only three of them successfully practiced in a parallel way to their blended beliefs.

Table 3

Espoused vs. Enacted Beliefs

Espoused-Enacted match

	Teacher-centered *(N)*	Blended *(N)*	Student-centered *(N)*
Espoused	2	18	0
Enacted	17	3	0
Matches	2	3	0
Mismatches	0	15	0

Instructional Design Practices

In its generic form, ADDIE constitutes the main elements of any instructional design process. In its modified forms, the elements can be seen in different organizations. The flow of elements may differ with regards to IDers. The following starting and ending points were observed from the participants' explanations of their instructional design practices:

1. Starting point: analyzing students (*N*=7); planning lesson (*N*=5); preparing materials (*N*=3)
2. Ending point: evaluation of the materials (*N*=2); evaluation of the students (*N*=2); evaluation of the process (*N*=1)

Novice designers did not show common ID patterns, but it was observed that they had their own unique ID process. Their definitions of problems, focus of attention, main considerations, solution approaches, and linearity of process differed in practice. Pre-service IT teachers with more constructivist pedagogical beliefs and practices showed a common pattern of ID: the focus of the learners. They specifically paid attention to the needs and previous experiences of the target learners. In addition, they considered putting alternative strategies into practice. From a larger perspective, those participants' analysis and design elements were different from the rest. Those with more traditional pedagogical beliefs and practices were obsessed with environmental issues, such as seating plans. Therefore, they had difficulties in the analysis and implementation elements of the ID process. According to the

participants, they all followed the ADDIE model because of its simplicity, but it was not clearly demonstrated in their reflections or during teaching practices.

Discussion and Conclusion

From the beginning of their university lives, pre-service IT teachers attend many classes and sometimes have the opportunity to practice their teaching skills. This study took place during their practices in real schools. The results clearly indicated that the pedagogical beliefs of pre-service IT teachers have been shaped over time. Although experiences as a student can contribute to their beliefs, in this study, most pre-service teachers explained how their way of teaching was shaped with the help of mentor observations. It is known that mentors can lead to belief changes and improvements (Grudnoff, 2011; Johnson, 2006; Yuan & Lee, 2014). The influence of the mentor is crucial, but his/her inappropriate strategies can sometimes transform pre-service teachers' teaching styles and beliefs. In our case, the pre-service IT teachers were frequently obsessed with providing silence, as they believed this facilitates control of the class (Chen, 2008). This was what has been observed and learned, in addition to complications with being a novice teacher (Kagan, 1992b, Yilmaz, Sendurur, & Sendurur, 2016). In that sense, the selection of mentors is very important. Of course, it is not always possible to find less traditional teachers. Pre-service teachers should be warned and guided about inappropriate styles. In addition to mentors, the pre-service teachers' own experiences can be quite effective in constructing beliefs. Teachers themselves are one of the most crucial factors in shaping one's own belief systems (Al-Amoush et al., 2014; Oleson & Hora, 2014). In other words, the way pre-service teachers are educated can influence the way they teach.

The majority of the participants' espoused beliefs tended to be blended, which is a combination of both student-centered and teacher-centered approaches. However, a few students explained their beliefs in a more traditional way. In modern teacher education programs, the value of student-centered approaches is a clear priority, but these IT teachers have been educated with more traditional approaches starting from elementary school. That might be the reason why they had difficulties practicing student-centered approaches. As a result, mismatches emerged between espoused and enacted beliefs, which is in line with other similar studies (Chen, 2008; Kul & Celik, 2017). Moreover, the participants explained that their ID practices were very linear, but in practice, what they conducted was not a complete ID process. It was observed that they nearly ignored evaluation aspects and focused on the analysis and design components. Considering both ID processes and teaching periods, the overall mismatch can be attributed to their inexperienced nature. Therefore, they might need another mentor to specifically guide the ID process, as novices generally need cognitive apprenticeship to gain ID expertise (Ertmer & Cennamo, 1995; Stefaniak, 2017). However, they cannot be considered complete novices, because during their education, they had opportunities to practice in the field. The data was collected

during their last semester, which means they should be ready to teach after three more months. However, some external factors might have caused this mismatch (Chen, 2008; Wolff et al., 2015). For example, some students mentioned crowded classes with inappropriate seating plans. Such a situation can discourage novices from trying modern methods due to the anxiety of silence and control. In other words, the mismatch between espoused beliefs and enacted beliefs can be a result of the mismatch between expected settings and real classes (Kagan, 1992b). The pre-service IT teachers expressed disappointment about the physical conditions, which might hinder their willingness to integrate modern methods.

To sum up, teacher education programs in Turkey usually emphasize constructivist theories and practices, but as this study showed, they might not be practiced in real settings. The results of the current study cannot be generalized due to the limited number of participants and culture/country specific conditions. Nevertheless, current conditions in IT teacher education programs might not function as educators expected. Policy makers should consider revisions of teacher education programs with respect to the integration of modern approaches to current school conditions, because these programs are crucial to shaping belief systems of teachers (Markic & Eilks, 2013; Tang et al., 2012). Since pre-service teachers attend school practice during the final year of university, they may feel isolated and unprepared for the application of student-centered approaches. That may be the reason why they put too much emphasis on classroom control. If they had been practicing from the beginning of their university studies, they would have left such anxieties behind. A comparative study clearly indicated that pre-service teachers who practice as they enter the university express more modern beliefs (Al-Amoush et al., 2014). University is the place where their career foundations are shaped; thus, it can also be the right place to change and shape their beliefs, as they can be shifted and improved before it is too late (Yuan & Lee, 2014). Finally, since the belief system includes more than one dimension, further studies may shed light on other constructs within both ID and teaching practices.

Acknowledgement

I acknowledge that Polat Şendurur had contributed as a co-coder during data analysis part to increase trustworthiness.

References

Al-Amoush, S., Markic, S., Usak, M., Erdogan, M., & Eilks, I. (2014). Beliefs about chemistry teaching and learning-A comparison of teachers and student teachers'beliefs from Jordan, Turkey, and Germany. *International Journal of Science and Mathematics Education, 12*(4), 767-792.

Amador, J. (2016). Professional noticing practices of novice mathematics teacher educators. *International Journal of Science & Mathematics Education, 14*(1), 217-241.

Becker, H.J. (2000). Findings from the teaching, learning, and computing survey: is Larry cuban right? *Educational Policy analysis Archives,* 8(51), 1-35.

Chen, C.H. (2008). Why do teachers not practice what they believe regarding technology integration? *The Journal of Educational Research, 102*(1), 65-75.

Chesley, G.M., & Jordan, J. (2012). What's missing from teacher prep. *Educational Leadership, 69*(8), 41-45.

Dick, W., & Carey, L. (1996). *The systematic design of instruction*. 4th ed. New York, NY: Harper Collin.

Ertmer, P.A. (2005). Teacher pedagogical beliefs: the final frontier in our quest for technology integration? *ETR & D, 53*(4), 25-39.

Ertmer, P.A. & Cennamo, K. (1995). Teaching instructional design: An apprenticeship approach. *Performance Improvement Quarterly, 8*(4), 43-58.

Ertmer, P.A., Gopalakrishnan, S., & Ross, E. M. (2001). Technology-using teachers: comparing perceptions of exemplary technology use to best practice. *Journal of Research on Computing in Education,* 33(5), 1-26.

Ertmer, P.A., Ottenbreit-Leftwich, A.T., Sadik, O., Sendurur, E., & Sendurur, P. (2012) Teacher beliefs and technology integration practices: A critical relationship. *Computers and Education, 59(2),* 423-435.

Fraenkel, J.R., & Wallen, N.E. (2006). *How to design and evaluate research in education.* 6th ed. McGraw-Hill: New York:NY.

Grudnoff, L. (2011). Rethinking the practicum: limitations and possibilities. *Asia-Pasific Journal of Teacher Education, 39*(3), 223-234.

Hardre´, P. L., Ge, X., & Thomas, M. K. (2006). A qualitative study of the development of instructional design expertise. *Performance Improvement Quarterly*, 19(4), 63–90.

Johnson, K.E. (2006). The sociocultural turn and its challenges for second language teacher education. *TESOL Quarterly, 40*(1), 235-257.

Kagan, D. M. (1992a). Implication of research on teacher belief. *Educational Psychologist, 27*(1), 65-90.

Kagan, D. M. (1992b). Professional growth among preservice and beginning teachers. *Review of Educational Research, 62*(2), 129-169.

Kemp, J.E., Morrison, G.R., & Ross, S.M. (1994). *Designing effective instruction*. Merrill: New Jersey.

Kul, U., & Çelik, S. (2017). Exploration of pre-service teachers' pedagogical beliefs in relation to mathematics teaching activities in classroom-based settings. *International Journal of Research in Education & Science, 3*(1), 245-257.

Markic, S., & Eilks, I. (2013). Prospective chemistry teachers' beliefs about teaching and learning-A cross-level study. *International Journal of Science and Mathematics Education, 11*, 979-998.

Molenda, M., Reigeluth, C.M., & Nelson, L.M. (2003). Instructional design. In L. Nadel (Ed.), *Encyclopedia of cognitive science* (Vol 2, pp. 574-578). London: Nature Publishing Group.

Oleson, A., & Hora, M.T. (2014). Teaching the way they were taught? Revisiting the sources of teaching knowledge and the role of prior experience in shaping faculty teaching practices. *Higher Education, 68*, 29-45.

Pajares, M.F. (1992). Teachers' beliefs and environmental research: clearing up a messy construct. *Review of Educational Research, 62*(3), 307-332.

Perez, R. S., & Emery, C. D. (1995). Designer thinking: How novices and experts think about instructional design. *Performance Improvement Quarterly*, 8(3), 80–95.

Seels, B., & Glasgow, Z. (1997). *Making instructional design decisions.* Columbus, OH: Prentice Merrill.

Stefaniak, J.E. (2017). The role of coaching within the context of instructional design. *TechTrends, 61*(1), 26-31.

Stürmer, K., Seidel, T., Müller, K., Haüsler, J., & Cortina, K.S. (2017). What is in the eye of preservice teachers while instructing? An eye-tracking study about attention process in different teaching situations. *Zeitschrift für Erziehungswissenschaft, 20*(Suppl. 1), 75-92.

Tang, E., Lee, J., & Chun, C. (2012). Development of teaching beliefs and the focus of change in the process of pre-service ESL teacher education. *Australian Journal of Teacher Education, 37* (5), 90–107.

Tekindal, S., Roehrig, A.D., Jakiel, L.M., Arrastia, M.C., Rawls, E.S., & Izci, B. (2017). Differences in American and Turkish preservice teachers' beliefs about the effectiveness of classroom management practices. *International Journal of Pedagogies & Learning, 12*(2), 101-116.

Tondeur, J., van Braak, J., Ertmer, P.A., & Ottenbreit-Leftwich, A. (2017). Understanding the relationship between teachers' pedagogical beliefs and technology use in education: a systematic review of qualitative evidence. *ETR&D, 65*(3), 555-575.

Willis, J., & Wright, K. E. (2000). A general set of procedures for constructivist instructional design: the new R2D2 model. *Educational Technology,* 40(2), 5–20.

Wolff, C.E., van den Bogert, N., Jarodzka, H., & Boshuizen, H.P.A. (2015). Keeping an eye on learning differences between expert and novice teachers' representations of classroom management events. *Journal of Teacher Education, 66*(1), 68-85.

Yanchar, S. C., South, J. B., Williams, D. D., Allen, S., & Wilson, B. G. (2010). Struggling with theory? A Qualitative investigation of conceptual tool use in instructional design. *Educational Technology Research and Development*, 58(1), 39–60.

Yilmaz, R., Sendurur, E., & Sendurur, P. (2016). Transition from theory to practice: a comparison of math and computer education and instructional technology teachers during school experience course. *Kastamonu Egitim Dergisi, 24*(3), 1517-1532.

Yuan, R., & Lee, I. (2014). Pre-service teachers' changing beliefs in the teaching practicum: Three cases in an EFL context. *System, 44,* 1-12.

Pedagojik İnançlar ve Öğretim Tasarımı Uygulamaları: Bilişim Teknolojileri Öğretmen Adaylarının Durumu

Atıf:

Sendurur, E. (2018). The pedagogical beliefs and instructional design practices: Pre-service-IT teachers' case. *Eurasian Journal of Educational Research*, 75, 59-80, DOI: 10.14689/ejer.2018.75.4

Özet

Problem Durumu: Kuramsal bilginin uygulamaya dönüştürülmesi, öğretmen eğitiminin en zorlu aşamaları arasındadır. Problem-temelli öğretim, mikro öğretim, vb. deneyimler öğretmenlik uygulamasındaki deneyimler kadar gerçekçiliği sağlayamayabilir. Öğretmen adayları, 4 yıl boyunca edindikleri hem kuramsal hem de uygulamalı bilgileri gerçek okul ortamlarında, gerçek öğrencilere, gerçek dersler işleyerek sunma fırsatını 4. yılın son döneminde elde etmektedirler. Her öğretmende olduğu gibi öğretmen adaylarının da hem eğitimlerinden hem çevresel faktörlerden hem de kendilerinden kaynaklanan pedagojik inançları mevcuttur. Bu inançların şekillenmesinde ilk öğretmenlik deneyimlerinin de önemli bir yeri olduğu yadsınamaz. Öğrenci, öğrenme, sınıf, konu, altyapı, vb. yapılarla ilgili içsel varsayımların hepsi pedagojik inanç tanımlarında yer almaktadır. Bu varsayımların ön yargı oluşturması veya tam tersi esneklik oluşturması, acemi bir öğretmen adayında gözlemlenme ihtimali yüksek olan durumlardır. Bilişim Teknolojileri (BT) öğretmen adaylarında bu durum oldukça kritiktir, çünkü ön yargı oluşturabilecek inançlar nedeniyle bilgisayar gibi teknolojilerin öğretim sürecine entegrasyonunda sıkıntılar yaşanabilir. Diğer öğretmen adaylarından farklı olarak BT öğretmen adayları öğretim tasarımı eğitimi de almaktadır ve yine bunu uygulamaya dönüştürmek için "Öğretmenlik Uygulaması" dersinde fırsat elde etmektedirler, çünkü kısa süreliğine de olsa (2-4 hafta) süreci yönetme şansı mevcuttur. Bu süreçte, öğrencileri, ortamı, altyapıyı, resmi işleri, ders içeriğini, ölçme-değerlendirmeyi, mevcut materyalleri ve daha birçok faktörü analiz ederek dersini planlama, uygulama ve sonuçlandırma (değerlendirme) pratiği yapabilmektedirler. Fakat çoğu zaman öğretmenlikteki acemilik durumu öğretim tasarımcısı olarak da mevcuttur. Birçok etkene göre şekillenmekte olan pedagojik inançların öğretim tasarımı yaklaşımları üzerine de etkileri olduğu söylenebilmektedir. Öğretme ve öğretim tasarımı süreçleri oldukça karmaşıktır ve henüz acemi olan öğretmen adaylarının gerçek okul ortamında bu becerilerini sergilemesi zaman alabilir. Bu çalışmanın odak noktası da BT öğretmen adaylarının durumlarına göz atmaktır, böylece hem öğretme deneyimlerinde hem de öğretim tasarımı süreçlerinde izledikleri yaklaşımlarla sahip oldukları (espoused) pedagojik inançların şekil bulmasını gözlemlemek hedeflenmiştir.

Araştırmanın Amacı: Alanyazında acemi öğretim tasarımcılarını ve pedagojik inançlarını ayrı ayrı inceleyen çalışmalar mevcut olmakla birlikte her ikisini BT öğretmenleri üzerinde inceleyen çalışmaya rastlanmamaktır. 4 yıl boyunca şekillenmeye devam eden pedagojik inançların, henüz acemi olan BT öğretmen adaylarında nasıl ortaya çıktığını ve bunların öğretim tasarımı süreçleriyle bağlantısını incelemek bu çalışmanın temel amaçlarındandır. Bu bağlamda şu araştırma sorularına cevap bulunmaya çalışılmıştır: (i) BT öğretmen adaylarının pedagojik inançları, gerçek bağlamdaki öğretim tasarımı uygulamalarında nasıl somut hale bürünmektedir? (ii) BT öğretmen adaylarının ortaya koydukları pedagojik inançlarıyla içsel olarak benimsedikleri pedagojik inançlarının birbiriyle örtüşme durumu teknoloji entegrasyonu, sınıf yönetimi, öğretim yöntemleri ve değerlendirme açısından nasıldır?

Araştırmanın Yöntemi: Bu çalışmanın yöntemi araçsal (instrumental) durum çalışmasıdır. 20 gönüllü BT öğretmen adayının "Öğretmenlik Uygulaması" dersinde yaşadıkları durumlar üzerinden hareket edilmiş ve portfolyo, görüşme ve gözlemler yoluyla veriler toplanmıştır. Birden fazla veri kaynağının kullanılmasının nedeni veri kaybını önlemek ve mümkün olduğunca gerçek durumu yansıtmaktır. Katılımcıların hepsi hem pedagojik formasyon derslerini hem de öğretim tasarımı dersini almıştır. Bir önceki dönem tamamen gözlem yapan katılımcılar, bu dönem aktif ders anlatımı sürecindedirler. Bunun yanı sıra haftalık yansıtma raporlarını da danışman öğretim üyesine teslim etmektedirler. BT öğretmen adayları diğer öğretmen adaylarından farklı olarak, okulun teknolojik altyapısını sürdürülebilirliğiyle ilgili işlere de dahil olmaktadırlar. Bir dönem boyunca hem doküman olarak hem de gözlem ve görüşmeler yoluyla elde edilen veriler içerik analiziyle analiz edilerek kod ve temalar ortaya çıkarılmıştır. Birden fazla araştırmacı bu işlemleri yaptığı için kod ve temalardaki uyuşmazlıklar tespit edilerek üzerinde tekrar çalışılmıştır.

Araştırmanın Bulguları: Portfolyolar ve görüşmelerden elde edilen analizler sonucunda katılımcıların büyük bir çoğunluğunun karma (blended), yani hem öğretmen hem de öğrenci merkezli, inançları benimsedikleri (espoused) ortaya çıkmıştır. İlginç bir şekilde, hiçbir katılımcı öğrenci-merkezli pedagojik inancı benimsememiştir. Bunun yanı sıra sadece 2 katılımcının benimsedikleri inanç öğretmen-merkezlidir. Katılımcıların ortaya koyduğu (enacted) pedagojik inançlarda da öğrenci-merkezli yaklaşım yokken, diğer yaklaşımlardaki durum benimsenen inançların tersi yönünde çıkmıştır. Yani, çoğunluğun ortaya koyduğu inançlar öğretmen-merkezli olarak tespit edilmiş ve benimsedikleri karma yaklaşımı ortaya koyamadıkları gözlemlenmiştir. Sadece 3 öğretmen adayında karma inançlara uygun yaklaşımlar gözlemlenmiştir. Benimsediği inancı öğretmen-merkezli olan adayların ortaya koydukları performans tutarlılığını korumuştur. Öğretim tasarımı süreci odaklı bakıldığında ise herhangi bir ortak desene rastlanmamıştır. Her öğretmen adayının kendi stili ve kombinasyonu vardır. Fakat yapılandırmacı yaklaşıma daha yakın inançları ve performansları olanlar arasında ortak bir nokta gözlemlenmiştir: öğrencilerin ihtiyaçlarına ve önceki bilgilerine hassasiyet göstererek alternatif stratejiler üretmeye çalışmışlardır. Diğer yandan geleneksel inançlara sahip olanların özellikle öğrencilerin oturma düzeninden dolayı analiz ve uygulama aşamalarında zorlandıkları tespit edilmiştir. Bütün katılımcılar, ADDIE modelin aşamalarını uyguladıklarını söyleseler de yansıtma raporlarında veya gözlemlerde bu durum net olarak gözlemlenememiştir.

Araştırmanın Sonuçları ve Öneriler: Çalışmada gözlemlenen öğretmen adayları aslında karma inançlara sahip olmakla birlikte bunların hayata geçirilmesinde (hem öğretme hem de öğretim tasarımı boyutunda) aksaklıklar yaşamaktadırlar. Bunun sebeplerinden biri yeterli tecrübeye sahip olmamaları olabilir. Bu bağlamda, öğretmen adaylarının daha erken dönemde gerçek ortamda öğretmenlik uygulamalarına dahil olmaları sağlanabilir. Böylece sessizliği sağlamak, oturma düzenini kontrol etmek ve endişe/korku gibi duyguları yönetmek erken dönemde sağlanmış olur. İlerleyen dönemlerde ise öğretmen adayının inançları daha esnek biçimde şekillenerek hayata geçirilebilir. Öğretmen adaylarının her hafta eşlik ettikleri kılavuz (mentor) öğretmenlerin seçimi de oldukça önemlidir. Her ne kadar

pedagojik inançları tersine de olsa zaman zaman gözlemledikleri öğretmeni model alarak inançlarının tersine hareket edebildikleri gözlemlenmiştir. Benimsedikleri ve ortaya koydukları inançlar arasındaki uyuşmazlığın nedeni de yine bağlamdan ve tecrübesizlikten kaynaklı olabilir. Bu durumun öğretmenliğin ilk yıllarında da ortaya çıkabileceği riski düşünüldüğünde, erken dönemde yani eğitimlerinin ilk yıllarında öğretmen adaylarının gerçek ortamlarda kılavuz eşliğinde deneyim kazanması gerektiği aşikardır. Böylece önyargılarından uzaklaşabilir kendi pedagojik inançlarıyla tutarlı olmayı başarabilirler. Öğretim tasarımı açısından da çizgisel yaklaşım sergileyen bu adayların daha esnek uygulamalar yapabilmeleri için daha çok uzmanlaşma ihtiyaçları olduğu tespit edilmiştir. Bu bağlamda uzman öğretim tasarımcılarının kılavuzluğunda staj imkanı sağlanabilir. Yani diğer öğretmenliklerden farklılaşan bu durum göz önüne alınarak BÖTE bölümlerine ek uygulamalar getirilebilir.

Anahtar Sözcükler: deneyimsiz öğretmen, öğretim tasarımcısı, öğretmen inançları, öğretmenlik uygulaması.

Eurasian Journal of Educational Research 75 (2018) 81-98

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

Cooperative Learning and Learning Achievement in Social Science Subjects for Sociable Students

HERPRATIWI[1], DARSONO[2], SASMIATI[3], PUJIYATLI[4]

ARTICLE INFO

Article History:
Received: 15 Nov. 2017
Received in revised form: 4 Apr. 2018
Accepted: 28 Apr. 2018
DOI: 10.14689/ejer.2018.75.5

Keywords
sociable motivation, cooperative learning, elementary school

ABSTRACT

Purpose: The research objective was to compare students' learning achievement for sociable learning motivation students in social science (IPS) using cooperative learning. **Research Methods:** This research used a quasi-experimental method with a pre-test/post-test design involving 35 fifth-grade students. The learning process was conducted four times in one semester. The social science (IPS) learning outcome was measured using an essay test comprising eight items. The data concerning sociable learning motivation were obtained from a questionnaire comprising 29 items, with $\alpha = 0.956$. **Findings**: Using a paired-sample t-test, the analysis showed that there was a significant increase in students' motivation after implementing cooperative learning. The results also showed a positive correlation between students' curiosity and their perseverance in doing the task. **Implications for Research and Practice**: The results of this research confirm that cooperative learning can significantly increase students' motivation. Teachers should attempt to implement cooperative learning in their classes to ensure students' motivation to learn.

[1] Corresponding Author: Faculty of Teacher and Training Education, Universitas Lampung, Indonesia. Email: herpratiwi.1964@fkip.unila.ac.id. ORCID: https://orcid.org/0000-0003-4226-2757
[2] Faculty of Teacher and Training Education, Universitas Lampung, Indonesia. ORCID: https://orcid.org/0000-0001-8031-4533
[3] Faculty of Teacher and Training Education, Universitas Lampung, Indonesia. ORCID: https://orcid.org/0000-0003-1097-6704
[4] Faculty of Teacher and Training Education, Universitas Lampung, Indonesia. ORCID: https://orcid.org/0000-0002-9228-5191

Introduction

Learning outcomes and activities for social science (IPS) are low in elementary school (SD) class V (students approximately ten years old) in Bandar Lampung, Indonesia. Further, students lack the confidence to express their opinions and answer teachers' questions. Previous studies concerned with teaching and learning processes for social science in elementary school found similar results (Nurjanah, 1912; Yulyani, 2016; Hendarwati, 2016). Learning outcomes are influenced by three factors: (a) internal factors, including physical and psychological factors, which inherently exist within the individual who is in the process of learning, (b) external factors, which exist outside the individual, including family, school and community factors. One of the most important internal factors is motivation (Ulstad, 2016; Jo & Park, 2016).

Motivation is defined as an attempt to achieve a goal, or an ability to meet individual needs (Robbins, 1996). Motivation is a process consisting of three parts. Firstly, it concerns something that makes someone move (arise) or do something. Secondly, it concerns the process of motivation as a direction to meet a need. Thirdly, it deals with something that maintains (maintenance) the chosen path so the need is met (Barron and Greenberg, 1998). Motivation is a factor that makes someone do something; it activates and energizes (Ulstad, 2016; Atak, 2016; Ball, 2016; Gabrielle, 2016; Nashar, 2004).

Learning motivation is the effort students make to reach the target to find meaningful learning activities that are valuable and educative (Atak, 2016). According to Pangesti (2014), there are four styles of student learning motivation: achiever, sociable, conscientious and curious. Achiever students tend to excel in competition; they are competitive and influenced by friends and family factors. Sociable students have a spirit of togetherness, non-competitiveness and are cooperative by nature. Students with this motivation enjoy mutual success to achieve learning outcomes and high productivity (Gillies, 2016; Johnson, 2002; Zhang, 2015; Reigeluth, 2016).

Cooperative learning focuses on the interaction among students and their cooperation to achieve mutual benefits (Zhang, 2017) and is highly systematic (Zhang. 2015). Cooperative learning fosters positive interdependence, individual and group accountability, and interpersonal skills to improve team effectiveness (Cheruvelil, 2014). Cooperative learning can improve students' achievements (Leasa, 2017; Casey, 2015), knowledge and skills, learning motivation, self-esteem Further, it can reduce anxiety and create a harmonious environment (Xue, 2018; Fernandez, 2017). In cooperative learning, knowledge is built through social interaction, (Jarvela, 2015, Huang, 2014). Students with a conscientious motivation style better perform activities if they have received clear guidance regarding the rules. Students who are motivated by curiosity are always inquisitive. They do not like the establishment, and they like scientific developments.

According to Goleman (2001), motivation in learning has six aspects: (1) pleasure, learning pleasure, attention and interest in learning, (2) orientation to the mastery of material, (3) curiosity, (4) tenacity in doing the tasks, (5) high involvement in tasks, and (6) orientation to challenging, difficult and new tasks. Students with different

levels of motivation require different learning approaches. It follows that achiever students might need different learning approaches than students with sociable, conscientious or curious motivations.

Considering these previous studies, we decided to examine whether cooperative learning had any influence on IPS achievement for elementary school students in Bandar Lampung City, Indonesia. This study focused on students who have a sociable learning motivation for achievement in IPS.

Method

Research Design

This study used quasi-experimental methods aimed at measuring the impacts and deducing the changes induced by treatment. It also aimed to discover any cause-and-effect relationships in non-deterministic ways. Rather, it is merely probabilities or increasing probabilities of occurrence (Cook & Campbell, 1979; Shadish, 1995; Shadis et al., 2002). The study sample was determined by a random sampling technique (Roscoe, 1975). The number of samples was determined using Isaac and Michael's table (Isaac, 1981) with a five percent error rate, resulting in a sample comprised of 35 students (20 female and 15 male) of grade V primary school with an average age of 10 years. The study was conducted four times in one semester.

Research Instruments

The data on student motivation were collected using a questionnaire consisting of six aspects (Goleman, 2001) namely, (1) pleasure, enjoyment in learning, indicated by paying attention to study, having an interest in learning, being happy to do the task (rated by 6 statements, item numbers 1–6), (2) orientation to mastery of the material, indicated by being capable of presenting the material, (rated by 4 statements, item numbers 7– 10), (3) curiosity, indicated by motivation to find out new things (rated by 6 statements, item numbers 11– 14), (4) tenacity in performing tasks, indicated by being fully focused on accomplishing the tasks, being tough (rated by 6 statements, item numbers 15–20), (5) high involvement in tasks, indicated by being diligent in completing the task, concentrating on tasks and taking time to learn (rated by 6 statements, item numbers 21– 26), and (6) orientation towards new and challenging tasks, indicated by being motivated to do the tasks (rated by 3 statements, item numbers 27– 29), see Table 1.

Table 1

Questionnaire on Learning Motivation and Number of Statements

Aspects Measured	Number of Items
Pleasure, the enjoyment of learning	6
Orientation to the mastery of the material	4
Curiosity	6
Tenacity in doing the task	6
High engagement on task	6
Orientation to new and challenging tasks	3
Total	29

Table 2 shows the reliability values for the questionnaire, where the following Cronbach's Alpha values were found: the pleasure indicator, the enjoyment of learning (0.89), the orientation to the mastery of the material (0.85), curiosity (0.81), tenacity in doing the task (0.86), high involvement in assignments (0.91) and orientation to new and challenging tasks (0.93). Based on this reliability test, it can be seen that all aspects have a Cronbach's Alpha value of 0.956. It was inferred that the items in the questionnaire were reliable and all tests were internally consistent because they had strong reliability (Maier, Wolf & Randler, 2016; Bonett & Wright, 2015; Rainsch, 2004).

Table 2

The Reliability Aspects of Learning Motivation

Aspects Measured	Cronbach's Alpha Value
Pleasure, the enjoyment of learning	0.89
Orientation to the mastery of the material	0.85
Curiosity	0.81
Tenacity in doing the task	0.86
High engagement on task	0.91
Orientation to new and challenging tasks	0.93
Total	0.956

The data on learning outcomes were collected using a self-explanatory test (essay) consisting of eight items developed by the authors and taken from the standard competence "The role of Indonesia in Southeast Asian countries" and the basic competencies of describing the background of the formation of Southeast Asian countries. The eight items covered remembering (numbers 1, 2, 5 and 6) and understanding (numbers 3, 4, 7 and 8) and were scored according to reliability, validity and the level of difficulty, scored as low, medium, and difficult. The results are shown in Table 3.

Table 3

Instrument Aspect of Learning Result Test of IPS

No	Target Indicator	Thinking Step	Question Number	Validity	Reliability	Different Score	Difficulty Level
1.	Describing the national historical artifacts of Hinduism, Buddhism, and Islam in Indonesia	C1	1, 2	0.82	0.87	high	easy
2.	Giving an example of the national historical artifacts of Hinduism, Buddhism, and Islam in Indonesia	C2	3, 4	0.80	0.92	high	medium
3.	Recounting the historical characters of Hinduism, Buddhism, and Islam in Indonesia	C1	5, 6	0.80	0.91	high	medium
4.	Giving an example of the historical characters of Hinduism, Buddhism, and Islam in Indonesia	C2	7, 8	0.86	0.94	high	difficult

Data Analysis

Table 4 shows that a one-sample Kolmogorov-Smirnov test confirmed the data for learning motivation were normally distributed (Yu Zheng, 2008).

Table 4

Test Results of Normality Data

Measurement Aspect	Kolmogorov-Smirnov	Significance Score
Pleasure, the enjoyment of learning	0.221	0.200
Orientation to the mastery of the material	0.248	0.200
Curiosity	0.318	0.075
Tenacity in doing the task	0.302	0.073
High engagement on task	0.267	0.200
Orientation to new and challenging tasks	0.257	0.200

A homogeneity test using one-way ANOVA (Donald, 2010) found a significance level of 0.100> 0.05, indicating that the sample was homogeneous.

The data were analyzed by a paired-samples t-test because it used a one-sample t-test design (Donald, 2010). The steps in the data analysis are shown in Table 5.

Table 5

Data Analysis Steps

Step	Purpose	Analysis
1	Reliability	Cronbach's Alpha test
2	Correlation between variables	Correlation analysis
3	Different test before and after acknowledgment	Paired-samples t-test

Results

The mean and standard deviation for each of the six aspects of learning motivation were compared. Table 6 shows that the highest average was for high-engagement on the task (26.49 ± 3.38), followed by pleasure and enjoyment of learning (26.06 ± 2.71), tenacity in performing tasks (26.00 ± 3.92), curiosity (25.89 ± 3.37), orientation to the mastery of the material (15.37 ± 3.51) and orientation to new and challenging tasks (12,17 ± 2,63). Of the six aspects, orientation towards new and challenging tasks was the lowest, and engagement on the task enjoyed the highest position.

Table 6

Mean and Deviation Standard of Learning Motivation

Pleasure, the pleasure to learn	Mean	Std. Deviation
Orientation to the mastery of the material	26.06	2.71
Curiosity	15.37	3.51
Tenacity in doing the task	25.89	3.37
High engagement on task	26.00	3.92
Orientation to new and challenging tasks	26.49	3.38
Pleasure, the enjoyment of learning	12.17	2.63

The mean and standard deviation for each of the four indicators of IPS learning outcomes were compared. Table 7 shows that the highest average score was for explaining the meaning of the formation of Southeast Asian countries (2.83 ± 0.00), followed by explaining Indonesian foreign policy (2.77 ± 0.00), giving examples of Indonesia's role in Southeast Asian countries (2.74 ± 0.71) and giving examples of Indonesia's foreign policy role in international regulations (2.71 ± 0.00).

Table 7

Mean and Standard Deviation of Learning Results

Learning Result	Mean	Std. Deviation
Describing national historical relics of Hinduism, Buddhism, and Islam in Indonesia	2.83	0.00
Giving an example of national historical relics of Hinduism, Buddhism, and Islam in Indonesia	2.74	0.71
Recounting the historical characters of Hinduism, Buddhism, and Islam in Indonesia	2.77	0.00
Giving an example of the historical characters of Hinduism, Buddhism, and Islam in Indonesia	2.71	0.00

The six aspects of student learning motivation were analyzed by correlation analysis. The results presented in Table 8 show that curiosity had a significant relationship with tenacity in doing the task ($r = 0.852$ and $p < 0.005$) The correlation analysis between these two aspects of learning motivation did not show a closer relationship among the other four aspects. Therefore, further analysis was needed regarding the correlation.

Table 8

Intercorrelation of Student Learning Motivation

	Pleasure in Learning		Orientation to mastery of the material		Curiosity		Tenacity in doing the task		High engagement on task		Orientation to new and challenging tasks	
	R	Sig	R	Sig	R	Sig	R	Sig	R	Sig	R	Sig
Pleasure, the desire to learn	-	-	.809	.000	.659	.000	.680	.000	.533	.001	.399	.018
Orientation to mastery of the material	.809	.000	-	-	.524	.001	.567	.000	.586	.000	.521	.001
Curiosity	.659	.000	.524	.001	-	-	.852	.000	.791	.000	.513	.002
Tenacity in doing task	.680	.000	.567	.000	.852	.000	-	-	.831	.000	.607	.000
High engagement on task	.533	.001	.586	.000	.791	.000	.831	.000	-	-	.540	.001
Orientation to new and challenging tasks	.399	.018	.521	.001	.513	.002	.607	.000	.607	.000	-	-

Regression analysis was carried out to ascertain whether the learning outcome for IPS students who have sociable learning motivation was predicted by the aspects of curiosity and tenacity in doing the task and to identify which of these two aspects was the stronger predictor of learning outcome. Regression analysis was done to determine the extent to which the aspect of desire wants to know and tenacity in doing the task.

Regression analyses were performed with either curiosity or tenacity in doing the task as the independent variable and learning outcomes as the dependent variable. Tables 9 and 10 show that curiosity explained 37.6 percent (significance 0.000 < 0.05), and tenacity explained 30.2 percent (significance 0.001< 0.05) of the learning outcome.

The regression equation with curiosity as the independent variable was determined as Y = 58.060 + 1.317X. With tenacity in doing the task as the independent variable, it was determined as Y = 65.679 + 1.016X; where Y is the learning outcome and X is the independent variable

Table 9

Constant Value of to Know Aspect and Tenacity in Doing the Task

Model	Unstandardized Coefficient		Standardized Coefficient	T	Sig
	B	Std. Error	Beta		
(Constant)	58.060	7.744		7.497	0.000
Curiosity	1.317	0.297	0.611	4.437	0.000
(Constant)	65.679	7.082		9.275	0.000
Tenacity in Doing the Task	1.016	0.269	0.550	3.778	0.0001

Table 10

The Amount of Desire to Know Aspects and Tenacity in Doing the Task

Model1	R	R Square	Adjusted R Square	Std. Error of the Estimate
Curiosity	0.611	0.374	0.355	5.841
Tenacity in Doing the Task	0.550	0.302	0.281	6.167

The data in Table 10 show how curiosity and tenacity in doing the task contributed to learning outcomes. The results of the analysis showed that these two aspects were predicted to be significant in determining the learning outcome. The most obvious contribution to the variance was curiosity (β = 611), followed by tenacity in doing the task (β = 551). Curiosity and tenacity in doing the task both had the potential to be a variable to improve learning outcomes.

To find out which of these two independent variables acted as a predictor of learning outcomes, multiple regression analysis with a stepwise approach was performed. As shown in Table 11, when the two independent variables were included in the equation, only the curiosity variable was statistically significant in predicting learning outcome. The analysis showed that curiosity was predicted to contribute to learning outcomes (β = 0.611, p < 0.01). It was found that curiosity contributed 37.4

percent and tenacity to do the task contributed 30.2 percent to learning outcomes. Specifically, curiosity and tenacity played an important role in predicting learning outcomes.

Table 11

Regression Analysis Aspects of Curiosity and Tenacity

Variable	B	SE B	Beta (β)	t	Sig
Curiosity	1.317	0.297	0.611	4.437	0.000
Tenacity in Doing the Task	1.018	0.269	0.550	3.778	0.001

Table 12 shows the results of the descriptive statistical analysis. The average pre-test score was 53.333 ± 11.610 with a standard error of 1.962, and the average post-test score was 76.191 ± 12.806 with a standard error of 2.164.

Table 12

Statistical Results of Paired Sample

	Mean	N	Std Deviation	Std Error Mean
Pair 1 Pre-test	53.333	35	11.610	1.962
Post-test	76.191	35	12.806	2.164

Table 13 shows the results of the correlation analysis between the two pairs of data (pre-test and post-test scores). The correlation coefficient of 0.811 was significant (0.000 < 0.05), indicating indicated that both pairs of data were correlated.

Table 13

Correlation Test Result of Paired Sample

	N	Correlation	Sig
Pair 1 Pre-test & Post-test	35	0.811	0.000

Table 14 shows the average difference between the pre-test and post-test scores. The calculated value of t was less than the t-table value (significance 0.000< 0.05); consequently, there was a significant difference between the pre-test and post-test scores. Thus, it was concluded that learning outcomes improved for grade V elementary school students, with higher scores obtained after IPS was taught with cooperative learning.

Table 14

Test of Paired Sample

	Paired Differences					T	df	Sig (2-tailed)
	Mean	Std Deviation	Std Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1 Pre-test-Post-test	-1.785	7.600	1.2846	-20.469	-15.247	-13.901	34	0.000

Based on the analysis, it appears that the learning outcome in social studies subjects for students who have sociable learning motivation was higher after being instructed using cooperative learning. However, the subject matter of social studies was tested only in the dimensions of the cognitive process of remembering and understanding and in the dimension of factual and conceptual knowledge (Anderson, 2001).

Group learning can improve learning outcomes for students with sociable learning motivation. With cooperative learning, students experience the process of diffusion and socialization and have an unlimited view of science. Students with sociable learning motivation have a spirit of togetherness and non-competitive cooperation (Pangesti, 2014). These students need a learning atmosphere that demands cooperation, not competition, and learning should enable students to socially interact (Costa, 2014). Cooperative learning is imbued with constructivist theory, where learning involves students building personal and social knowledge. This is in accordance with Vygotsky's social reconstruction theory, which places students in the closest zone of child development or Zone of Proximal Development (ZPD) (Salomon, 1989; Clapper, 2015; Lantolf, 2015; Gommans, 2015). This theory attempts to persuade students to learn in their proper position according to the level of child development and to guide learners at the beginning of the learning phase, and then reduces their guidance when learners have started to take responsibility for learning. This stage is often called scaffolding (Rojas-Drummond, 2015; Gibbons, 2002; Smit, 2013; Wilson, 2014). Based on the theory of social reconstruction, learners are instructed with an applicable situation in their daily lives where values, knowledge, and skills in social life are central to education (Taghibaygi, 2015). In this process, the students are in their respective development zones and are guided at the beginning of the lesson. They are gradually given responsibility for completing tasks themselves in their study groups. This allows students to construct and reconstruct their understanding of the material (Maulidi, 2016).

Teachers can use cooperative learning daily to support students' learning in every subject, from basic skills to complex problem-solving. Characteristics of cooperative learning are positive interdependence among students, face-to-face interaction (educational interaction), personal responsibility to groups and skills in communicating in groups. Cooperative learning emphasizes group achievement. The

purpose of cooperative learning is to give students the knowledge, concepts, abilities, and understanding they need. In cooperative learning, students are placed in small heterogeneous groups (4 to 5 students) to complete group tasks prepared by teachers and followed with individual assistance for students who need it.

Group heterogeneity includes gender, race, religion (if possible), skill levels (high, medium, low), and so on (Slavin, 2014). Cooperative learning enables students to work in teams and assume responsibility for managing, checking, helping each other in the face of problems, encouraging each other forward and getting scaffolding from teachers and friends (Sumaryati, 2013). The results prove that cooperative learning effectively increases students' motivation and performance, as stated by researchers such as Slavin (2014), Slavin (2015), and Hertz-Lazarowitz (2013).

According to Albers (2008), when students have an opportunity to interact with others, they succeed in interpreting solutions in learning. Experience in communicating can provide a potential source of knowledge about learning. Constructive interactions that include knowledge of the purpose and implementation guidance can build up an individual's thinking and generate new knowledge. New knowledge will be stored in long-term memory if the students are directly involved in the process of understanding and constructing their own concepts or knowledge. Students with sociable motivation will be able to apply the knowledge that has been obtained in new situations (Carin, 1993).

There is a need for a learning strategy that aims to assist students in linking theory to its implementation in everyday life so they have a mindset for understanding logic. Students should not just spend their time listening and completing tedious exercises. Exams should not only test understanding and measure students' ability to memorize facts without them knowing what they are being asked. Discussing, finding out, thinking critically, engaging in real work projects and problem-solving are important for the learning process (Johnson, 2002).

Six aspects shape student learning outcomes for sociable motivation: (1) pleasure and enjoyment of learning, (2) orientation to mastery the of material, (3) curiosity, (4) tenacity in doing the task, (5) high involvement in the task, and (6) orientation towards challenging, difficult and new tasks (Goleman, 2001). Table 8 shows that curiosity and tenacity in doing tasks are more dominant in influencing learning outcomes than the other four aspects.

The data also showed that the relationship between curiosity and tenacity in doing the task was higher than the relationships with or between the other aspects. This relationship between curiosity and tenacity to do the task affects students' learning outcomes in social studies. The learning outcomes of students in social studies who have sociable motivation are influenced by curiosity and tenacity to do the task and can be conditioned by teachers through cooperative learning (Muldayanti, 2013; Dadds, 2002; McKeachie, 1990; Ginsberg, 2012; de Oliveira, 2016). Curiosity is an attribute one can develop, in this case with cooperative learning. By learning with friends, one can increase one's curiosity and make it a daily habit to become more

intelligent and knowledgeable. A sense of curiosity can be used to find solutions to difficult tasks or situations.

The findings of this study support the results of Gillies (2004) and Gillies (2016), that cooperative learning will accelerate learning outcomes. Students who participate actively in class and learn with friends will more quickly understand what is learned. Students dialogue with each other and take a role-play it because learning is not individual (Chen, 2013). Students with sociable learning motivation require efficient learning, and cooperative learning can help students solve problems and examine study themes. Sociable children cannot compete with other children, because they need scaffolding from their peers. Learning assistance from peers can eliminate awkwardness; peer language is more easily understood. With peers, there is less reluctance and embarrassment in learning to obtain better results and ask for help. Interactions in peer groups do not exclude the possibility of students helping each other. Peers provide cognitive, affective and psychomotor thinking solutions in an atmosphere of cohesive learning activities, which result in innovative and productive learning changes in the form of improved problem-solving skills and learning achievement (Purnomo, 2013; Fauzi, 2013; Rahmawati, 2016).

Teachers' commitment to using cooperative learning contributes greatly to the achievement of student learning outcomes. Cooperative learning will encourage sociable students to work together to acquire ideas in solving problems or themes and collectively conveying solutions obtained. The learning outcomes for sociable students in non-cooperative learning environments (competitive learning) are less likely to be successful because the learning is on an individual basis (Johnson, 2002; Uhamista, 2016; Soebiyanto, 2016; Dudija, 2016; Pratiwi, 2015; Huda, 2016). A class with heterogeneous students must be taught by cooperative learning (Cohen, 2015).

Discussion and Conclusion

Cooperative learning is an intervention for improving learning outcomes in the field of social studies for primary school students who have sociable learning motivation. Students who have sociable motivation are more suited to learning that prioritizes cooperation instead of competition.

Student learning outcomes are built on curiosity and tenacity in doing the task. Therefore, it is necessary to design and re-formulate the syntax of cooperative learning to recognize it as an important variable in improving learning outcomes. It is also necessary to develop further research, especially questionnaires used to measure students' standardized motivation, which contains a more comprehensive aspect of learning motivation.

This analysis was not able to determine whether student learning outcomes from cooperative learning change over time for students with sociable learning motivation. Also, this study was not able to determine how the six aspects of motivation are related to learning outcomes. The relationship between the six aspects described in this study

is identified to recognize how aspects of learning motivation can produce significant learning outcomes.

The research results will provide a meaningful contribution for elementary school teachers providing social science learning to students who have sociable motivation. Until now, teachers have given equal treatment to all students regardless of the student's type of motivation. Thus, teachers should identify each student's type of motivation before engaging in the learning process. Doing so will help ensure effective learning because it is suited to the students' needs. It would be better if the school cooperates with certain parties, such as education researchers or government bodies who have instruments to measure students' motivation so teachers will have easy access when they require data about their students' motivation.

Acknowledgment

Thank you to the Department of Education, the principals, the teachers and primary school students in Bandar Lampung City for their participation in this research. We would also like to thank the Faculty of Teacher Training and Education, Universitas Lampung, for funding to conduct this research.

References

Albers, C. (2008). Improving pedagogy through action learning and scholarship of teaching and learning. *Journal of International Teaching Sociology*, 79-86.

Anderson, L.W., & Krathwohl, D.R.(2001). *Taxonomy learning, teaching, and assessing*. NY: Longman.

Atak, M., Atik, I., & Ceylan, İ. (2016). A research on the effect of career and job getting perception on learning motivation and career development efforts. *International Journal of Educational Researchers*, *7*(1), 1-16.

Ary, D.,. Jacobs, L.C., & Sorensen, C. (2010). *Introduction to research in education: Eight edition*. Canada: Nelson Education ltd.

Baron, R.A., & Greenberg, J. (1998). *Behavior in organization: Understanding and managing the human side of work*. Allyn and Bacon, New York.

Ball, A., Bowling, A., & Bird, W. (2016). A case study of learning, motivation, and performance strategies for teaching and coaching CDE teams. *Journal of Agricultural Education*, *57*(3), 115-128.

Bonett, D.G., & Wright, T.A. (2015). Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning. *Journal of Organizational Behavior*, *36*(1), 3-15.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.

Carin, A.A. (1993). *Teaching science through discovery.* USA: Macmillan Publishing Company.

Casey, A., & Goodyear, V. A. (2015). Can cooperative learning achieve the four learning outcomes of physical education? A review of literature. *Quest*, *67*(1), 56-72.

Chen, Y.L. (2013). A missing piece of the contemporary character education puzzle: The ındividualisation of moral character. *Studies in Philosophy and Education*, *32*(4), 345-360.

Cheruvelil, K. S., Soranno, P. A., Weathers, K. C., Hanson, P. C., Goring, S. J., Filstrup, C. T., & Read, E. K. (2014). Creating and maintaining high-performing collaborative research teams: the importance of diversity and interpersonal skills. *Frontiers in Ecology and the Environment*, *12*(1), 31-38.

Clapper, T.C. (2015). Cooperative-based learning and the zone of proximal development. *Simulation & Gaming*, *46*(2), 148-158.

Cohen, E.G., & Press, C. (2015). *Designing groupwork: Strategies for the heterogeneous classroom*. USA: Hawker Brownlow Education.

Cook, D.T., & Campbell, D.T. (1979). *Quasi experimentation: Design & analysis for field settings.* Boston: Houghton Mifflin Company.

Costa, R. (2014). Effective teaching methods ın the master's degree: learning strategies, teaching-learning processess, teacher training. *European Scientific Journal*, (1), 106-120.

Dadds, M. (2002). Taking curiosity seriously: The role of awe and wanda inresearch-based professionalism. *Educational Action Research*, *10*(1), 9-26.

De Oliveira, M.L.B., Bezerra, I.M.P., Machado, M.D.F.A.S., Sorpreso, I.C.E., Raposo, F.A.G., Pinasco, G.C., & de Abreu, L.C. (2016). Health education: Curiosity as a parameter of the freirian model in primary care. *International Archives of Medicine*, *9*(1).

Dudija, N. (2016). Cooperative vs competitive: Filosofi keseimbangan "yin-yang" dalam hubungan ınterdependency. *Buletin Psikologi*, *23*(2), 65-81.

Fauzi, A. (2013). *Hubungan antara keharmonisan keluarga dan dukungan sosial teman sebaya dengan perilaku prososial remaja di sman 2 jombang*. (Doctoral Dissertation). UIN Sunan Ampel Surabaya.

Fernandez-Rio, J., Sanz, N., Fernandez-Cando, J., & Santos, L. (2017). Impact of a sustained Cooperative Learning intervention on student motivation. *Physical Education and Sport Pedagogy*, *22*(1), 89-105.

Gibbons, P. (2002). *Scaffolding language, scaffolding learning: Teaching second language learners in the mainstream classroom*. Portsmouth, NH: Heinemann.

Gillies, R.M. (2016). Cooperative learning: Review of research and practice. *Australian Journal of Teacher Education*, *41*(3), 3.

Ginsberg, A.E. (2012). *Embracing risk in urban education: Curiosity, creativity, and courage in the era of" no excuses" and relay race reform*. R&L Education.

Goleman, D. (2001). Emotional intelligence: Issues in paradigm building. In C.Cherniss and D. Goleman (Ed.), *The emotionally ıntelligent workplace*. Jossey-Bass.

Gommans, R., Segers, E., Burk, W.J., & Scholte, R.H. (2015). The role of perceived popularity on collaborative learning: A dyadic perspective. *Journal of Educational Psychology*, *107*(2), 599.

Hertz-Lazarowitz, R., Kagan, S., Sharan, S., Slavin, R., & Webb, C. (Eds.). (2013). *Learning to cooperate, cooperating to learn*. Springer Science & Business Media.

Huang, Y. M., Liao, Y. W., Huang, S. H., & Chen, H. C. (2014). A jigsaw-based cooperative learning approach to ımprove learning outcomes for mobile situated learning. *Journal of Educational Technology & Society*, *17*(1).

Huda, N., & Buditjahjanto, I.G.P.A. (2016). Pengaruh model pembelajaran kooperatif tipe jigsaw dan model pengajaran langsung terhadap hasil belajar siswa pada mata pelajaran teknik elektronika dasar kelas x tei di smkn 1 jetis mojokerto. *Jurnal Mahasiswa Teknologi Pendidikan*, *5*(1).

Hendarwati, E. (2016). Pengaruh pemanfaatan lingkungan sebagai sumber belajar melalui metode ınkuiri terhadap hasil belajar siswa SDN I sribit delanggu pada pelajaran IPS. *Pedagogia*, *2*(1), 59-70.

Isaac, S., & Michael, W.B. (1981). *Handbook in research and evaluation for education sciences (2nd ed.)*. San Diego 92107: EdITS Publishers.

Järvelä, S., Kirschner, P. A., Panadero, E., Malmberg, J., Phielix, C., Jaspers, J., ... & Järvenoja, H. (2015). Enhancing socially shared regulation in collaborative learning groups: designing for CSCL regulation tools. *Educational Technology Research and Development*, *63*(1), 125-142.

Johnson, D., & Johnson, R. (2002). Learning together and alone: Overview and meta analysis. *Asia Pacific Journal of Education, 22*, 95-105.

Johnson, D.W., Johnson, R.T., Stanne, M., & Garibaldi, A. (1990).The impact of leader and member group processing on achievement in cooperative groups. *Journal of Social Psychology, 130*, 507-516.

Johnson, D. W., & Johnson, R. T. (1994). *Learning together and alone: Cooperative, competitive, and individualistic learning*. Boston: Allyn and Bacon.

Jo, J., Park, J., Ji, H., Yang, Y., & Lim, H. (2016). A study on factor analysis to support knowledge based decisions for a smart class. *Information Technology and Management*, *17*(1), 43-56.

Lantolf, J.P., Thorne, S.L., & Poehner, M.E. (2015). Sociocultural theory and second language development. *Theories in Second Language Acquisition: An Introduction*, 207-226.

Leasa, M., & Corebima, A. D. (2017, January). The effect of numbered heads together (NHT) cooperative learning model on the cognitive achievement of students with different academic ability. In *Journal of Physics: Conference Series* (Vol. 795, No. 1, p. 012071). IOP Publishing.

Liu, Y., & Gao, Y. (2011). *Cooperative learning strategy*. Beijing Normal University Press.

Maier, U., Wolf, N., & Randler, C. (2016). Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types. *Computers & Education*, *95*, 85-98.

Maulidi, A.Y., Hamid, A., & Leny, L. (2016). Komparasi hasil belajar kognitif siswa antara penggunaan lks berbasis learning cycle 3 fase dan non lks pada materi reaksi redoks kelas x sma negeri 10 banjarmasin. *Quantum*, *5*(2).

McKeachie, W.J. (1990). Research on college teaching: The historical background. *Journal of Educational Psychology*, *82*(2), 189.

Muldayanti, N.D. (2013). Pembelajaran biologi model STAD dan TGT ditinjau dari keingintahuan dan minat belajar siswa. *Jurnal Pendidikan IPA Indonesia (Indonesian Journal of Science Education)*, *2*(1).

Nashar. (2004). *Peranan motivasi dan kemampuan awal dalam kegiatan pembelajaran*. Jakarta: Delia Press.

Nurjanah, S.I., & Ragil, W.A. (2012). Model kooperatif tipe TGT untuk meningkatkan pemahaman konsep koperasi pada mata pelajaran IPS. *Didaktika Dwija Indria*, *2*(2).

Pangesti, A. (2014). *Pengaruh motivasi terhadap prestasi belajar pada siswa mts bahrul ulum kecamatan ukui kabupaten pelalawan.* (Doctoral Dissertation). Universitas Islam Negeri Sultan Sarif Kasim Riau.

Pratiwi, Y., Mulyani, S., & Ashadi, A. (2015). Upaya peningkatan prestasi belajar dan ınteraksi sosial siswa dengan menggunakan metode pembelajaran Teams Games Tournament (TGT) dilengkapi media peta konsep pada materi kelarutan dan hasil kali kelarutan siswa kelas Xi IPA SMA Batik 1 surakarta. *Jurnal Pendidikan Kimia*, *4*(1), 182-188.

Purnomo, K.I., Murti,B., & Suriyasa, P. (2013). Perbandingan pengaruh metode pendidikan sebaya dan metode ceramah terhadap pengetahuan dan sikap pengendalian hiv/aids pada mahasiswa fakultas olahraga dan kesehatan universitas pendidikan ganesha. *Jurnal Magister Kedokteran Keluarga*, *1*(1), 49-56.

Rahmawati, E.D. (2016). Pengaruh pergaulan teman sebaya dan konsep diri terhadap kecerdasan emosional siswa kelas V SD Negeri Se-Gugus III kecamatan tegalrejo yogyakarta tahun ajaran 2014/2015. *Basic Education*, *4*(14).

Rainsch, S. (2004). *Dynamic strategic analysis: Demistyfying simple succes strategies.* Wiesbaden: Deutcscher Universitasts-verlag.

Reigeluth, C., Myers, R., & Lee, D. (2016). The learner-centered paradigm of education. *Instructional-Design Theories and Models, Volume IV: The Learner-Centered Paradigm of Education*, 5.

Robbins, S.P. (1996). *Organizational behaviour*. New York: Prentice Hall.

Roscoe, J.T. (1975). *Fundamental research statistic for the behavior sciences. (2nd,ed), Holt.* New York: Rinehart and Winston.

Salomon, G., Globerson, T., & Guterman, E. (1989). The computer as a zone of proximal development: Internalizing reading-related metacognitions from a Reading Partner. *Journal of Educational Psychology*, *81*(4), 620.

Soebiyanto, S., Masykuri, M., & Ashadi, A. (2016). Pembelajaran kimia menggunakan model Student Teams Achievement Division (STAD) dan Team Games Tournament (TGT) ditinjau dari kemampuan awal dan gaya belajar. *Jurnal Kimia dan Pendidikan Kimia*, *1*(1), 52-66.

Shadish, W.R. (1995) *Philosophy of Science an the Quantitatve-Qualitative Debates: Thirteen common errors. Evaluation and Program Planning*, *18*(1), 63-75.

Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and quasi-experimental design for generalized causal ınference*. Houghton Mifflin Company: Boston.

Slavin, R. E. (1996). Research on cooperative learning and achievement: what we know, what we need to know. *Contemporary Educational Psychology, 21*(1), 43-69.

Slavin, R. (2014). Cooperative learning and academic achievement: Why does groupwork work? *Anales De Psicologia, 30*, 785-791.

Slavin, R.E. (2015). Cooperative learning in elementary schools. *Education 3-13*, *43*(1), 5-14.

Smith, J., van Eerde, H., & Bakker, A. (2013). A conceptualisation of whole-class scaffolding. *British Educational Research Journal*, *39*(5), 817-834.

Sumaryati, E. (2013). Pendekatan ınduktif-deduktif disertai strategi think-pair-square-share untuk meningkatkan kemampuan pemahaman dan berpikir kritis serta disposisi matematis siswa SMA. *Infinity Journal*, *2*(1), 26-42.

Taghibaygi, M., Rafe, M., & Moosavi, S.A. (2015). Analysis students' motivation in vocational schools and agricultural training centers in kermanshah province toward studying the field of agriculture. *International Journal of Advanced Biological and Biomedical Research*, *3*(1), 105-114.

Uhamista, U. (2016). Pengaruh pendekatan belajar kooperatif dan belajar kompetitif serta kemampuan motorik terhadap pengembangan self esteem melalui kegiatan olahraga permainan pada siswa sekolah dasar. *Journal of Physical Education and Sport*, *1*(2), 34-44.

Ulstad, S.O., Halvari, H., Sørebø, Ø., & Deci, E.L. (2016). Motivation, learning strategies, and performance in physical education at secondary school. *Advances in Physical Education*, *6*(01), 27.

Wang, T. (2002). The basic concept of cooperative learning.*Educational Research,* (2), 68-72.

Wang, T. (2005). The analysis of theoretical basis of cooperative learning. *Curriculum,Teaching Material and Method(1)*, 30-35.

Wilson, K., & Devereux, L. (2014). Scaffolding theory: High challenge, high support in Academic Language and Learning (ALL) contexts. *Journal of Academic Language and Learning*, *8*(3), A91-A100.

Xue, G., & Lingling, L. (2018). A comparative study on cooperative learning in multimedia and network environment used by english majors between China Mainland and Taiwan. *Advances in Language and Literary Studies*, *9*(1), 127-135.

Xu Xueping. (2012). Research the theory of cooperative learning in English teaching. *Journal of Gansu Normal Colleges*, 17(1), 119-120.

Yulyani, S.D., Sundari, N., & Hamid, S.I. (2016). Meningkatkan kemampuan berpikir kritis siswa melalui model cooperative learning teknik send a problem dalam pembelajaran IPS SD. *Jurnal PGSD Kampus Cibiru*, *4*(3).

Yu, H., Zheng, Z., & Zheng, W. (2008). *Understanding user behavior in large scale video on demand systems*. In L. Song (Ed.). Collaboration: New York.

Zhang, Z., Coutinho, E., Deng, J., & Schuller, B. (2015). Cooperative learning and its application to emotion recognition from speech. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, *23*(1), 115-126.

Zhang, X., Meng, Y., de Pablos, P. O., & Sun, Y. (2017). Learning analytics in collaborative learning supported by Slack: From the perspective of engagement. *Computers in Human Behavior,* 1-9. Doi: https://doi.org/10.1016/j.chb.2017.08.012

Eurasian Journal of Educational Research 75 (2018) 99-114

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

A Comparison of Difficulty Indices Calculated for Open-Ended Items According to Classical Test Theory and Many Facet Rasch Model

Mustafa ILHAN [1], Nese GULER [2]

ARTICLE INFO

Article History:
Received: 23 Jan. 2018
Received in revised form: 09 Mar. 2018
Accepted: 26 Mar. 2018
DOI: 10.14689/ejer.2018.75.6

Keywords
open ended question, item difficulty index classical test theory, many facet rasch model

ABSTRACT

Purpose: This study aimed to compare difficulty indices calculated for open-ended items in accordance with the classical test theory (CTT) and the Many-Facet Rasch Model (MFRM). Although theoretical differences between CTT and MFRM occupy much space in the literature, the number of studies empirically comparing the two theories is quite limited. Therefore, this study is expected to be a substantial contribution to the literature.

Research Methods: The research data were collected through three teachers rating the answers given by 375 eighth grade students to ten open-ended questions in a mathematics test. The difficulties of the items in the test were calculated according to CTT and MFRM by using the obtained data, and the consistency between the difficulty indices estimated based on the two theories was tested. While the Microsoft Excel program was used in the analyses for CTT, the FACETS package was employed in the analyses for MFRM.

Findings: The research findings showed that CTT and MFRM yielded similar results in terms of difficulty indices of open-ended questions. It was found that, according to both theories, the ten items in the achievement test were ranked as I2, I3, I1, I4, I7, I6=I8, I5 and I9, from easiest to most difficult.

Implications for Research and Practice: It may be said that estimating item difficulties according to either CTT or MFRM will not cause any notable differences in terms of the items to be included or excluded in the development of an achievement test with open-ended questions.

[1] Dicle University, TURKEY, e-mail: mustafailhan21@gmail.com, ORCID: orcid.org/0000-0003-1804-002X
[2] İzmir Democracy University, TURKEY, e-mail: gnguler@gmail.com, ORCID: orcid.org/0000-0002-2836-3132

Introduction

The majority of constructs intended to be measured in education and in psychology are abstract and cannot be directly observed. For this reason, stimuli to transform the constructs into observed outcomes are needed to measure such constructs. The stimuli expected to uncover certain types of responses in individuals are called *items*, and the process of selecting the items to stimulate only the properties of individuals intended to be measured on the basis of certain criteria is referred to as *test development*. In this context, selecting the appropriate items in the test development process is the pre-requisite to accurately measuring a property. The appropriate items are selected through item analysis. Item difficulty and discrimination indices are calculated, and efforts are realized to determine the functioning of items in item analysis. Item analysis is a common procedure in all test development processes. However, the following stages can differ according to the theory of measurement used. There are two main theories used to estimate item statistics: the Classical Test Theory (CTT) and the Item Response Theory (IRT).

Classical Test Theory (CTT)

CTT, which is also called the True Score Theory, is described with the concepts of true score, observed score, and random error. According to CTT, a value found in consequence of a measurement operation represents the observed score for the measured property, and the score is composed of the true score and random error (Kline, 2005). Therefore, whether the measured property reaches its true value depends on if the random error in measurement is zero. Nevertheless, it is inevitable for measurements to contain a certain amount of error, no matter how carefully the measurement is performed. Therefore, it is impossible to reach the true score in measurement activities, and true scores is estimated by means of observed scores. The estimation is based on the assumptions that true scores and error scores are uncorrelated, that there are no systematic patterns between the error scores obtained from the parallel applications of the same measurement tool, and that the expected value of the error scores is zero (Hambleton & Jones, 1993). CTT does not have many assumptions, which is considered as an advantage as it is easier to apply CTT to several test situations (Kelecioğlu, 2001). In addition, the mathematical operations it requires are not difficult, and it can be used with small samples (Schumacher, 2010). In addition to its positive aspects, which make it possible to use CTT in a wide range of areas, the theory also has certain limitations forcing researchers to search for new methods. First, item parameters are dependent on the group to which a test is administered, and the individuals' ability levels are dependent on the items available in a test in CTT (Hambleton, 2004). In addition, it is not possible to make an evaluation on individuals' performance at the level of items in CTT, since it is dependent on the total scores received from a test (Hambleton, Swaminathan, & Rogers, 1991). Other weaknesses of CTT include yielding only one standard error value for all individuals to whom a test is administered, difficulties performing measurements of high reliability with a small number of items, and that the reported data are in the ordinal scale (Embretson & Reise, 2000). These limitations of CTT have paved the way for new methods; thus, IRT was proposed, claiming that it could offer solutions to the abovementioned limitations.

Item Response Theory (IRT)

IRT is a theory of measurement that is based on the probabilistic relations between responses to items in a test and the construct a test aims to measure (Schultz & Whitney, 2005). The construct that is intended to be measured with a test but is not directly observed is called a *latent trait* in IRT. For this reason, the Latent Trait Theory is another name for IRT (De Ayala, 2009). Differing models were suggested throughout the historical development of IRT. The first model suggested within the framework of IRT was the Rasch model, which was developed for items rated in two categories and contains only difficulty parameter (DeVellis, 2003). A two-parameter model was developed with the inclusion of a discrimination parameter in the Rasch model, and a three-parameter model was developed with the inclusion of a guessing parameter in the two-parameter model (Furr & Bacharach, 2008). As can be understood, the first factor influential in the emergence of different models in the development process of IRT was the number of estimated item parameters. The second factor was the response categories in relation to items. IRT was first developed for items that were rated dichotomously. However, later, the use of the theory was not limited to dichotomously rated items and, thus, models for polytomous items (nominal response model, partial credit model and graded response theory) were also included in IRT (Harvey & Hammer, 1999; van der Linden, 2005). IRT is divided into two categories, parametric and non-parametric models, in terms of approaches considered in estimating the item characteristic curve. While parametric IRT models assume that the item characteristic curve has normal ogive or logistic properties, non-parametric models do not have an assumption limiting the item characteristic curve to a certain form (Takano, Tsunoda & Muraki, 2015). Another element distinguishing IRT models from each other is dimensionality. IRT is considered as unidimensional and multi-dimensional in this respect (Reckase, 2009). And finally, IRT can be considered in two groups, two-facet models and many-facet models, in terms of the number of variability sources, which can be influential in measurement results. The sources of variability that can affect measurement results are limited to -items and persons- in the two-facet model. On the other hand, other sources of variability (such as raters) apart from items and persons can also affect measurement results in the many-facet model. Today, IRT has only one model containing more than two sources of variability. The model, which is based on the Rasch analysis, is referred to as the Many-Facet Rasch Model (MFRM).

Many-Facet Rasch Model (MFRM)

MFRM was developed as an extension of the partial credit model by Linacre in 1989. MFRM is a model in which all sources of variability, such as raters, items and persons, which have the potential to influence measurement results, are considered simultaneously (Lunz & Stahl, 1990). In this respect, it differs from other IRT models that have two sources of variability labeled as items and persons, and it becomes a model that is primarily preferable in analyzing data from open-ended items (Mulqueen, Baker & Dismukes, 2000). Using MFRM in the analysis of subjectively rated tests enables researchers to compare all facets, such as raters, persons and items considered in analyses, on a common metric. It also enables researchers to detect rater errors (such as Halo effects, central tendency, bias, range restriction etc.) and assures that measurements for raters are also taken into consideration in estimations

for item difficulty and individuals' ability levels (Lynch & McNamara, 1998). Furthermore, MFRM includes all the advantages common to the other IRT models. In other words, abilities can be estimated for individuals independently of item parameters, and item parameters can be estimated independently of individuals' ability levels in MFRM, similarly to other IRT models when the model-data fit has been attained (Sudweeks, Reeve, & Bradshaw, 2005). Additionally, the data in the ordinal scale are brought to the level of those in the interval scale in MFRM, and separate error values are reported for each element in facets of measurement, in contrast to CTT, which yields only one value of standard error (Prieto & Nieto, 2014). Therefore, MFRM offers more advantages than CTT in those aspects. Nevertheless, discussions on how the differences between MFRM and CTT are reflected into the analysis results based on the two theories still occupy a significant place in the literature of measurement and evaluation.

Studies Comparing CTT and MFRM

Although MFRM entered the literature approximately 30 years ago, empirical studies comparing MFRM with CTT started after the 2000s. The first study to compare the results obtained through CTT to those obtained through MFRM was performed by MacMillian (2000), and an increase of similar studies was observed in the following years. Studies concerning a comparison between CTT and MFRM available in the literature are summarized in Table 1.

Table 1

Studies Concerning a Comparison between CTT and MFRM available in the Literature

Study Tag	Purpose of Study
MacMillian (2000)	The study examines the consistency between results reported in CTT, MFRM and generalizability theory in rater reliability.
Haiyang (2010)	Reliability coefficients calculated for an English test including open-ended questions according to CTT and MFRM are compared.
Kadir (2013)	Findings on analyses on the basis of CTT, MFRM, and the generalizability theory are compared in assessing writing performance in English.
Huang, Guo, Loadman, and Low (2014)	Item difficulty parameters calculated in CTT and MFRM and reliability estimated according to the two theories are compared.
İlhan (2016)	Ability estimations made according to CTT and MFRM are compared in terms of relative agreement, absolute agreement, and criterion-related validity.

As is clear from Table 1, the issue most frequently considered in comparing CTT and MFRM is the extent to which reliability values calculated according to the two theories agree. It is evident that studies that were conducted more recently compared the two theories in terms of ability estimation and item difficulty parameters. Upon examining the studies comparing the reliability values reported in CTT and MFRM, studies were found to differ in terms of designs used. Accordingly, some of those studies (Güler & Gelbal, 2010; Haiyang, 2010) used crossed designs, in which all the students' responses were assessed by the same group of raters. Others (Huang et al., 2014; MacMillian; 2014), however, employed nested designs, in which different groups of raters were utilized in the assessment process. Thus, the available studies in the literature presented comprehensive information on the degree to which reliability values calculated on the basis of the two theories are in agreement. However, the same cannot be said about the comparison of calculated item difficulty indices in accordance with CTT and MFRM. This is because only one study comparing item difficulty in accordance with CTT and MFRM was found in the literature (Huang et al., 2014), and that study used a nested design. No studies comparing the item difficulties calculated according to the two theories by using a crossed design were encountered in relevant literature. In addition, while the study conducted by Huang et al. (2014) estimated item difficulty according to CTT by basing it on measurements for the top 25% and bottom 25% groups, the study used measurements for all individuals in estimating item difficulty according to MFRM. Such a difference is thought to be important for a study comparing the item difficulty values calculated according to both theories. In this sense, comparing item difficulty values calculated according to CTT and MFRM with different measurement conditions from those mentioned in the literature would be considered valuable.

Purpose and Significance of the Study

This study aimed to compare item difficulty indices calculated according to CTT with those calculated according to MFRM for open-ended questions. The study employed a crossed design in which students' responses to open-ended items were assessed by the same group of raters. Furthermore, since measurements for all individuals were taken into consideration in estimating item difficulty according to MFRM, item difficulty indices in CTT were also calculated by including all the individuals in the analysis, and not on the basis of the top and bottom groups. Thus, this study differs from Huang et al. (2014) in this respect. For this reason, it may be said that the study is original and could contribute to the literature. The fact that it can function as a resource calculating item difficulty for open-ended questions on the basis of CTT is another property of this study that is expected to be a valuable contribution. Calculation of item difficulty according to CTT is generally restricted to multiple-choice tests in studies in Turkish literature. This study offers a detailed description on calculating CTT-based item difficulty. Therefore, it is thought to serve as an important resource in calculating difficulty indices for open-ended items in measurement activities where one or more than one rater is available.

Method

Research Model

This study aimed to compare difficulty indices calculated for open-ended items according to two different theories of measurement, which qualifies it as basic research. Basic research is concerned with generating new knowledge, unlike applied studies, which focus on the use of knowledge (Bickman & Rog, 2009). Therefore, studies aiming to develop a theory or compare the theories available in the literature are defined as basic research (Connaway & Powell, 2010).

Study Group

This study was conducted in Diyarbakır city center in the spring semester of the 2016-2017 academic year. The participants were 375 eighth graders, of which 183 (48.80%) were girls and 192 (51.20%) were boys, and three mathematics teachers who rated the students' responses to open-ended mathematics questions.

Data Collection Tool

The data were collected through an achievement test of open-ended questions and a holistic rubric used to grade the students' responses to the test items. The achievement test used in the study contained ten open-ended mathematics questions and was developed by Ilhan (2016a). The rubric developed by Ilhan (2016b) was employed in marking the responses to the open-ended items. The rubric has four categories: *inadequate*, *needs improvement*, *good,* and *very good*. The students' responses to the items were rated between 1 and 4, based on the four categories listed in the rubric. After grading, analyses for the validity and reliability of the measurements were realized.

The arithmetic mean was calculated for the grades given by three raters for each item within the scope of CTT-based validity and reliability analyses. Then, exploratory factor analysis was executed and Cronbach's alpha internal consistency coefficient was calculated. Accordingly, it was found that the explained variance was 70.60% in the factor analysis, and that the test items had one factor with factor loads ranging between .68 and .93. Cronbach's alpha internal consistency coefficient was found to be .95. Inter-rater correlation coefficients were also calculated for the estimation of rater reliability according to CTT, and the correlation values were found as .75 (rater1-rater2), .65 for (rater1-rater3), and .60 (rater2-rater3).

The psychometric properties of the collected data were analyzed not only on the basis of CTT, but also on the basis of MFRM. Table 2 shows the findings reported for reliability and model-data fit in MFRM. As is apparent from Table 2, the infit and outfit statistics are in the suggested interval of .5 and 1.5 for all three person, item, and rater facets (Wright & Linacre, 1994). These values for fit indices demonstrate that the model-data fit was attained and that the measurements are valid.

Table 2

Findings Reported for Reliability and Model-Data Fit in MFRM

	Person	Item	Rater
Infit	.99	.99	.99
Outfit	1.02	1.02	1.02
Separation ratio	4.45	10.57	39.11
Reliability	.95	.99	1.00
df	374	9	2
Chi-square	6467.3*	1000.1*	3063.3*

*p<001

According to Table 2, the chi-square value for the rater facet is significant and the reliability coefficient and the separation ratio are high. This result indicates that the raters differed in severity/leniency. Despite the differences mentioned, the values reported for item and person facets show that the measurements are reliable. This is clear from Table 2 that the chi-square values for person and item facets are significant, the reliability coefficients are above .80, and the separation ratios are above 2 (Linacre, 2012). In other words, the students' performance on different items of the test were marked independently of each other, and students with differing ability levels were distinguished from each other effectively.

Data Analysis

The data were analyzed at two stages. First, CTT-based item difficulties were found. The following formula was used in calculating item difficulty indices for open-ended items:

Difficulty Index = (x – y) / (z – y)

x: Mean scores received from the item
y: The minimum score receivable from the item
z: The maximum score receivable from the item

The formula can be directly used in cases in which there is only one rater. However, when there is more than one rater, certain procedures should be followed prior to using the formula. The first step taken here was to calculate the mean scores assigned by different raters to students' responses to each item. The second step was to divide the sum total of the scores students had received from the items into the number of participants (separately for each item) to attain mean scores for the items. After that, the abovementioned formula was used. In other words, the difference was found for each item by subtracting the mean scores received from an item from the minimum score receivable from an item, and then the difference was divided into the item score range to attain the CTT-based item difficulty indices. The Microsoft Office Excel program was used in all operations for estimating item difficulty according to CTT.

Second, the MFRM analysis was executed in a design containing three sources of variability as persons, items, and raters. Before interpreting the analysis outcomes, whether the assumptions of MFRM were met was tested. As was stated above under Data Collection Tool, the fit statistics suggested that the model-data fit was attained.

Another indicator that the model-data fit had been attained, in addition to the infit and outfit statistics, was the standardized residual reported in MFRM. A model is considered to fit the data when the number of standardized residuals remaining outside ±2 interval is not above 5% of the total number of data, and when the number of standardized residuals remaining outside ±3 interval is not above 1% of the total amount of data in consequence of analyses (Linacre, 2014). While the proportion of the standardized residual outside the ±2 interval to the total number of data was 0.22%, according to the MFRM outcomes [25 out of 11250 (375x10x3) data], there were no data found with standardized residuals remaining outside the ±3 interval. Accordingly, it may be said that there is a high fit between the model and the data. Since the Rasch analysis was based on unidimensional data, the high fit between the model and the data indicated that the assumption of unidimensionality was met. As the assumption of unidimensionality functioned in parallel to local independence (Hambleton et al., 1991), attaining unidimensionality indicated that local independence – another assumption of MFRM – was also met. Having found that the assumptions were met, the measurement reports for the item facet were examined to determine the difficulty indices calculated in MFRM. The FACETS package program was used in the analyses for MFRM in this study.

After calculating item difficulty indices according to CTT and MFRM, correlations between difficulty indices estimated according to both theories were checked. Furthermore, a chart for the correlation between item difficulty indices found for CTT and for MFRM was created to visually express the correlation. Microsoft Excel was used in operations for calculating correlation coefficients and in forming the chart as in CTT-based item difficulty analyses.

Results

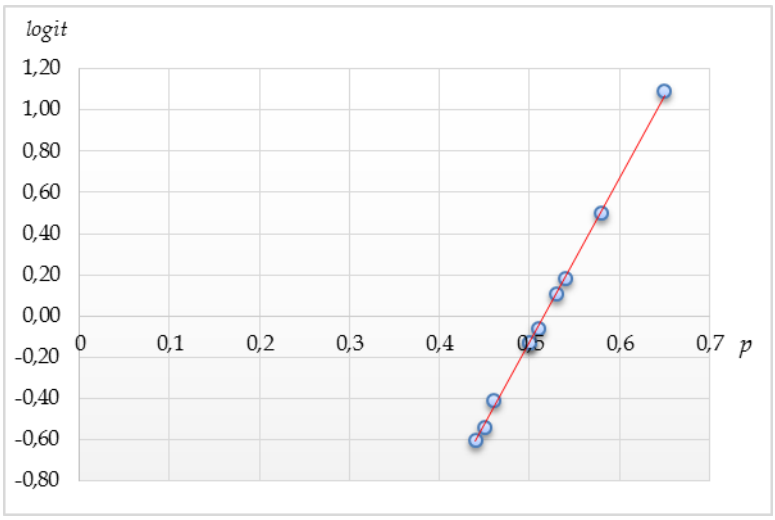
Difficulty indices calculated according to CTT and MFRM for the ten open-ended items in the achievement test used in this study are shown in Table 3. An item difficulty index close to zero in CTT demonstrates that an item is difficult, whereas a value close to 1 indicates that the item is easy. The way item difficulty indices are interpreted differs according to whether the item facet is positively or negatively oriented in MFRM. When an item facet is described as positively oriented, items become increasingly more difficult as one moves from the negative end of the logit scale to the positive end of it. On the contrary, when an item facet is described as negatively oriented, it is said that items with high logit values are easier and that the ones with low logit values are more difficult. Therefore, the item facet was defined as negatively oriented in the Rasch analysis to accurately compare the difficulty indices in CTT and in MFRM.

Table 3

Item Difficulty Indices Calculated in CTT and MFRM

Items	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10
CTT	.54	.65	.58	.53	.45	.50	.51	.50	.44	.46
MFRM	.18	1.09	.50	.11	-.54	-.13	-.06	-.13	-.60	-.41

As is clear from Table 3, the items are ranked from easiest to most difficult as I2, I3, I1, I4, I7, I6=I8, I10, I5 and I9 in both CTT and MFRM. Thus, it may be said that there is a complete agreement between the item difficulties calculated according to both theories. This is also evident from the chart below showing the correlation between item difficulty indices calculated in CTT and MFRM.



r = .999, p<.001

Chart 1. Correlation between Item Difficulty Indices Calculated in CTT and MFRM

As is clear from Chart 1, there is a linear correlation between item difficulty indices estimated according to the two theories. Chart 1 includes 9 points, although there are ten items in the achievement test, because I6 and I8 have equal difficulty indices in both CTT and MFRM. Spearman's correlation coefficient shown at the bottom of Chart 1 suggests that there is a positive and perfect correlation between item difficulty indices estimated in CTT and MFRM.

Discussion and Conclusion

This study aimed to compare difficulty indices for open-ended items calculated according to CTT and MFRM. The results obtained in this study suggested that there was a high level of agreement between difficulty indices estimated according to the two theories. Upon ranking the items according to their difficulty, the rankings were found to be identical in both theories. This was a similar result as that obtained by Huang et al. (2014), which compared item difficulty indices in CTT with those in MFRM by using a nested design. In Huang et al. (2014), 124 competitive grant applications were rated according to a six-point graded rubric with 24 items. Sixty-four experts rated the proposals, and each of the 124 proposals was assessed by only three experts; therefore, each expert rated approximately six different proposals. However, this current study, instead used a crossed design in which all of the 375

student responses to the ten open-ended mathematics questions were assessed by the same group of raters. Therefore, upon considering the results obtained in Huang et al. (2014) and those obtained in this study, the CTT and the MFRM yielded similar results in terms of item difficulty indices for open-ended questions, no matter which design – crossed or nested – was used.

Upon comparing the findings obtained in Huang et al. (2014) with those obtained in this study, it may be inferred that item difficulty indices in CTT – whether they are calculated by comparing the top and bottom groups or by including all individuals in analyses – agree with those reported in MFRM. This is because the item difficulties in CTT were calculated by comparing the top and bottom groups in Huang et al. (2014), but they were estimated by including all the individuals in the analyses in this study. Despite this difference, the difficulty indices calculated according to CTT and MFRM were found to agree in both studies.

It may be stated, based on the results of this study, that estimating item difficulty according to CTT or MFRM does not cause a difference in terms of the items to be included or excluded in the development of an achievement test with open-ended items. In the Mead and Meade (2010) simulation study, it was concluded that test construction using either CTT or IRT procedures lead to empirically similar exams. Thus, other properties, such as ease of use and the comprehensiveness of the reported results, should be prioritized in making decisions on whether to use CTT or MFRM in developing an achievement test containing open-ended questions. For instance, the fact that CTT is a more frequently used theory and that the analyses for this theory can easily be performed by using Microsoft Excel can cause researchers/practitioners to consider CTT as a more practical way to develop an open-ended test. In spite of those positive characteristics, MFRM also has advantages compared to CTT. For example, synchronically calculating the validity and reliability of measurements, item difficulties, individuals' ability levels, and raters' severity and leniency; comparing all the facets used by putting them on the same logit; and analysis outcomes having test information function, category statistics and unexpected responses – all of which have no counterparts in CTT – make MFRM a more preferable model to CTT, even though it yields similar results in terms of item difficulty indices.

Recommendations

A review of the literature showed that the item parameters estimated in CTT and IRT were mostly restricted to using multiple-choice tests. This current study, however, compared difficulty indices of open-ended items calculated on the basis of CTT with those calculated on the basis of MFRM-model based on IRT. It is thought that the study will contribute to the literature in this respect. Nevertheless, this study – as all scientific studies – also has some limitations. The restrictions, which also imply suggestions for further research, are related to the external validity of the study. Studies comparing different theories can contain effects stemming from the data set (Engelhard, 1984) and limit the generalizability of the conclusions reached in the study. Therefore, conducting similar studies with different data sets is important in raising generalizability of the conclusions reached.

References

Bickman, L., & Rog, D.J. (2009). *Applied social research methods*. Los Angeles: Sage.

Connaway, L.S., & Powell, R.R. (2010). *Basic research methods for librarians*. Santa Barbara, CA: Libraries Unlimited.

De Ayala, R.J. (2009). *The theory and practice of item response theory*. New York: The Guilford.

DeVellis, R.F. (2003). *Scale development: Theory and applications*. Thousand Oaks, CA: Sage.

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

Engelhard, G. (1984). Thorndike, Thurstone, and Rasch: A comparison of their methods of scaling psychological and educational tests. *Applied Psychological Measurement*, *8*(1), 21-38. http://dx.doi.org/10.1177/014662168400800104

Furr, R.M., & Bacharach, V.R. (2008). *Psychometrics: An Introduction*. Thousand Oaks, CA: Sage.

Haiyang, S. (2010). An application of classical test theory and many facet Rasch measurement in analyzing the reliability of an English test for non-English major graduates. *Chinese Journal of Applied Linguistics*, *33*(2), 87–102.

Hambleton, R.K. (2004). Theory, methods, and practices in testing for the 21st century. *Psicothema*, *16*(4), 696-701.

Hambleton, R.K., & Jones, R.W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice, 12*(3), 38-47. http://dx.doi.org/10.1111/j.1745-3992.1993.tb00543.x

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

Harvey, R.J., & Hammer, A.L. (1999). Item response theory. *The Counseling Psychologist*, *27*(3), 353-383. http://dx.doi.org/10.1177/0011000099273004

Huang, T.W., Guo, G.J., Loadman, W., & Low, F.M. (2014). Rating score data analysis by classical test theory and many-facet Rasch model. *Psychology Research*, *4*(3), 222-231.

İlhan, M. (2016a). A comparison of the results of many-facet Rasch analyses based on crossed and judge pair designs. *Educational Sciences: Theory & Practice*, *16*(2), 579-601. http://dx.doi.org/10.12738/estp.2016.2.0390

İlhan, M. (2016b). A comparison of the ability estimations of classical test theory and the many facet Rasch model in measurements with open-ended questions. *Hacettepe University Journal of Education*, *31*(2), 348-358. http://dx.doi.org/10.16986/HUJE.2016015182

Kadir, K.A. (2013). Examining factors affecting language performance: A comparison of three measurement approaches. *Pertanika Journal of Social Sciences & Humanities*, *21*(3), 1149-1162.

Kelecioğlu, H. (2001). The relationship between b and a parameters in latent trait theory and p and r statistics in classical test theory. *Hacettepe University Journal of Education*, *20*, 104–110.

Kline, T.J.B. (2005). Classical test theory: Assumptions, equations, limitations, and item analyses. In T.J.B. Kline (Ed.), *Psychological testing: A practical approach to design and evaluation* (pp. 91–105). Thousand Oaks, CA: Sage. http://dx.doi.org/10.4135/9781483385693.n5

Linacre, J.M. (1989). *Many-facet Rasch measurement*. Chicago: MESA Press.

Linacre, J. M. (2012). *Many-facet Rasch measurement: Facets tutorial2*. Retrieved December 19, 2018, from http://www.winsteps.com/a/ftutorial2.pdf

Linacre, J.M. (2014). *A user's guide to FACETS Rasch-model computer programs*. Retrieved December 18, 2018, from http://www.winsteps.com/a/facets-manual.pdf

Lynch, B.K., & McNamara, T.F. (1998). Using G-theory and Many-facet Rasch measurement in the development of performance assessments of the ESL speaking skills of immigrants. *Language Testing*, *15*(2), 158–180. http://dx.doi.org/10.1191/026553298674579408

Lunz, M.E., & Stahl, J.A. (1990, April). *Severity of grading across time periods*. Paper presented at the Annual Meeting of the American Educational Research Association, Boston. Retrieved January 11, 2018, from https://files.eric.ed.gov/fulltext/ED317602.pdf

MacMillan, P.D. (2000). Classical, generalizability and multifaceted Rasch detection of interrater variability in large sparse data sets. *The Journal of Experimental Education*, *68*(2), 167-190. http://dx.doi.org/10.1080/00220970009598501

Mead, A.D., & Meade, A.W. (2010, April). *Item selection using CTT and IRT with unrepresentative samples*. Paper presented at the twenty-fifth annual meeting of the Society for Industrial and Organizational Psychology in Atlanta, GA. Retrieved December 22, 2018, from http://mypages.iit.edu/~mead/Mead_and_Meade-v10.pdf

Mulqueen C., Baker D., & Dismukes, R.K. (2000, April). *Using multifacet Rasch analysis to examine the effectiveness of rater training*. Presented at the 15th Annual Conference for the Society for Industrial and Organizational Psychology (SIOP). New Orleans. Retrieved December 14, 2018, from https://www.air.org/sites/default/files/downloads/report/multifacet_rasch_0.pdf

Prieto, G., & Nieto, E. (2014). Analysis of rater severity on written expression exam using many faceted Rasch measurement. *Psicológica*, *35*, 385-397.

Reckase, M.D. (2009). *Multidimensional item response theory*. New York, NY: Springer.

Schumacker, R.E. (2010). *Classical test analysis*. Applied measurement associates LLC. Retrieved December 12, 2018, from http://appliedmeasurementassociates.com/ama/assets/File/CLASSICAL_TEST_ANALYSIS.pdf

Shultz, K.S., & Whitney, D.J. (2005). *Measurement theory in action: Case studies and exercises*. Thousand Oaks, CA: Sage.

Sudweeks, R.R., Reeve, S., & Bradshaw, W.S. (2005). A comparison of generalizability theory and many-facet Rasch measurement in an analysis of college sophomore writing. *Assessing Writing*, *9*(3), 239–261. https://doi.org/10.1016/j.asw.2004.11.001

Takano, Y., Tsunoda, S., & Muraki, M. (2015). Mathematical optimization models for nonparametric ıtem response theory. *Information Science and Applied Mathematics*, *23,* 1-16.

van der Linden, W.J. (2005). Item response theory. In K. Kempf-Leonard (Ed.), *Encyclopedia of social measurement* (Vol. 2, pp. 379-387). San Diego, CA: Academic Press.

Wright, B.D. & Linacre, J.M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, *8*, 370. Retrieved November 23, 2018 from http://www.rasch.org/rmt/rmt83b.htm

Açık Uçlu Maddelerde Klasik Test Kuramı ile Çok Yüzeyli Rasch Modeline göre Hesaplanan Güçlük İndekslerinin Karşılaştırılması

Atıf:

Ilhan, M., & Guler, N. (2018). A comparison of difficulty ındices calculated for open-ended ıtems according to classical test theory and many facet Rasch model. *Eurasian Journal of Educational Research*, 75, 99-114, DOI: 10.14689/ejer.2018.75.6

Özet

Problem Durumu: Klasik test kuramı (KTK) ve çok yüzeyli Rasch modeli (ÇYRM) arasındaki kuramsal farklılıklar alanyazında geniş bir yer tutmasına rağmen bu iki kuramı ampirik açıdan karşılaştıran araştırmaların oldukça sınırlı olduğu görülmektedir. KTK ve ÇYRM'nin karşılaştırılmasına yönelik çalışmalarda üzerinde en fazla durulan konu iki kurama göre hesaplanan güvenirlik değerlerinin ne derece tutarlı olduğudur. Daha yakın zamanda yapılan araştırmalarda ise iki kuramının yetenek kestirimleri ile madde güçlük parametreleri açısından karşılaştırıldığı anlaşılmaktadır. KTK ve ÇYRM'de rapor edilen güvenirlik değerlerinin

karşılaştırıldığı araştırmalar incelendiğinde, bu çalışmaların kullanılan desen açısından farklılık gösterdiği saptanmıştır. Çalışmaların bir kısmında açık uçlu maddelere verilen öğrenci cevaplarının tamamının aynı puanlayıcı grubu tarafından değerlendirildiği çapraz bir desen kullanılmıştır. Bazılarında ise değerlendirme sürecinde birbirinden farklı puanlayıcı gruplarının görev aldığı yuvalanmış bir desen tercih edilmiştir. Dolayısıyla, konu ile ilgili alanyazındaki mevcut çalışmalar iki kurama göre hesaplanan güvenirlik değerlerinin ne derece tutarlı olduğuna ilişkin kapsamlı bir bilgi sunabilmektedir. Ancak aynı şeyi KTK ile ÇYRM'de hesaplanan madde güçlük indekslerinin karşılaştırılmasına yönelik araştırmalar için söylemek güçtür. Çünkü alanyazında KTK ve ÇYRM'de hesaplanan madde güçlüklerinin karşılaştırıldığı yalnızca bir araştırmaya rastlanmış ve bu çalışmada yuvalanmış bir desenin kullanıldığı belirlenmiştir. İki kurama göre hesaplanan madde güçlüklerinin çapraz bir deseninin kullanıldığı ölçme koşulları altında karşılaştırıldığı bir çalışmaya ise alanyazında rastlanmamıştır. Ayrıca alanyazındaki sözü edilen araştırmada, KTK'ya dayalı madde güçlükleri %25'lik alt ve üst gruba ait ölçümler esas alınarak kestirilirken; ÇYRM'ye ilişkin madde güçlük kestiriminde tüm bireylere ait ölçümler kullanılmıştır. Böylesi bir farkın iki kurama göre hesaplanan madde güçlüklerinin karşılaştırıldığı bir çalışma için önemli olabileceği düşünülmektedir. Bu anlamda, ölçme koşulları açısından alanyazındaki bahsi geçen araştırmadan farklılık gösteren bir çalışma ile KTK ve ÇYRM'de hesaplanan madde güçlüklerinin karşılaştırılması önemli görülmektedir.

Araştırmanın Amacı: Bu araştırmada, açık uçlu maddelerde klasik test kuramı (KTK) ile çok yüzeyli Rasch modeline (ÇYRM) göre hesaplanan güçlük indekslerinin karşılaştırılması amaçlanmıştır.

Araştırmanın Yöntemi: Araştırmanın verileri, sekizinci sınıfa devam eden 375 öğrencinin açık uçlu 10 maddeye verdiği cevabın üç öğretmen tarafından puanlanmasıyla elde edilmiştir. Puanlamalarda dörtlü derecelemeye sahip bütüncül bir rubrik kullanılmıştır. KTK'ya dayalı madde güçlüklerinin hesaplanmasındaki ilk adım, öğrencilerin her bir maddeye verdikleri cevaplar için farklı puanlayıcılar tarafından atanan puanların ortalamasının alınması olmuştur. İkinci adımda tüm maddeler için ayrı ayrı olmak üzere, öğrencilerin maddelerden aldıkların puanların toplamı çalışmadaki katılımcı sayısına bölünmüş ve bu şekilde maddelere ilişkin puan ortalamaları hesaplanmıştır. Daha sonra her bir madde için, ilgili maddeden alınan puanların ortalaması ile maddeden alınabilecek en düşük puan arasındaki fark bulunmuştur. Bulunan bu farkın madde puan ranjına bölünmesiyle KTK dayalı madde güçlük parametrelerine ulaşılmıştır. Madde güçlüklerinin KTK'ya göre hesaplanmasında Microsoft Excel'den yararlanılmıştır. KTK'ya ilişkin analizlerin ardından ÇYRM'ye yönelik analizlere geçilmiştir. Bu kapsamda, FACETS paket programı kullanılarak puanlayıcı, madde ve öğrenci şeklinde üç yüzeyli bir desen ile Rasch analizi gerçekleştirilmiştir. Analiz çıktılarında, madde yüzeyine ilişkin ölçüm raporları incelenerek ÇYRM'ye dayalı madde güçlük parametreleri elde edilmiştir. Madde güçlük indekslerinin KTK ve ÇYRM'ye göre hesaplanmasını takiben, iki kurama göre kestirilen güçlük değerleri arasındaki tutarlılığa bakılmıştır.

Araştırmanın Bulguları: Araştırmadan elde edilen bulgular, iki kurama göre kestirilen güçlük indeksleri arasında yüksek bir tutarlılık olduğunu göstermiştir. Maddeler güçlük düzeyleri açısından bir sıralamaya tabi tutulduğunda KTK ile ÇYRM'de

ulaşılan sıralamaların özdeş olduğu saptanmış ve iki kurama göre kestirilen güçlük indeksleri arasında pozitif yönde, güçlü ve anlamlı bir korelasyonun (r=.999, p<.001) bulunduğu belirlenmiştir. Her iki kurama göre de başarı testindeki 10 maddenin kolaydan zora doğru; M2, M3, M1, M4, M7, M6=M8, M10, M5 ve M9 şeklinde sıralandığı sonucuna ulaşılmıştır.

Araştırmanın Sonuç ve Önerileri: Araştırma sonuçlarından hareketle, açık uçlu maddeler içeren bir başarı testi geliştirme sürecinde, madde güçlüklerinin KTK veya ÇYRM'ye göre kestirilmiş olmasının teste alınacak ya da test dışında tutulacak maddeler ile ilgili bir farklılık yaratmayacağı söylenebilir. Dolayısıyla açık uçlu maddelerin bulunduğu bir başarı testi geliştirirken KTK ile ÇYRM'den hangisinin tercih edilmesi gerektiğine dair verilecek kararlarda kullanım kolaylığı ve rapor edilen sonuçların ne derece ayrıntılı olduğu gibi kuramlara ilişkin diğer özelliklerin ön plana çıkacağı düşünülmektedir. Örneğin, KTK'nın birçok kişinin daha aşina olduğu bir kuram olması ve bu kurama ilişkin madde analizlerinin Microsoft Excel'de kolaylıkla gerçekleştirilebilmesi araştırmacıların/uygulayıcıların açık uçlu test geliştirme sürecinde KTK'yı daha pratik bir yol olarak görmesine sebep olabilir. KTK'yı ÇYRM'ye göre daha kullanışlı hale getiren bu özelliklerine karşın ÇYRM'nin de KTK'ya kıyasla daha avatanjlı olduğu bazı yönleri bulunmaktadır. Ölçümlerin geçerliği ile güvenirliğinin, madde güçlüklerinin, bireylerinin yetenek düzeylerinin ve puanlayıcıların katılık/cömertliklerinin eş zamanlı olarak hesaplanması, analizde işlem gören tüm yüzeylerin ortak bir metrik (logit) üzerine yerleştirilerek birbiriyle karşılaştırılabilmesi ve analiz çıktıları arasında KTK'da karşılığı olmayan test bilgi fonksiyonunun, kategori istatistiklerinin ve beklenmedik yanıtların yer alması madde güçlük indeksleri açısından benzer sonuçlar üretmesine rağmen ÇYRM'yi KTK'ya göre daha tercih edilebilir bir model haline getirebilecek özelliklerdir.

Anahtar Kelimeler: Açık uçlu maddeler, madde güçlük indeksi, klasik test kuramı, çok yüzeyli Rasch modeli.

Eurasian Journal of Educational Research 75 (2018) 115-136

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

Adapting a Residential Course to Web-Based Blended Learning

Busra OZMEN[1], Tansel TEPE[2], Hakan TUZUN[3]

ARTICLE INFO

Article History:
Received: 13 Mar. 2016
Received in revised form: 30 Jun. 2017
Accepted: 20 Apr. 2018
DOI: 10.14689/ejer.2018.75.7

Keywords
ADDIE instructional design model
distance education
foreign language learning
instructional design
learning management systems.

ABSTRACT

Purpose: This study describes the design of the Basic English 1 and Basic English 2 courses taught in the Foreign Languages School (FLS) of a large public university in Turkey as blended learning in a mostly distance education system. **Research Method:** This study was structured as a case study describing an instructional design effort. The study used the ADDIE instructional design model to convert these courses into the distance education format. In the analysis phase, the designers held discussions with the FLS lecturers to determine their needs and contextual limitations. These discussions revealed the knowledge and skills of the FLS students, the expectations of the lecturers and students, and the course objectives. The design process focused on determining Basic English 1 and 2 learning activities and presentation formats for these activities. Each course has a duration of 15 weeks and includes the skills of grammar, reading comprehension, writing, listening comprehension, and speaking. In the development phase, the designers added all the tools from the design phase to the Moodle environment. In the implementation phase, the designers conducted usability tests with the students and identified deficiencies of the design phase. In the evaluation phase, the designers improved the content based on student feedback from the usability tests. **Findings:** The designers conducted a planned and systematic design process based on the ADDIE instructional design model. At the end of the design process, the course content was transferred to the Moodle environment, and a usability test was administered to the students. The designers revised the environment's design based on student feedback. Eventually, the study presented an environment that was suitable for blended learning educational activities that mainly used distance education. **Implications for Research and Practice:** Designing blended learning opportunities for formal learning environments is a rising trend, and the community can benefit from the documented design process.

[1] Fırat University, TURKEY, e-mail: bozmen@firat.edu.tr, ORCID: https://orcid.org/0000-0003-3761-6215
[2] Kilis 7 Aralık University, TURKEY, e-mail: tanseltepe@kilis.edu.tr, ORCID: https://orcid.org/0000-0003-3576-6172
[3] Corresponding Author: Hacettepe University, TURKEY, e-mail: htuzun@hacettepe.edu.tr, ORCID: https://orcid.org/0000-0003-1153-5556

Introduction

With the use of special design software together with increasing access to pedagogical content, the Internet has begun to offer educational environments that enhance the distribution of information. At the same time, distance education increases the quality and strength of education since it is carried out over the Internet. Distance education also has the advantages of supporting teachers and students in the exchange of information independent from time and location, using multimedia, notes and materials; facilitating individual learning; and even reaching more people (Peachey, 2015). However, there are some important issues to be considered when preparing distance education environments. Sales (2010) indicated that online education could produce effective results when maintained in a successful, fast, and cheap way. The uninterrupted online provision of content prepared for a specific purpose at a low cost increases the quality of distance education. Quality online courses can be conducted quickly and economically thanks to the Learning Management Systems (LMSs).

The main disadvantages of online education have been its inabilities to provide learning materials that are not asynchronous and to share and discuss materials in different ways, to keep student, lecturer, and system records and to meet needs such as reporting. LMSs have recently been introduced to eliminate these deficiencies. Designers might use a variety of instructional strategies and design models to make courses effective and productive through LMSs (Cinar & Tuzun, 2016). There are many LMSs available on the market. However, Moodle is one of the most useful LMSs due to its ease of use (Ueda & Nakamura, 2016). Moodle is an open-source and free LMS. One of the greatest advantages of Moodle is that it offers many components—including course content, activities, examinations, surveys, discussion environments, and calendars—anytime and anywhere in both synchronous and asynchronous ways (Kudryashova, Gorbatova, & Rozhkova, 2016; Limongelli, Sciarrone, & Vaste, 2011). Moodle makes it possible to track students' participation within the system and determine how active they are during practice.

Instructional design examines everything related to planning instructional activities. It is described as the systematic development of instructional elements applying teaching and learning theories to the purpose of maintaining the quality of education (Berger & Kam, 1996). In the instructional design process, teaching activities and materials are developed, and teaching and learning activities are tested and evaluated. The theories that provide guidance for improving people's learning and development are called instructional design models (Reigeluth & Frick, 1999). In other words, instructional design models are approaches that help the systematic and effective progress of the instructional design processes based on a plan (Tuzun, 2001). Distance education can be implemented to determine the educational needs of a specific target group and to shape the development of effective learning systems to

satisfy these needs. Therefore, it is important to make use of the instructional design models to obtain the learning outcomes that are expected from distance education learning environments.

This study reviewed the relevant studies on the design of web-based distance education courses and found that Tuzun (2001) proposed an instructional design model including nine phases: preliminary efforts to the design, creating a center for the development of the course, doing the analysis, developing instructional strategies, providing administrative structure, doing the design, eliminating technological barriers, administering student assessment, and conducting course evaluation. Tuzun and Cinar (2016) improved this model by proposing an experience-based e-course design model consisting of seven stages: forming a design team, doing a preliminary search, conducting analysis, creating an instructional and technical design, integrating the designs, performing tests, and making improvements. Power (2009) tried to help university instructors who taught their courses using traditional methods to transfer their courses to web-based distance education. Power also described personal experiences in ten case studies based on notes taken over three years. Based on this work, Power (2009) proposed an instructional design model with six steps: conducting analysis, developing the modules, developing teaching activities, developing activities that support teaching, developing assessment tools, and performing ongoing revision. Balci (2010) also proposed an instructional design model with eight steps: deciding on the redesign of a course for web-based distance education, developing the program, administering the course, providing support, preparing the technical structure, testing the program, implementing changes, and updating and maintaining the e-learning program.

Blended learning is a form of learning that combines face-to-face learning and online learning. According to Garrison and Kanuka (2004), "At its simplest, blended learning is the thoughtful integration of classroom face-to-face learning experiences with on-line learning experiences" (p. 96). Blended learning offers self-learning opportunities and flexible online learning opportunities for students. With this learning approach, students are able to control learning content, speed, time, and location to a certain extent (Powell et al., 2015). Real-time data provided through digital technology to blended learning helps teachers organize the teaching process according to the progress of the student (Hilliard, 2015). In foreign language domain, blended learning is defined as a language teaching approach that combines the most effective features of face-to-face learning activities and online collaborative learning activities (So & Lee, 2013).

This study reviewed the research on the design and effectiveness of blended learning environments in foreign language learning and found that Dogan et al. (2011) conducted a study to convert the “Human-Computer Interaction” course into distance education. They used the ADDIE instructional design model (Analysis, Design,

Development, Implementation, and Evaluation), conducted a usability test and tried to demonstrate the effectiveness of the system based on feedback. Tomlinson and Whittaker (2013) wrote a book that compiled 20 case studies of blended learning in English language teaching. They tried to provide guidance for the design of blended learning environments and their implementation processes by balancing classroom learning with e-learning. Bueno-Alastuey and Lopez Perez (2014) created a blended learning environment to determine students' perception of the benefits of using Information and Communication Technologies (ICTs) in English-language teaching. Although all the participants had weak ICT skills in language learning at the beginning, most of them experienced positive changes in their perception of ICT's use in language learning. The study revealed the useful outcomes of ICT use in foreign language teaching. Sun and Qui (2016) presented a framework of a blended learning model for English teaching. They observed that the model is useful for developing participants' English proficiency. Additionally, participants earned higher scores on the final test at the end of the study. Challob, Bakar, and Latif (2016) investigated the effects of the blended learning approach on the writing anxieties and writing performance of foreign language learners. A 13-week study was conducted with secondary school pupils in Malaysia. Students participated in both face-to-face and online writing activities through in-class discussions, blogs, and online discussion tools. As a result, there was a significant improvement in students' anxiety about and performance in writing.

Emelyanova and Voronina (2017) conducted a study to determine students' attitudes and perceptions towards blended learning English classes and then designed an online course using an institutional LMS. According to the results obtained from the data collected before and after the lesson, it was revealed that there was a remarkable increase in the perception and attitudes of the students towards the use of the blended learning in foreign language teaching. Pinto-Llorente, Sanchez-Gomez, Garcia-Penalvo, and Casillas-Martin (2017) conducted a study to determine students' perceptions of the technological tools used in blended learning environments in foreign language teaching. The results of the study showed that the positive attitude of students about the effectiveness of technological tools contributed to the development of English grammar skills. Additionally, Wright (2017) investigated the preferences and reasons for students to take online and/or face-to-face lessons in an English grammar course conducted online at the undergraduate level in Malaysia. Prescott, Bundschuh, Kazakoff, and Macaruso (2017) implemented a blended learning program for reading and writing in English at a primary school, until the fifth year of school. With this study, it was found that the reading and writing performances of the students were positively affected. Other studies of blended learning models in foreign language teaching have also been conducted. These studies investigated the influence of blended learning environments on students' grammatical competence (Emelyanova & Voronina, 2017; Pinto-Llorente et al., 2017; Wright, 2017), reading (Pinto-Llorente,

Sanchez-Gomez, & Garcia-Penalvo, 2016; Prescott et al., 2017), speaking (Yang, Chuang, Li, & Tseng, 2013), writing (Cahyono & Mutiaraningrum, 2016; Challob, Bakar, & Latif, 2016; Eydelman, 2015; Ferriman, 2013; Pinto-Llorente et al., 2016), listening (Cigdem, Ozturk, & Topcu, 2016; Emelyanova & Voronina, 2017; Yang et al., 2013), pronouncing (Al Zumor, Al Refaai, Eddin, & Al-Rahman, 2013), learning foreign languages (Hinkelman & Gruba, 2012; Kudryashova, Gorbatova, & Rozhkova, 2016) and teaching foreign languages (Alpala & Florez, 2011; Peachey, 2015; Shaykina, 2015). Overall, the findings of these studies concluded that blended learning environments have positive effects on students' foreign language learning.

Although there are many studies of blended learning environments for foreign language learning, there are few Turkish studies of this subject. This study aimed to convert the Basic English 1 and 2 courses taught at the Foreign Languages School (FLS) of a large public university in Ankara into blended learning using the ADDIE instructional design model and to transfer these courses to a web-based environment. The web-based learning environment prepared for this purpose covered the content of the lessons that were being taught in the classroom environment and was suitable for the blended learning environment that combined classroom activities with the online learning environment. To summarize, the aim of the study was to adapt a classroom course to blended learning and transfer it to the web-based environment. Designing blended learning opportunities for formal learning environments is a rising trend (Bilgic & Tuzun, 2015a), and the community can benefit from the documented design process.

Method

Research Design

This is a case study research project describing a design effort. In a case study, the researcher conducts a profound analysis of one event or multiple events in a given timeframe and describes either the events or the themes associated with these events. The researcher collects information using data collection tools that include multiple sources such as observation, interviews, documents, and reports (Creswell, 2007). Interviews along with document analysis were utilized to collect data in this study in order to guide the design efforts.

Study Group

The study was conducted with students and lecturers from the Foreign Languages School at a large public university in Ankara, Turkey. The study group was chosen through a convenience sampling method. The students and lecturers participated in the study on a voluntary basis.

The Design Process

The study converted the Basic English 1 and 2 classroom courses to the distance education format and transferred them to the web-based environment. Its underlying model was the ADDIE instructional design model. The ADDIE model is one of the most common models used in the design of e-learning systems. It is a systematic design model that consists of five steps (Analysis, Design, Development, Implementation, and Evaluation) (Driscoll, 1998) and contains the components of the other models (Akkoyunlu, Altun, & Soylu, 2008). In this model, all steps are interconnected with each other and are designed based on the outcomes of the previous step.

Analysis.In the first phase of the design, designers held semi-structured interviews with the lecturers in the FLS to identify their needs and contextual limitations. The designers obtained information about the knowledge and skills of the students who attended these courses, the expectations of the lecturers and students, and the various course objectives. The FLS has a preparatory year program and a modern languages program. The students who attend the preparatory year programs are those who are enrolled in a program with a preparatory year because they failed to obtain the required score on the Foreign Language Proficiency Exam at the beginning of the academic year. The students who cannot pass this exam are admitted to courses based on their examination scores. The modern languages program students are enrolled in the university's colleges and need to take mandatory Basic English 1 and 2 courses. A majority of the students fail and enroll in the Basic English 1 and 2 courses again, and there are also students who enroll in these courses several times. The level of attendance in these courses is very low.

In context of the analysis of the lessons, the designers obtained the textbooks and CDs used in the lessons, as well as curricula, course descriptions, course content, general and specific course attainments, and assessment criteria from the lecturers. The designers closely examined these resources and shaped the design process using this information. Basic English 1 and 2 are mandatory three-credit courses taught by the Modern Languages Unit. To attend Basic English 2, students must complete Basic English 1. Each lesson is taught for four hours a week (two theoretical and two practical hours). In the FLS, the practical lessons are taught on DynED software due to a shortage of classrooms and lecturers. Two written mid-term exams and a final exam are administered to assess the students' achievement.

Course content includes

- In both courses: instruction in reading comprehension, writing, listening comprehension and speaking, and beginner-level English grammar.
- In Basic English 1: the verb "to be" in positive, negative and interrogative sentences, the plural suffix and plural verbs, question sentences starting with "what", possessive pronouns and possessive suffixes, simple present tense in positive, negative, and interrogative sentences, countable and uncountable nouns (using "a," "an," "some" and "any"), possessive expressions ("have got"/"has got") and gerunds.
- Basic English 2: the expressions "there is" and "there are," frequency adverbs ("always," "usually," "often," "sometimes," "never"), prepositions ("in," "on," "above," "in front of," "under," "opposite," "next to"), present continuous tense in positive, negative, and interrogative sentences, the usage of "can" and "cannot" to express ability, past simple tense (the verb "to be" in positive, negative, and interrogative sentences with regular and irregular verbs).

The general objective of these courses is "to improve students' beginner-level English language skills." The general and specific objectives of the courses were provided by the lecturers, and the courses were designed based on them.

Design. After completing the analysis phase, the designers conducted the design phase, which included designing learning activities, assessments, methodology, and environments. This phase primarily focused on determining the Basic English 1 and 2 learning activities and adapting them to new presentation formats. The course content was changed by the FLS during the study, which made it necessary to create a flexible and general design for use in different contexts. When designing the learning units, the designers decided to use the subject format instead of Moodle's default weekly format.

In context of the course, the program supported grammar, reading comprehension, writing, listening comprehension, and speaking skills. Accordingly, the designers decided to provide texts with basic grammar rules for each subject and added videos to help improve grammar skills. Additionally, the design process focused on reading passages and corresponding exercises to improve reading comprehension skills. It was assumed that providing writing exercises with the help of open-ended questions after the presentation of various sample cases would support the improvement of writing skills. Subject-relevant audio files were included to enhance listening comprehension skills. Moreover, quizzes were added to allow students to assess themselves at the end of each subject. These quizzes were not graded, but the students could see their results

and the correct responses to the questions. After completing all the subjects, the designers prepared a general assessment examination to evaluate the students.

The planned duration of the courses was 15 weeks. The assessments were scheduled to be conducted every three weeks with two mid-terms and a final examination. It was determined that the first week's lessons should be conducted in the classroom environment to allow the lecturer and students to meet and share information related to the course. A week before the final exam, lessons would again be taught in the classroom for review purposes. For the rest of the course's 10 weeks, the lessons were planned as distance education. Moreover, the content of all weeks is not shared at the beginning of the term but is unveiled week by week to prevent students from seeing all the content when they first enter the environment as this may prevent them from returning.

This design used Moodle, an open-source Learning Management System, as the web-based learning environment. Moodle is a free LMS software that is widely used in higher education. It is easy to use and offers a broad range of tools for teaching on the web, including file sharing, hyperlinks, blogs, wikis, examinations, and discussion forums.

Development. After the design phase, the designers conducted the development phase where all teaching materials and multimedia components were developed and the format was evaluated. All the tools planned for use in the design phase were uploaded to the Moodle environment (Figure 1).

The lessons were recorded and uploaded as videos to convey the content to students. Since the lessons were continuing during the project design process, only the lectures of the last three weeks (weeks 8, 9, and 10) could be recorded. A brief grammar text related to each subject was also provided at the beginning of each week, particularly for the weeks when no video recording was available.

Porter (1997) claims that it is necessary to create suitable interactive environments in this phase to encourage the learners to be creative and enthusiastic about their future studies. Accordingly, the study made use of forums to plan discussions and provide interaction. A variety of videos and games were also included in the environment to make learning more entertaining, to increase students' motivation and help the students who do not attend the classroom lessons to get involved in activities in the web-based environment.

When designing the environment, the designers sought to make the subject titles understandable by using short and clear expressions. All links were embedded and opened in the same window as a blank page to foster both safety and usability. Eventually, the designers concluded that the design was simple enough and not distracting. The content was also presented in a similar fashion for consistency.

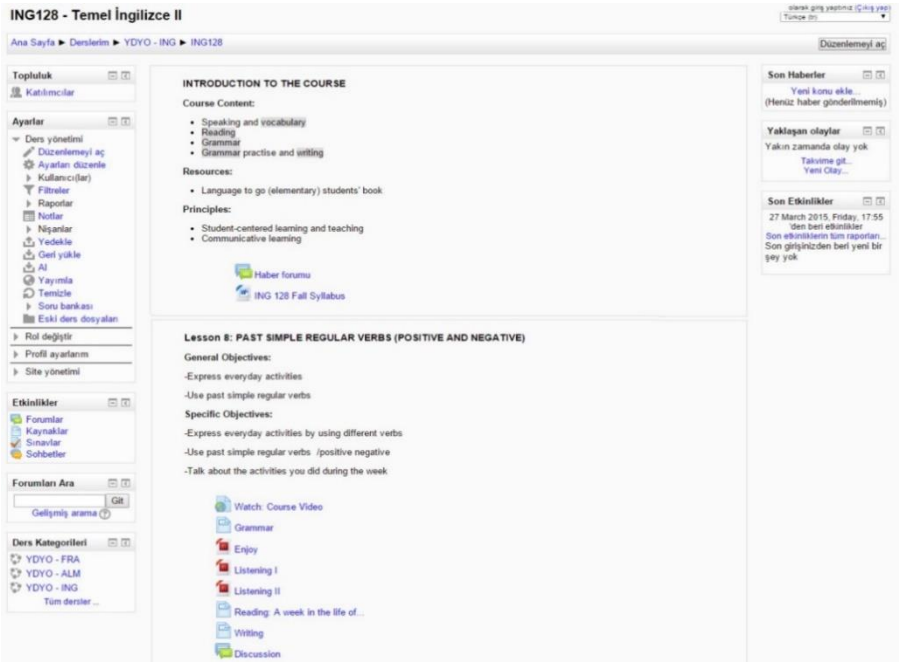


Figure 1. The Moodle homepage for Basic English 2

Implementation. After the teaching materials of Basic English 1 and 2 were transferred to the Moodle environment, a usability test was administered to three students who had previously attended these courses. Two of the students were male and one of them was female. The male participants were 19 and the female participant was 18 years old. The participants had not been involved in any web-based learning before. The participants had Internet access at their homes.

The aim of the usability test was to evaluate the environment from the users' point of view, develop an environment that is easy to use, increase users' satisfaction with the environment, and improve the environment based on their feedback (Rubin, 1994). Initially, the designers made plans for the usability test. Authentic tasks were created for the student testers, and the following were selected for the Moodle usability test:

1. Go to the address moodle.fls.university.edu.tr, and log into Basic English 1 with the username=XXX and the password= XXX.
2. Lesson 6 in Basic English 1: Answer the first three questions on the PRESENT SIMPLE (QUESTIONS AND NEGATIVES) quiz.
3. Lesson 8 in Basic English 1: Watch the lecture video for the subject HAVE/HAS GOT.

4. Lesson 7 in Basic English 1: Find the discussion topic A/AN, SOME/ANY, and participate in the discussion.
5. Lesson 9 in Basic English 2: Complete the writing task for PAST SIMPLE IRREGULAR VERBS.
6. Lesson 6 in Basic English 2: Listen to the audio in the listening section of the CAN FOR ABILITY subject.

The aim of the test was shared with the students at the beginning of the usability study. It was explained that the total duration of the test was 10 minutes and that they could leave the test if they wished. The students were also reminded that they were supposed to think aloud when performing the tasks in the environment. Thus, the participants made observations about the positive and negative aspects of the system as well as suggesting improvements. The designers noted the duration of each task. Their accomplishments of the tasks were graded from 1 to 5 (Table 1), 1 being not solved and 5 being solved.

The problems with the tasks were also noted. After the tasks were completed, the evaluation survey, which was created by the designers, was administered to obtain the users' positive and negative opinions about the system. The evaluation survey had items with scaled responses and open-ended questions. Table 2 presents these data.

The usability test produced remarkable data about the environment. The students had difficulty finding the button for comments in the forum in the fourth task (Basic English 1, Lesson 7: Find the discussion topic A/AN, SOME/ANY and participate in the discussion). To eliminate this problem, the "standard forum with general purpose" setting was replaced with "in-blog standard forum" on the forum settings. This made it easier for the students to find the button to write comments. One student opened the activity video instead of the lecture video in the third task (Basic English 1, Lesson 8: Watch the lecture video HAVE/HAS GOT). To solve this problem, the activity video, which was listed first, and the lecture video changed places, and the name of the lecture video was changed. The students stated that they had difficulty submitting their writing passages in the writing section. It was noted that there was no "Submit" button that students could use to send their writing passages, and the "Next" button, which had the same function, was perceived differently by the students. The "Next" button was renamed as "Submit" to address this problem.

Table 1

Usability Test Task Performance and User Experience Results

Task		***Task Duration***	***Accomplishment level of the tasks 1=Not solved 5=Solved***				
	User		***1***	***2***	***3***	***4***	***5***
	1	21 sec.					x
Task 1	2	28 sec.					x
	3	52 sec.				x	
	1	42 sec.			x		
Task 2	2	72 sec.		x			
	3	19 sec.					x
	1	12 sec.					x
Task 3	2	15 sec.					x
	3	9 sec.					x
	1	80 sec.			x		
Task 4	2	57 sec.		x			
	3	126 sec.		x			
	1	28 sec.			x		
Task 5	2	27 sec.			x		
	3	30 sec.			x		
	1	19 sec.					x
Task 6	2	9 sec.					x
	3	15 sec.					x

Table 2

Evaluation Survey and User Experience Results

	User	***Totally Disagree***	***Disagree***	***Neither Agree nor Disagree***	***Agree***	***Totally Agree***
1. I could use the Moodle System with no difficulties.	1			x		
	2			x		
	3			x		
2. The system helped me to learn the English lessons.	1		x			
	2		x			
	3				x	
3. The content of the system was understandable and suitable for me.	1		x			
	2			x		
	3				x	
4. The environment was sufficiently clear.	1			x		
	2	x				
	3			x		
5. I think the location of the menus in the system was suitable.	1				x	
	2		x			
	3		x			

Evaluation. In this final phase, the web-based learning environment created by the study was evaluated based on the feedback from the participants in the implementation phase, and the environment was altered to address their difficulties and suggestions. The participants used the environment in the usability test, which included authentic tasks. The problems, which designers did not recognize until the usability test, were identified by the students in this phase. At the end of the usability test, the designers reviewed and evaluated these problems, offered the most appropriate solutions for the content of the environment and prepared the environment for use.

Discussion and Conclusion

This study presented the design process of the Basic English 1 and 2 courses taught at the Foreign Languages School (FLS) of a large-scale public university in Ankara, Turkey. The study aimed to offer solutions to problems of classroom teaching including time, place, and number of lecturers. The designers conducted a planned and systematic design process based on the ADDIE instructional design model. At the end of the design process, the course content was transferred to the Moodle environment, and a usability test was administered to the participants to determine its usability. The designers revised the environment's design based on participants' feedback. Eventually, the study presented an environment that was suitable for blended learning educational activities that mainly used distance education.

In this study, blended learning was used in foreign language teaching. In foreign language teaching, blended learning provides an ideal and flexible learning environment that affects the learning process positively (Cahyono & Mutiaraningrum, 2016; Chan, 2014; Pinto-Llorente et al., 2016; Prescott et al., 2017; Wright, 2017). Krasnova and Ananjev (2015) point out that blended learning has many advantages over traditional foreign language learning methods. They state that blended learning provides self-learning, collaborative work, instant feedback, and interaction flexibility. Additionally, blended learning offers many motivational and meaningful learning opportunities for students (Rybushkina & Krasnova, 2015). Emelyanova and Voronina (2017) emphasize that students are mostly positive about blended learning and are willing to use a distance learning platform in foreign language lessons. Pinto-Llorente et al. (2017) noted that perceptions and attitudes towards blended learning tools (podcasts, videocasts, online tests, online dictionaries, and forums) have changed quite positively despite the fact that most of the students had not used them before. It is also stated that the blended learning environment is more efficient than the traditional learning approach. In this respect, Wright (2017) observed that the development of language skills was a significant difference in the language performance of the group in which the blended learning environment was used.

Studies have been carried out on the development of foreign language grammar, reading comprehension, writing, listening comprehension, and speaking skills, especially in the blended learning environment. Many studies in the literature of the field show that blended learning environments are effective in the development of these skills. For example, the study of Pinto-Llorente et al. (2017) shows that the blended learning model in Moodle contributes to the development of grammar skills (parts of speech, sentence types, and vocabulary). Similarly, Emelyanova and Voronina (2017) emphasized that the limited duration of faculty teaching provided a major obstacle to the development of a complex skill such as listening. On this topic, they emphasized that the use of blended learning in foreign language teaching improves the listening skills of the students by giving them the opportunity to reach the materials related to the listening activities at any time and place and to learn at their own pace.

Awan, Azher, Anwar, and Naz (2010) indicated that anxiety about learning a language can be a significant barrier to successful language acquisition. In this respect, it is very important that the anxiety factor not be ignored and that the students are not left alone to confront this problem. In this context, blended learning environments encourage learning autonomously by improving self-learning skills and increasing internal motivation (Liu, 2013). Hence, blended learning environments can be an effective learning tool for increasing students' beliefs in their abilities and helping teachers to raise awareness of the students' own potential.

At the beginning of the design period, the designers interviewed the instructors who worked in the FLS to determine user needs. While some of the instructors indicated that the course should be conducted entirely online, some have indicated that the distance education system should be used as support for classroom instruction. The instructors who expressed the opinion that the course should be conducted entirely online cited student success as a basis for this. Accordingly, the success of the students taking the course is very low and many students re-take these courses. Using blended learning environments that combine online and classroom instruction is important for student achievement, satisfaction, and participation. In this direction, a blended learning environment has been prepared. Similarly, Means, Toyama, Murphy, Bakia, and Jones (2009) found that teaching with a combination of online and classroom elements was more advantageous for learner achievement than teaching only in the classroom or only online.

The design team worked with the students and lecturers who would be using the system to take into account their needs and recommendations to shape the design. Since there were no subject-matter experts on the design team, the designers obtained support from the experts in the FLS. Additionally, designers sought expert opinions on distance education and human-computer interaction during the design process. Furthermore, participants were given the opportunity to use the environment with a

usability test where they could perform certain authentic tasks. The obtained data were used to revise the design. Similar studies (Dogan et al., 2011; Tuzun et al., 2011; Tuzun and Cinar, 2016) emphasize that co-operation between designers, experts, trainers, and students will save time and resources. Cagiltay (2011) and Bilgic and Tuzun (2015b) suggest that support should be provided about how instructional technologies can be used and designed, that various activities should be organized to make teachers and students use technological facilities effectively and productively, and that instructional technology support offices should be established, specifically in Turkey, to allow for new studies be conducted. Cagiltay (2011) also stresses that it is also important to analyze the studies conducted in foreign countries to attain these goals.

The designers in this research discussed the course content and format with the instructors at the beginning of the design process. The data obtained from the interviews were used in the design process. At the beginning of the design process, educational needs should be evaluated, and it should be determined whether they can be satisfied by distance education. Limitations including budget, time, and cost should clearly be identified at the beginning of the process. Tuzun and Cinar (2016) emphasized that conducting needs analysis and preliminary studies are important in terms of teaching effectiveness prior to designing a distance education course. Sales (2010) emphasized that it is very difficult to design an e-learning product that is at once cheap, fast, and high quality and that typically only two of these three criteria are satisfied. The popularization of LMSs, the creation of learning objects repositories, and the development of learning environments make it possible for instructors to prepare web-based courses using e-learning environments without having programming skills. Instructors use these tools to produce fast and cheap learning products. However, it has been documented that quality learning products cannot be created in these environments. This reduces students' enthusiasm for using them. Sales (2010) made the following suggestions about the process of designing a quality e-learning product: a) an instructional design model should be the basis, b) a systematic approach should be used, c) not only subject-matter experts, but also e-learning designers should be involved, d) instead of developing cheap and low quality content, quality content should be developed at a little more cost, and e) cost-benefit analyses should be conducted to ensure a balance of quality, cost, and effectiveness.

In this study, synchronous and asynchronous communication tools were used in order to provide teacher-student and student-student communication and encourage collaboration among students. In order to increase the effectiveness of teaching and get students to participate in learning, it is important that discussion panels, chat rooms, forums, and blogs be used and relevant activities be designed to help students communicate with teachers and other students and access the content and the environments to study cooperatively (Means et al., 2009). Challob, Bakar, and Latif (2016) emphasize that one of the most important factors for the success of blended

learning environments in foreign language teaching is the creation of a collaborative learning environment. Accordingly, in a blended learning environment, cooperative learning tools are used to help learners give feedback to each other's missing or incorrect information and, in so doing, feel more independent and confident. This also encourages students to share their knowledge and learning experiences.

It is important to stipulate that the system be improved according to feedback given after students/learners use it. The design process is iterative, and the product should be improved continuously. Accordingly, Tuzun et al. (2011) emphasized that the use of the newly-created environment should be monitored, opinions should be obtained from students and instructors about the process, and the course design should be conducted in an iterative way with continuous improvements.

This study has some limitations to consider when planning future research. One is the size of the sample group, and it is recommended that larger participant groups be used in future studies. Another limitation is the instructional design model used in the design process. This study followed the ADDIE model and future studies might utilize a different instructional design model and document its effectiveness. Also, the involvement of a subject-matter expert in the design team in future studies about course design in distance education will make it possible to carry out a more effective and rapid design process. The support offices that will be established will not only help teachers use the instructional technologies effectively but will also enhance their desire and willingness to participate in e-learning and make them and their students more inclined to use technologies in classroom activities. A project member who is a technical expert should also be included in the design team to reduce technical difficulties with the design process. Activities should be organized to increase students' communication with their instructors and other students. Discussion tools should be included in the design since students will not have the chance to meet each other and the instructor, especially if all the lessons are taught through distance education. Using tools that provide synchronous communication, such as video conferencing, in addition to asynchronous communication tools will increase students' communication with the instructor and their peers.

Acknowledgements

The authors would like to thank the students and the lecturers from the Foreign Languages School for participating in the study and the director of the Foreign Languages School for permitting the data collection in context.

References

Akkoyunlu, B., Altun, A., & Soylu, M. Y. (2008). *Ogretim tasarimi*. Ankara: Maya Akademi.

Alpala, C. A. O., & Florez, E. E. R. (2011). Blended learning in the teaching of English as a foreign language: An educational challenge. *HOW Journal, 18*(1), 154-168.

Al Zumor, A. W. Q., Al Refaai, I. K., Eddin, E. A. B., & Al-Rahman, F. H. A. (2013). EFL students' perceptions of a blended learning environment: Advantages, limitations and suggestions for improvement. *English Language Teaching*, *6*(10), 95.

Awan, R.-U. N., Azher, M., Anwar, M. N., & Naz, A. (2010). An investigation of foreign language classroom anxiety and its relationship with students' achievement. *Journal of College Teaching & Learning, 7*(11), 33-40.

Balci, B. (2010). E-ogrenme programi tasarim surecleri. In G. Telli-Yamamoto, U. Demiray, & M. Kesim (Ed.), *Turkiye'de e-ogrenme: Gelismeler ve uygulamalar* (pp. 83-110). Ankara, Cem Web Ofset.

Berger, C., & Kam, R. (1996). *Definitions of instructional design*. Retrieved from http://www.umich.edu/~ed626/define.html.

Bilgic, H. G., & Tuzun, H. (2015a). Web tabanli uzaktan egitim programlarinin yuksekogretim kurumlarinda acilmasi sureci. In B. Akkoyunlu, A. Isman, & H. F. Odabasi (Ed.), *Egitim teknolojileri okumalari 2015* (pp. 469-491). Ankara, Turkey: Ayrınti Basim Yayin.

Bilgic, H. G., & Tuzun, H. (2015b). Yuksekogretim kurumlari web tabanli uzaktan egitim programlarinda yasanan sorunlar. *Acikogretim Uygulamalari ve Arastirmalari Dergisi, 1*(3), 26-50.

Bueno-Alastuey, M. C., & Lopez Perez, M. V. (2014). Evaluation of a blended learning language course: Students' perceptions of appropriateness for the development of skills and language areas. *Computer Assisted Language Learning, 27*(6), 509-527.

Cagiltay, K. (2011). Universite ve ogretim teknolojileri destek ofisi: Mevcut durum ve Turkiye universiteleri icin gerekliligi. *Uluslararası Yuksekogretim Kongresi: Yeni yonelisler ve sorunlar (UYK-2011), İstanbul* (pp. 1597-1603).

Cahyono, B. Y., & Mutiaraningrum, I. (2016). Indonesian EFL teachers' familiarity with and opinion on the internet-based teaching of writing. *English Language Teaching, 9*(1), 199.

Challob, A. I., Bakar, N. A., & Latif, H. (2016). The influence of blended learning on EFL students' writing apprehension and writing performance: A qualitative case study. *European Journal of Multidisciplinary Studies, 1*(2), 253-265.

Chan, W. M. (2014). Video podcasting as a supplementary language learning tool. A study of its use, student access and perceptions. *Electronic Journal of Foreign Language Teaching, 11*(1), 183-206.

Cigdem, H., Ozturk, M., & Topcu, A. (2016). Vocational college students' acceptance of web-based summative listening comprehension test in an EFL course. *Computers in Human Behavior, 61*, 522-531.

Cinar, M., & Tuzun, H. (2016). Web-based course design models. In M. M. Pinheiro & D. Simões (Ed.), *Handbook of research on engaging digital natives in higher education settings* (pp. 374-395). Hershey, PA: Information Science Reference.

Creswell, J. W. (2007). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Dogan, D., Tuzun, H., Daghan, G., Altintas, A., Ilgaz, H., Ozdinc, F., Kayaduman, H., & Ozpala, N. (2011). Uzaktan egitimde ders tasarımı: Yuz yuze verilen bir dersin uzaktan egitim surecine hazir hale getirilmesi [Course design for distance education: Getting a class thought face to face ready for distance education]. *5th International Computer Education and Instructional Technologies Symposium Proceedings* (pp. 1037-1043).

Driscoll, M. (1998). *Web-based training: Using technology to design adult learning experience*. San Francisco, CA: Jossey-Bass Pfeiffer.

Emelyanova, N., & Voronina, E. (2017). Introducing blended learning in the English language classroom: Students' attitudes and perceptions before and after the course. *Knowledge Management & E-Learning, 9*(1), 33–49.

Eydelman, N. (2015). A blended English as a foreign language academic writing course. In B. Tomlinson & C. Whittaker (Ed.), *Blended learning in English language teaching: Course design and implementation* (pp. 43-62). London: England: British Council.

Ferriman, N. (2013). The impact of blended e-learning on undergraduate academic essay writing in English (L2). *Computers & Education, 60*(1), 243-253.

Garrison, R. D., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *Internet and Higher Education, 7*, 95-105.

Hilliard, A. T. (2015). Global blended learning practices for teaching and learning, leadership and professional development. *Journal of International Education Research, 11*, 179-187.

Hinkelman, D., & Gruba, P. (2012). Power within blended language learning programs in Japan. *Language Learning & Technology, 16*(2), 46-64.

Krasnova, T., & Ananjev, A. (2015). Students' perception of learning in the online discussion environment. *Mediterranean Journal of Social Sciences*, *6*(6 S1), 202-207.

Kudryashova, A. V., Gorbatova, T. N., & Rozhkova, N. E. (2016). Developing a blended learning based model for teaching foreign languages in engineering institutions. *SHS Web of Conferences, Tomsk, Russia* (pp. 1-4).

Limongelli, C., Sciarrone, F., & Vaste, G. (2011). Personalized e-learning in Moodle: The moodle LS system. *Journal of E-Learning and Knowledge Society, 7*(1), 49-58.

Liu, M. (2013). Blended learning in a university EFL writing course: Description and evaluation. *Journal of Language Teaching and Research*, *4*(2), 301-309.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. Washington, DC: US Department of Education, Office of Planning, Evaluation, and Policy Development.

Peachey, N. (2015). A blended learning teacher development course for the development of blended learning in English Language Teaching. In B. Tomlinson & C. Whittaker (Ed.), *Blended learning in English language teaching: Course design and implementation* (pp. 65-74). London: England: British Council.

Pinto-Llorente, A. M., Sanchez-Gomez, M. C., & Garcia-Penalvo, F. J. (2016). Assessing the effectiveness of interactive and collaborative resources to improve reading and writing in English. *International Journal of Human Capital and Information Technology Professionals (IJHCITP), 7*(1), 66-85.

Pinto-Llorente, A. M., Sanchez-Gomez, M. C., Garcia-Penalvo, F. J., & Casillas-Martin, S. (2017). Students' perceptions and attitudes towards asynchronous technological tools in blended-learning training to improve grammatical competence in English as a second language. *Computers in Human Behavior*, *72*, 632-643.

Porter, L. (1997). *Creating the virtual classroom: Distance learning with the internet.* Toronto: John Wiley & Sons.

Power, M. (2009). *A designer's log: Case studies in instructional design*. Athabasca, AB: Athabasca University Press.

Powell, A., Watson, J., Staley, P., Patrick, S., Horn, M., Fetzer, L., & Verma, S. (2015). *Blended learning: The evolution of online and face-to-face education from 2008-2015. Promising practices in online learning.* Vienna, VA: International Association for K-12 Online Learning.

Prescott, J. E., Bundschuh, K., Kazakoff, E. R., & Macaruso, P. (2017). Elementary school-wide implementation of a blended learning program for reading intervention. *The Journal of Educational Research*, 1-10.

Reigeluth, C. M., & Frick, T. W. (1999). Formative research: A methodology for creating and improving design theories. In C. M. Reigeluth (Ed.), *Instructional design theories and models* (Vol. 2, pp. 633-652). Mahwah, NJ: Lawrence-Erlbaum Associates.

Rubin, J. (1994). *The handbook of usability testing: How to plan, design and conduct effective tests*. New York, NY: John Wiley.

Rybushkina, S., & Krasnova, T. (2015). Key factors to use blended learning in teaching foreign languages in Russian engineering universities. EDULEARN_15 Proceedings, 6886-6892.

Sales, G. C. (2010). Can modern e-learning development follow a process that ensures quality? *Contemporary Educational Technology, 1*(1), 93-94.

Shaykina, O. I. (2015). Blended learning in English language teaching: Open educational resources used for academic purposes in Tomsk Polytechnic University. *Mediterranean Journal of Social Sciences, 6*(3), 255-260.

So, L., & Lee, C. H. (2013). A case study on the effects of an L2 writing instructional model for blended learning in higher education. *Turkish Online Journal of Educational Technology (TOJET)*, *12*(4), 1-10.

Sun, Z., & Qui, X. (2016). Developing a blended learning model in an EFL class. *International Journal of Continuing Engineering Education and Life Long Learning, 27*(1-2), 4-21.

Tomlinson, B., & Whittaker, C. (2013) *Blended learning in English language teaching: Course design and implementation*. London: British Council.

Tuzun, H. (2001). Guidelines for converting existing courses into web-based format. *Annual Proceedings of Selected Research and Development. Papers Presented at the National Convention of the Association for Educational Communications and Technology* (pp. 360-370).

Tuzun, H., Bilgic, H. G., Kalayci, E., Cinar, M., Akinci, A., Yildirim, D., & Yuksel, Y. (2011). Yerlesik bir dersin web-tabanli uzaktan egitim icin yeniden tasarimi. In B.B. Demirci, G. Telli-Yamamoto, & U. Demiray (Ed.), *Turkiye'de e-ogrenme: Gelismeler ve uygulamalar II* (pp. 175-199). Eskisehir, Turkey: Anadolu Universitesi Yayinlari.

Tuzun, H., & Cinar, M. (2016). Guidelines for transferring residential courses into web. *International Review of Research in Open and Distributed Learning, 17*(4), 145-165.

Ueda, H., & Nakamura, M. (2016). GakuNinMoodle: Toward robust e-learning services using Moodle in Japan. *Procedia Computer Science, 96*(2016), 1710-1719.

Wright, B. M. (2017). Blended learning: Student perception of face-to-face and online EFL Lessons. *Indonesian Journal of Applied Linguistics*, *7*(1), 64-71.

Yang, Y. T. C., Chuang, Y. C., Li, L. Y., & Tseng, S. S. (2013). A blended learning environment for individualized English listening and speaking integrating critical thinking. *Computers & Education, 63*, 285-305.

Yüz Yüze Verilen Bir Dersin Web Tabanlı Karma Öğrenmeye Uyarlanması

Atıf:

Ozmen, B., Tepe, T., & Tuzun, H. (2018). Adapting a residential course to web-based blended learning. *Eurasian Journal of Educational Research, 75*, 115-136, DOI: 10.14689/ejer.2018.75.7

Özet

Problem Durumu: Uzaktan eğitim yoluyla yürütülen bir dersin hedeflenen öğrenme çıktılarının elde edilmesi amacıyla tasarlanmasında, öğretim tasarımı modellerinden faydalanmak önem taşımaktadır. Bu kapsamda, öğretim tasarımcıları öğretim etkinliklerinin tasarlanması için var olan modelleri kullanmanın yanı sıra öğrenenlerin bireysel gereksinimlerini karşılamak, etkili ve verimli bir öğretim süreci düzenlemek amacıyla farklı modellere ihtiyaç duyabilmektedir. Yüz yüze verilen yabancı dil derslerinin karma öğrenme ortamı kullanılarak web tabanlı formata dönüştürülmesinde tasarım sürecine ilişkin çalışmaların azlığı göze çarpmaktadır.

Araştırmanın Amacı: Bu çalışmada Ankara'da bir Devlet Üniversitesinin Yabancı Diller Yüksek Okulu'nda (YDYO) yürütülmekte olan Temel İngilizce I ve Temel İngilizce II derslerinin ADDIE öğretim tasarımı modeli kullanılarak uzaktan eğitim ile verilebilecek biçime dönüştürülmesi ve web tabanlı ortama aktarılması amaçlanmıştır. Web tabanlı karma öğrenme ortamı, halihazırda yüz yüze sınıf ortamında yürütülmekte olan Temel İngilizce I ve Temel İngilizce II derslerinin içeriklerini kapsayacak şekilde ve sınıf içi uygulamaların çevrimiçi öğrenme ortamıyla birleştirildiği karma öğrenme yaklaşımına uygun bir biçimde tasarlanmıştır.

Araştırmanın Yöntemi: Bu çalışma bir tasarım çabasını betimleyen durum çalışması olarak desenlenmiştir. Çalışmada, bir Yabancı Diller Yüksek Okulu'nda yüz yüze yürütülmekte olan Temel İngilizce I ve Temel İngilizce II derslerinin uzaktan eğitim yoluyla verilebilecek biçimde dönüştürülerek web tabanlı ortama aktarılması gerçekleştirilmiştir. Bu doğrultuda, çalışmanın tasarım sürecinde ADDIE öğretim tasarımı modeli temel alınmıştır. Çalışma kapsamında, ADDIE modeline uygun olarak analiz aşamasında, ihtiyaç ve sınırlılıkları belirlemek amacıyla YDYO'da görev yapan öğretim elemanlarıyla görüşmeler yapılmıştır. Bu görüşmeler ile söz konusu dersleri alan öğrencilerin sahip olduğu bilgi ve beceriler, öğretim elemanlarının ve öğrencilerin beklentileri, dersin amaçlarına ilişkin bilgiler elde edilmiştir. Tasarım aşamasında, Temel İngilizce I ve Temel İngilizce II derslerine ilişkin öğrenme etkinliklerinin belirlenmesi ve bunların sunulma biçimleri üzerinde yoğunlaşılmıştır. Her iki dersin içeriği 15'er hafta olacak şekilde dilbilgisi, okuma-anlama, yazma, dinleme-anlama ve konuşma becerilerini kapsayacak şekilde tasarlanmıştır. Geliştirme aşamasında

tasarım aşamasında kullanılması planlanan tüm araçlar Moodle ortamına eklenmiştir. Bazı haftalarda video çekimleri yapılarak derslerin haftalık içerikleri video izleme, dilbilgisi konu anlatımı, okuma, dinleme, oyun, kısa testler ve tartışma forumları konularını kapsayacak şekilde sıralanmıştır. Uygulama aşamasında öğrencilerle kullanılabilirlik testleri yapılmış ve tasarım sürecindeki eksiklikler belirlenmiştir. Tasarım sürecinin son basamağı olan değerlendirme basamağında ise kullanılabilirlik testi sonucu öğrencilerden gelen dönütlere yönelik içeriklerde iyileştirmelere gidilmiştir.

Araştırmanın Bulguları: Kullanılabilirlik testi çalışması sonucunda oluşturulan ortama ilişkin önemli veriler elde edilmiştir. Dördüncü görevde (Temel İngilizce I dersinde Lesson 7: A/AN, SOME/ANY tartışma konusunu bulup katılım sağlayınız) öğrencilerin foruma yorum yazma butonunu bulmakta sorun yaşadıkları gözlemlenmiştir. Bu sorunun giderilmesi için forumun ayarlarından, var olan ayar "genel amaçlı standart forum" değiştirilerek "blog içinde standart forum" yapılmıştır. Bu sayede öğrencilerin yorum yazma butonunu bulması daha kolay hale getirilmiştir. Bir öğrencinin üçüncü görevde (Temel İngilizce I dersinde Lesson 8: HAVE/HAS GOT ders videosunu izleyiniz) ders videosunu açmak yerine etkinlik videosunu açtığı gözlemlenmiştir. Bunun çözümü olarak ilk sırada yer alan etkinlik videosu ile ders videosu yer değiştirilerek etkinlik videosunun ismi değiştirilmiştir. Öğrenciler yazma (writing) bölümünde yazdıklarını göndermede sorun yaşadıklarını belirtmişlerdir. Öğrencilerin yazdıklarını gönderirken "gönder" butonunun olmadığı, onun işlevi olarak "sonraki" butonunun olmasının öğrencilerde farklı bir algı oluşturduğu anlaşılmıştır. "Sonraki" butonu "gönder" olarak isimlendirilmiştir.

Araştırmanın Sonuçları ve Önerileri: Tasarım sürecinin başında öncelikle var olan durum ve ihtiyaçlar göz önünde bulundurularak nelere ihtiyaç duyulduğuna ve söz konusu eğitim ihtiyacının uzaktan eğitim ile giderilip giderilemeyeceğine karar verilmesi önem taşımaktadır. Bu kapsamda, sürecin başında bütçe, zaman ve maliyet gibi sınırlılıkların net bir biçimde belirlenmesi gerekmektedir. ÖYS'lerin yaygınlaşması, öğrenme nesnesi depolarının oluşturulması ve öğrenme çevrelerinin geliştirilmesi ile programlama becerisi gerekmeden öğretim elemanlarının e-öğrenme ortamlarını kullanarak web tabanlı dersler hazırlayabilmelerine olanak sağlamaktadır. Öğretim elemanları bu araçları hızlı ve ucuz öğrenme ürünleri üretmede kullanmaktadır. Ancak, bu ortamlarda kaliteli öğrenme ürünlerinin ortaya konulamadığı göze çarpmaktadır. Bu durum ise, öğrencilerin söz konusu ortamları kullanma konusundaki isteklerini azaltmaktadır. Tasarlanacak ortamın kullanımıyla birlikte alınacak dönütler ile sistemin iyileştirilmesi ve tasarım sürecinin aslında birikimli bir süreç olduğu ve ortaya çıkacak ürünün sürekli geliştirilmesi gerektiğinin kabul edilmesi önem arz etmektedir. İleride kurulacak destek ofisleri, hem öğretim teknolojilerinin etkin kullanımının sağlanmasında öğretmenlere katkı sağlayabilecek hem de öğretmenlerin bu konudaki gönüllülük ve istekliliklerini arttırarak

öğretmenlerin ve dolayısıyla öğrencilerin bu teknolojileri sınıf içi etkinliklerde kullanımlarına ilişkin yatkınlıklarını arttırabilecektir. Ayrıca tasarım sürecinde yaşanabilecek teknik aksaklıkları azaltmak amacıyla tasarım ekibine bir de teknik açıdan donanımlı bir proje üyesi dahil edilmesi önerilmektedir. Öğrencilerin öğretim elemanı ve diğer arkadaşlarıyla iletişimini arttıracak etkinlikler düzenlenmesi gerekmektedir. Özellikle derslerin tümü uzaktan eğitim yoluyla verildiğinde öğrencilerin birbirlerini ve öğretim elemanını tanıma fırsatı olmayabileceğinden tartışma ortamlarının tasarımda bulundurulmasına özen gösterilmelidir. Bu kapsamda, eş zamansız iletişimin sağlanabileceği araçların yanı sıra video konferans araçları gibi eş zamanlı iletişimin sağlanacağı araçların kullanılması da öğrencilerin öğretim elemanı ve akranlarıyla iletişimini arttırabilir.

Anahtar kelimeler: ADDIE öğretim tasarımı modeli, uzaktan eğitim, yabancı dil öğrenme, öğretim tasarımı, öğrenme yönetim sistemleri.

Eurasian Journal of Educational Research 75 (2018) 137-154

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

The Relationship between Psychological Capital and Stress, Anxiety, Burnout, Job Satisfaction, and Job Involvement

Selcuk DEMIR [1]

ARTICLE INFO

Article History:
Received: 02 Sep. 2016
Received in revised form: 21 Dec. 2017
Accepted: 03 Feb. 2018
DOI: 10.14689/ejer.2018.75.8

Keywords
Individual outputs
teacher attitudes
productivity
mediation effect

ABSTRACT

Purpose: It is extremely important for teachers to have the psychological capital to enhance their performance outputs in educational organizations. Teachers having psychological capital will positively influence all the members of the school. This kind of teacher will develop and support colleagues' and students' positive attitudes such as competence, hope, motivation, and productivity. The purpose of this research is to determine the relationship between psychological capital and stress, anxiety, burnout, job satisfaction, and job involvement. **Research Methods:** The sample of this study consists of 335 teachers in 27 schools that were selected randomly from the schools at Hatay's Kırıkhan district in the 2014–2015 academic year. The Psychological Capital Scale, Stress Scale, Anxiety Scale, Burnout Scale, Job Satisfaction Scale, and Job Involvement Scale were used to collect the data. **Findings:** As teachers' psychological capitals increase, their stress levels decrease. Psychological capital has a negative impact on anxiety levels of teachers through the full mediation effect of stress. Psychological capital negatively affects teachers' burnout levels through the partial mediation effects of stress and anxiety. Psychological capital has a positive effect on teachers' job satisfaction through the partial mediation effects of stress, anxiety, and burnout. Psychological capital has a positive effect on the job involvement levels of teachers through the partial mediation effects of stress, anxiety, burnout, and job satisfaction. **Implications for Research and Practice:** It is essential for school administrators to contribute to the development and strengthening of teachers' psychological capital if they wish for teachers to feel negative attitudes less frequently and to develop more positive attitudes at work.

[1] Gaziantep University, Institute of Educational Sciences, Educational Administration, Supervision, Planning and Economy TURKEY, e-mail: selcuk_demirs@hotmail.com https://orcid.org/0000-0003-2904-6443

Introduction

It is crucial for teachers to have positive attitudes to enhance their performance outputs in educational organizations. Previous studies revealed that psychological capital influences teachers' attitudes and behaviors positively and prevents emotional distress. This kind of teacher will help colleagues and students to develop positive attitudes such as competence, hope, motivation, and productivity. This research is designed to clarify how psychological capital increases terminal attitudes and behaviors and decreases un-solicited attitudes and behaviors in educational organizations.

Teachers are expected to experience higher levels of work-related stress than workers in white collar professions (Sorenson, 2007). There have problems caused by teacher-student conflict, teacher-family conflict, student discipline problems, insufficient support, lack of promotion, disgruntled colleagues, and insufficient participation in school decision-making (Cokluk, 2014; Hock, 1988; Sorenson, 2007). These problems cause teachers to experience negative effects such as stress, anxiety, job dissatisfaction, intent to leave, and physical and psychological distress (Bradley, 2014; Cokluk, 2014). These discomforts suffered by teachers are supposed to reflect on students so the problems increase more and more in the education system.

A large number of studies have revealed a relationship between the psychological capital of employees and terminal and undesirable outcomes. Recent publications have highlighted the importance of psychological capital level on employees' work-related behaviors and attitudes (Bradley, 2014; Burke, 2000; Karakus & Demir, 2015; Luthans, Avolio, Avey, & Norman, 2007; Ocak, Guler, & Basim, 2016; Schulz, Luthans, & Messersmith, 2014; Tuzun, Cetin, & Basim, 2014; Wang, Chang, Fu, & Wang, 2012). Psychological capital influences teachers' attitudes positively and prevents distressed emotions (Karakus & Demir, 2015; Ocak et al., 2016; Ocak & Guler, 2017). Also, teachers with psychological capital will positively affect all the members of the school. This kind of teacher will develop and support colleagues' and students' positive attitudes such as competence, hope, motivation and productivity. Despite the growing importance and interest in the effects of psychological capital on various individual constructs, there is a shortage of multivariate studies on the subject, especially in relation to educational organizations. This study has been done to develop our knowledge of the concept of psychological capital.

M. Seligman and his colleagues have performed many studies to make the ignored functions of psychological capital a current issue. These functions are important: to heal human life and to develop superior abilities (Polatci, 2014). Luthans and his colleagues started to carry out related studies in the field of organizational behavior. Their research points out that psychological capital is an important approach to understand a human's strengths and capacities, to support personal development, and to have a positive impact on motivations. In today's workplace, adopting a positive approach to develop and manage human resources helps us to have positive outcomes at individual and organization levels. According to this point of view, psychological capital as a psychological resource is a necessary

force needed to develop and to manage all organizational settlements (Luthans et al., 2007).

This study is designed to clarify how psychological capital increases terminal attitudes and behaviors and decreases unsolicited attitudes and behaviors. It is crucial for teachers to have positive emotions and attitudes in educational organizations. If teachers have positive emotions and attitudes, they will communicate effectively with each other and with all the internal and external members of the school (Demir, 2018). Since emotions are contagious (Robbins & Judge, 2012), all the individuals (internal and external) of a school could be affected by teachers who have positive or negative work-related attitudes. Teachers are the most important aspects of students' learning environment (Celik, 2005).

Conceptual Framework

Psychological capital. Luthans (2002), using a positive approach at a micro level, points out that employees' strengths and positive capacities are not fixed and can be developed. Positive psychology, called "positive organizational behavior", focuses on employees' strengths instead of their weaknesses. Psychological capital can be expressed as an individual's positive psychological state of development. Psychological capital is a psychological resource for increasing individuals' performance and making organizations more successful (Luthans & Youssef, 2004).

Psychological capital has four dimensions that can be measured and developed: hope, optimism, resilience, and self-efficacy (Luthans, 2002). Hope is the motivational state to persevere towards desired goals (to attain the desired goals); optimism is making a positive attribution to success; resilience is the capacity, when being unsuccessful, to sustain and bounce back to achieve success; self-efficacy is having the confidence to succeed in a challenging task (Luthans et al., 2007).

Stress. Stress is a concept that can hardly be defined. According to Genc (2012), stress is a state of psychological tension, affecting individuals physiologically and psychologically. It also affects people's behaviors, work outcomes, and relationships with other people in a negative way. Simsek and Celik (2012) define stress as a physiological and emotional reaction to potential threats in the environment.

Teachers always face pressure from work overloads, through high stakes testing, demanding administrators, disgruntled colleagues, and irate parents. They have to complete tasks and avoid errors (Sorenson, 2007). Teachers are supposed to cope with problems, overcome conflict, set goals, manage time, make ethical decisions, and deal with organizational politics in addition to directing in one or more subject matter (Bradley, 2014). Having too much workload promotes teacher's work-related stress levels (Cokluk, 2014; Hock, 1988). Stress is also related to undesirable emotions like anxiety, burnout, depression, loneliness, fear, boredom, and tension (Simsek & Celik, 2012). A number of studies have explored how work-related stress has negative effects. Because of these effects, stress causes negative outcomes and negative work experiences for people (Bradley, 2014; Cokluk, 2014).

Anxiety. Anxiety is an ongoing dominant psychological state causing an ambiguous fear or a feeling that something bad will be experienced. An individual having a high anxiety level can experience shortness of breath, flying into a temper, tension, sweltering, stomache ache, and heart throbbing (Simsek & Celik, 2012). Thus, a high level of anxiety causes undesired physical and psychological situations.

The degree and severity of anxiety is mostly related to the intensity of stress. Anxiety causes insomnia and attention failure. It is one of the most undesired reactions of stress (Altintas, 2016).

Burnout. Burnout is defined as … failure, exhaustion, energy, and loss of power or an exhausted state emerging from desires that are not met (Jones, 1981). As mentioned in Maslach and Jackson (1981), burnout is a syndrome including emotional exhaustion, depersonalization, and decreased individual accomplishment. Burnout refers to physical, mental, and emotional exhaustion and develops silently (Hock, 1988). According to the generally accepted tendency in the literature, burnout is a psychological process resulting from chronic job stress. An alienation from work and duty occurs in this process, and this state has been identified as an occupational danger by Freudenberger (Dolgun, 2012).

Burnout has been common in the sectors involving face-to-face communication such as health, education, and security (Maslach & Jackson, 1981). Burnout causes various negative results, such as turnover, intent to leave, and dissatisfaction from the perspective of the individual. However, it decreases economic and service quality by causing the loss of skilled employees in an organization (Cokluk, 2014).

The negative sides of teaching such as disciplinary problems, student apathy, work overload, inadequate salaries, demanding or unsupportive parents, and lack of administrative support are among the stressors that affect teachers (Sorenson, 2007). As a consequence of these stressful sides of teaching, burnout occurs among teachers; that is to say, it causes physical problems such as headaches and peptic ulcers; psychological problems such as depression and anger; and behavioral problems such as deterioration in work performance and absenteeism (Russell, Altmaier, & Van Velzen, 1987). When teachers experience burnout in schools, they tend to have undesired attitudes and behaviors for them and their organizations (Kyriacou, 2000).

Job satisfaction. As Locke (1976) indicated, job satisfaction is a positive psychological state that an individual gets from his / her job or job experience. It is also related to how an individual likes her / his job. Employees having high job satisfaction think about their jobs in a positive way. But employees having low job satisfaction have negative emotions about their jobs (Robbins & Judge, 2012). Job satisfaction can be expressed as an affective state resulting from the evaluation of an individual's personal work experience. Therefore, it can only be observed through individual's behaviors. Job satisfaction happens when employees struggle to get the rewards that they value or exceed their own previous accomplishments (Kaplan & Bickes, 2013).

Job involvement. Job involvement has gained importance lately because of the economic crisis, globalization, and increased competition. In spite of this fact, there is no consensus on either the explanation or the operationalization of job involvement (Yuksel & Akdag, 2011). Individuals who have job involvement are addicted to their jobs. These individuals see their job as the center of their lives (Haymon, 1992). Job involvement has positive and negative associations in the related literature. Some studies viewed job involvement as a positive attitude (Burke, 2000), but most of the research viewed job involvement as a negative behavior (Burke, Burgess, & Fallon, 2006; Haymon, 1992; Potter, 1996).

The Relationship between Psychological Capital and Stress, Anxiety, Job Satisfaction, and Job Involvement

Psychological capital increases positive attitudes and decreases negative attitudes in organizations (Tuzun et al., 2014; Karakus & Demir, 2015). Previous studies revealed that psychological capital has developed individuals' positive attitudes more easily and also has prevented individuals from developing negative attitudes. Karakus and Demir (2015) found that there is a positive relationship between psychological capital and positive attitudes such as job satisfaction, organizational commitment, and motivation. They also found that psychological capital has a negative effect on teachers' intent to leave through the full mediation effect of organizational commitment. Related research found a negative relationship between psychological capital and negative attitudes like stress (Bradley, 2014), anxiety (Liu et al., 2013), and burnout (Ocak & Guler, 2017; Wang et al., 2012); and a positive relationship between psychological capital and positive attitudes like job satisfaction (Akcay, 2012; Erkus & Findikli, 2013; Cetin & Basim, 2011; Kaplan & Bickes, 2013; Karakus & Demir, 2015; Luthans et al., 2007; Luthans & Youssef, 2004; Ocak et al., 2016; Schulz et al., 2014) and job involvement (Burke, 2000; Yuksel & Akdag, 2011).

Purpose of the Study

It is necessary to investigate negative and positive attitudes with psychological capital in a single study to obtain more conceptual knowledge. The purpose of this research is to determine the relationship between psychological capital and stress, anxiety, burnout, job satisfaction, and job involvement with mediation effects. This study also aims at constructing an integrated framework to reveal the impacts of psychological capital on these positive and negative outputs.

Method

Research Design

This study used a survey-based correlational design that explains the relationships between two or more variables. This design is also important for giving an idea about the cause-and-effect relation between variables (Robson, 2015). After psychological capital, the stress, anxiety, burnout, job satisfaction, and job

involvement levels of the teachers were determined through surveys and the relationships among these variables were determined.

Research Sample

The population of the study is comprised of the 1,145 teachers that work at Hatay's Kırıkhan district during the 2014–2015 academic year. Using the cluster sampling method, a total of 27 schools were selected randomly and the scales were administered to all the teachers at these schools. A sample of 335 teachers agreed to participate in this study. According to Field (2009), this sample size is enough, at a 95% confidence interval for this population. The teachers' response rate from the selected schools was approximately 90%.

While 60.6% of the teachers participating in this study were male (*n*=203), 39.4% were female (*n*=132). Of the participants, 78.2% were married (*n*=262) and 21.8% were single (*n*=73). The most frequent age range of the participants is 33–44 years (*n*=144), with a percentage of 43%. The most frequent tenure range of the participants is between six to ten years (*n*=93) with a percentage of 27.8%.

Research Instruments and Procedure

Data for this study was obtained using a five-point Likert-type scale with options differentiated as "1: I strongly disagree," "2: I disagree," "3: I partially agree," "4: I agree," and "5: I strongly agree." Study variables were measured with six scales as follows: The *Organizational Psychological Capital Scale* was developed by Luthans et al. (2007) and adapted to Turkish by Cetin and Basim (2012). The *Stress* scale was developed by Karakus (2013). The *Anxiety* scale was among the scales of International Personality Item Pool (Goldberg, 1999; Goldberg et al., 2006) and adapted to Turkish by Karakus (2013). The *Burnout Syndrome Inventory Short Version* was prepared by Pines (2005) and adapted to Turkish by Tumkaya, Cam, and Cavusoglu (2009). To measure *Job Satisfaction*, a global measure of job satisfaction was used rather than a facet measure. A five-factor measure that was developed by Griffin, Hogan, Lambert, Tucker-Gail, and Baker (2010) was adapted to Turkish by Karakus and Demir (2015). The *Job Involvement* scale was developed by Griffin et al. (2010) and adapted to Turkish by the researcher.

Validity and Reliability

The Psychological capital scale has four dimensions: hope, self-efficacy, optimism and resilience. The four factorial structure presented a good fit to the data (χ^2 = 68.22, df = 40, P = 0.00036, RMSEA= 0.041, GFI = 0.98, AGFI = 0.96, CFI = 0.99, IFI= 0.98). Cronbach's alpha of the scale was 0.82. The Stress scale is a single-factor scale consisting of four items that presented a good fit to the data (χ^2 = 0.08, df = 2, P-value = 0.8796, RMSEA = 0.000, GFI = 1.00, AGFI = 1.00, CFI = 1.00, IFI = 1.00). Cronbach's alpha of the scale was 0.80. The Anxiety scale is a single-factor scale consisting of five items that fit the data well (χ^2 = 0.07, df = 2, P-value = 0.9123, RMSEA = 0.000, GFI = 1.00, AGFI = 1.00, CFI = 1.00, IFI = 1.00). Cronbach's alpha of the scale was 0.91. The Burnout scale is a single-factor scale that presented a good fit to the data (χ^2 = 22.34, df = 14, P-value = 0.00045, RMSEA = 0.032, GFI = 0.98, AGFI = 0.99, CFI = 0.99, IFI =

0.97). Cronbach's alpha of the scale was 0.88. The Job satisfaction scale is a single-factor scale containing five items that fit the data well (χ^2 = 0.12, df = 4, P-value = 0.7863, RMSEA = 0.000, GFI = 1.00, AGFI = 1.00, CFI = 1.00, IFI = 1.00). Cronbach's alpha of the scale was 0.79. The Job involvement scale is a single-factor scale consisting of three items that presented a good fit to the data (χ^2 = 0.06, df = 1, P-value = 0.9842, RMSEA = 0.000, GFI = 1.00, AGFI = 1.00, CFI = 1.00, IFI = 1.00). Cronbach's alpha of the scale was 0.91.

Data Analysis

After collecting the data, the missing values were eliminated, and skewness and kurtosis coefficients were checked. The data were normally distributed. The sums of these scales were taken, and path analysis was performed with the maximum likelihood method through AMOS. The root mean square error of approximation (RMSEA), goodness of fit index (GFI), adjusted goodness of fit index (AGFI), comparative fit index (CFI), incremental fit index (IFI), and chi-square/degrees of freedom (X2/sd =CMIN/DF) and level of significance (p) fit indexes were considered in the evaluation of the model for goodness of fit. The RMSEA value being between 0.05 and 0.08 revealed acceptable fit indexes. The X2/sd (CMIN/DF) value was between 0 and 3; the values of GFI, AGFI, CFI and IFI were between 0.95 and 1.00 and revealed good fit indexes (Byrne, 2010; Kline, 2011).

Results

Descriptive statistics and correlation matrix of the variables in the study are given in Table 1.

Table 1

Descriptive Statistics and Correlation Matrix of the Variables in the Study

Variables	$\overline{X}$	*Sd*	1	2	3	4	5	6
1. PsyCap	3.90	.43	1					
2. Stress	2.82	.93	-.20**	1				
3. Anxiety	3.02	1.02	-.17**	.37**	1			
4. Burnout	2.53	.88	-.28**	.52**	.31**	1		
5. JSat	4.15	.69	.41**	-.17**	-.06	-.23**	1	
6. JInv	3.40	.94	.29**	-.20**	-.01	-.26**	.30**	1

**p<.01, Notes: PsyCap: Psychological capital, JSat: Job satisfaction, JInv: Job involvement

According to a five-point Likert type scale, Table 1 shows descriptive statistics of the teachers who participated in this paper. The findings show that the psychological capital and job satisfaction levels of teachers are moderately high (4). Also, their stress, anxiety, burnout, and job involvement levels are at a moderate level (3).

According to the correlation matrix, psychological capital (sum) is positively correlated with job satisfaction and job involvement and negatively correlated with stress, anxiety, and burnout. Stress, anxiety and burnout are positively correlated

with each other. The relationship between anxiety and job satisfaction and job involvement is not significant. Both stress and burnout are negatively correlated with job satisfaction and job involvement. There is a positive correlation between job satisfaction and job involvement.

Structural Equation Model

The final structural equation model with standardized path coefficients is given in Figure 1. Because of their insignificant path coefficients, the paths between psychological capital → anxiety (β = -0.97, p = 0.058), anxiety → job satisfaction (β = 0.060, p = 0.263), stress → job satisfaction (β = -0.054, p = 0.368), and stress → job involvement (β = -0.105, p = 0.083) were deleted from the model. According to the final model, which yields the best fit indices, as teachers' psychological capital levels increase, their stress levels decrease. Psychological capital negatively affects the anxiety levels of teachers through the full mediation effect of stress. Psychological capital has a negative impact on teachers' burnout levels through the partial mediation effects of stress and anxiety. Psychological capital has a positive effect on teachers' job satisfactions through the partial mediation effects of stress, anxiety, and burnout. Psychological capital has a positive effect on the job involvement levels of teachers through the partial mediation effects of stress, anxiety, burnout, and job satisfaction.

The final structural equation model with standardized path coefficients is given in Figure 1.

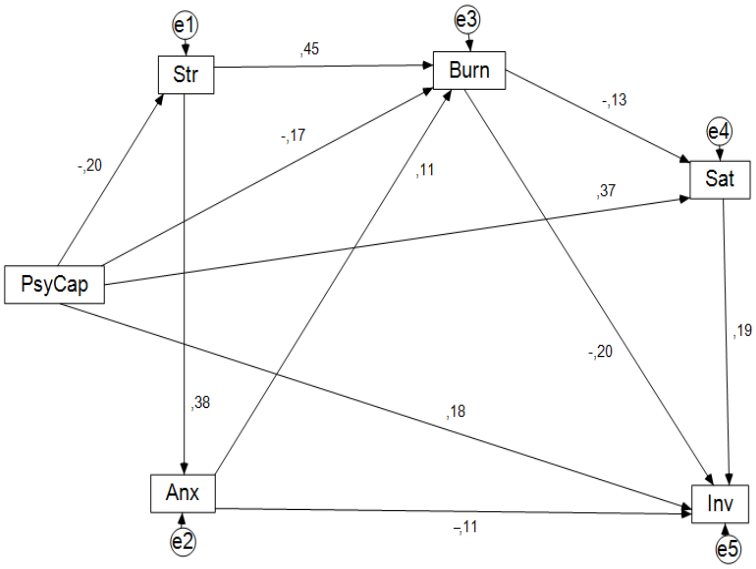


Figure 1. *The final structural equation model with standardized path coefficients*

Notes: PsyCap: Psychological capital, Str: Stress, Anx: Anxiety, Burn: Burnout, Sat: Job satisfaction, Inv: Job involvement. Fit indices: $x^2 = 8.203$, sd = 4, $x^2/sd = 2.051$, GFI = .992, AGFI = .957, CFI = .987, IFI = .987, RMSEA = .056.

Discussion and Conclusion

This research tried to reveal the effect of psychological capital on teachers' attitudes. The researcher reasoned that if teachers' psychological capital affected their stress, anxiety, burnout, job satisfaction, and job involvement, then this would provide support regarding the importance of investing in psychological capital. According to the structural model, teachers' psychological capital reduces their stress, anxiety, and burnout and enhances their level of job satisfaction and job involvement. A great deal is known about the positive effects of psychological capital and the ways in which psychological capital contributes to positive attitudes and personal outcomes for individuals. The literature review and subsequent data analysis for this paper attempts to examine psychological capital in terms of desired and undesired attitudes in education.

The findings of this research showed that psychological capital has a negative effect on stress. Similarly, Bradley (2014) revealed that teachers attribute positive meanings to their stressful work experiences thanks to their positive psychological capital capacities (hope, self-efficacy, resilience, and optimism). It has been shown that teachers, under pressure, can get positive work outputs by using stress with the help of their positive psychological capital competence. Thus, Bradley (2014) has revealed the concept of positive stress on the basis of positive psychological capital.

This study shows that psychological capital has a positive effect on teachers' job satisfaction through the partial mediation effects of stress, anxiety, and burnout. Similarly, other researchers found that there is a positive relationship between psychological capital and job satisfaction (Akcay, 2012; Erkus & Findikli, 2013; Cetin & Basim, 2011; Kaplan & Bickes, 2013; Karakus & Demir, 2015; Luthans et al., 2007; Luthans & Youssef, 2004; Schulz et al., 2014). Ocak et al. (2016) revealed that only the optimism factor of psychological capital has positive effects on organizational commitment and job satisfaction. Also, Larson and Luthans (2006) found that there is a positive relationship between job satisfaction and the hope and psychological resilience dimensions of psychological capital. Besides, Schulz et al. (2014) indicated that there is a positive relationship between psychological capital and the intent to leave through the mediation effects of job satisfaction and organizational commitment. Unlike these studies, it has been discovered in this paper that psychological capital decreases the stress, anxiety, and burnout levels of teachers while it increases their job satisfaction.

Cokluk (2014) states that being exposed to stressful environments continuously and dramatically decreases teachers' physical and emotional resources, and this prevents an individual to cope with stress. Also, Kyriacou (2000) asserts that an individual having trouble to cope with stress constantly experiences disappointment, tension, and anxiety, or stress turns into the burnout syndromes since it goes on increasingly. Russell et al. (1987) found that stressful events increase the burnout levels of teachers.

In this study, stress has a positive effect on burnout, and burnout has a negative effect on job satisfaction. Moreover, psychological capital has a positive effect on job satisfaction with the partial mediation effects of stress and burnout. Similar findings in the literature support the findings of this study. Cokluk (2014) found that burnout in the teaching occupation is related to the variables of stress and job satisfaction. Kyriacou (2000) revealed that teachers being exposed to stress continuously feel burnout. Arslantas and Aslan (2015) have found that the more that primary and secondary school teachers have high levels of burnout, the less job satisfaction they have. Hock (1988) revealed a negative relationship between the burnout levels of teachers and their stress and dissatisfaction.

Findings show that teachers' psychological capital has a negative effect on their burnout levels. Wang et al. (2012) found that psychological capital is a positive psychological resource to overcome burnout and that psychological capital is a mediator in the relationship between work-family conflict and burnout. Ocak and Guler (2017) revealed that optimism dimension of psychological capital negatively influences feelings of inefficacy and resilience dimension of psychological capital negatively influences the depersonalization dimension of burnout. This study also shows that psychological capital has a negative effect on the anxiety levels of teachers through the full mediation effect of stress. Similarly, Liu et al. (2013) found that psychological capital decreases the symptoms of anxiety and depression.

The findings of this study reveal that psychological capital has a positive effect on the job involvement levels of teachers through the partial mediation effects of stress, anxiety, burnout, and job satisfaction. The psychological capital of individuals makes them more involved and engaged in their work thanks to their positive attitudes at work. Corroboratively with the findings of this study, Simons and Buitendach (2013) showed a positive relationship between psychological capital and work engagement. On the other hand, De Waal and Pienaar (2013) found that psychological capital does not predict work engagement.

Teachers with higher levels of psychological capital look positively to the events around them and are less affected by negative experiences. Teachers having higher levels of psychological capital have lower levels of stress and anxiety, and this causes them to feel less burnout. Teachers feeling less stress, anxiety and burnout have higher levels of job satisfaction in their work, thus they feel higher levels of job involvement.

The results of this research reveal that the teachers with higher levels of psychological capital also have more powerful psychological resources and so they are less vulnerable to the negative effects of the experiences they face at work. Therefore, they feel negative attitudes in their work environment less frequently, and they develop more positive attitudes towards work and school. It is crucial for school administrators to invest in developing and strengthening teachers' psychological capital if they want teachers to feel negative attitudes less frequently and to develop more positive attitudes at work. To create more positive educational environments, it would be beneficial to take psychological capital into consideration at all stages of

human resource management, especially during pre-service and in-service training and in the recruitment and selection of teachers.

This paper contributes to the literature in the following ways: Firstly, based on the literature review, this study constructs an integrated framework to determine the effects of psychological capital on stress, anxiety, burnout, job satisfaction, and job involvement. Secondly, this study presents a conceptual framework that examines the relationship between psychological capital, stress, anxiety, burnout, job satisfaction, and job involvement in educational organizations. Thirdly, this paper examines the relationship between psychological capital and stress, anxiety, burnout, job satisfaction, and job involvement with mediation effects.

References

Altıntas, E. (2014). *Stress yönetimi* [Stress management]. Ankara: Ani yayincilik.

Arslantas, H.I., & Aslan, I. (2015). Ilk ve ortaokul ogretmenlerinin mesleki tukenmislik düzeyleri ile is doyumları arasindaki iliskinin incelenmesi (Sırnak ili örnegi) [The examining of the relationship between primary and secondary school teachers's levels of occupational burnout and job satisfaction (an Example of Şırnak, Turkey)]. 10. Ulusal Egitim Yonetimi Kongresi 7-9 Mayis [10. National Educational Administration Congress 7-9 May]. Ankara: Pegem Akademi.

Bradley, K. (2014). *Educator's positive stress responses: Eustress and psychological capital*. (Unpublished doctoral dissertation). DePaul University, Chicago, United States.

Burke, R. J. (2000). Workaholism in organizations: Psychological and physical well-being consequences. *Stress Medicine*, 16, 11-16.

Burke, R.J., Burgess, Z., & Fallon, B. (2006). Workaholism among Australian female managers and professionals, job behaviors, satisfactions and psychological health. *Equal Opportunities International*, 25(3), 200-213.

Byrne, B. M. (2010). *Structural equation modeling with AMOS*. New York: Routledge.

Celik, V. (2005). Liderlik. [Leadership]. In Y. Özden (Ed.), *Egitim ve okul yoneticiligi el kitabi* [Educational and school administration handbook] (pp. 187-215). Ankara: Pegem Akademi.

Cetin, F., & Basim, H.N. (2012). Orgutsel psikolojik sermaye: Bir olcek uyarlama calismasi [Organizational psychological capital: A scale adaptation study]. *Amme Idaresi Dergisi, 45*(1), 121-137.

Cetin, F., & Basim, H.N. (2011). Psikolojik dayanikliligin is tatmini ve orgutsel baglilik tutumlarindaki rolu [The role of resilience in the attitudes of job satisfaction and organizational commitment]. *"IS, GUC" Endustri Iliskileri ve Insan Kaynaklari Dergisi, 13*(3), 79-94.

Cokluk, O. (2014). Orgutlerde tukenmislik. Icinde C. Elma ve K. Demir (Ed.) [Burnout in organizations]. In C. Elma & K. Demir (Ed.), *Yonetimde cagdas yaklasimlar uygulamalar ve sorunlar* [Modern approachments practises and problems in administration] (pp. 117-143). Ankara: Ani Yayincilik.

Demir, S. (2018). Okul yoneticilerinin motivasyonel dili ile ogretmen oz yeterligi arasındaki iliski uzerine bir calısma [A study on the relationship between school adminisrators' motivational language and teacher self-efficacy]. *Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi, 6*(2), 177-183. DOI: 10.18506/anemon.384848

Dolgun, U. (2012). Tukenmislik sendromu. [Burnout syndrome]. In D.E. Ozler (Ed.). *Orgutsel davranista guncel konular* [Recent subjects in organizational behavior] (pp. 287-310). Bursa: Ekin Basim Yayin Dagitim.

Erkus, A., & Findikli, M.A. (2013). Psikolojik sermayenin is tatmini, is performansi ve isten ayrilma niyeti uzerindeki etkisine yonelik bir arastirma [A research on the impact of psychological capital to job satisfaction, job performance and intention to quit]. *Istanbul Universitesi Isletme Fakultesi Dergisi, 42*(2), 302-318.

Field, A. (2009). *Discovering statistics using SPSS*. Dubai: Oriental Press.

Genc, N. (2012). *Yonetim ve organizasyon* [Management and organization]. Ankara: Seckin Yayincilik.

Goldberg, L.R., Johnson, J.A., Eber, H.W., Hogan, R., Ashton, M.C., Cloninger, C.R., & Gough, H.C. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84–96.

Griffin, M.L., Hogan, N.L., Lambert, E.G., Tucker-Gail, K.A., & Baker, D.N. (2010). Job involvement, job stress, job satisfaction, and organizational commitment and the burnout of correctional staff. *Criminal Justice and Behavior, 37*, 239-255.

Haymon, S. W. (1992). *The relationship of work addiction and depression, anxiety and anger in college males.* (Unpublished doctoral dissertation). Florida State University, USA.

Hock, R.R. (1988). Professional burnout among public teachers. *Public Personel Management*, 17(2), 167-189.

Jones, J.W. (1981). *The burnout syndrome: Current research, theory, interventions*. London: London House Press.

Kaplan, M., & Bickes, D.M. (2013). The relationship between psychological capital and job satisfaction: A study of hotel businesses in Nevşehir. *Yonetim ve Ekonomi, Celal Bayar Universitesi IIBF, 20*(2), 233-242.

Karakus, M. (2013). Emotional intelligence and negative feelings: A Gender specific moderated mediation model. *Educational Studies, 39* (1), 68-82.

Karakus, M., & Demir, S. (2015). Psikolojik sermaye ile is doyumu, orgutsel baglilik, motivasyon ve isten ayrilma niyeti arasindaki iliski [The relationship between psychological capital and job satisfaction, organizational commitment, motivation and intent to leave]. 10. Ulusal Eğitim Yonetimi Kongresi 7-9 Mayıs [10. Educational Administration Congress 7-9 May]. Ankara: Pegem Akademi.

Kline, R.B. (2011). *Principles and practice of structural equation modeling*. New York: The Guilford Press.

Kyriacou, C. (2000). *Stress-busting for teachers*. London: Nelson Thornes Ltd.

Larson, M., & Luthans, F. (2006). Potential added value of psychological in predicting work attitudes. *Journal of Leadership and Organizational Studies, 13*, 44-61.

Liu, L., Pang, R., Sun, W., Wu, M., Qu, P., Lu, C., & Wang, L. (2013). Functional social support, psychological capital, and depressive and anxiety symptoms among people living with hiv/ Aids employed full-time. *Liu et al. BMC Psychiatry, 13*:324.

Locke, E.A. (1976) The nature and causes of job satisfaction. In M.D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 1297-1343). Chicago: Rand McNally.

Luthans, F. (2002). The need for and meaning of positive organizational behavior. *Journal of Organizational Behavior, 23*(6), 695-706.

Luthans, F., & Youssef, C. M. (2004). Human, social and now positive psychological capital management: Investing in people for competitive advantage. *Organizational Dynamics*, *33*(2), 143-160.

Luthans, F., Avolio, B.J., Avey, J.B., & Norman, S.M. (2007). Positive psychological capital: Measurement and relationship with performance and satisfaction. *Personnel Psyhology, 60*, 541-572.

Luthans, F., & Youssef, C. M. (2007). Emerging positive organizational behavior. *Journal of Management, 33*, 321-349.

Maslach, C., & Jackson, S.E. (1981). The measurement of experienced burnout. *Journal of Occupational Behaviour, 12*, 99-113.

Ocak, M., & Guler, M. (2017). Psikolojik sermayenin tukenmislik uzerine etkisi: Gorgul bir arastırma **[**The effects of psychological capital on burnout: An empirical research]. *Erciyes Universitesi Iktisadi ve Idari Bilimler Fakultesi Dergisi, 49*, 117-134.

Ocak, M., Guler, M., & Basim, H.N. (2016). Psikolojik sermayenin orgutsel baglılık ve is tatmini tutumları uzerine etkisi: Bosnalı ogretmenler uzerine bir arastırma **[**The effects of psychological capital on the attitudes of commitment and satisfaction: A research on bosnian teachers]. *Cankırı Karatekin Universitesi IIBF Dergisi*, http://dx.doi.org/10.18074/cnuiibf.274

Pines, A. M. (2005). The burnout measure short version (BMS). *International Journal of Stress Management, 12*(1), 78-88.

Polatcı, S. (2014). Psikolojik sermayenin gorev ve baglamsal performans uzerindeki etkileri: Polis teskilatinda bir arastirma **[**The Effects of psychological capital on task and contextual performance: A research on police organization**]**. *Ege Akademik Bakıs, 14*(1), 115-124.

Porter, G. (1996). Organizational impact of workaholism: Suggestions for researching the negative outcomes of excessive work. *Journal of Occupational Health Psychology, 1*, 70-84.

Robbins, S.P., & Judge, T.A. (2012). *Orgutsel davranıs* [Organizational behavior] (14th Edition, I. Erdem trans. ed.). Istanbul: Nobel yayincilik.

Robson, C. (2015). *Bilimsel arastırma yontemleri: Gercek dunya arastırması* [Real world research] (S. Cınkır & Nihan D., Ed. of trans.). Ankara: Anı yayıncılık.

Russell, D.W., Altmaier, E., & Van Velzen, D. (1987). Job-related stress, ssocial support, and burnout among classroom teachers. *Journal of Applied Psychology, 72*(2), 269-274.

Schulz, S.A., Luthans, K.W., & Messersmith, J.G. (2014). Psychological capital a new tool for driver retention. *International Journal of Physical Distribution & Logistics Management, 44*(8/9), 621-634.

Simons, J.C., & Buitendach, J.H. (2013). Psychological capital, work engagement and organisational commitment amongst call centre employees in South Africa. *SA Journal of Industrial Psychology/SA Tydskrif vir Bedryfsielkunde*, *39*(2), Art. #1071, 12 pages. http://dx.doi.org/

Simsek, M., & Celik, A. (2012). *Yonetim ve organizasyon* [Management and organization]. Konya: Egitim Akademi.

Sorenson, R. D. (2007). Stress management in education: Warning signs and coping mechanisms. *Management in Education, 21*(3), 10-13.

Tumkaya, S., Cam, S., & Cavusoglu, I. (2009). Tukenmislik olcegi kisa versiyonu'nun Turkceye uyarlama, gecerlik ve guvenirlik çalışması [Turkish adaptation of the burnout snydrome inventory short version]. *C.U. Sosyal Bilimler Enstitusu Dergisi, 18* (1), 387–398.

Tuzun, I.K., Cetin, F., & Basim, H.N. (2014). The role of psychological capital and supportive organizational practices in the turnover process. *METU Studies in Development, 41*(2), 85-103.

De Waal, J.J., & Pienaar, J. (2013). Towards understanding causality between work engagement and psychological capital. *SA Journal of Industrial Psychology/SA Tydskrif vir Bedryfsielkunde*, *39*(2), Art. #1113, 10 pages. http:// dx.doi.org/10.4102/sajip. v39i2.1113.

Wang, Y., Chang, Y., Fu, J., & Wang, L. (2012). Work-family conflict and burnout among Chinese female nurses: the mediating effect of psychological capital. *Wang et al. BMC Public Health, 12*:915.

Yuksel, M., & Akdag, F. (2011). Framing workaholism and locus of control as psychological capital in an emergent economy. *China-USA Business Review, 10*(5), 389-398.

Psikolojik Sermaye ile Stres, Kaygı, Tükenmişlik, İş Doyumu ve İşe Sargınlık Arasındaki İlişki

Atıf:

Demir, S. (2018). The relationship between psychological capital and stress, anxiety, burnout, job satisfaction and job involvement. *Eurasian Journal of Educational Research, 75*, 137-154, DOI: 10.14689/ejer.2018.75.8

Özet

Problem Durumu: Bu araştırma, psikolojik sermayenin eğitim örgütlerinde istenen davranışları nasıl arttırdığı, istenmeyen davranışları ise nasıl azalttığı yönünde önemli bilgilere ışık tutacak şekilde dizayn edilmiştir. Eğitim örgütlerinde öğretmenlerin olumlu tutumlar sergilemeleri oldukça önemlidir. Öğretmenlerin duyguları olumlu olduğunda öğrencileriyle, birbirleriyle kısacası okulun tüm üyeleriyle daha sağlıklı iletişim kurabilirler. Ayrıca duygular bulaşıcı olduğu için, öğretmenin işe karşı geliştirdiği olumlu ya da olumsuz tutumları içinde bulunduğu örgütün iç ve dış üyelerini etkilemektedir. Öğretmenin daha olumlu duygulara sahip olması ve olumsuz duygularının azaltılması ya da önlenmesi eğitimin kalitesi açısından oldukça önemlidir. Zira öğrencinin öğrenme ortamının en önemli öğesi öğretmenlerdir. Eğitim örgütlerinde psikolojik sermaye ile ilgili olarak yapılan çalışmalar sınırlı sayıdadır. Ayrıca eğitim örgütlerinde psikolojik sermaye, stres, kaygı, tükenmişlik, iş doyumu ve işe sargınlık değişkenlerinin bir arada işlendiği bir çalışmaya ilgili alan yazında rastlanmamıştır. Psikolojik sermayenin olumlu ve olumsuz duygularla bir arada ilişkisini ortaya çıkaracak böyle çok değişkenli bir çalışmaya ihtiyaç duyulmuştur.

Araştırmanın Amacı: Bu çalışmada psikolojik sermaye ile stres, kaygı, tükenmişlik, iş doyumu ve işe sargınlık arasındaki ilişkinin tespit edilmesi amaçlanmaktadır.

Yukarıdaki temel amaç bağlamında şu sorulara cevap aranmıştır:

1) Okul ortamında öğretmen tarafından algılanan psikolojik sermayenin stres, kaygı, tükenmişlik, iş doyumu ve işe sargınlık üzerinde etkisi var mıdır? Varsa stres, kaygı, tükenmişlik, iş doyumu ve işe sargınlığı hangi yönde etkilemektedir?

2) Okul ortamında öğretmen tarafından algılanan stresin, kaygı, tükenmişlik, iş doyumu ve işe sargınlık üzerinde etkisi var mıdır? Varsa kaygı, tükenmişlik, iş doyumu ve işe sargınlığı hangi yönde etkilemektedir?

3) Okul ortamında öğretmen tarafından algılanan kaygının, tükenmişlik, iş doyumu ve işe sargınlık üzerinde etkisi var mıdır? Varsa kaygı, tükenmişlik, iş doyumu ve işe sargınlığı hangi yönde etkilemektedir?

4) Okul ortamında öğretmen tarafından algılanan tükenmişliğin, iş doyumu ve işe sargınlık üzerinde etkisi var mıdır? Varsa iş doyumu ve işe sargınlığı hangi yönde etkilemektedir?

5) Okul ortamında öğretmen tarafından algılanan iş doyumunun, işe sargınlık üzerinde etkisi var mıdır? Varsa işe sargınlığı hangi yönde etkilemektedir?

Araştırmanın Yöntemi: Bu çalışmada iki ya da daha fazla değişken arasındaki ilişkileri açıklayan ilişkisel tarama deseni kullanılmıştır. Araştırmanın çalışma evreni 2014-2015 eğitim öğretim yılında Hatay ili Kırıkhan ilçesindeki okullarda görev yapan 1145 öğretmendir. Bu araştırmada tek tek bireylerin değil, seçkisiz olarak belirlenen grupların örneklem için seçilmesi olarak tanımlanan oransız küme örnekleme yöntemi kullanılmıştır. Kırıkhan'daki her okul bir küme kabul edilip asgari ihtiyaç duyulan şekilde okullar tesadüfî olarak seçilmiştir. Örnekleme alınan 27 okulda görev yapmakta olan 335 öğretmen bu araştırmanın örneklemini oluşturmaktadır. Örneklem sayısı belirlemede, ana kütle sayısı belli olan büyüklükler için hazırlanan formüle göre alınan 335 öğretmenin %95 güven düzeyi için yeterli olduğu görülmüştür. Araştırmanın verileri anketler yoluyla elde edilmiştir. Veri toplama aracı olarak öğretmenlere verilen anketlerde; psikolojik sermaye ölçeği, stres ölçeği, kaygı ölçeği, tükenmişlik ölçeği, iş doyumu ölçeği ve işe sargınlık ölçeği kullanılmıştır. Bu çalışmada beşli likert tipi anketler kullanılmıştır.

Araştırmanın Bulguları: Öğretmenlerin, araştırma kapsamında uygulanan ölçme araçlarındaki maddelere katılım düzeylerini gösteren aritmetik ortalama ve standart sapma değerlerine göre psikolojik sermaye genel, psikolojik sermaye alt boyutları ve iş doyumu düzeyleri "katılıyorum (4)" düzeyindedir. Ayrıca öğretmenlerin stres, kaygı, tükenmişlik ve işe sargınlık düzeyleri ise "kısmen katılıyorum (3)" düzeyindedir. Öğretmenlerin psikolojik sermaye düzeylerinin genel ortalaması; stres, kaygı ve tükenmişlik değişkenleriyle negatif korelasyon içindedir. Öğretmenlerin psikolojik sermaye düzeyleri; iş doyumu ve işe sargınlık değişkenleriyle pozitif korelasyon içindedir. Ayrıca stres ile kaygı ve tükenmişlik değişkenleri ve iş doyumu ile işe sargınlık değişkeni pozitif korelasyon içindedir. İş doyumu ile stres ve tükenmişlik, işe sargınlık ile stres ve tükenmişlik negatif korelasyon içindedir. Yapısal eşitlik modellemesi sonucunda ortaya çıkan ve en iyi uyum indislerini üreten yapısal modele göre; öğretmenlerin sahip oldukları psikolojik sermaye düzeyi arttıkça stres düzeyleri düşmektedir. Psikolojik sermaye,

stres değişkeninin tam aracılık etkisiyle öğretmenlerin kaygı düzeylerini düşürmektedir. Psikolojik sermaye, stres ve kaygı değişkenlerinin kısmî aracılık etkisiyle öğretmenlerin tükenmişlik düzeylerini düşürmektedir. Psikolojik sermaye; stres, kaygı ve tükenmişlik değişkenlerinin kısmî aracılık etkisiyle öğretmenlerin iş doyumu düzeylerini artırmaktadır. Psikolojik sermaye; stres, kaygı, tükenmişlik ve iş doyumu değişkenlerinin kısmî aracılık etkisiyle öğretmenlerin işe sargınlık düzeylerini artırmaktadır.

Sonuç ve Öneriler: Bu araştırmanın sonuçlarına göre öğretmenler; sahip oldukları psikolojik sermaye düzeyi arttıkça, psikolojik açıdan daha güçlü hale gelmekte ve iş içinde yaşadıkları olumsuz deneyimlerden daha az etkilenmektedirler. Bu yüzden iş içinde olumsuz duyguları daha az hissetmekte, işe ve okula ilişkin daha olumlu tutumlar geliştirmektedirler. Öğretmenlerin iş içinde olumsuz duyguları daha az hissetmelerini ve daha olumlu tutumlar geliştirmelerini sağlamak isteyen okul yöneticilerinin, psikolojik sermayeyi güçlendirme ve geliştirme çalışmaları yararlı görülmektedir. Eğitim örgütleri öğretmenlerin daha huzurlu, mutlu, şevkli, iş doyumları yüksek, işe sargınlıkları fazla ve verimli olmalarını sağlayıcı fiziksel, sosyal ve psikolojik ortamlar ve politikalar belirlemeli ve uygulamalıdır. Benzer bir çalışma farklı örneklemlerle ve farklı araştırma desenleri kullanılarak yapılabilir. Bu araştırmada belirlenen ilişkilerin nedenleri derinlemesine incelenerek tespit edilebilir. Farklı yerleşim birimlerinde, farklı örgütlerde ve daha geniş örneklemlerle benzer araştırmalar yapılabilir.

Anahtar sözcükler: Bireysel çıktılar, öğretmen tutumları, verimlilik, aracılık etkisi.

Eurasian Journal of Educational Research 75 (2018) 155-178

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

Education for Syrian Refugees: Problems Faced by Teachers in Turkey*

Pelin TASKIN[1], Ozge ERDEMLI[2]

ARTICLE INFO

Article History:
Received: 15 Dec. 2017
Received in revised form: 23 Mar. 2018
Accepted: 22 Apr. 2018
DOI: 10.14689/ejer.2018.75.9

Keywords
Syrian students, educational challenges, immigrants' right to education

ABSTRACT

Purpose: Providing educational services for Syrian refugee children is a new fact of life in Turkey, and the teachers who work at public schools and temporary education centers encounter some difficulties. The main purpose of this paper is to describe the problems faced in the process of educating Syrian refugee children as well as the solutions to such difficulties recommended by the teachers. **Research Methods:** The study was carried out using qualitative research methods and the phenomenological model to investigate the problems faced in the process of educating the Syrian students'. The purposive sampling method was used in order to identify the participants. Thereafter; the plan was to interview the participants consisting of nine teachers of Syrian students. Interview forms were designed and developed by the researchers and used in the study. The content analysis technique was used during the analysis of the data. The Nvivo 10 package program was used. **Findings**: In the study, the problems faced by the teachers were identified as a language barrier, cultural problems and discipline problems. While the teachers at the public schools appear to think that the students are on good terms with their friends, teachers and the school principal, the teachers at the temporary education centers seem to have a completely opposite view on the matter. Teachers state that they do not receive adequate support in the process of educating Syrian students and add that the students are provided with limited books and additional class support. Last, to solve these problems, the teachers recommend that the language barrier be overcome by teaching the Syrians the Turkish language, that materials and a curriculum for Turkish language education be developed, that the Syrian students be taught separately from the Turkish students, that the teachers be provided with in-service seminars, and that class populations be reduced. **Implications for Research and Practice**: Teaching so many refugees is a new phenomenon for Turkey. In order to properly overcome this issue, the MoNE must be in close contact with the teachers and educational institutions working in this process. Also, for the healthy functioning of this process, the need for support from both the general public and national and international organizations should not be ignored.

* This article was presented at ECER 2016, "Leading Education: The Distinct Contributions of Educational Research and Researchers", in Dublin, Ireland.

[1] Ankara University Faculty of Educational Sciences, TURKEY. E-mail: ptaskin@education.ankara.edu.tr. ORCID: orcid.org/0000-0001-8860-579X

[2] Ankara University Faculty of Educational Sciences, TURKEY. E-mail: oerdemli@ankara.edu.tr. ORCID: 0000-0002-8004-020X

Introduction

Migration, a phenomenon as old as the history of humanity, appears to be one of the major issues that has recently been the focus of worldwide discussion, particularly with the recent increases in global population movements. While some people immigrate to other countries voluntarily for a better life, others are forced to immigrate to other countries for extraordinary reasons such as war.

The phenomenon of migration was first discussed in the context of men and labor migration, then in the context of women after the 1980s, and in the context of children after the 2000s. Particularly, the increased number of refugee children in several countries has marked children as a new immigrant population (Topcuoglu, 2012). This situation has required many countries to take children into consideration in their immigration policies. Therefore, education policies for refugee children have also begun to be discussed.

Compared to other immigrant populations, refugee children are more vulnerable to serious problems related to malnourishment, diseases, physical injuries, brain damage, and sexual or physical abuse (Neugebauer, 2013). In this respect, access to education may be suggested to be critically important for refugee children. For, refugee children who do not benefit from the right to attend a school are at higher risk of abuse and ill treatment, demonstrate post-traumatic stress disorder symptoms more often, and are less likely to complete their psychological development (UNESCO, 2011). Also, crises experienced in the education of refugees lead to child labor and early marriage. Furthermore, young people who have missed out on the opportunity to be educated are at risk of being led into radical groups (Watkins & Zyck, 2014). Schools enable refugee children to normalize their lives and give them hope for the future (Beste, 2015). This being the case, in order for refugee children to have a better life, it is important to ensure their access to education and identify and solve the barriers to education before they begin their education process.

The conflict between the governing power and the dissidents has displaced many people from Syria to neighboring countries, European countries and countries in the Americas. According to UNHCR (2018) report there are 5,645,914 registered refugees as of 2018, 3,586,679 of whom are in Turkey. Facilitating their access to the right to education is probably the most important arrangement made for refugee children. Turkey has taken significant steps and offered all of these children of compulsory education age the opportunity to receive free education in public schools. However, instead of using the word "refugee", Turkey tends to call these people "temporary asylees" relying on the reservation in the 1967 Protocol Relating to the Status of Refugees. After 2011 with Law 6458, Turkey granted "temporary protection" status to the Syrians (Uzun, 2015, 112).

Although the status of the Syrian asylum seekers seems to be important only for countries that are neighbors of Syria, the increased immigration movements, particularly as of 2015, also concern the European countries (Deutsche Welle, 2015). This is because several neighboring countries are regarded by some refugees as a transit point to Europe. Therefore, the problems faced regarding the Syrian refugees

must be addressed and solved on a global scale. The countries in Europe follow different policies on the education of the Syrian children and youth.

When the education policies implemented by Turkey for the Syrian children are considered, pursuant to article 22 of the UN Convention on the Rights of the Child to which Turkey is a contracting party, Turkey is required to take necessary measures to ensure that all children, including refugee or asylum-seeking children in the country, benefit from all of the rights set forth in the Convention. Also, the Law on the Protection of Children stipulates that every child can benefit from protection, irrespective of their national origin. However, since there was no special regulation on the education of these children, the Ministry of National Education (MoNE) published the Circular 2014/21 of September 23, 2014 to regulate the Educational and Learning Services for Foreigners. The following options are available in Turkey for the education of the Syrian children under temporary protection:

Camp education centers. In Turkey, the Disaster and Emergency Administration (AFAD) established a total of 26 camps/temporary accommodation centers in 10 provinces (General Directorate of Immigration Administration, 2016). All of the camps have education centers offering education from primary school through high school (Dillioglu, 2015). More than 90% of the Syrian children living in the camps receive education (UNICEF, 2014).

Public schools. According to the figures, most children living in the camps receive education while very few children living outside the camps receive education (Ozturk, 2015). Therefore, in order to fulfill its international obligations and provide the Syrian children living outside the camps with access to education, the MoNE published the Circular on "Educational and Learning Services for Foreigners" in September 2014. With this Circular, the Syrian children have had the opportunity to receive Turkish education in the public schools, under the MoNE, with a "foreigner identification document" and without the requirement of having a "residency permit".

According to the regulation, the Syrian students have the right to receive education in the same class with the Turkish students. In these schools, the Turkish education curriculum is implemented for the Syrian students as for all students. The Syrian children do not go through a preparatory stage and are admitted to formal classes taught in Turkish as soon as they enroll. In contrast, in November 2016, the MoNE created the project, "Integration of the Syrian Students with the Turkish Education System" with the European Union for the Syrian students in Turkey to learn Turkish better. Under this project, teachers will be specifically employed to teach the Turkish language to the Syrian students (MoNE, 2016). A total of 36,655 Syrian students enrolled in the primary, middle and high schools in the public education system in Turkey in the 2014–2015 academic year (Human Rights Watch Report, 2015, 12).

Temporary education centers (TECs). TECs are those in and outside the camps established by the AFAD and offer primary and secondary education in Arabic by adhering to the Syrian curriculum for school aged Syrian children and youth (Emin, 2016). In TECs, non-formal educational courses in demand can also be opened by non-formal educational institutions.

In the present study, some information was also received from the principal and assistant principal of the first TEC opened to provide education to the Syrian students living outside the camps. According to the information provided by the school administrators, Turkish students receive education in the morning, and the Syrian students receive education in the afternoon in these schools. The Syrian curriculum approved by the Board of Education is implemented in these schools where six hours of education is provided. Syrian teachers work in these schools, and only the Turkish classes are taught by Turkish teachers. While the Turkish teachers are employed on a contractual basis by the MoNE, the Syrian teachers are found by the Turkish Religious Foundation, and the funding is provided by UNICEF. However, not all of the Syrians working in these schools have a teaching degree. Syrian nationals who practiced a profession before the war in Syria such as medicine, engineering and law teach in these schools. While co-education is implemented in the first four years in these schools, male and female students receive education separately in the subsequent years. Nonetheless, another study conducted in Istanbul suggests the opposite. In that study, researchers visited several schools where all current Syrian teachers were university graduates and an overwhelming majority of them (97.9%) were accredited teachers in Syria, with only a small group of individuals from other backgrounds, such as law (Aras & Yasun, 2016), which points to the fact that there is a lack of consistency in the education of the Syrian refugees in Turkey.

Private initiatives. Another alternative for the education of the Syrian children is private schools established with the help of philanthropists or non-governmental organizations (Dillioglu, 2015). Educational materials and school supplies are donated. In these private schools, a revised version of the Syrian curriculum (the curriculum eliminating the sections regarding Assad) is implemented and subjects of Science, Social Sciences, Mathematics and Turkish are taught (Milliyet, 2014).

As seen, there are various initiatives that have been taken by Turkey to provide the refugee children with access to education. Naturally, despite such initiatives, there are several challenges facing the Syrian children's education today. One of the major problems is ensuring children's access to education. According to a study conducted in Lebanon, most of the Syrian children cannot attend a school yet. The study suggests financial problems, i.e. poverty, as the primary barrier (Mayer, 2015). Turkey, now home to a considerable number of Syrian immigrant children, strives to create various opportunities both inside and outside the camps to overcome this issue. With Circular 2014/21, the Syrian children have been granted the right to enroll in public schools and thus access free education. Also, course books are provided by the MoNE free of charge to all Syrian students enrolled in a public school, which is an attempt to eliminate the financial difficulties constituting a major barrier to education.

Even in cases where the immigrant children's access to education has been achieved, there are yet several problems faced in the education process, to name but a few, language barrier, capacity issues, discrimination, bullying, sectarian tensions, child marriage, and school dropouts (Mayer, 2015). Although language barrier is the primary problem, lack of sufficient educational materials, concerns about safety, lack

of transportation fees, emotional trauma, and ill treatment at schools are among the other problems these students face (Sirin & Rogers-Sirin, 2015).

Although the concept of refugee is not a new term, it is impossible to find a universal solution to the problems of refugees and apply it across all situations and conditions leading to refugee status. However, it would be helpful to benefit from the past actions taken to improve the status of refugees while seeking a solution. There are valuable lessons to be learned from the activities of the International Rescue Committee (IRC) for the education of the refugees in Liberia and Sierra Leone between 1990 and 1991. Said program is a good example of a well-structured education program. The program may prove to be an important resource for Turkey in the education of Syrian refugees in respect of the attention given by the program to the selection and guidance of the teachers, effective management and the inclusion of messages and activities in the education program for adolescents and youth displaced due to this civil war (UNESCO, 2017).

The literature includes studies focusing on the process of educating the Syrian children living in Turkey. One of them found that the qualities of teachers at TECs were not deeply vetted by the MoNE and that refugee education might require additional skills for teachers in order to cope with student trauma and provide psychological support to students (Aras & Yasun, 2016, 7). However, no study has been conducted in Turkey like the present study focusing directly on the evaluation of the process from the teachers' perspective. The studies conducted in Turkey focused on the educational activities in the camps or compiled and evaluated the education policies for refugees in Turkey. The admission of Syrians into public schools outside the camps with the Circular 2014/21 has made it necessary to identify the problems experienced in the education process in the new situation.

Acting on this necessity, the objective of this research is to determine the difficulties faced by teachers working in public schools and TECs in Ankara in the process of education of Syrian students and to recommend solutions to such difficulties. Responses to the following questions will be sought in line with the overall objective of the research: According to the teachers' views:

1) What are the problems faced by teachers in the process of educating Syrian students?
2) How is the relationship between Syrian students and their friends, teachers and principals in the school?
3) What support is provided to teachers and students during the integration of the Syrian students within the school and their classes?
4) What are their recommendations for the solution to the problems faced during the process of educating Syrian students?

Method

Research Design

The study was carried out using qualitative research methods. A qualitative study is defined as a "research investigating the words and actions of the participants in a descriptive and expressive way to better describe the situations experienced by the participants" (Maykut & Morehouse, 1994, 3). Specifically, a phenomenological investigation examines the lived experiences of individuals as they relate to the phenomenon under investigation (Creswell, 2012). In other words, phenomenological studies investigate daily events from the perspective of the people experiencing them (Seah & Wilson, 2011). This study adopted the phenomenological model to investigate the problems faced in the process of educating Syrian students.

Participants

The purposive sampling method was used in order to identify the participants. Thereafter, interviews with the participants, consisting of nine teachers who teach Syrian students at public schools in Ankara, were planned. Both the teachers working in public schools and the teachers working in a TEC were interviewed to achieve maximum diversity as recommended by Meriam (2013). Five of the teachers work in a public school, and four of them work in a TEC. Of the teachers, seven are female, and two are male.

All of the teachers working in the public schools are primary school teachers, and their length of service varies between 9 and 16 years. All of the teachers working in the TEC teach the Turkish Language and Literature class. The teachers working in this school have yet to complete their first year in the profession and are employed on a contract basis by the MoNE. None of these teachers has attended a training or seminar provided with respect to the Syrian students under temporary protection.

Research Instruments and Procedures

Interview forms were designed and developed by the researchers and used in the study. Semi-structured interviews were conducted with the teachers working in the public schools. The participants from the TECs preferred the interviews to be conducted in the form of a focus group. In phenomenological investigations, it is important to provide participants with an environment that makes them feel comfortable (Ersoy, 2016). Glesne (2013) also suggests that group interviews are particularly appropriate for studies in which participants can express multiple perspectives on similar experiences such as the implementation of a specific policy or an educational program. Considering that all of the teachers from the TECs are a Turkish Language and Literature teacher, focus groups were thought likely to achieve a richer data set on the Turkish education of Syrian students. Therefore, in line with the preference of the participants, it was decided that a focus group interview would be conducted at this school. As a result, both the semi-structured interview and focus group methods were used to collect data.

In line with the information received from the MoNE, the principals of the public schools and the TECs where the Syrian students are educated were contacted by phone to identify the teachers with Syrian students in their class in order to schedule interviews. During the interviews, the participants were asked for their permission to record the audio of the session to avoid any data loss. The views of the participants who rejected audio recording were noted. Each interview lasted about 25-30 minutes. The focus group interview took 45 minutes.

Validity and Reliability

The semi-structured interview form prepared within the scope of the study has been presented to experts for content validity. Through the corrections and suggestions received from the experts, the interview form has been finalized and then the interviews have been carried out.

In order to ensure the validity and reliability of the study, characteristics of participants of the research were clearly defined, and detailed information with data collection and data analysis were provided about processes as recommended by Yildirim and Simsek (2011). Also, the participants' views were directly quoted in order to maximize the reliability of the study.

Data Analysis

The content analysis technique was used during the analysis of the data. The Nvivo 10 package program was used. In this context, the audio recording was initially decoded. The transcribed and noted data were transferred fully to a computer program. The coded data was grouped under appropriate themes in line with the objectives of the research. Data analysis was performed by the two researchers collectively. The views directly quoted from the teachers are provided in a smaller font along with the code of the participant in parenthesis. Each participant was assigned a code (the teachers working in the public schools were assigned a number between 1 and 5, and the teachers working in the TEC between 6 and 9) and the letter F (female) or M (male) was added next to the code based on the gender of the participant.

Results

The Syrian students under temporary protection in Turkey have the option to receive off-camp education either in the public schools or in the TECs. These two options offered to the Syrian refugees are explored in line with the research questions. The findings were shaped around four themes explained under the respective headings.

The Views of the Participants on Difficulties Experienced in the Process of the Syrian Students' Education

Difficulties experienced by the teachers in the process of educating the Syrian students, as provided in Table 1, are grouped under three themes including language barriers, cultural problems and discipline problems.

Table 1

Difficulties Experienced by the Teachers in the Process of Educating Syrian Students

	Themes	*Representative Comments*
Theme 1	*Language barriers*	They have learned how to read and write but have not been able to learn the language. That's why, they cannot understand what they read (3F). Two or three students speak Turkish well. We use those speaking Turkish as translators (6F). We mostly communicate through body language (9F). There are those who do not want to learn Turkish, those who use full efforts to learn Turkish, and also those who are not allowed by their families to learn Turkish (7F).
Theme 2	*Cultural problems*	They are more hesitant toward the male students and don't want to talk to them. They are on better terms with the girls (3F). When I first arrived, they would be saying salavat when the call to prayer was recited, which seemed very strange to me. They want to pray in the classroom and everywhere (8F). The gender factor is also important. I communicate better with the female students. Since I don't wear a headscarf, there are male students who will not look at my face at all. The small children ask me why I don't wear a headscarf. Some have even written warning letters to me (6F). Since it is more comfortable to communicate with the female students, they learn more easily, and boys find learning harder (9F).
Theme 3	*Discipline problems*	We have a lot of discipline problems. They were raised by teachers more inclined to violence. We just do with yelling, and that's why they don't listen to what we say (7F). We have discipline problems because different age groups are in the same classroom [. . .] so the older age group uses violence on the younger age group (6F).

Language barrier. According to the teachers, the primary problem experienced in the process of educating the Syrian children is the "language barrier". Under the language barrier, the most important difficulty was stated to be the lack of Turkish literacy. In schools implementing the Turkish curriculum, the Syrian students first learn how to read and write in Turkish. Some of the students learn how to read and write in Turkish, but according to teacher remarks, their proficiency in Turkish is not satisfactory; thus, they have difficulty understanding what they read and write. In other words, according to the teachers, these students copy and write a text/sentence they see, but do not know what it means (1M, 2M, 3F, 4F, 5F). Therefore, the teachers stated that they were teaching the Turkish students the Turkish curriculum at the same time as they were striving to teach the Syrian students how to read and write in Turkish. This may disrupt the harmony of the class as well as prevent the Syrian students from integrating with the rest of the class population. It may also lead to the teacher slowing down and falling behind the syllabus. The teachers also added that they didn't have the opportunity to teach other subjects to the Syrian students because of their poor level of Turkish. The teachers said that they could only teach them how to read and write and that these students had more difficulty when they moved into the regular program after having learned how to read and write (1M, 2M).

The teachers said that they were using the students with a higher level of Turkish abilities as translators, which is how they try to communicate with the class. According to what the teachers said, there are students willing to learn Turkish, but there are also those resisting the lessons. These students are particularly not allowed to learn Turkish by their families with the concerns that they would lose their culture.

Cultural problems. According to the teacher statements under this theme, one of the cultural problems is perceptions regarding gender. Female teachers said that they were able to communicate better with the female students due to the gender factor. The female teachers also said that the male students did not want to engage with them because of their gender. Among the participants, those women not wearing a headscarf said that some of the male students would not communicate with them. A teacher in a headscarf (8F) said that she was able to communicate comfortably with the male students, which may be because the students consider teachers wearing a headscarf close to their culture and religious beliefs. Additionally, the teachers stated that the female students were quite shy toward males.

Another cultural problem expressed by the teachers is the difference between the religious practices. The majority of the Turkish people are Muslim. However, secularism is the fundamental principle of the Turkish Education System. Therefore, no religious service is performed at schools. The Turkish teachers were surprised to see that the students stood up to pray during a class when the call to prayer is recited. Such cultural problems regarding religious practices adversely affects the course of the classes in the education process because students wish to perform their religious service instead of conducting the class during prayer times. Also, this cultural difference affects the view the students have of the teachers. For example, one teacher stated that the students asked her why she was not wearing a headscarf and sent her warning letters (6F).

Discipline problems. The teachers stated that they experienced discipline problems. The most important discipline problem regarding the students is the use of violence. Particularly, the participant teachers observe that the older children use violence on the younger children. According to the teachers, the reason for this discipline problem is that students from different age groups are in the same classroom.

Another discipline problem expressed by the teachers relates to negative behavior and conduct. The participants stated that the Syrian students listened to what the Syrian teachers told them to do but did not heed what the Turkish teachers told them. According to the participants, the reason for this is the frequent use of physical punishment by the Syrian teachers in classroom management. Therefore, the students behave better in the Syrian teachers' classes out of fear.

Views of the Participants Regarding the Relationship between the Syrian Students, and Their Friends, Teachers and Principals in the School

The views of the participants regarding the relationship of the Syrian students with each other, Turkish students, teachers and school administration are summarized in Table 2.

Table 2

Relationship of the Syrian Students with Each Other, Turkish Students, Teachers and School Administration

Communication Groups	*School Type*	*Communication Status*	*Representative Comments*
With each other	Public schools	Well	The Syrian students treat each other well and have a strong communication and dialog among themselves during intermissions and lunch breaks (2M).
	TECs	Based on violence	Their relationship among themselves is based on violence. Even their jokes contain some fight. They are very cruel to each other (8F).
With Turkish students	Public schools	Well	The Syrian students communicate very well with the Turkish students [. . .] Even if they cannot speak Turkish, children somehow understand each other (4F).
	TECs	Based on violence	The Syrian and Turkish students don't like each other (6F). Last week, the Turkish students and the Syrian students got into an armed fight (9F).
With teachers	Public schools	Respectful	They are very respectful to teachers and me. They ask for permission for everything (for bathroom and other needs) (2M).
	TECs	Disrespectful	They are disrespectful of and disregardful toward Turkish teachers.
With school administration	Public schools	Respectful	They are as respectful to the principal as the other students are (2M).
	TECs	Respectful	The relationship of the Syrian students with the school administration is unproblematic and respectful.
School-family relationship	Public schools	Very little communication	The parents speak very little Turkish. Therefore, I can't directly communicate with the family. The family has Turkish neighbors who have students attending this school, and I try to communicate with them through these neighbors (4F).
	TECs	No communication	We have never established communication with their families (6F, 7F, 8F, 9F).

The teachers from the public schools stated that the students were on good terms with both each other (1M, 2M, 3F, 5F) and the Turkish students (1M, 2M, 3F, 4F). The teachers in the TECs stated that the Syrian students' relationship with each other was based on violence. The students in these centers receive education separately from the

Turkish students, and therefore see the Turkish students only at the beginning and end of school. However, according to what the teachers said, even during such short encounters, the Turkish students and the Syrian students fight with each other. As seen, while the relationship of the Syrian students with each other and Turkish students is positive in the public schools, that in the TECs is based on violence. The reason for this may be the very low number of Syrian students in public schools versus the high number thereof in TECs. The Syrian students display a reserved attitude in public schools because of their low number.

When the relationship of the Syrian students with the teachers and the school administration is examined, the teachers from the public schools state that they are highly respectful toward both the teachers and the school administration (1M, 2M, 3F, 4F, 5F). However, the teachers from the TECs state that the Syrian students are disrespectful and disregardful toward them (6F, 7F, 8F, 9F). The participants attribute the reason for this situation, as mentioned above, as being that the relationship between the Syrian teachers and students was based on violence and punishment.

All of the teachers stated that the relationship of the families of the Syrian students with the teachers and the school administration in the public schools was very limited because the families could not speak Turkish (1M, 2M, 3F, 4F, 5F). However, the teachers try to communicate with the Syrian families through the families' Turkish-speaking relatives or neighbors (1M, 3F, 4F, 5F), which shows that the school-family relationship is achieved through the individual efforts of the teachers. The Turkish teachers in the TECs state that they are not able to communicate with the families of the Syrian students because teachers do not speak Arabic (6F, 7F, 8F, 9F).

The Views of the Participants Regarding the Support provided to Teachers and Students during the Integration of the Syrian Students within the School and Classes

The views of the participants regarding the support provided to teachers and students during the integration of the Syrian students with the school and classes are summarized in Table 3.

The support provided to the Syrian students in the education process is explored in two dimensions including the support provided to the teachers and the support provided to the students. It should be noted that the Turkish teachers both in the public schools and the TECs are funded by the MoNE. The Syrian teachers working in the TECs are funded by UNICEF. The teachers in the public schools stated that they received support from neither the school administration nor the non-governmental organizations.

They experienced difficulties due to the lack of adequate support particularly for language learning, adequate school infrastructure (such as lack of Internet connection and counseling services), and adequate school materials for language learning. Non-governmental organizations provide various support to the teachers in the TECs. For example, UNICEF provided cards for Turkish language education to the teachers for use during the classes (7F, 9F). Also, Yunus Emre Institute sent books to the teachers for Turkish language education (6F, 7F, 8F, 9F).

Table 3

The Support Provided to the Syrian Students in the Education Process

	Supporter	*School Type*	*Support*
Support provided to teachers	MoNE	Public Schools	Funding of the Turkish teachers
		TECs	Funding of the Turkish teachers
	School Administration	Public Schools	-
		TECs	-
	Non-Governmental Organizations	Public Schools	-
		TECs	Funding of the Syrian teachers (UNICEF), language cards (UNICEF) and Books for Turkish language education (Yunus Emre Institute)
Support provided to Syrian students	MoNE	Public Schools	Books
		TECs	-
	School Administration	Public Schools	Support education
		TECs	-
	Non-Governmental Organizations	Public Schools	-
		TECs	Books School supply expenses (UNICEF)

As for the support provided to the Syrian students in their education process, the Syrian students receive books provided by the MoNE free of charge to all students. The teachers said that there was no counseling service available for even the Syrian students who had fled from the war. The teachers stated that the school administration directed teachers whose schedule was free to provide support education to those students in one-on-one Turkish language learning sessions; in other words, the only support the school administration provided was for Turkish language learning. Also, the teachers provided reading books to the Syrian students to learn Turkish. In the TECS, on the other hand, the students' books and school expenses are paid by the non-governmental organizations. The Turkish teachers in the TECs use different books, materials and colored drawings to improve the Syrian students' engagement in their classes (6F, 8F, 9F). The teachers motivate the students with trinkets and have them listen to Turkish children's songs (6F, 7F). One of the participant teachers (8F) states, "[. . .] I also use creative drama techniques in the class".

The Recommendations of the Participants for Solutions to the Problems Faced in the Syrian Students' Education

The teachers from the public schools provided some recommendations for solutions to the problems faced in the Syrian students' education. These recommendations are categorized by theme and presented in Table 4.

Table 4

Recommended Solutions to the Problems Faced in the Syrian Students' Education

	Themes	*Representative Comments*
Theme 1:	*Turkish language education for Syrians*	The students definitely need to learn Turkish because they don't understand what is meant even if they can read. The other problems cannot be solved without Turkish (4F).
Theme 2:	*Preparation of materials and a curriculum for Turkish education*	A formal Turkish Language curriculum should be developed as we do not have one (6F). The Turkish study book currently in use is inadequate and intended for adults. Appropriate study books must be prepared for these children [. . .] (8F).
Theme 3:	*Arrangement of separate classes from Turks*	Either a separate classroom must be created for these students, or they must attend separate schools (1M). Syrian students must be taught in an environment where there are no Turkish students (3F).
Theme 4:	*Seminars for Teachers*	We (Turkish teachers) must also be provided with seminars/trainings (7F, 8F).
Theme 5:	*Reduced class population*	The class populations may be reduced. There are at least 40 students in classes (8F, 9F).
Theme 6:	*Other views*	Groups can be arranged based on age (6F). They (Syrians) must be taught by teachers who speak their language (2M).

As seen in Table 4, teachers first said that the Syrian students should definitely be taught the Turkish language, and then should be placed in the same class environment with the Turkish students. Secondly, the teachers pointed to the material and curriculum problems regarding Turkish language teaching and stressed that a special curriculum and materials for use in classes should be developed for Turkish teaching. The third recommendation by the teachers is that separate classrooms must be created for the Syrian students in the schools where they are taught according to a program specifically designed for them.

Teachers also seemed to think that they should definitely be provided with trainings and seminars on Turkish language education as a second language (6F, 7F). The teachers recommended that the class population should be decreased because there were at least 40 students in the current classes (8F, 9F). In addition to the above recommendations, it was recommended that the Syrian students be taught by Arabic speaking teachers (2M), and that the Syrian students be placed in classes based on their age and not on their academic levels (6F).

Discussion and Conclusion

This research aims to identify the difficulties faced by the teachers working in the public schools and TECs in the process of educating Syrian students. The primary difficulty faced by the teachers in the education process was identified to be the language barrier. Hence, many studies have identified the Syrian students' lack of language skills to be a major problem in the education process (Human Rights Watch Report, 2015; Bilgi University Report 2015; Dincer et al, 2013; Bircan & Sunata, 2015, 235; Nayir, 2017). Therefore, placing the Syrian students in formal education classes regardless of their Turkish language skills reflects adversely on the education process. Instead, the students may first be placed in classes created as preparatory classes based on their Turkish language levels as in Germany, Sweden and the Netherlands. After achieving proficiency in the Turkish language to a satisfactory level degree, the students may then be placed in formal education classes. Such practices would enable the Syrian students to better comprehend the subjects taught in the other subjects. Hence, the teachers who participated in the research stated that the Syrian students should first be taught the Turkish language as the solution to the difficulties faced in their education. Besides, it may be useful to employ translators at schools educating Syrian students to overcome the language barrier.

Another difficulty faced in the process of the education of these students was said to be the cultural problem. Differences between the Turkish and Syrian education systems, and the cultural structures, such as the fact that girls and boys receive education separately in Syria and that the Syrian families want their girls to wear a burqa in school, were noted in previous studies (Nielsen & Grey, 2013; UNICEF, 2014) to be among the cultural and political barriers to education for Syrian students. However, according to the findings of the present study, the differences between the education systems cause the Turkish teachers to experience a cultural conflict. The differences between the Turkish and Syrian cultures therefore impact the education process adversely. In order for the Syrian and Turkish students and teachers to learn about each other's culture, cultural mediators may be employed in schools as in Italy, and integrative activities may be organized in schools with the help of such experts.

As far as the relationships of the Syrian students in the school are concerned, the students in the public schools appear to be on good terms with each other, their classmates and the teachers. However, those in the TECs appear to be on good terms with neither each other nor the teachers. In that case, it may be suggested that teaching the Syrian students separately from

the Turkish students causes them to become more isolated from society while continuing in their old habits, rather than helping them adapt to society. Hence, Krakow (2016) criticizes teaching refugees separately from their German peers in Germany and suggests that the refugees could learn the language and culture faster, have a stronger belief that they are wanted by society, and adapt to society more quickly if they were taught together with the Germans. Therefore, it may be more helpful if the Syrian students are taught together with the Turkish students for them to better adapt to Turkish society.

It is very important that teachers receive support to enable the Syrian students to adapt to the school and classes. However, during the interviews, the teachers working in the public schools stated that they received support from neither the school nor the MoNE nor any non-governmental organization. Teachers in the TECS are provided with various support by non-governmental organizations for language teaching (language cards, etc.); however, according to the teachers, such support is insufficient. The study conducted by Aras and Yasun (2016) also revealed that the teachers lacked qualifications particularly for teaching Turkish to non-native speakers. The teachers stated that it would be helpful if the Turkish Language and Literature teachers teaching the Syrian students were provided with seminars on how to teach Turkish to non-native speakers. This is corroborated by another study that stated that the Turkish teachers needed professional development and support to work with the Syrian students (Cinkir, 2015, 53). Therefore, as expressed by the participants, developing a curriculum for Turkish language education and preparing and sharing appropriate books and materials with teachers would contribute to the improvement of the education process. Another study conducted by the Norwegian Refugee Council (2013; as cited by Abu-Amsha, 2014, 33) also revealed that the Lebanese teachers did not receive any support for the education of Syrian students. Whereas it will be helpful if the teachers teaching the Syrian students receive training on psychological support and protection, special education and child-focused pedagogic practices (World Bank 2013b; as cited by Abu-Amsha, 2014, 33). Therefore, it is important to provide support to teachers in order to ensure a quality educational experience for all students.

When the support provided to the students was examined, the students in public schools were provided with free course books by the MoNE, and the school administration in these schools provides support education to the students for Turkish language learning. However, according to the teachers, the Syrian students attending public schools do not receive adequate support. For example, the teachers stated that there were no counseling services

available for the Syrian students who had fled from the war. Bircan and Sunata (2015, 235) stressed the importance of the matter in their research by stating, "the needs of Syrian child refugees in psychological support are still at issue." Book and school supply expenses of the students in TECs are paid by non-governmental organizations.

Turkish language teaching of is at the top of the list of the recommendations of the teachers to solve the problems faced in the process of educating Syrian students as, according to the participants, reading and writing cannot be effectively achieved because of the students' lack of Turkish language skills. However, education provided to the Syrians should not be regarded solely in terms of teaching them how to read and write because, as Matthews (2008, 42) said, "literacy is critically important, but schools are not simply literacy delivering machines." Learning environments and areas must be created at schools where all students can participate, and efforts must be made to enhance the students' engagement in the education process and their intercommunication and friendships.

In conclusion, the steps taken by Turkey for the education of the Syrian children arriving in the country are crucially important for the future lives of these children. Naturally, teaching so many refugees is a new phenomenon for Turkey. In order to properly overcome this issue, the MoNE must be in close contact with the teachers and educational institutions working in this process. Also, in order to ensure healthy functioning of this process, the need for support from both the general public and the national and international organizations should not be ignored.

References

Abu-Amsha, O. (2014). *The resilience of Syrian refugee children in Lebanon: What do they need to acquire quality education? Education Resilience Approaches (ERA) program; systems approach for better education results (SABER).* Washington, DC: World Bank Group. Retrieved from http://documents.worldbank.org/curated/en/262571468091188809/The-resilience-of-Syrian-refugee-children-in-Lebanon-What-do-they-need-to-acquire-quality-education (accessed 12 January 2016).

Aras, B., & Yasun, S. (2016) *The educational opportunities and challenges of Syrian refugee students in Turkey: Temporary education centers and beyond.* Istanbul Policy Center-Sabancı University-stiftung mercator initiative. Retrieved from http://ipc.sabanciuniv.edu/publication/the-educational-opportunities-and-challanges-of-syrian-refugee-students-in-turkey-temporary-education-centers-and-beyond/?lang=en

Beste, A. (2015). Education provision for Syrian refugees in Jordan, Lebanon and Turkey Preventing a "Lost Generation". Retrieved from https://gcm.unu.edu/publications/policy-reports/education-provision-for-syrian-refugees-in-jordan-lebanon-and-turkey-preventing-a-lost-generation.html

Bircan, T., & Sunata, U. (2015). Educational assessment of Syrian refugees in Turkey. *Migration Letters, 12*(3), 226 – 237.

Cinkir, S. (2015). Turkey. In *Curriculum, accreditation and certification for Syrian Children in Syria, Turkey, Lebanon, Jordan, Iraq and Egypt* (pp. 40-55). UNICEF.

Creswell, J. W. (2012). *Qualitative ınquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Crul, M., Keskiner, E., Schneider, J., Lelie, F., & Ghaeminia, S. (2016). *No lost generation: Education for refugee children a comparison between Sweden, Germany, The Netherlands and Turkey.* Retrieved from http://www.eui.eu/Projects/TheForum/Documents/Keynote-lecture-education.pdf (accessed 18 October 2016).

Deutche Welle. (2015). *Syrians are permanent in Turkey.* Retrieved from http://www.dw.com/tr/suriyeliler-t%C3%BCrkiyede-kal%C4%B1c%C4%B1/a-18654827. (accessed 13 January 2016).

Dincer, O.B., Federici, V., Ferris, E., Karaca, S., Kirisci, K., & Ozmenek Carmikli, E. (2013). *Suriyeli mülteciler krizi ve Türkiye- Sonu gelmeyen misafirlik (Syrian refugees crisis and Turkey- Non-ending guest).* Retrieved from https://www.brookings.edu/wp-content/uploads/2016/06/USAKBrookings-report-final-version14November13.pdf (accessed 14 January 2016).

Dillioglu, B. (2015). *Suriyeli multecilerin entegrasyonu egitim ve istihdam politikalari.* Ankara: Uluslararası Terörizm ve Güvenlik Araştırmaları Merkezi.

Emin, M.N. (2016). *Turkiye'deki Suriyeli cocuklarin egitimi temel egitim politikalari* (*Education of Syrian children in Turkey basic education policies*). Istanbul: SETA.

Erguven, N.S., & Ozturanli, B. (2013). Uluslararasi multeci hukuku ve Turkiye (International refugee law and Turkey). *Ankara Universitesi Hukuk Fakultesi Dergisi, 62*(4), 1007-1061.

Ersoy, F. (2016). Fenomenoloji. In A. Saban & A. Ersoy (Ed.), *Egitimde nitel arastirma desenleri* [Qualitative research models in education] (pp. 51-110). Ankara: Ani Publishing.

Glesne, C. (2013). *Nitel arastirmaya giris [Becoming qualitative researchers: An introduction]* (A. Ersoy & P. Yalcinoglu, Trans. Ed.). Ankara: Ani Publishing.

Harb, C., & Saab, R. (2014). *Social cohesion and intergroup relations: Syrian refugees and Lebanese nationals in the Bekaa and Akkar.* Retrieved from http://www.alnap.org/resource/12632 (accessed 25 October 2016).

Hoel, M. (2015). *The European Union's response to the Syrian refugee crisis: An analysis of the response of Member States and EU institutions* (Master's thesis in European Studies). Norwegian University of Science and Technology. Retrieved from https://brage.bibsys.no/xmlui/bitstream/id/413167/Hoel,%20Maria.pdf

Human Rights Watch Report. (2015). *'When I picture my future, I see nothing' barriers to education for Syrian refugee children in Turkey.* Retrieved from https://www.hrw.org/et/node/283247

Istanbul Bilgi University Child Studies Unit Report. (2015). Suriyeli multeci cocuklarin Turkiye devlet okullarindaki durumu politika ve uygulama onerileri (Situation of Syrian refugee children in Turkish state schools, policy and suggestions for application). Retrieved from http://cocuk.bilgi.edu.tr/wp-content/uploads/2015/09/Suriyeli-Cocuklar-Egitim-Sistemi-Politika-Notu.pdf

Ingleby, D., & Kramer, S. (2012). The Netherlands. In Nonchev, A. & N. Tagarov (Ed.), *Integrating refugee and asylum seeking children in the educational systems of EU member states* (pp. 250-279). Sofia: CSD.

Italia, A., & Bellesi, L. (2016). Italy. In Nonchev, A. & N. Tagarov (Ed.), *Integrating refugee and asylum seeking children in the educational systems of EU member states* (pp. 136-154). Sofia: CSD.

Krakow, K. (2016). *Are current efforts to educate immigrants in Germany living up to the task?* Retrieved from http://www.humanium.org/en/educating-immigrants-in-germany/(accessed 17 October 2016).

Matthews, J. (2008). Schooling and settlement: Refugee education in Australia. *International Studies In Sociology of Education, 18*(1), 31-45.

Mayer, R. (2015). Research and analysis paper: Education for Syrian Children -The imperative of improving Syrian refugee children's access to education in Lebanon.

Maykut, P., & Morehouse, R. (1994). *Beginning qualitative research.* London and Washington, DC: The Falmer Press.

Merriam, S.B. (2013). *Nitel arastirma desen ve uygulama icin bir rehber [Qualitative research: A guide to design and ımplementation]* (S. Turan, Trans. Ed.). Ankara: Nobel Publishing.

Milliyet. (2014). *Turkiye'deki Suriyelilere ozel okullar (Private schools for Syrians in Turkey).* Retrieved from http://www.milliyet.com.tr/turkiye-deki-suriyelilere-ozel/gundem/ydetay/1828752/default.htm.

MoNE (2016). *Suriyeli ogrenciler icin 4 bin 200 ogretmen alimi yapilacak (4 thousand 200 teachers will be purchased for Syrian students).* Retrieved from http://www.meb.gov.tr/suriyeli-ogrenciler-icin-4-bin-200-ogretmen-alimi-yapilacak/haber/12204/tr (accessed 15 November 2016).

Nayir, F. (2017). Problems and suggested solutions during refugee education process. In *Innovation and Global Issues in Social Sciences Congress publications* (pp. 120-122). 27-29 April 2017.

Neugebauer, R. (2013). *Bleak outlook for children who are refugees*. Report on Children Who Are Refugees.

Nielsen, Y., & Grey, M.A. (2013). *Schooling in a crisis: The case of Syrian refugees in Turkey.* Humanitarian Practice Network (HPN)59. Retrieved from http://odihpn.org/magazine/schooling-in-a-crisis-the-case-of-syrian-refugees-in-turkey/(accessed 20 June 2016).

Ozturk, S. (2015). *Effects of Syrian migration.* GSA 2015 International Conference: Living the Global (Roehampton University, London). Retrieved from http://docplayer.biz.tr/3859709-Suriye-den-gocun-etkileri-gsa-2015-global-i-yasamak-uluslar-arasi-konferansi-roehampton-universitesi-londra-ingiltere-bildirisi.html.

Rothschild, N. (2016). On the frontline of integration: How Swedish schools are helping refugees. Retrieved from https://www.theguardian.com/teacher-network/2016/jun/26/on-the-frontline-of-integration-how-swedish-schools-are-helping-refugees (accessed 18 October 2016).

Seah, C.H., & Wilson, A. (2011). A Phenomenological study of university students' grieving experiences. *Illness, Crisis & Loss, 19*(1), 3-25.

Sirin, S. R., & Rogers-Sirin, L. (2015) *The educational and mental health needs of Syrian refugee children.* Washington, DC: Migration Policy Institue.

Topcuoglu, R. (2012). *Turkiye'de gocmen cocuklarin profili, sosyal politika ve sosyal hizmet onerileri hizli degerlendirme arastirmasi.* Retrieved from http://www.turkey.iom.int/documents/Child/IOM_GocmenCocukRaporu_tr_03062013.pdf (accessed 08 March 2017).

UNESCO. (2011). *The hidden crisis: Armed conflict ad education. No lost generation: Protecting the future of children affected by the crises in Syria.*

UNESCO. (2017). *Refugee education in Guinea: Development of curricula and examinations oriented to repatriation.* Retrieved from http://www.unesco.org/education/educprog/emergency/casestudy/guine e.htm (accessed 01 March 2017).

UNHCR. (2018). Syria regional refugee response. Retrieved from https://data2.unhcr.org/en/situations/syria (accessed 15 May 2018).

UNICEF. (2014). *No lost generation, protecting the futures of children affected by the crisis in Syria.* Retrieved from https://www.unicef.org/appeals/files/No_Lost_Generation_Strategic_Overview__January_2014.pdf(accessed 16 July 2016).

Uzun, A. (2015). Gunumuzun sosyal ve ekonomik sorunu olan Suriyelilerin multeci ve ekonomi hukuku bakimindan degerlendirilmesi (The evaluation of the Syrians as Turkey's current social and economic problem under refugee law and law and economics). *Ankara Barosu Dergisi, 1*, 105-120.

Watkins, K., & Zyck, S.A. (2014). *Living on hope, hoping for education-The failed response to the Syrian refugee crisis.* Retrieved from https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/9169.pdf(accessed 24 November 2016).

Weise, Z. (2016). *Refugee children get little education in Europe.* Retrieved from http://www.politico.eu/article/refugee-children-education-europe-volunteers-greece-migrant-camp/ (accessed 17 October 2016).

Yildirim, A., & Simsek, H. (2011). *Sosyal bilimlerde nitel arastirmaya giris [Qualitative research methods in the social sciences]*. (8th Edition). Ankara: Seckin Publishing.

Suriyeli Öğrencilerin Eğitimi: Türkiye'de Öğretmenlerin Karşılaştığı Sorunlar

Atıf:

Taskin, P., & Erdemli, O. (2018). Education for syrian refugees: problems faced by teachers in turkey. *Eurasian Journal of Educational Research, 75,* 155-178, DOI: 10.14689/ejer.2018.75.9

Özet

Problem Durumu: İnsanlık tarihi kadar eski bir olgu olan göç konusunun özellikle küresel olarak artan nüfus hareketliliği ile son zamanlarda tartışılan en önemli konulardan biri olduğu görülmektedir. İnsanlar daha iyi bir yaşam elde etmek için gönüllü olarak başka ülkelere göç edebileceği gibi, savaş gibi olağanüstü sebeplerden dolayı zorunlu olarak da başka ülkelere göç edebilmektedir.

Göç olgusu öncelikle erkekler ve emek göçü bağlamında tartışılırken, 1980'lerden sonra kadınlar, 2000'lerde ise çocuklar bağlamında tartışılmaya başlanmıştır. Özellikle birçok ülkede mülteci çocuk nüfusunda yaşanan artışlar, yeni bir göçmen grup olarak çocukları ortaya çıkarmıştır (Topçuoğlu, 2012). Bu durum ise ülkelerin göçmen politikalarında çocukları dikkate almasını gerekli kılmıştır. Dolayısıyla göçmen çocuklara yönelik eğitim politikaları da tartışılmaya başlanmıştır.

Diğer göçmenlerle karşılaştırıldığında, mülteci çocukların yetersiz beslenme, hastalık, fiziksel yaralanmalar, beyin hasarı ve cinsel veya fiziksel taciz ile ilişkili ciddi problemlere sahip olma olasılığı daha yüksektir (Neugebauer, 2013). Bu açıdan eğitime erişimin mülteci çocuklar için kritik bir öneme sahip olduğu söylenebilir. Çünkü okula gitme hakkından faydalanamayan mülteci çocuklar; daha fazla istismar ve kötü muamele riski altında bulunmakta, travma sonrası stres bozukluğu semptomlarını daha fazla göstermekte ve fiziksel ve psikolojik gelişimlerini tamamlamayı başaramamaktadırlar (UNESCO, 2011). Ayrıca mültecilerin eğitiminde yaşanan krizler çocuk işçiliğine ve erken evliliğe neden olmaktadır. Eğitim fırsatını kaybeden gençlerin radikal gruplara sürüklenme riski de doğmaktadır (Watkins & Zyck, 2014). O halde mülteci çocukların eğitime erişiminin sağlanması ve bu çocukların eğitim sürecini engelleyen sebeplerin tespit edilerek çözülmesi onların iyi bir hayata sahip olması açısından önemlidir.

Araştırmanın Amacı: Bu araştırmanın amacı, Ankara ilinde yasal olarak kayıtlı bulunan Suriyeli öğrencilerin eğitim gördüğü devlet okullarında ve geçici eğitim merkezlerinde görev yapan öğretmenlerin, geçici koruma kapsamındaki Suriyeli öğrencilerin eğitim sürecinde karşılaştıkları güçlükleri tespit etmek ve çözüm önerileri getirmektir.

Araştırmanın bu genel amacı doğrultusunda aşağıdaki sorulara yanıt aranacaktır: Öğretmen görüşlerine göre;

1) Geçici koruma kapsamındaki Suriyeli öğrencilerin eğitimi sürecinde yaşanan sorunlar nelerdir?

2) Geçici koruma kapsamındaki Suriyeli öğrencilerin okul ortamındaki ilişkileri nasıldır?

3) Geçici koruma kapsamındaki Suriyeli öğrencilerin okula ve derslere uyumları sürecinde öğretmenlere ve öğrencilere sağlanan destekler nelerdir?

4) Geçici koruma kapsamındaki Suriyeli öğrencilerin eğitimleri sürecinde karşılaşılan sorunların çözümüne ilişkin öneriler nelerdir?

Araştırmanın Yöntemi: Bu çalışma nitel araştırma modeli benimsenerek yapılmıştır. Araştırmada nitel araştırma desenlerinden fenomenoloji deseni kullanılmıştır. Araştırma kapsamında hem devlet okulunda hem de geçici eğitim merkezinde görev yapan öğretmenlerle görüşme yapılmıştır. Çalışma grubunun belirlenmesinde amaçlı örnekleme tekniğinden yararlanılmıştır. Bu kapsamda sınıfında Suriyeli öğrenci bulunan 9 öğretmen çalışmaya dâhil edilmiştir. Öğretmenlerin beşi devlet okulunda, dördü geçici eğitim merkezinde görev yapmaktadır. Öğretmenlerin yedisi kadın, ikisi erkektir. Eğitim durumu açısından ise öğretmenlerin altısı lisans, üçü yüksek lisans mezunudur. Devlet okullarında görev yapan öğretmenlerin hepsi sınıf öğretmenidir ve kıdemleri 9 yıl ile 16 yıl arasında değişmektedir. Bu öğretmenlerden hiçbiri daha önce geçici koruma kapsamındaki Suriyeli öğrencilerle ilgili herhangi bir eğitime ya da seminere katılmamıştır. Geçici eğitim merkezinde görev yapan öğretmenlerin hepsi Türk Dili ve Edebiyatı öğretmenidir. Bu okullarda görev yapan öğretmenler mesleklerinde henüz bir yıllarını doldurmamış ve Milli Eğitim Bakanlığı (MEB)

tarafından sözleşmeli öğretmen statüsünde istihdam edilmektedir. Ayrıca öğretmenlerin ilk görev yerleri bu okuldur. Öğretmenler benzer şekilde daha önce geçici koruma kapsamındaki Suriyeli öğrencilerle ilgili herhangi bir eğitime ya da seminere katılmamıştır.

Bu çalışmada veriler öğretmenlerle yapılan görüşmeler aracılığıyla toplanmıştır. Çalışmada iki farklı görüşme tekniğinden yararlanılmıştır. Devlet okullarında görev yapan öğretmenlerle yarı yapılandırılmış görüşme, geçici eğitim merkezinde görev yapan öğretmenlerle ise odak grup görüşmesi yapılmıştır. Araştırmacılar tarafından yarı yapılandırılmış görüşme formu geliştirilerek görüşmeler bu doğrultuda yapılmıştır. Her bir görüşme yaklaşık 25-30 dakika sürmüştür. Odak grup görüşmesi ise 45 dakika sürmüştür. Katılımcılardan toplanan veriler içerik analizi tekniği ile analiz edilmiş ve Nvivo 10 Paket Programından yararlanılmıştır. Veri analizi iki araştırmacı tarafından beraber yapılmıştır. Öğretmenlerden doğrudan alıntı yapılan görüşler küçük puntoyla aktarılmış ve katılımcının kodu parantez içerisinde sunulmuştur. Katılımcılara kod verilirken sırayla kendilerine numara verilmiş (devlet okullarında görev yapan öğretmenlere 1-5 arasında, geçici eğitim merkezinde görev yapan öğretmenlere 6-9 arasında numara verilmiştir) ve numaranın yanına cinsiyetine göre F (female) veya M (male) harfi eklenmiştir.

Araştırmanın Bulguları: Katılımcıların Suriyeli öğrencilerin eğitimi sürecinde karşılaştıkları sorunlara ilişkin görüşleri "dil bariyeri", "kültürel problemler" ve "disiplin problemleri" olarak üç tema altında toplanmıştır. Öğretmenlerin görüşlerine göre Suriyeli çocukların eğitimi sürecinde yaşanan birincil sorun "dil bariyeri"dir.

Geçici koruma kapsamındaki Suriyeli öğrencilerin okul ortamındaki ilişkileri Suriyeli öğrencilerin birbirleriyle, Türk öğrencilerle, öğretmenlerle ve okul yönetimi ile ilişkileri şeklinde dört ana tema altında değerlendirilmiştir. Suriyeli öğrencilerin eğitim sürecinde sağlanan destekler, "öğretmenlere sağlanan destekler" ve "öğrencilere sağlanan destekler" olmak üzere iki boyutta incelenmiştir.

Devlet okullarındaki öğretmenler, Suriyeli öğrencilerin eğitim sürecinde karşılaştığı sorunlara çözüm önerileri getirmiştir. İlk olarak öğretmenler Suriyeli öğrencilere mutlaka Türkçe öğretilmesi daha sonra Türk öğrencilerle birlikte sınıf ortamına alınması gerektiğini belirmişlerdir. İkinci olarak, öğretmenler Türk dili öğretimi ile ilgili materyal ve müfredat sorunlarına dikkat çekmiş ve Türkçe öğretimi için özel bir müfredat ve derslerde kullanılacak materyallerin geliştirilmesi gerektiğini vurgulamışlardır. Öğretmenlerin üçüncü önerisi, Suriyeli öğrenciler için okullarda ayrı sınıfların oluşturulması ve bu öğrenciler için özel olarak hazırlanmış bir programa göre eğitim verilmesidir.

Araştırmanın Sonuçları ve Önerileri: Suriyeli öğrencilerin eğitim sürecinde karşılaştıkları güçlükleri belirlemeyi amaçlayan bu araştırmada, öğretmenlerin eğitim sürecinde birincil güçlük olarak dil problemi ile karşılaştıkları tespit edilmiştir. Nitekim yapılan birçok çalışmada (Human Rights Watch Raporu, 2015; Bilgi Üniversitesi Raporu 2015; Dinçer vd., 2013; Bircan & Sunata, 2015, 235; Nayır, 2017) da Suriyeli öğrencilerin dil bilmemelerinin eğitim sürecinde önemli bir problem olduğu saptanmıştır.

Dil probleminin yanı sıra öğrencilerin eğitim sürecinde yaşanan diğer bir önemli güçlük kültürel problem olarak belirlenmiştir. Suriye'de kız ve erkeklerin ayrı ayrı eğitim görmesi, Suriyeli ailelerin kızlarının okulda çarşaf giymelerini istemeleri ya da okula devam etmek yerine evlenmelerini tercih etmeleri gibi Türkiye ve Suriye'nin eğitim sistemi ve kültürel yapısındaki farklılıklar önceki çalışmalarda (Nielsen and Grey, 2013; UNICEF 2014) Suriyeli öğrencilerin eğitime erişmelerinin önünde kültürel ve politik engeller olarak algılanmaktaydı. Ancak bu çalışmanın bulgularına göre, eğitim sistemlerindeki farklılıklar Türk öğretmenlerin kültürel çatışma yaşamasına yol açmaktadır. Bu nedenle Türk ve Suriye kültürü arasındaki farklılıklar eğitim sürecini olumsuz etkilemektedir. Suriyeli ve Türk öğrenci ve öğretmenlerin birbirlerinin kültürlerini tanıyabilmeleri için, İtalya'da olduğu gibi okullarda cultural mediators istihdam edilebilir; bu uzmanların da yardımıyla okullarda kaynaştırıcı etkinlikler düzenlenebilir. Atanan bu mediator'lar sayesinde Arapça bilmeyen Türk öğretmenler ile Türkçe bilmeyen Suriyeli aileler arasında iletişim kurulabilir ve böylelikle Suriyeli öğrencilerin akademik başarıları artabilir.

Suriyeli öğrencilerin okul ortamındaki ilişkilerine yönelik öğretmen görüşleri değerlendirildiğinde, devlet okullarında öğrencilerin hem birbirleriyle, hem sınıf arkadaşlarıyla hem de öğretmenleriyle ilişkilerinin iyi olduğu belirlenmiştir. Ancak geçici eğitim merkezinde öğrencilerin hem birbirleriyle hem de öğretmenleriyle ilişkilerinin pek iyi olmadığı tespit edilmiştir. O halde Suriyeli öğrencilerin Türk öğrencilerden ayrı bir şekilde öğrenim görmelerinin onların topluma uyum sağlamaktan ziyade toplumdan daha fazla izole olmalarına ve eski alışkanlıklarını devam ettirmelerine sebep olduğu söylenebilir.

Suriyeli öğrencilerin okula ve derslere uyumlarını sağlamak için öğretmenlerin destek görmeleri oldukça önemlidir. Ancak yapılan görüşmelerde, devlet okullarında görev yapan öğretmenler ne okul yönetiminden ne de MEB ve sivil toplum kuruluşlarından bir destek gördüklerini belirtmişlerdir. Geçici eğitim merkezlerinde görev yapan öğretmenlere, dil öğretimi için sivil toplum kuruluşları tarafından çeşitli destek sağlanmaktadır (dil kartları vb.). Ancak, öğretmenlere göre, bu destek yeterli değildir. Bu kapsamda Suriyeli öğrencilere eğitim verecek Türk Dili ve Edebiyatı öğretmenlerine yabancılara Türkçe öğretimi konusunda seminer verilmesinin faydalı olacağı katılımcı öğretmenler tarafından dile getirilmiştir.

Sonuç olarak, Suriyeliler için açık kapı politikasını benimseyen Türkiye'nin, ülkeye gelen Suriyeli çocukların eğitime erişimi için attığı adımlar bu çocukların gelecek yaşamları için hayati önem taşımaktadır. Tabi ki Türkiye için bu kadar çok sayıda mülteciye eğitim vermek yeni bir meseledir. Bu meselenin hakkıyla üstesinden gelebilmek için bu süreçte görev alan öğretmenler ve eğitim kurumlarıyla MEB yakın ilişki içinde olmalıdır. Ayrıca sürecin sağlıklı bir şekilde işleyişi için hem kamuoyunun hem de ulusal ve uluslararası kuruluşların desteğine ihtiyaç duyulduğu göz ardı edilmemelidir.

Anahtar sözcükler: Suriyeli öğrenciler, eğitim zorlukları, göçmenlerin eğitim hakkı.

Eurasian Journal of Educational Research 75 (2018) 179-196

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

The Effects of Sample Size and Missing Data Rates on Generalizability Coefficients

Sumeyra SOYSAL[1] , Haydar KARAMAN[2] , Nuri DOGAN[3]

ARTICLE INFO

Article History:
Received: 1 Oct. 2016
Received in revised form: 5 Sep. 2017
Accepted: 4 Apr. 2018
DOI: 10.14689/ejer.2018.75.10

Keywords
Reliability, G coefficient, phi coefficient, zero imputation, MCAR, generalizability theory, matrix of missing data

ABSTRACT

Purpose of the Study: Missing data are a common problem encountered while implementing measurement instruments. Yet the extent to which reliability, validity, average discrimination and difficulty of the test results are affected by the missing data has not been studied much. Since it is inevitable that missing data have an impact on the psychometric properties of measurement instruments, it was considered necessary to investigate this topic. Depending on the identified need, a simulative study was conducted on the effects of missing data on reliability. The reliability estimates were discussed in terms of generalizability theory (G theory). **Research Methods**: Depending on the research questions, complete data sets having different sample sizes (100, 200, 400, 1000) in weak and strong one-dimensional structures under normal distribution were produced. Missing data sets were created by deleting data at different rates (5%, 10%, 20%, 30%) randomly from the complete sets. **Findings and Results:** When the estimates obtained by missing and complete data sets were compared, it was found that G and phi coefficients were significantly affected for the weak one-dimensional design when the missingness was 20% and more. However, for the strong one-dimensional design, those coefficients were negligibly affected even when the missingness was 30%. Moreover, it was also found that the estimates obtained by missing coded incorrect in particularly weak one-dimensional data were lower than the estimates from missing data matrix. Also error statistics of the weak one-dimensional data based on missing coded incorrect were significantly higher than their strong one-dimensional data counterparts, especially at the rates of 20% and 30% missingness. **Implications for Research and Practice** Thus, missing coded incorrect is not suggested to be used as a missing data treatment method in reliability estimations. Instead, generalizability theory, which allows us to conduct analysis with missing data in matrices, might be recommended.

[1] Corresponding author, Hacettepe University, TURKEY, e-mail: sumeyrasoysal@hotmail.com, ORCID: https://orcid.org/ 0000-0002-7304-1722
[2] Karamanoğlu Mehmet Bey University, TURKEY, e-mail: karaman28haydar@gmail.com, ORCID: https://orcid.org/0000-0002-8212-5565
[3] Hacettepe University, TURKEY, e-mail: nuridogan2004@gmail.com, ORCID: https://orcid.org/ 0000-0001-6274-2016

Introduction

One of the problems often encountered in research data collection and analysis is missing data. Missing data can be defined as the difference between the planned set of data and the obtained set of data (Longford, 2005, p. 13). The data set having no missing data is called a complete data matrix, while the data set with missing data is called an incomplete data matrix. The results obtained with statistical methods applied to a complete data matrix can be quite different from the ones obtained with the same statistical methods applied to an incomplete data matrix (Enders, 2010). This is called the missing data problem in statistics. Missing data are an important problem for all branches of science concerned with collecting numerical data. The problem of missing data is more manifest especially in cases where data collection has considerably high costs (Rubin, Witkiewitz, Andre, & Reilly, 2007). Because the missing data in a data matrix is likely to spoil the structure of the matrix, statistical analyses will yield erroneous results and/or biased statistical estimates. Thus, missing (or lacking) values reduce the quality of the data and may risk the reliability of statistical analyses. Methods to eliminate the problem of missing data should be used in order to raise the quality of measurement results containing missing data (Aydilek, 2013; Howell, 2008).

Two elements playing significant roles in the effects of missing data on statistical estimations is the rate of missing data and the design of missing data. It is expected according to research findings that estimation bias increases as the proportion of missing data in the total data increases (Bakis & Goncu, 2015; Cool, 2000). In parallel to the decrease in the ratio of missing data to the total data, the effects on statistical estimations can also be negligible. On the other hand, missing at random (MAR), missing at completely random (MCAR), or not missing at random (NMAR) are categories of missing data (Enders, 2010; Schlomer, Bauman, & Card, 2010). Since the way missing data behaves in terms of distribution is considerably influential in statistical estimations (Enders, 2010; Schafer & Graham, 2002; Zhu, 2014), statistical tests have been developed to determine the distribution of missing data (Little, 1988).

Although studies concerning missing data were started in the early 1900s, they accelerated with such studies as "Inference and Missing Data" by Rubin (1976) and "Statistical Analysis with Missing Data" by Little and Rubin (1987). A review of studies concerning missing data demonstrates that the studies mostly focus on the effects of statistical analysis results on missing data and that they also focus on the effects of value assignment to missing data in different methods on statistical analysis results. A great number of studies conducted by scientists of different branches on the effects of missing data on estimated statistics as well as studies about the effects of missing data assignment methods on statistical analysis results are available in the literature.

For instance, Kose and Oztemur (2014) compared the techniques of multiple coding, listwise deletion and pairwise deletion in their study concerning the variance analysis of missing data and its effects on t test results. Gu and Matloff (2015) also used the same three techniques to study the effect of missing data on regression analysis. They concluded that the method of deletion according to matching had better performance than the other two methods. Cool (2000) investigated the effects of deletion and

averaging for missing data on regression estimations and concluded that deletion methods reduced the power of statistical analyses since they shrank sample size. Bakis and Goncu (2015), in their study about biased estimations made by using two coding methods for the incomplete data about flow rate measurements of a stream, concluded that an excessive rate of missing data increased the level of bias in both methods. Ser and Bati (2015), in their study of repeated data in animal husbandry, pointed out that they used a multiple coding method for the missing data in their analysis of general linear mixed model and obtained results similar to the ones obtained through complete data. Shang, Liu, Cheng and Cheng (2016) researched the effects of missing data on the results of component analysis. In a study performed by Yilmaz (2014) on missing data in the field of medicine it was found that coding missing data through closest neighborhood and random forests methods could produce similar solutions in problems of statistical classification, and that the random forests method was preferable in highly related data sets. Some researchers compared the classical techniques with current coding techniques for missing data and they found that multiple coding and maximum likelihood were more advantageous than classical techniques (Allison, 2001; Aydilek, 2013; Baraldi & Enders, 2009; Graham, 2009; Graham, 2012; Nakai, & Ke, 2011; Piggot, 2001; Sari, 2012; Schlomer, Bauman, & Card, 2010). Statistical approaches such as ANOVA, longitudinal development models, structural equation models, regression, logistic regression, general linear models and classification models formed the basis for the comparisons of missing data coding in the abovementioned studies. Allison (2001), Horton and Clainman (2007), Soley-Bori (2013), Whang, Zhang and Tong (2014), on the other hand, considered 9 different software programs and 13 missing data coding techniques in their study introducing the techniques for missing data coding and the statistical software to apply the techniques.

Peng, Harwell, Liou and Ehman (2002), in their study aiming to identify which methods had been used for missing data in articles published in journals of education, examined the studies published in 11 journals in the period between 1998 and 2002. The researchers pointed out that missing data were available in 54% of the studies in the 11 journals, whereas there was no information about data in 18% of the studies. They also found that listwise deletion was used in 87% of the studies, with pairwise deletion used in 7% of the studies, no explanations offered in 3%, and five different methods of coding used in the remaining 3% of the studies.

On reviewing the literature, it was found that the number of studies concerning the effects of missing data on the psychometric properties of measurement instruments used in education and in psychology were limited. In one such study, Weaver and Maxwell (2014) researched the effects of coding missing data on the basis of expectation maximization technique on exploratory factor analysis and on the results of reliability and found it more useful than average coding on the basis of data deletion. Demir (2013) and Cum and Gelbal (2015) researched the effects of missing data coding on confirmatory factor analysis model-data fit values and obtained evidence that relatively new missing data coding methods yielded better results. Nartgun (2015) compared the methods of deletion based on a list, series mean, mean of nearby points, multiple coding and regression coding, which were used in solving the problem of missing data under

such conditions as completely random missing mechanism, normal distribution, one-dimensional structure, different sample sizes (n=150; n=650) and different rates of missing data (5%; 10%; 20%). Comparisons were made through the psychometric properties of the scales (eigenvalue, explained variance and Cronbach's alpha) and through statistics calculated from the scores.

Although the literature review showed that the problem of missing data was a common problem encountered in implementing measurement instruments, the review also made it clear that the direct effects on the psychometric properties of measurement instruments were not often considered. In particular, without any methods of missing data imputation and missing data deletion, the issues of how and to what extent reliability and validity of measurement results and such statistics as average discrimination and difficulty are influenced by missing data were not investigated in any depth.

It is common for participants in quantitative studies not to give a response due to various reasons when they are given achievement tests, attitude scales, questionnaires, etc. Participants may sometimes leave a question unanswered due to such reasons as having no idea, failing to find an appropriate answer, skipping a question inadvertently, or not marking the answer correctly. However, as the number of answers to measurement instruments decreases or as missing data increases, the amount of information gathered will decrease and the validity and reliability of measurement results will be expected to fall. It is inevitable that missing data will influence the psychometric properties of measurement instruments used in education and psychology. Therefore, it is believed that the effects of missing data on the psychometric properties of measurement instruments need to be researched. Due to this need, a decision was made to study the effects of missing data on reliability – a psychometric property of measurement instruments. The current study differs from others in this respect.

A second and more important aspect of this study is that it analyzes the effects of the rate of missing data on the generalizability (G) and phi (reliability) coefficients. Brennan (2001) demonstrated that the generalizability and reliability coefficients could be calculated from measurement results having missing data on the basis of generalizability theory with appropriate formulae without deleting a responder from the data. Yet the effects of missing data on the G and phi coefficients were not considered by any researchers. The effects of missing data on the G and phi coefficients represents the main question of this study. In addition, a method most frequently used by researchers encountering the problem of missing data in dual scored data is to regard missing data as incorrect answers and to code them blank. The effects of such an approach on reliability estimations constitutes a second question to which this study seeks answers. Thus, the current study searches for answers to this question: What are the effects of missing data on the reliability of measurement results obtained with one-dimensional measurement instruments? The fundamental research question was considered according to the following subproblems:

1. How is the reliability of measurement results having weak one-dimensional structure obtained with blank coding and incorrect coding influenced by varied sample sizes and by the rates of missing data?

2. How is the reliability of measurement results having strong one-dimensional structure obtained with blank coding and incorrect coding influenced by varied sample sizes and by the rates of missing data?

Method

Research Design

This study has a correlational survey design that aims to determine the presence or degree of co-variance between two or more variables (Karasar, 2004). It is also a simulative study.

Data and Conditions

This study analyzes the effects of missing data on the reliability of one-dimensional measurement results under the condition of varied sample sizes and missing data rates.

Differing recommendations are available in the literature for studies regarding G theory and reliability estimations. Kline (1986) states that sample size should be at least 200 in reliability calculations, while Nunnally and Bernstein (1994) point out that sample size should be at least 300 to reduce the amount of errors stemming from samples. Segal (1994), however, states that sample size of 300 would not be adequate and that it would be small in reliability calculations. Charter (2003), on the other hand, says that sample size of 400 could be adequate. Atilgan (2013) points out that the G and phi coefficients can be estimated in a sufficiently unbiased way if sample size is 50, 100, 200 and 300 in calculating the G coefficients but that the G and phi coefficients will be more precise and stable if sample size is 400. By considering the studies in the literature, sample size was determined as N= 100, 200, 400 and 1000 in this study.

On reviewing the studies concerning the effects of missing data, we found that they were often concerned with differing rates of missing data and structures of missing data. Nartgun (2015) and Kose and Oztemur (2014) conducted their research at a completely random mechanism at the rates of 5%, 10% and 20% missing data. Cheng (2016) analyzed the effects of the presence of 20% missing data at a random missing mechanism. Cum and Gelbal (2015) created data sets containing completely random missing data at the rates of 20% and 30% and not completely random missing data at the rate of 20%. Schafer and Olsen (1998) used a real data set under the condition of MAR and at the rates of 35% and 45%. Shang et al. (2016) used data sets containing missing data at the rates of 10%, 20% and 50% under the conditions of MCAR and MAR. Tabachnick and Fidell (2001) state that missing data at the rate of 5% or above at random do not cause serious problems. Therefore, by considering the

studies in the literature, the rates of missing data were determined as 5%, 10%, 20% and 30% in this study.

Depending on sample size, two different types of data sets containing 20 items with normal distribution and representative of strong and weak one-dimensionality were created. Item factor loads were manipulated between 0.50-0.85 in the first type, representing strong one-dimensionality, whereas they were free in the second type, representing weak one-dimensionality. The factor structures of both types of data were analyzed according to the unweighted least squares method; and the factor loads for sample sizes of 100, 200, 400 and 1000 were estimated at the intervals of 0.592-0.824, 0.598-0.808, 0.691-0.820 and 0.765-0.832, respectively, in the strong one-dimensional data, while they were estimated at the intervals of 0.058-0.684, 0.064-0.667, 0.046-0.699 and 0.077-0.677, respectively, for sample sizes of 100, 200, 400 and 1000 in the weak one-dimensional data. The fact that the created data sets had one-dimensional structure was confirmed through analyses by using Factor 10.3 software. Data were deleted from these two complete data sets in an MCAR manner (missing at completely random) at the rates of 5%, 10%, 20% and 30%. The transaction of deleting data was repeated 30 times and finally 30 data sets containing missing data for each sample size were obtained.

Analysis of Research Data

Reliability of measurement results obtained with measurement instruments is calculated in different ways depending on probable sources of error, such as raters, time, test forms, items and task, which may be contained in measurement. This study seeks answers to the research questions through G theory, which enables researchers to assess simultaneously the reliability coefficients obtained in different senses depending on the sources of error.

G theory is based on variance analysis, which ensures that inconsistencies that are present or may be present in observed scores are determined with powerful statistical analysis (Brennan, 2001). G theory divides variability in measurement results into categories according to their sources, and it aims to generalize the observed scores of variable or variables that are the object of measurement into population scores accurately. G theory also removes the traditional difference between validity and reliability to a certain extent. There are relative evaluations and absolute evaluations in education and in psychology, and G theory calculates the generalizability (G) coefficient for relative evaluations and dependability (phi) coefficient for absolute evaluations (Brennan, 2001). This study also examines the change in G and phi coefficients under the condition of sample size and missing data rates.

The G and phi coefficients were calculated in this study from direct missing data matrices and from matrices obtained by using the method of missing data coded incorrect respectively for data sets. Calculations were made manually in Excel because the EduG program was sensitive to missing data and the urGENOVA program could not analyze data containing more than 5% missing data. The calculations made in Excel were performed on the basis of Brennan's example (2001;

p. 227) for pxi design containing missing data and of the analyses for the example. Brennan (2001) employed ANOVA, which uses linear equations obtained by equalizing the expected values of squares averages to estimate variance components in balanced designs. Henderson (1953) recommends two methods for variance and co-variance estimations in unbalanced designs. Brennan (2001) used the method called Henderson Method 1 to calculate the G and phi coefficients in designs containing missing data. Based on this method, a T statistics called sum of squared mean scores is used instead of squares average statistics as in ANOVA. Since the sum of squares is a linear combination of squares average, Brennan (2001) points out that a similar estimation of variance components can be made by equalizing the sum of squares to the expected values. Variance estimations based on T statistics are shown in Table 1,

Table 1

Variance Estimations based on T Statistics for Person x Item Design

Source of Variance	df	*T*	Sum of Squares
person	n_b -1	$\sum_b \breve{n}_b \bar{X}_b^2$	T(b) - T(μ)
item	n_m -1	$\sum_m \breve{n}_m \bar{X}_m^2$	T(m) - T(μ)
person x item	$n_+ - n_b - n_m + 1$	$\sum_b \sum_m X_{bm}^2$	T(bm) - T(b) - T(m) + T(μ)
μ	**1**	$\boldsymbol{n_+ \bar{X}^2}$	

Variance estimations of the data created in this study based on T statistics were made and the G and phi coefficients were calculated with the help of these estimations. Absolute and relative error variances were obtained with the following formulas, respectively:

$$\sigma^2(\Delta) = \frac{\sigma^2(m)}{\ddot{n}_m} + \frac{\sigma^2(bm)}{\ddot{n}_m}$$

$$\sigma^2(\delta) = \left(\frac{\sum_b (\bar{X}_b - \overline{X})^2}{n_b - 1}\right) - \sigma^2(\Delta)$$

where $\sigma^2(\delta)$ represents relative error variance, $\sigma^2(\Delta)$ represents absolute error variance. $\ddot{n}_m$ is the harmonic average of n_m . All other calculations used in variance estimations can be found in Brennan (2001; pp. 225-237).

Complete data matrices for each condition were initially created in analyzing the data and the G and phi coefficients were calculated for these complete data matrices. After that, the G and the phi coefficients were calculated separately by the method of

missing data incorrect and by missing data design of pxi for each incomplete data set created and averages were found for 30 replications. Finally, the averages found for these two methods of estimation were compared with the results obtained for complete data. The root mean square of errors (RMSE) and bias values of error statistics were then calculated and interpreted.

Results

The findings are presented below according to the research questions.

Research question 1: How is the reliability of measurement results having weak one-dimensional structure obtained with blank coding and zero coding influenced by varied sample sizes and by the rates of missing data?

The G and phi coefficients estimated from matrices obtained from weak one-dimensional complete data matrices by the method of blank coding and incorrect coding (zero coding) are shown in Table 2 below.

Table 2

Averages for the G and Phi Coefficients Estimated from Weak One-Dimensional Data

Sample Size	*Type of Matrix*	*Complete Data*		*5%*		*10%*		*20%*		*30%*	
		G	*Phi*	*G*	*Phi*	*G*	*Phi*	*G*	*Phi*	*G*	*Phi*
N=100	Blank	0.617	0.581	0.601	0.567	0.584	0.551	0.554	0.525	0.519	0.493
	Incorrect			0.577	0.546	0.535	0.508	0.473	0.451	0.405	0.389
N=200	Blank	0.547	0.51	0.531	0.496	0.518	0.484	0.48	0.45	0.439	0.414
	Incorrect			0.506	0.474	0.47	0.442	0.4	0.38	0.339	0.324
N=400	Blank	0.626	0.594	0.613	0.582	0.597	0.568	0.562	0.536	0.522	0.5
	Incorrect			0.591	0.562	0.553	0.527	0.484	0.465	0.42	0.405
N=1000	Blank	0.623	0.592	0.608	0.578	0.592	0.564	0.559	0.534	0.519	0.497
	Incorrect			0.586	0.559	0.564	0.538	0.482	0.464	0.411	0.398

Blank: Missing data coded blank; incorrect: Missing data coded incorrect

According to Table 2, when the rate of missing data is 5%, 10%, 20% and 30% for sample size of 100; the G coefficients estimated from missing data matrices (blank coding) are 0.02, 0.05, 0.10 and 0.16, respectively, and the phi coefficients are estimated lower at the rates of 0.02, 0.05, 0.10 and 0.15, respectively. When the rate of missing data is 5%, 10%, 20% and 30% for sample size of 200; the G coefficients estimated are approximately 0.03, 0.05, 0.12 and 0.20, respectively, and the phi coefficients are approximately 0.03, 0.05, 0.12 and 0.18, respectively, which are low. This is similar to the other samples where the rate of missing data is 5%, 10%, 20% and 30% for sample sizes of 400 and 1000 and the G coefficients are 0.02, 0.05, 0.10 and 0.17, respectively, whereas the phi coefficients are approximately 0.02, 0.05, 0.10 and 0.16, respectively.

As is clear from Table 2, the G coefficients estimated from matrices obtained through incorrect coding for sample sizes of 100 and 200, according to complete data, are approximately 0.06, 0.13, 0.23 and 0.34, respectively, whereas the phi coefficients

are approximately 0.06, 0.12, 0.22 and 0.33, respectively, which are low. When the rate of missing data is 5%, 10%, 20% and 30% for sample size of 400, the G coefficients estimated for complete data are approximately 0.06, 0.16, 0.23, and 0.33, while the phi coefficients are approximately 0.06, 0.09, 0.22 and 0.33, which are low. The G and phi coefficients estimated through incorrect coding have been estimated lower than the G and phi coefficients estimated through missing data matrices (blank coding) for all rates of missing data and for all sample sizes.

Bias values and RMSE calculated for the G and phi coefficients from matrices obtained by blank coding and incorrect coding from weak one-dimensional missing data matrices are shown in Table 3.

Table 3

Error Statistics Calculated for Weak One-Dimensional Data

Sample Size	*Type of Matrix*	*Error*	*5%*		*10%*		*20%*		*30%*	
			G	*Phi*	*G*	*Phi*	*G*	*Phi*	*G*	*Phi*
N=100	Blank	RMSE	0,021	0,016	0,038	0,032	0,072	0,062	0,106	0,093
		Bias	0,019	0,013	0,036	0,029	0,066	0,055	0,101	0,087
	Incorrect	RMSE	0,045	0,041	0,085	0,078	0,151	0,137	0,221	0,201
		Bias	0,04	0,035	0,082	0,073	0,144	0,13	0,212	0,192
N=200	Blank	RMSE	0,021	0,017	0,035	0,029	0,074	0,065	0,115	0,101
		Bias	0,019	0,014	0,032	0,027	0,07	0,06	0,111	0,096
	Incorrect	RMSE	0,046	0,041	0,081	0,072	0,152	0,136	0,215	0,193
		Bias	0,041	0,037	0,077	0,068	0,147	0,13	0,208	0,186
N=400	Blank	RMSE	0,019	0,011	0,034	0,024	0,07	0,056	0,11	0,092
		Bias	0,017	0,009	0,033	0,022	0,068	0,054	0,108	0,09
	Incorrect	RMSE	0,037	0,034	0,074	0,068	0,143	0,131	0,208	0,191
		Bias	0,035	0,032	0,073	0,066	0,142	0,13	0,206	0,189
N=1000	Blank	RMSE	0,013	0,013	0,029	0,027	0,062	0,057	0,102	0,094
		Bias	0,012	0,012	0,028	0,026	0,061	0,056	0,101	0,093
	Incorrect	RMSE	0,037	0,034	0,071	0,065	0,141	0,129	0,213	0,196
		Bias	0,037	0,033	0,071	0,064	0,141	0,128	0,212	0,194

Blank: Missing data coded blank; Incorrect: Missing data coded incorrect

According to Table 3, the RMSE and bias error values for the G and phi coefficients estimated from missing data matrices for all conditions of sample size increase in parallel to the increase in the rate of missing data. Also, the situation is similar for the G and phi coefficients estimated from matrices obtained through

incorrect coding. It is observed that although error values calculated from both missing data matrices and through incorrect coding for all conditions of the rate of missing data are constant in some cases, as the size of sample increases, the error values decrease at least at minimal levels. Besides, it is also evident on comparing the data sets having and not having incorrect coding, regardless of their sample size, that the RMSE and bias values increase in data sets having incorrect coding.

Research question 2: How is the reliability of measurement results having strong one-dimensional structure obtained with blank coding and zero coding influenced by varied sample sizes and by the rates of missing data?

The G and phi coefficients estimated from matrices that are obtained from strong one-dimensional complete data matrices by the method of blank coding and method of incorrect coding are shown in Table 4 below.

Table 4

Averages for the G and Phi Coefficients Estimated from Strong One-Dimensional Data

Sample Size	*Type of Matrix*	*Complete Data*		*5%*		*10%*		*20%*		*30%*	
		G	*Phi*	*G*	*Phi*	*G*	*Phi*	*G*	*Phi*	*G*	*Phi*
N=100	Blank	0.951	0.951	0.950	0.948	0.947	0.945	0.940	0.938	0.933	0.930
	Incorrect			0.941	0.941	0.931	0.930	0.906	0.906	0.877	0.876
N=200	Blank	0.936	0.936	0.934	0.933	0.931	0.929	0.923	0.921	0.913	0.910
	Incorrect			0.926	0.926	0.915	0.915	0.890	0.890	0.861	0.861
N=400	Blank	0.933	0.933	0.930	0.930	0.926	0.926	0.917	0.917	0.905	0.905
	Incorrect			0.922	0.922	0.910	0.910	0.881	0.881	0.848	0.848
N=1000	Blank	0.942	0.942	0.938	0.939	0.935	0.936	0.927	0.928	0.917	0.918
	Incorrect			0.932	0.932	0.921	0.921	0.897	0.897	0.866	0.866

Blank: Missing data coded blank; Incorrect: Missing data coded incorrect

According to Table 4, when the sample size is 100, the G and phi coefficients obtained from complete data sets are the same as those obtained from matrices obtained by blank coding as data and having 5% and 10% missing data. When the rate of missing data is 20% and 30%, the G and phi coefficients were estimated lower than the complete data set and are 0.01 and 0.02, respectively. These findings are also similar for sample size of 200. Similar G and phi coefficients were estimated for the complete data set with sample sizes of 400 and 1000 and for missing data of 5% and 10%, whereas the coefficients were estimated lower (0.02 and 0.03, respectively) for data sets with 20% and 30% missing data. As is evident from Table 3, equal G and phi coefficients were estimated for all conditions of missing data with sample size of 400. On the other hand, the phi coefficients were estimated higher than the G coefficients for sample size of 1000.

Still according to Table 4, the G coefficients estimated from matrices obtained through incorrect coding for sample sizes of 100, 200 and 1000, according to complete data, are approximately lower at the rates of 0.01, 0.02, 0.05 and 0.08 When the rate of missing data is 5%, 10%, 20% and 30% for sample size of 400, the G and phi

coefficients estimated according to complete data are approximately lower at the rates of approximately 0.01, 0.03, 0.06 and 0.09. The G and phi coefficients estimated through incorrect coding are lower than those estimated through matrices of missing data for all rates of missing data and all sample sizes.

Bias values and RMSE calculated for the G and phi coefficients from matrices obtained by blank coding and incorrect coding from strong one-dimensional missing data matrices are shown in Table 5 below.

Table 5

Error Statistics Calculated for Strong One-Dimensional Data

Sample Size	*Type of Matrix*	*Error*	*5%*		*10%*		*20%*		*30%*	
			G	*Phi*	*G*	*Phi*	*G*	*Phi*	*G*	*Phi*
N=100	Blank	RMSE	0.005	0.003	0.020	0.022	0.029	0.030	0.040	0.041
		Bias	0.004	0.003	0.020	0.022	0.028	0.030	0.038	0.041
	Incorrect	RMSE	0.010	0.010	0.021	0.021	0.045	0.046	0.075	0.075
		Bias	0.010	0.010	0.020	0.021	0.045	0.045	0.074	0.075
N=200	Blank	RMSE	0.004	0.003	0.006	0.007	0.015	0.015	0.025	0.026
		Bias	0.003	0.003	0.005	0.007	0.013	0.015	0.023	0.026
	Incorrect	RMSE	0.011	0.011	0.021	0.021	0.046	0.046	0.076	0.075
		Bias	0.010	0.010	0.021	0.021	0.046	0.046	0.075	0.075
N=400	Blank	RMSE	0.004	0.003	0.008	0.007	0.017	0.016	0.029	0.028
		Bias	0.003	0.003	0.007	0.007	0.016	0.016	0.028	0.028
	Incorrect	RMSE	0.011	0.011	0.023	0.024	0.052	0.052	0.085	0.085
		Bias	0.011	0.011	0.023	0.023	0.052	0.052	0.085	0.085
N=1000	Blank	RMSE	0.005	0.003	0.007	0.006	0.015	0.014	0.026	0.024
		Bias	0.004	0.003	0.007	0.006	0.015	0.014	0.025	0.024
	Incorrect	RMSE	0.010	0.010	0.020	0.020	0.045	0.045	0.076	0.076
		Bias	0.010	0.010	0.020	0.020	0.045	0.045	0.076	0.076

Blank: Missing data coded blank; Incorrect: Missing data coded incorrect

An examination of Table 5 makes it clear that the RMSE and bias error values for the G and phi coefficients estimated from missing data matrices (blank coding) for all conditions of sample size increase in parallel to the increase in the rate of missing data. The situation is similar for the RMSE and bias values of the G and phi coefficients estimated from matrices, which are obtained through incorrect coding. It is observed that although error values calculated from both missing data matrices and through incorrect coding for all conditions of the rate of missing data are constant in some cases. The error values decrease at least at minimal levels as the size

of sample increases. Moreover, it is also evident on comparing the data sets having and not having incorrect coding regardless of their sample size that the RMSE and bias values increase in data sets having incorrect coding.

Discussion and Conclusion

This study examined the effects of missing data on measurement results and also considered reliability estimates from the aspect of generalizability theory. Studies available in the literature mostly approached the problem of missing data from the aspect of methods for treating missing data and made evaluations by comparing the results for complete data with the ones for treating missing data. However, they did not analyze the psychometric effects of missing data measurement results on statistical analyses without methods for treating missing data.

This study investigated how the rates of missing data in data sets with differing sample sizes and normal distribution influenced the generalizability and phi coefficients when a method for missing data coding was not used.

First, the effects of missing data rates in weak one-dimensional data on G and phi coefficients according to sample sizes were examined, and similar results were obtained for both of these coefficients in weak one-dimensional data. On comparing the estimates made from complete data and the ones made from data with missing data, it was found that the greatest fall was in the data with 20% missing data and especially in the data with 30% missing data. A further conclusion was that the estimates had not been affected greatly by sample sizes. The rate of missing data for a weak one-dimensional set of data having a rate of missing data of 20% and above affected the G and phi coefficients considerably.

Second, the effects of missing data rates and sample sizes in strong one-dimensional tests on G and phi coefficients were investigated, and it was found that the estimates made from missing data were minimally lower than those made from complete data, even in cases with 30% missing data. Thus, it was concluded that sample size did not affect estimates for strong one-dimensional data substantially either.

Estimation errors for the G and phi coefficients obtained from missing data matrices of strong and weak one-dimensional data were analyzed in terms of RMSE and bias statistics. It was found that as the rate of missing data for each condition of sample size increased, error values increased more in weak one-dimensional data and that it increased at minimal levels in strong one-dimensional data. It was also found that RMSE and bias values either did not change or decreased at minimal levels as sample size increased in both weak and strong one-dimensional data for each condition of missing data rates. On evaluating all these conditions together, it was found that error statistics for weak one-dimensional data were bigger than those for strong data.

A method that researchers frequently employ when they encounter missing data in binary data matrices is to regard missing data as incorrect answers and to code them zero. This study also examined the effects of this method and concluded that estimates made by incorrect (zero) coding, especially in weak one-dimensional data, were lower than those made through missing data matrices. On comparing the RMSE and bias values for the G and phi coefficients estimated from missing data matrices with those for matrices obtained by incorrect coding, it was found that the errors based on incorrect coding were higher, which was a remarkable finding. In a similar vein, the error statistics for weak one-dimensional data based on incorrect coding were found to be significantly higher than those for strong one-dimensional data, especially at 20% and 30% rates of missing data. Based on this research finding, it may be said that the incorrect coding method should not be used as a method for treating missing data since reliability estimates with incorrect coding yields biased results. Instead, by considering the fact that the G coefficient obtained in one-faced designs is equal to Cronbach's alpha, G theory, which enables one to perform analyses with missing data matrices in calculating the reliability of measurement results, is highly recommended.

Another remarkable result obtained in this study was that the G and phi coefficients grew ever closer as sample size increased in strong one-dimensional designs and that the phi coefficient was estimated to be bigger than the G coefficient when the sample size was 1000. Yet the phi coefficient is mathematically smaller than (or equal to) the G coefficient in generalizability analyses for balanced designs. Brennan (2001) states that this situation stems from using different quadratic forms to calculate the T statistics in unbalanced designs.

This study, which aimed to draw the reader's attention to the fact that the reliability of measurement results could be calculated with G theory, was conducted with binary data. Besides repeating the existing analyses with polytomous data, they can also be performed at differing levels of the conditions in a study. The effects of methods for treating missing data on the reliability of measurement results, which was one of the research problems here, can be analyzed separately in the context of generalizability theory.

References

Allison, P. D. (2001). *Missing data*. Thousands Oaks, CA: Sage Publiation.

Atilgan, H. (2013). Sample size for estimation of G and phi coefficients in generalizability theory. *Eurasian Journal of Educational Research, 51*, 215-227.

Aydilek, İ. B. (2013). *Veri kumelerindeki eksik degerlerin yeni yaklasimlar kullanilarak hesaplanmasi.* (Unpublished doctoral dissertation). Selcuk University, Institute Of Science, Konya.

Bakis, R., & Goncu S. (2015). Akarsu debi olcumlerinde eksik verilerin tamamlanmasi: Zap suyu havzasi ornegi [Completion of missing data in

rivers flow measurement: case study of zab river basin]. *Anadolu University Journal of Science and Technology A - Applied Sciences and Engineering*, *16*(1), 63–79.

Baraldi, A.N., & Enders, C.K. (2009). An introduction to modern missing data analyses. *Journal of School Psychology*, 48, 5-37.

Brennan, R. L. (2001). *Generalizability theory*. New York: Springer-Verlag

Cheng, H. (2016). Principle components analysis with missing values and outliers. Retrieved April 19, 2016, from http://citeseerx.ist.psu.edu/viewdoc/versions?doi=10.1.1.4.6605&version=3

Cool, A.L. (2000). A review methods for dealing with missing data. Paper presented annual meeting of the Southwest Educational Research Association, Dallas, January 28, TX.

Cum, S., & Gelbal, S. (2015). Kayip veriler yerine yaklasik deger atamada kullanilan farkli yontemlerin model veri uyumu uzerindeki etkisi [The effects of different methods used for value imputation instead of missing values on model data fit statistics]. *Mehmet Akif Ersoy University Journal of Education Faculty*, 35, 87-111.

Demir, E. (2013). Kayip verilerin varliginda coktan secmeli testlerde madde ve test parametrelerinin kestirilmesi: SBS ornegi [item and test parameters estimations for multiple choice tests in the presence of missing data: the case of SBS]. *Journal of Educational Sciences Research, 3* (2), 47-68.

Enders, C.K. (2010). *Applied missing data analysis*. New York: The Guilford Press.

Graham, J. W. (2009). Missing data analysis: making it work in the real world. *The Annual Review of Psychology*, *60*, 549-566.

Graham, J. W. (2012). *Missing data: Analysis and design, statistics for social and behavioral sciences*. New York: Springer.

Gu, X., & Matloff, N. (2015). A different approach to the problem of missing data. Retrieved March 20, 2016, from http://arxiv.org/abs/1509.04992

Horton, N. J., & Clainman, K. P. (2007). Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models. *The American Statistician, 61* (*1*), 79-90.

Howell, D. C. (2008). The treatment of missing data. In Outwaite, W. & Turner, S. (Ed.), *The analysis of missing data*. London: Sage Publicaton.

Kose, I. A., & Oztemur, B. (2014). Kayip veri ele alma yontemlerinin t-testi ve ANOVA parametreleri uzerine etkisinin incelenmesi [Examining the effect of missing data handling methods on the parameters of t-test and ANOVA]. *Abant İzzet Baysal University Journal of Faculty of Education*, *14*(*1*), 400-412.

Little, R. J. A., & Rubin D. R. (1987). *Statistical Analysis with Missing* Data. New York: John Wiley and Sons.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(*404*), 1198-1202.

Longford, N.T. (2005). *Missing data and small area estimation*. New York: Springer.

Nakai, M., & Ke, W. (2011). Review of the methods for handling missing data in longitudinal data analysis. *İnternational Journal of Math Analysis*, *5*(*1*), 1-13.

Nartgun, Z. (2015). Comparison of various methods used in solving missing data problems in terms of psychometric features of scales and measurement results under different missing data conditions. *İnternational Online Journal of Educational Sciences*, 7(4), 252-265.

Peng, C, J, Harwell, M., Liou, S., & Ehman, L. H., (2002*). Advances in missing data methods and implication for educational research.* Retrieved *March 20, 2016, from www.indiana.edu/~leeehman/missrerfin.pdf*

Piggot, T. D. (2001). *A review of methods for missing data. Educational Research and Evaluation, 7(4), 353-383.*

Rubin, D. R. (1976). Inference and missing data. *Biometrika*, *63(3*), 581-592.

Rubin, L. H., Witkiewitz, K., St. Andre, J., & Reilly, S. (2007). Methods for handling missing data in the behavioral neurosciences: don't throw the baby rat out with the bath water. *The Journal of Undergraduate Neuroscience Education (JUNE)*, *5*(2), A71-A77.

Sari, İ. K. (2012). *Karma ayristirma analizinde kayip gozlem tahmin yontemlerinin değerlendirilmesi* [Evalution of missing value estimation methods for mixture discriminant analysis]. (Unpublished doctoral dissertation). Selcuk University, Institute Of Science, Konya.

Schafer, J. L., & Olsen, M. K. (1998). Multiple imputation for multivariate missing-data problems: a data analyst's perspective. *Multivariate Behavioral Research*, *33*(4), 545-571.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*., *7*(2), 147-177.

Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology*, *57*(*1*), 1-10.

Ser, G., & Bati, C. T. (2015). Eksik veri analizinde coklu atama yonteminin degerlendirilmesi: Hayvancilikta tekrarli olcum verisi uzerine bir uygulama [Evaluation of multiple ımputation in missing data analysis: an application on repeated measurement data in animal science]. *Turkish Journal of Agriculture - Food Science and Technology*, *3*(*12*), 926-932.

Shang, F., Liu, Y., Cheng, J., & Cheng, H.(2014). Robust principle component analysis with missing data. İn *Proceedings of the 23rd ACM İnternational Conference on Conference on İnformation and Knowledge Management*, New York, USA, 1149-1158. DOİ=http://dx.doi.org/10.1145/2661829.2662083.

Soley-Bori, M. (2013). Dealing with missing data: Key assumptions and methods for applied analysis. Technical Report no:4. Boston University, School of Public Health, Department of Helath Policy and Managment. Retrived March 20, 2016, from www.bu.edu/sph/files/2014/05/Marina-tech-report.pdf

Tabachnick, B.G., & Fidel L.S. (2001). *Using multivariate statistics (4th ed.).* Needham, Heights, MA: Allyn& Bacon.

Weaver, B., & Maxwell, H. (2014). Exploratory factor analysis and reliability analysis with missing data: A simple method for SPSS users. *The Quantitative Methods for Psychology, 10*(2), 143-152.

Whang, L., Zhang, Z., & Tong, X. (2015). Mediation analysis with missing data through multiple imputation and bootstrap. *Quantitative Psychology Research*, *140*, 341-355.

Yilmaz, H. (2014). *Random Forests yonteminde kayip veri probleminin incelenmesi ve saglik alaninda bir uygulama* [studying the missing data problem in random forests method and an application in health field]. (Unpublished doctoral dissertation). Eskisehir Osmangazi University, Eskisehir.

Zhu, X.P. (2014). Comparison of four methods for handing missing data in longitudinal data analysis through a simulation study. *Open Journal of Statistics*, *4*, 933-944.

Örneklem Büyüklüğünün ve Kayıp Veri Oranının Genellenebilirlik Katsayılarına Etkisi

Atıf:

Soysal, S., Karaman, H., & Dogan, N. (2018). The effects of sample size and missing data rates on generalizability coefficients. *Eurasian Journal of Educational Research, 75, 179-196*, DOI: 10.14689/ejer.2018.75.10

Özet

Problem Durumu: Veri toplama ve bu verilerin analiz edilmesinin temele alındığı araştırmalarda karşılaşılma olasılığı olan en önemli problemlerden biri kayıp veridir. Kayıp veri planlanan veri kümesi ile elde edilen veri kümesi arasındaki fark olarak tanımlanabilir. Alan yazın incelendiğinde kayıp veri problemi ölçme araçlarının uygulanması sonucu karşılaşılan yaygın bir problem olmasına karşın, ölçme

araçlarının psikometrik özelliklerine etkisi üzerinde pek durulmamıştır. Özellikle ölçme sonuçlarının güvenirliğinin, geçerliğinin, ortalama ayırıcılık ve güçlük gibi istatistiklerin kayıp verilerden nasıl ve ne düzeyde etkilendiği konusu pek araştırılmamıştır. Başarı testi, tutum ölçeği, anketler vb. katılımcılara uygulandığı zaman çeşitli sebeplerle bazı katılımcıların cevap vermemesi durumuna çok sık rastlanır. Katılımcılar ölçme araçlarındaki soruları bir fikri olmaması, uygun bir cevap bulamaması, yanlışlıkla soruyu cevaplamadan atlaması veya cevabını doğru bir şekilde işaretlememesi nedenleriyle boş bırakabilmektedir. Ancak ölçme araçlarına gelen cevaplar azaldıkça ya da kayıp veri arttıkça toplanan bilgi azalacak ve ölçme sonuçlarının geçerliği ve güvenirliğinin düşmesi beklenecektir. Kayıp verilerin eğitim ve psikolojide kullanılan ölçme araçlarının psikometrik özelliklerini etkilemesi kaçınılmaz bir durumdur. Dolayısıyla kayıp verilerin ölçme araçlarının psikometrik özellikleri üzerindeki etkisinin araştırılmasına ihtiyaç olduğu düşünülmektedir.

Araştırmanın Amacı: Belirlenen bu ihtiyaca bağlı olarak kayıp verinin ölçme araçlarının psikometrik özelliklerinden güvenirliğe etkisi üzerinde çalışılmasına karar verilmiştir. Bu yönü ile çalışma diğer çalışmalardan farklılık göstermektedir. Çalışmanın ikinci ve daha önemli bir yönü ise kayıp veri oranının genellenebilirlik (G) ve phi (güvenirlik) katsayısına olan etkisini incelemesidir. Brennan (2001), Genellenebilirlik kuramına dayalı olarak kayıp veriye sahip ölçme sonuçlarından uygun formüllerle herhangi bir cevaplayıcıyı verilerden silmeden genellenebilirlik ve güvenilirlik katsayılarının hesaplanacağını göstermiş ancak kayıp verinin G ve Phi katsayısına olan etkisi herhangi bir araştırmacı tarafından incelenmemiştir. Kayıp verilerin G ve Phi katsayısına etkisi bu araştırmanın temel sorusunu oluşturmaktadır. Ayrıca ikili puanlanan verilerde kayıp veri sorunu ile karşılaşan araştırmacıların en sık başvurdukları yöntemlerden biri kayıp verileri yanlış cevap olarak kabul edip sıfır puan ataması yapmaktır. Bu yaklaşımın güvenirlik kestirimine etkisi, bu çalışmayla cevaplamaya çalışılan bir başka sorudur. Dolayısıyla, bu araştırmada normal dağılım altında zayıf ve güçlü tek boyutluluk özelliği gösteren kayıp verili ve sıfır atamayla elde edilen ölçme sonuçlarının güvenirliğinin değişen örneklem büyüklüleri ve kayıp veri oranlarından nasıl etkilendiği sorusuna yanıt aranmıştır.

Araştırmanın Yöntemi: Güvenirlik kestirimleri, hata kaynaklarına bağlı olarak farklı anlamlarda elde edilen güvenirlik katsayılarını aynı anda değerlendirmeyi sağlayan Genellenebilirlik Kuramı açısından ele alınmıştır. Araştırma sorularına bağlı olarak öncelikle normal dağılım gösteren zayıf ve güçlü tek boyutlu yapılarda farklı örneklem büyüklüğüne (N=100, 200, 400, 1000) sahip tam veri setleri üretilmiştir. Bu setlerden tamamıyla seçkisiz olacak şekilde farklı kayıp veri oranlarında (%5, %10, %20,%30) veriler silinerek kayıp verili setler oluşturulmuştur. Araştırma sonuçları tam veri setleri ile kayıp ve sıfır atama yapılmış veri matrislerinden elde edilen G ve phi katsayılarının ortalamaları karşılaştırılarak değerlendirilmiştir. Ayrıca değerlendirmeleri daha isabetli yapabilmek için hata istatistiklerinden hataların kareleri ortalamasının karekökü (RMSE) ve yanlılık (bias) değerleri hesaplanarak yorumlanmıştır.

Araştırmanın Bulguları: Tam veri ile kayıp veri setlerinden elde edilen kestirimler karşılaştırıldığında, zayıf tek boyutlu desenler için kayıp veri oranının %20 ve daha fazla olduğu durumlarda G ve Phi katsayılarının önemli derecede etkilendiği ancak güçlü tek boyutlu desenler de kayıp veri oranın %30 olduğu durumda dahi bu katsayıların minimal düzeyde etkilendiği bulunmuştur. Örneklem büyüklüğünün her bir koşulu için kayıp veri oranı artıkça hata değerlerinin zayıf tek boyutlu verilerde daha fazla arttığı; güçlü tek boyutlu verilerde ise minimal düzeyde arttığı gözlenmiştir. Kayıp veri oranının her bir koşulu için zayıf ve güçlü tek boyutlu verilerin her ikisinde de örneklem büyüklüğü arttıkça hata ve yanlılık değerlerinin ya değişmediği ya da minimal düzeyde azaldığı görülmüştür. Bütün koşullar bir arada değerlendirildiğinde zayıf tek boyutlu verilere ait hata istatistiklerinin güçlü tek boyutlu verilerden elde edilenlere göre daha büyük olduğu gözlenmiştir. Ayrıca özellikle zayıf tek boyutlu verilerde sıfır atama sonucu elde edilen kestirimlerin kayıp veri matrisinden elde edilen kestirimlerden daha düşük ve sıfır atama yöntemine dayalı olarak zayıf tek boyutlu verilerin hata istatistiklerinin güçlü tek boyutlu verilerin hata istatistiklerinden, özellikle %20 ve %30 kayıp veri oranlarında, önemli derecede yüksek olduğu bulunmuştur.

Araştırmanın Sonuçları ve Önerileri: Dolayısıyla sıfır atama yöntemi ile elde edilen güvenirlik kestirimleri yanlı sonuçlar verdiğinden bu yöntemin güvenirlik kestirimlerinde kayıp veri ile baş etme yöntemi olarak kullanılmaması; bunun yerine ölçme sonuçlarının güvenirliğinin hesaplanmasında kayıp veri matrisleri ile analiz yapmaya olanak sağlayan Genellenebilirlik kuramının kullanılması önerilebilir. Ayrıca kayıp veri matrisleriyle ölçme sonuçlarının güvenirliğinin Genellenebilirlik kuramı ile hesaplanabileceğine dikkat çekmek istenilen bu çalışma iki kategorili veriler ile yürütülmüştür. Mevcut analizler çok kategorili veriler için tekrarlanabileceği gibi araştırmada incelenen koşulların farklı düzeylerinde de gerçekleştirilebilir. Bir başka araştırma problemi olan kayıp veri ile baş etme yöntemlerinin ölçme sonuçlarının güvenirliğine etkisi Genellenebilirlik kuramı bağlamında ayrıca incelenebilir.

Anahtar Kelimeler: Güvenirlik, G katsayısı, phi katsayısı, sıfır atama, MCAR, genellenebilirlik kuramı, kayıp veri matrisi.

Eurasian Journal of Educational Research 75 (2018) 197-212

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

Effectiveness of Project Based Learning in Statistics for Lower Secondary Schools

Tatag Yuli Eko SISWONO[1], Sugi HARTONO[2], Ahmad Wachidul KOHAR[3]

ARTICLE INFO

Article History:
Received: 10 Jan. 2018
Received in revised form: 21 May. 2018
Accepted: 21 May. 2018
DOI: 10.14689/ejer.2018.75.11

Keywords
project-based learning, conventional learning, statistics learning

ABSTRACT

Purpose: This study aimed at investigating the effectiveness of implementing Project Based Learning (PBL) on the topic of statistics at a lower secondary school in Surabaya city, Indonesia, indicated by examining student learning outcomes, student responses, and student activity. **Research Methods**: A quasi experimental method was conducted over two months involving two classes of seventh grade students, consisting an experiment class (PBL) and a control class (conventional learning). Data were collected through student activity observation sheets, student responses, and pretest-posttests. Data were analyzed by employing covariance analysis (ANCOVA). **Findings**: Based on ANCOVA, student learning outcomes in PBL are higher than those in conventional learning. In addition, based on the results of a descriptive analysis, results of the student learning outcomes in PBL obtained more than the minimum standard score (MSC), the students' responses in learning were positive, and the students were active in the class activities. Thus, PBL is effective in statistical learning. **Implications for Research and Practice**: Based on the results of the research, it can be concluded that PBL was effective in statistical learning. These findings suggest that the students were enthusiastic in working the given project and actively discussed with other students in the class. We suggest teachers apply PBL on other mathematics topics so students can be enthusiastic in mathematics learning in the class. Alternatively, teachers can use PBL with technology-assisted learning to make learning more interesting for students.

[1] Universitas Negeri Surabaya, INDONESIA, e-mail:tatagsiswono@unesa.ac.id, ORCID: orcid.org/0000-0002-7108-8279
[2] Universitas Negeri Surabaya, INDONESIA, e-mail:sugihartonounesa@gmail.com, ORCID: orcid.org/0000-0002-0959-7013
[3] Universitas Negeri Surabaya, INDONESIA, e-mail:ahmadw.kohar@gmail.com, ORCID: orcid.org/0000-0002-2813-8584
Corresponding Author: Tatag Yuli Eko SISWONO,Universitas Negeri Surabaya, Indonesia, Mathematics Departement, Faculty of Mathematics and Natural Science, e-mail: tatagsiswono@unesa.ac.id

Introduction

One mathematics topic considered difficult by junior high school students is statistics (Garfield & Ahlgren, 1988). In relation to data presentation, some research found that many examples of incorrect choices of the graph in statistical projects were experienced by secondary students (Li & Shen, 1992). For example, some used a line graph with qualitative variables or a bar graph to represent the evolution of an index number through a sequence of years. Consequently, student statistical ability and interest decreased. Also, they were found to quickly get bored and lazy in learning (Basbay & Ales, 2009). Therefore, an interesting learning process in statistics in the presence of research is needed to solve this problem.

The teaching and learning process should place the teacher at the center to engage students in an active learning process. One interesting learning approach is Project Based Learning (PBL). PBL, which is recommended in the Indonesian curriculum, is touted as superior to teaching methods in improving problem solving and thinking skills (Siswono et al., 2017a) and engaging students in their learning (Berends, Boersma & Weggemann, 2003). PBL is expected to achieve the standard of learning. Thus, teachers are required to create an active student learning atmosphere by constructing, locating, and developing their knowledge (Erdem & Demirel, 2002; Siswono et al., 2017b). In this context, project-based learning (PBL) is one of the recommended lessons for enhancing active students' engagement and material understanding (Krajcik & Blumenfeld, 2006), even bringing new skills to teachers and students and improving their existing skills (Basbay, 2010).

PBL allows students to learn by doing, implementing their ideas while performing activities in the real world through investigating questions, proposing hypotheses and explanations, discussing their ideas, and eventually developing solutions or outcomes (Diffily, 2002; Krajcik & Blumenfeld, 2006). It therefore helps them in developing problem solving strategies, making authentic products such as models, stories, or presentations (Özdemir&Ubuz, 2006; Kaldi, et al., 2011; Korkmaz, 2002; Saracaloğlu et al., 2006; Robinson, 2009), working in groups, and learning social skills, interaction, cooperation, responsibility, social and democratic behavior, critical thinking, and decision making (Blumenfeld et al., 1991; Dede & Yaman, 2003; Demirhan, 2002; Ozden et al., 2009; Saracaloğlu et al., 2006; Thomas, 2000).

According to Demirhan (2002), PBL is defined and described as an approach that (1) requires interdisciplinary study, (2) encourages students to be responsible for groups or individuals and collaborative studies in real-life issues based on pre specified topics and their personal interests and skills, (3) provides teacher roles to facilitate the learning process as well as guide the students, (4) yields to students' authentic products or presentations outcome, and (5) integrates different approaches within themselves. Leviatan (2008) also explains in his research that project-based learning is an innovative learning that emphasizes complex activities with the goal of solving the problems based on inquiry activities. In addition, according to a study from Miswanto (2011), when the project-based learning model is applied, student learning outcomes on the topic of linear programming increases.

This research was conducted to explore the effectiveness of PBL as an alternative learning strategy that can be introduced to students both in teaching and learning mathematics. Students are taught in the school environment, especially when conventional learning strategies are preoccupied with theories, examples, and exercises (Soedjadi, 2001), but their application is limited in unusual situations such as solving real-life problems. This contrasts with PBL strategies, wherein students who are taught PBL strategies will be given the opportunity to develop their skills and adjust and change methods as they are in new situations. Furthermore, students who follow PBL teaching have a greater chance to engage in real-world activities through questioning, hypothesizing and explaining, discussing their ideas, and eventually developing solutions or outcomes (Diffily, 2002; Krajcik & Blumenfeld, 2006).

According to Markham et al. (2003), there are six aspects needed in PBL: 1) authentic, real-world challenges, 2) rigorous academics; 3) applying learning by using high-performances skills; 4) active exploration; 5) interacting and making adult connections, and 6) formal and informal assessment practices. Besides that, PBL also allows students to investigate questions, propose hypotheses and explanations, discuss their ideas, challenge the ideas of others, and try out new ideas (Krajcik and Blumenfeld, 2006).

In recent years there have been many studies relating to the effects and effectiveness of project-based learning in science education (Ladewski, Krajcik, and Harvey, 1994; Krajcik, et al., 1998; Dede & Yaman, 2003; Ozden et al., 2009). However, there have only been a few studies in Indonesia that examined the effectiveness of PBL on mathematical topics. Miswanto (2011), for instance, shows student improvements regarding learning outcomes on linear programming topics for secondary school students. However, these results show the effectiveness of PBL only on the basis of learning outcomes. In fact, Slavin (1997) states that the effectiveness of learning is not only determined by the quality of instruction (learning outcomes), but also the appropriate level of instruction, incentive, and time. Eggen and Kauchak (1998) also suggested that learning effectiveness is characterized by the students' active engagement. If the students are more active in the learning process, then this learning is more effective. In addition, Mudhofir (1987) shows that the effectiveness can be measured by observing student interest. This interest affects the learning process. If students are not interested in learning then they cannot be expected to succeed in this learning.

Based on the description above, the objectives of this study are to examine the effectiveness of PBL on statistical learning for secondary schools from three aspects: learning outcomes, students' responses, and students' activities.

Method

Research Design

This study is an experimental quasi-research using a factorial of non-equivalent pretest-posttest control group design. The experimental class was taught by applying project-based learning, while the control class was taught with conventional learning. The experimental class was divided into small groups, wherein each group has heterogeneous (consisting of low, medium, and high mathematical ability students). In addition to teacher information, the grouping was also based on the test score before project-based learning was carried out. The students carried out the project for two months.

Research Sample

The participants of the study were from two classes, selected using cluster random sampling from 10 classes in the same grade from a junior high school in Surabaya, Indonesia. The participants were divided into two groups: experimental and control (Cohen et al., 2005). The control group consisted of 38 students and the experimental group consisted of 37 students. All students were in grade seven and aged between 12 - 13 years.

Research Instruments and Procedures

The research instruments developed were a student activity observation sheet, student response questionnaire, and pretest posttests. The observation sheet and response questionnaire were adapted from previous research. Meanwhile, the pretest and posttest were developed by the researchers. The pretest and posttest have the same questions, each of which consist of nine essay items. All the items were then tested to examine validity and reliability to a class that was different from the control and experiment classes. Eight of the nine essay items were rated using a score of 0-10. The scores were as follows: understanding: 0-3; implementing strategy: 4-6; conclusion: 7-10. The scoring of the remaining item was: understanding: 0-5; implementing strategy: 6-15; conclusion: 16-20. Of the nine items, three tested about line graphs, three about bar charts, and three about pie charts. Therefore, the total score pretest and posttest obtained by each subject ranged from 0-100.

In agreement with the multiple coding procedures of the scoring test, we calculated the interrater reliability for each essay item, which resulted in Cohen's Kappa of 0.71-0.85, which indicates that the coding of scores varies from substantial to perfect agreement (Landis & Koch, 1977). To confirm the validity and reliability of the pretest and posttest we applied a product moment correlation and alpha coefficient test (Cronbach's, 1951), each shown in Table 1.

Table 1
The Reliability and Validity of Instrumental Test

Test	**Item**								
Validity	1	2	3	4	5	6	7	8	9
(r_{xy})	0.48	0.50	0.56	0.44	0.49	0.71	0.45	0.51	0.43
Reliable	0.6								

Based on the test analysis conducted on 37 students, all the items tested are significantly valid, with the coefficient validity of each of the items is interpreted as at least medium ($r_{xy} > 0.40$), and the reliability coefficient (0.60) as medium as well. Data were collected through student activity observation sheets, student response questionnaires, and pre and posttests in learning. Observational data of student activity were observed during the learning process. In our study, the student activity categories observed include: 1) listening or paying attention to the teacher or friend's explanations; 2) observing, listening to, or viewing problems, events, or explanations in the student worksheet; 3) discussing or solving the student worksheet or finding ways and answers in the student worksheet; 4) presenting the results of the discussion and providing feedback in groups; 5) asking about the results of the discussion or observations from friends or teachers; and 6) making conclusions or summarizing the learning materials in groups or with teachers.

In this research, the steps of PBL in statistics, as follows:

Table 2

Steps of PBL in Statistics

No	PBL steps	Learning Activity
1	*Start with the essential question*	a. Teachers initiate learning by providing PBL problems that are demonstrated through video. b. This video contains statistics problems in daily life: "The headmaster of a school plan to build a new building in the school. Then, he asks a mathematics teacher to collect and present some types of data: (1) students visiting the school medical room or library from the years 2009-2014, (2) student hobbies/sports in the 7th graders, and (3) how the 7th graders go to school. The data are used for consideration of the construction of new buildings that will soon be made. c. Then, students start with the essential question about this video; for example, how to help the mathematics teacher in collecting and presenting the data.
2	*Design a plan for the project*	a. The students in the class are divided into three groups and each group must choose one type of data b. Teachers and students collaboratively create a plan for solving the problem.
3	*Create a schedule*	a. Teachers and students arrange the schedule of the project activity based on the plan they created. The schedule includes: 1) the students create a timeline (time allocation) to solve the project; 2) the students also create a project deadline; 3) the teacher guides the students when they use a method unrelated to the project; and 4) the teacher asks the students to explain an alternative method
4	*Monitor the students and the progress of the project*	a. Teacher is responsible for monitoring student activities during solving the project. b. Teacher becomes a mentor for student activities. c. Students implement and document projects that have been designed.
5	*Assess the outcome*	Students present the results of the project in front of the class to discuss with other groups.
6	*Evaluate the experience*	a. At the end of the lesson, teachers and students reflect on the activities of PBL and the outcomes of PBL that have been implemented. b. The reflection process is done both individually and in groups.

The results of the data tests were collected by giving the pretest before learning and the posttest given after learning in the experimental and control classes. Meanwhile, students' responses were collected by using a questionnaire in the experimental and control classes, with the aim to determine the student response to the learning. This was done after the learning process was completed.

Data Analysis

This study applies two data analysis techniques: inferential statistical analysis and descriptive statistical analysis.

*Inferential statistical analysis.*The purpose of this analysis is to see the differences in student learning outcomes following project-based learning or conventional learning on statistical materials. Pretest-posttest data were analyzed by the inferential statistics of ANCOVA. To perform ANCOVA, it was necessary to satisfy assumptions regarding normality, the equation of the variance, and the equation of the regression lines. Based on analyses performed to satisfy the assumptions, this was determined with the Shapiro-Wilk Test because the sample was less than 50 students, showing a significance value of 0.173 for the experimental group and of 0.329 for the control group. This finding suggested that the distribution of the data among the experimental group and the control group was normal. The homogeneity test of variance showed a significance value of 0.465, which indicated that the data had the same (homogeneous) variance. Descriptive statistical analysis was used to compare the average score of the pretest and posttest for the experimental and control groups.

The ANCOVA test was conducted to reveal the effect of PBL in the statistical learning for students. The determined significance level was $\alpha = 0.05$.

Descriptive statistical analysis. This analysis was used to analyze the effectiveness of project-based learning on the topic of statistics. The data were collected from student activity data, student responses data, and student learning outcomes data. Student activities are said to be effective if the percentage of every aspect observed at each meeting is in the ideal time range of student activity. Student responses are said to be positive if the answers of students who choose the positive category for each aspect is more than 80%. Data analysis of learning outcomes and student learning outcomes descriptively aims to describe the student learning outcomes based on the tests implemented. A student has mastery of learning individually if their score is at least 70 with a maximum score of 100.

Results

This study found the results of inferential statistical data analysis and descriptive data analysis obtained during the research process. All data were analyzed using Statistics Package for Social Science (SPSS 22) and the results are presented as follows:

Inferential Statistical Data Analysis

Prior to statistical tests, normality tests were performed first to determine whether the sample data taken followed normal distribution. Normality is important for inferential statistics that aims to generalize the results of the analysis of sample data. The data reveals:

Table 3

Test the Normality of the Experimental Class Data and the Control Class by using the Shapiro Wilk test

	Tests of Normality		
	Shapiro-Wilk		
	Statistic	df	Sig.
x	.958	37	.173

a. Lilliefors Significance Correction

	Tests of Normality		
	Shapiro-Wilk		
	Statistic	df	Sig.
x	.967	37	.329

*. This is the lower bound of the true significance

a. Lilliefors Significance Correction

Table 4

Levine Test Results for Homogeneity of Variance

Levine Statistic	df1	df2	Sig.
.540	1	72	.465

As shown in Table 3, the Shapiro-Wilk Test reveals the following findings: the critical values for the variables x and y (coefficient x and y ratio) are 0.173 and 0.329 > significant (α) = 0.05. Hence, it can be concluded that Ho is accepted, meaning that the experiment class data (x) and data of control class (y) ratio come from a normally distributed population. The Levine test also showed a significance value of 0.465 (in Table 4), which indicated that the data had homogeneity of variance. Since the data is normally distributed and has homogeneity of variance, the independence test can be conducted to influence students' initial ability (x) to students' learning outcomes for each experimental group and control group; the linearity test determines whether the linear model obtained can be applied to show the effect on students' learning result. Based on the data obtained the results are as follows:

Table 5

Independence Test of Experimental Class and Control Class Data

Dependent variable: y					
Source	Type I Sum of Squares	df	Mean Square	F	Sig.
z	740.670	1	740.670	11.183	.001
x	6850.800	1	6850.800	103.439	.000
Corrected Model	7591.471[a]	2	3795.735	57.311	.000

a R Squared = .618 (Adjusted R Squared = .607)

From the table above the variable x significance is 0.000. Because the value of significance is far below 0.05, it can be concluded that student initial ability influences student learning outcomes. Furthermore, for the corrected model, the significant value is 0.000 < significant level (α) = 0.05, thus, it can be concluded that the learning model in the experimental and control classes also influences student learning outcomes.

Furthermore, the ANCOVA test was administered to see whether the student learning outcomes in the experimental and control classes were different after being given conventional and PBL lessons. The results of the data analysis are as follows:

Table 6

ANCOVA Test of Experiment and Control Class Data

Test of Between-Subjects Effects

Dependent Variable: y

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected	7591.471[a]	2	3795.735	57.311	.000
Model	29972.490	1	29972.490	452.548	.000
Intercept	2463.687	1	2463.687	37.199	.000
z	6850.800	1	6850.800	103.439	.000
Error	4702.367	71	66.231		
Total	446432.000	74			
Corrected Total	12293.838	73			

a. R Squared = .618 (Adjusted R Squared = .607)

Based on the Tests of Between-Subjects Effects table above, it can be seen that F test results for x show the value of 103.439 with Sig. of 0,000. Since Sig value < significant (α) = 0.05, it can be concluded that there are differences in learning outcomes between students taught by PBL learning strategies and students taught with conventional learning strategies. Thus, using a regression test shows which of the two learning models is better. The analysis of data is shown as follows:

Table 7.1

Test of Experiment Class Data Regression

Coefficients[a]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
	B	Std. Error	Beta			Zero-order	Partial	Part
1 (Constant)	71.784	3.448		20.821	.000			
X	.423	.103	.568	4.084	.000	.568	.568	.568

a. Dependent Variable: y

Table 7.2.

Test of Control Class Data Regression

Coefficients[a]

Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig.	Correlations		
	B	Std. Error	Beta			Zero-order	Partial	Part
1	48.630	4.620		10.526	.000			
(Constant)	.512	.114	.603	4.474	.000	.603	.603	.603
x								

a. Dependent Variable: y

From the above regression test, the constants obtained from the test of the experiment class data regression are 71.784 and the constant of a test of the control class data regression is 48.630. Both regression models show that the constant regression line for the experimental class is greater than the constant regression line of the control class, while the two regression lines are parallel. This shows a significant difference. Geometrically, the regression line for the experimental class is the regression line for the control class. This means student learning outcomes when PBL learning is applied are higher than the student learning outcomes when conventional learning is employed on statistical topics. This shows that PBL is effective in statistical learning for students.

The Results of Descriptive Statistical Data Analysis

Analysis of data obtained on the implementation of PBL and conventional learning is as follows.

Table 8

Comparison of Student Learning Results in Experiment and Control Classes

Information	Experiment Class	Control Class
Average of student learning outcomes	84,9	68,3
Percentage of students who mastered learning	100%	48,6%

Regarding mastery of learning outcomes, the result of table 7 shows there are 37 students in the experimental class who pass in learning (who obtained more than the minimum standard score (MSC), with MSC = 70). Thus, the percentage of student learning mastery is 100%. This is compared to the control class, which had 19 students who failed in the learning process; we can thus conclude their students are not passing in learning.

Observation of student activity during three meetings at the PBL class shows that during every five minute observation, the student could do at least one category of student activity. Most of the students were in the activity of category 3 (discussing or solving the student worksheet or finding ways and answers in the student worksheet). Since the students did PBL activities every 5 minutes instead of irrelevant activities, the students in each meeting can be considered active; thus, it can be concluded that the student activity category is effective. However, in the conventional class some students did irrelevant activities instead of PBL activities, for as long as 75 minutes, such as sleeping and playing. Therefore, the control class did not meet the effective category.

The result of a questionnaire for student responses to the learning shows that the number of students who selected response items in the positive category is 78.02%. Thus, the student responses are positive. Meanwhile, in conventional learning the student responses are positive but much lower, with only 55% selecting positive responses.

Discussion and Conclusion

Based on analysis of inferential statistical data, the results of the analysis showed ANCOVA test $F = 103.4339$ against $p = 0.000$. Because $p < 0.05$, there is a significant difference between project-based learning in the experimental class and conventional learning in the control class.

In addition, from the regression test results, it is seen that the constant of the regression line for the experimental class is higher than the constant of the regression line for the control group. This shows that student learning outcomes in PBL learning are higher than the student learning outcomes in conventional learning. This might be because the learning strategies in the experimental and control classes affected the student learning outcomes. According to Miswanto's findings (2011), after students are engaged in project-based learning, student learning outcomes increase compared to the previous test (pretest). Similarly, Dede & Yaman's (2003) study explains that learning projects are effective in science and mathematics. Moreover, Özdemir (2006) also confirms project learning is effective in grade seven geometry. In line with these results, Ay (2013) notes that PBL is highly considered as a process in which learners' heterogeneity is beneficial in their learning and development.

Meanwhile, in the descriptive analysis, student learning outcomes that followed project based learning got an average score of 84.9% and mastery learning students achieved a score of 100%. Therefore, it can be concluded that this learning is effective.

This is line with the finding of Morrison, et al. (2010), in which measurement of effectiveness can be ascertained from test scores.

Positive results were also obtained regarding the students activities in the experimental class; from the first meeting to the third meeting, it is indicated that the students' activity was in the effective category. At the beginning of the meeting, the students felt unfamiliar with this learning situation, but in the second and third meetings, the implementation of learning came better, wherein students already knew what was done in a learning activity. The students were enthusiastic about working the given project, and actively discussed and presented their group work to other groups. This is in accordance with Barab & Luehmann (2002) in which PBL principles are deduced from the constructivist perspective, emphasizing active learning and higher order thinking skills. Similarly, Green (1998) emphasized that participants in project-based learning learn better and are more actively acting in their learning. In relation to the student response to the implementation of project-based learning, the results of the analysis showed that 85.83% of the students agreed with all statements in the questionnaire, which means that the students responded positively to the implementation of the lesson.

Based on the research results it can be concluded that PBL in learning study is effective. The student learning outcomes in PBL obtained a greater score than of MSC, student activity was active, and student response to learning was positive. In addition, the experimental classes wherein the project based learning was applied were more effective than the control classes that used conventional learning. Thus, PBL demonstrated higher learning outcomes than conventional groups.

Further studies should be conducted to discover the importance of project-based learning and relevance in mathematics. This research study may be repeated with a different topic in mathematics. Furthermore, experimental studies may be carried on to discover the dynamics of project-based learning to compare individual work with group work. Alternatively, the use of certain technologies might be changed to determine ways in which it affects the process, if at all.

References

Ay, S. (2013). Öğretmen adaylarının proje tabanlı öğrenme ve geleneksel öğretime ilişkin görüşleri [Trainee teachers' views on project-based learning and traditional education]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 28(1), 53-67.

Barab, S. A., & Luehmann, A. L. (2002). Building sustainable science curriculum: Acknowledging and accommodating local adaptation. *Science Education*, 87, 454–467.

Berends, H., Boersma, K., & Weggeman, M. (2003). The structuration of organizational learning. *Human Relations*, 56(9), 1035-1056.

Basbay, M. Y. (2010). Proje tabanlı öğrenme [Project based learning]. In Ö. Demirel (Ed.), *Eğitimde yeni yönelimler* (4. baskı) (pp. 67-80). Ankara: Pegem A Yayıncılık.

Basbay, M., & Ales, A. (2009). The reflections of student teachers on project based learning and investigating self evaluation versus teacher evaluation. *Procedia social and Behavioural Sciences*, 1, 242-247.

Blumenfeld, P., Soloway, E., Marx, R., Krajcik, J., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, 26 (3&4), 369-398.

Cohen, L., Manion, L., & Morrison, K. (2005). *Research methods in education (5th ed.)*. UK: Taylor & Francis e-Library.

Cronbach, L., J. (1951). Coefficient alpha and the internal structure of tests. *Psyclzometrika*, 16, 297-334.

Dede, Y., & Yaman, S. (2003). Fen vema tematik eğitimin deprojeç alışma larını nyeri, önemi vede ğerlen diril mesi [The place, importance and assessment of project work in science and mathematics education]. *Gazi Eğitim Fakültesi Der*gisi, 23(1), 117-132.

Demirhan, C. (2002). *Program geliştir mede projet emelli öğren meyaklaşımı [Project basedlearning in program development*]. (Unpublished doctoral dissertation). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.

Diffily, D. (2002). Project-based learning: Meeting social studies standards and the needs of gifted learners. *Gifted Child Today*, 25(3), 40-59.

Eggen, P.D., & Kauchak, D.P. (1998). *Strategies for the teacher: Teaching Content and Thinking Skill*. Allyn and Bacon: Boston

Erdem, E., & Demirel, Ö. (2002). Constructivism in curriculum development. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 23, 81-87.

Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for research in Mathematics Education*, 44-63.

Green, A. M. (1998). *Project-based-learning: moving students toward meaningful learning*. ERIC Data base, ED422466

Kaldi, S., Filippatou, D., & Govaris, C. (2011). Project-based learning in primary schools: effects on pupils' learning and attitudes. *Education* 3–13, 39(1), 35–47.

Korkmaz, H. (2002). *Fen eğitiminde projet emelli öğrenmenin yaratıcı düşünme, problem çöz meve akademik risk alma düzeylerineetkisi [The effects of project based learning in science*

education on creative thinking, problem solution and academic risk taking levels]. (Unpublished doctoral dissertation). Hacettepe Üniversitesi, Ankara.

Krajcik, J.S., & Blumenfeld, P. (2006). Project-based learning. In Sawyer, R. K. (Ed.), *The Cambridge handbook of the learning sciences.* New York: Cambridge

Krajcik, J. S., Blumenfeld, P. C., Marx, R. W., Bass, K. M., Fredricks, J., & Soloway, E.(1998). Inquiry in project-based science classrooms: Initial attempts by middle school students. *The Journal of the Learning Sciences*, 7, 313-350.

Ladewski, B. G., Krajcik, J. S., & Harvey, C. L. (1994). A middle grade science teacher's emerging understanding of project-based instruction. *The Elementary School Journal*, 94(5), 498-515.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.

Leviatan, T. (2008). *Innovative Teaching and assessment Method: QBI and project based learning*. *Mathematics Education Research Journal,* 10(2), 105 – 116.

LI, K. Y., & Shen, S. M. (1992). Students' weaknesses in statistical projects. *Teaching Statistics,* 14(1), 2-8.

Markham, T. (2003). *Project based learning handbook: A guide to standards-focused project based learning for middle and high school teachers*. Novato, CA: Buck Institute for Education.

Miswanto. (2011). Penerapan model pembelajaran berbasis proyek pada materi program linier siswa kelas X SMK Negeri 1 Singosari [Application of project-based learning model on linear program topic of students of class X in SMK Negeri 1 Singosari]. *Jurnal Penelitian Dan Pemikiran pendidikan*, 1, 1-7.

Morrison, G. R., Ross, S. M., Kemp, J. E., & Kalman, H. (2010). *Designing effective instruction*. John Wiley & Sons.

Mudhofir. (1987). *Teknologi instruksional [Instructional technology]*. Bandung: Remaja Rosdakarya

Özdemir, E. (2006). *An ınvestıgatıon on the effects of project-based learnıng on students' achuevement ın and attıtude towards geometry*. (Unpublished dissertation). Mıddle East TechnıcalUnıversıty, Ankara.

Özdemir, E., & Ubuz, B. (2006). *Projet emelli öğren meninö ğrencil erin geometriye yönelik tutum larınaet kisi [Effects of project based learning on student attitudes towards geometry]*. VII. Ulusal Fen Bilim lerive Matematik Eğitimi Kongresi, (7-9 Eylül). (pp.225), Gazi Üniversitesi, Ankara.

Özden, M., Aydın, M., Erdem, A., & Ekmekçi, S. (2009). Öğretmenlerin proje tabanlı fen öğretimi konusunda görüşlerinin değerlendirilmesi [Evaluation of teacher views about project based learning in science education]. *Elektronik Sosyal Bilimler Dergisi*, 8(30), 92-102.

Robinson, D. L. (2009). *Project-based hıgh school geometry*. (Unpublished Thesis). B.S. Louisiana State University.

Saracaloğlu, A. S., ÖzyılmazAkamca, G., & Yeşildere, S. (2006). İlköğretimde proje temelli öğrenmenin yeri [The place of project based learning in elementary education]. *Gazi Türk Eğitim Bilimleri Dergisi*, 4(3), 241-260.

Siswono, T. Y. E., Kohar, A. W., & Hartono, S. (2017a). Secondary teachers' mathematics-related beliefs and knowledge about mathematical problem-solving. *In Journal of Physics: Conference Series* (Vol. 812, No. 1, p. 012046). IOP Publishing.

Siswono, T. Y. E., Kohar, A. W., Rosyidi, A. H., & Hartono, S. (2017b). Primary school teachers' beliefs and knowledge about mathematical problem-solving and their performance on problem-solving task. *World Transactions on Engineering and Technology Education*, 15, 2.

Slavin, R. E. (1997). *Educational psychology theory and practice*. (Fifth Edition). Boston: Allyn and Baco.

Soedjadi. (2001). *Kiat Pendidikan Matematika Indonesia [Tips for Indonesian Mathematics Education]*. Jakarta: Dirjen dikti depdikbud.

Thomas, J. W. (2000). *A review of research on project-based learning executive summary*. San Rafael, CA: The Autodesk Foundation.

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 YAYINCILIK'A 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

		Kaynakça'nın 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

Individual */ Kişi Aboneliği .....: 80 TL*

Institution */ Kurum Aboneliği .....: 250 TL*

Abroad */ Yurtdışı Aboneliği:.....: 250 $*

Student */ Öğrenci Aboneliği:.....: 70 TL*

For single issue / Tek sayılar için

Individual */ Kişiler .....: 25 TL*

Institution */ Kurumlar.....: 75 TL*

Payment / Ödeme

İş Bankası Kızılay Şubesi, Anı Yayıncılık 4214-0962561

Account number/Numaralı hesap