EURASIAN JOURNAL OF EDUCATIONAL RESEARCH

Bimontly Peer-Reviewed Journal, Year: 19 Issue: 80 / 2019
İki Ayda Bir Yayımlanan Hakemli Dergi, Yıl: 19 Sayı: 80 / 2019

EURASIAN JOURNAL OF EDUCATIONAL RESEARCH

Bimontly Peer-Reviewed Journal, Year: 19 Issue: 80 / 2019
İki Ayda Bir Yayımlanan Hakemli Dergi, Yıl: 19 Sayı: 80 / 2019

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
Özgür GÜLER

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: 20.03.2019
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 80th ISSUE
80. Sayı Hakemleri

Arda Arıkan

Cansu Ayan

Deha Doğan

Dekant Kıran

Didin Widyartono

Funda Nayır

Gökhan Atik

Hüseyin Anılan

İsmail Karakaya

K. K. Chan

Mesut Türk

Nafiye Cigdem Aktekin

Nguyen Vu

Racheal Poh

Saadet Kuru Çetin

Seçil Ömür Sünbül

Serkan Özel

Sultan Sultan

Şenel Poyrazlı

Truong Vien

Tufan Aytaç

Yasemin Taş

Eurasian Journal of Educational Research 80 (2019) 1-24

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

What is the Predictive Power of Visual Mathematics Literacy Perception and Its Sub-dimensions for Geometry Success?

Aziz ILHAN[1], Tayfun TUTAK[2], Halil Coskun CELIK[3]

ARTICLE INFO

Article History:
Received: 05 July 2017
Received in revised form: 09 June 2018
Accepted: 08 Mar. 2019
DOI: 10.14689/ejer.2019.80.1

Keywords
visual mathematics literacy, mathematics literacy, geometry success, mathematics education, teacher candidates.

ABSTRACT

Purpose: In this study, it was aimed to examine the relationship between the visual mathematics literacy perceptions and its sub-dimension for geometry success levels of prospective teachers. It was also aimed to examine to what extent visual mathematics literacy perception and its sub-dimensions predicted geometry success.

Research Methods: This study was designed in accordance with a qualitative, scanning model. The research was carried out with 232 (97 males and 135 females) prospective teachers who studied in Mathematics Education Program at Firat University, Education Faculty, and were selected by simple random sampling method. "Visual Mathematics Literacy Perception Scale" and "Geometry Success Test" were employed in this study as data collection tools. Correlation analysis and multiple regression analysis were used for analyzing the data.

Findings: After the data were analyzed, it was determined that there was a positive relationship between the visual mathematics literacy perception and geometry success of prospective teachers. It was also determined that visual mathematics literacy perception is a meaningful predictor of geometry success.

Implications for research and practice: Providing trainings on these concepts in the direction of the results, to investigate the concept of perception on other learning areas of mathematics and conducting experimental studies on the perception of visual mathematics literacy are suggested.

*This study was made by İlhan (2015) from a part of his master thesis and was partly presented at the International Conference on Mathematics and Mathematics Education Congress in Sanliurfa, 11-13 May, 2017.

[1] Corresponding Author, Munzur University, Çemişgezek Vocational School, Tunceli/Turkey, ailhan@munzur.edu.tr, Orcid ID: https://orcid.org/0000-0001-7049-5756.

[2] Fırat University, Education Faculty, Elazığ/Turkey, tayfuntutak@hotmail.com, Orcid ID: https://orcid.org/0000-0002-0277-6377.

[3] Siirt University, Education Faculty, Siirt/Turkey, hcoskun.celik@gmail.com, Orcid ID: https://orcid.org/0000-0003-0056-5338.

Introduction

The standards put forward by the National Council of Teachers of Mathematics (NCTM) in the USA aim to train prospective teachers as mathematics literate, while one of the primary goals of mathematics education in primary school mathematics curriculum published by the Turkish Ministry of National Education (MoNE) in 2013 is to achieve this (Ministry of National Education [MoNE], 2013). Traditionally, mathematics education has been very procedure-based. Over the past decade, international mathematics reform has placed more emphasis on literacy competencies and their links with mathematics learning. This shift, which has influenced both pedagogy and curriculum expectations, has increased the overlap between literacy skills and mathematics learning in instructional practice (Ontario Ministry of Education [OME], 2004). Children's and teens' ability to learn mathematics, and their awareness of mathematical thoughts may only be achieved by verbal, numerical, visual, symbolic and written communication in mathematics. In fact, "mathematics for everyone", "mathematics literacy perception" and "improvement in mathematics" have gone beyond being just slogans, and become one of the primary goals, and this constituted a field of education and research that every community should invest it.

Mathematics Literacy

Literacy was described by Karunaratne (2000) as the individuals' ability to maintain their life in their community, having enough reading-writing skills to communicate within the community, and being able to apply basic mathematical operations. It is seen that the concept of literacy is related to the area of mathematics by its definition. Therefore, the concept of literacy has gained a place in the mathematics literature.

Mathematics literacy is described as individuals' awareness and understanding of the role of mathematics in the real world, and having judgements based on solid foundations and usage of mathematics to meet their needs as a sensitive citizen (MoNE, 2013). In that case, it can be said that mathematical literacy is not only knowing mathematical concepts and solving routine problems, but also identifying oneself with mathematics (Colak, 2006). It was aimed that students should be trained as mathematically literate by standards asserted by NCTM of America and main objectives of mathematics education in the 2005 Primary Education Mathematics Education Program (NCTM, 2000). Individuals with mathematical literacy competence can keep the mathematical concepts in mind, transfer mathematical skills into daily life, and use the mathematical information in analysis and synthesis situations. In this view, for individuals to have mathematical literacy competence, some basic competency and skills about mathematics should be gained (Bekdemir & Duran, 2012). Mathematical literacy is closely related to literacy. While literacy is the foundation for all learning, mathematical literacy is also necessary if we are to fully understand information that surrounds us in modern society (OME, 2004, p. 23). Mathematics learning cannot occur without a strong literacy background. Early recognition of written and spoken letters and numbers are important stepping-stones for later development of math abilities (Cappelli, 2015). In this context, mathematical

literacy can be defined as a type of literacy that involves the use of critical thinking, reasoning, and mathematical thinking skills in solving real-life problems and strengthening the individual in mathematics field. When the literature is examined, it is possible to come across studies that indicate the importance of mathematics literacy for teaching processes and learners. Gatabi, Stacey and Gooya (2012) see mathematical modelling as the key process in mathematical literacy. In some of the studies conducted in the field of mathematics literacy, researchers focused on mathematics literacy levels of mathematics teachers (Demir & Altun, 2018; Tekin & Tekin, 2004). In other studies, the effect of mathematics literacy on variables such as gender (OECD, 2004; Ozgen & Bindak, 2011) or academic achievement (Kocaarslan & Celikturk, 2013; Tat, 2018) was examined.

Math and Visual Perceptions

Visual perception is the ability to notice visual stimulants, distinguish them and decipher these stimulants by associating them with previous experiences (Frostig, 1968). According to Kavale (1982), visual perception is related to the ability of the individual to organize his/her skills and interpret them. When visual perception skills are analyzed, it is seen that they are generally divided into sub-categories such as visual discrimination, visual figure ground, visual closure and visual memory (National Educational Psychological Service [NEPS], 2015). Additionally, it is seen that other skill types such as establishing spatial relationships as well as visual memory and visual discrimination are in a strong relationship with mathematics skills (Olkun, Altun & Deryakulu, 2009). According to Coley and Gelman, perceptual stability, which is closely related to both mathematical and visual perception, is a subject that is important throughout the lives of human beings. This is because human beings are faced with a constantly changing perspective. When we get closer to objects, the size and shape of the objects change according to the changes in our position. When we turn or raise our head, our harmony with our environment also changes. Our eye movements cause similar changes, but the environment is perceived as stable thanks to perceptual stability. The information that a triangle is a triangle, even if its location or position is changed, is related to both mathematical and visual perception. When mathematical skills are analyzed, it is seen that visual perception processes are necessary for usage and improvement of many skills (Erden & Akman, 1995; Sigmundsson, Anholt, & Talcott, 2010).

Geometry Learning Field

Although geometry is one of the primary fields of mathematical learning, the cognitive processes underlying the academic success in geometry are not examined in detail. Visual literacy and visual mathematics literacy concepts have emerged as a result of the relationship between the geometry learning area and cognitive literacy skills. Beauchamp, Braden and Baca (1994) argue that visual literacy is the fourth element of the general education in the modern world following reading, writing and arithmetic. Additionally, the reason why visual literacy is of great importance in terms of general education is explained in four items: First, the concept of visual literacy requires using the right hemisphere of the brain, which is of great importance for the

development of human beings. Thus, usage of both hemispheres of the brain in thinking process reinforces holistic thinking. Second, the concept of visual literacy makes it possible to understand abstract thoughts in the left hemisphere of the brain better by making them alive, persuasive, intense and known. Third, visual literacy provides the ability to process the same ideas in different ways. Finally, visual literacy enables individuals to read and understand the visual environment around them, so that they can make their own decisions instead of being influenced by the natural and man-made environment around them. Beauchamp et al. (1994) argue that students should be trained in three different aspects within the scope of visual literacy and geometry education: visualization through imagery, reading and interpreting visual elements, and designing visual materials. Although there is not a unity and solidarity in visual literacy, which improves rapidly and has a wide scope, it is seen that visual literacy education is necessary in the 21st century, in which people are bombarded with visuals. Geometry education to be planned in this regard should help students maximize the benefits of the visual materials among the course materials.

Visual communication and visual learning can be observed in many learning fields, especially in the geometry learning field. The use of visual materials for enriching learning processes is a teaching technique approved by educators. According to Dwyer (1978), visuals (such as TV, images, slide presentations, diagrams, graphics, etc.) are influential on teaching the facts and concepts and using valid methods. Levie (1987) argues that visuals materialize the abstract information, make them thinkable in an imaginative way, and are useful in cooperative reasoning (analogical reasoning). The abilities of teachers and other training personnel related to visual literacy and other literacy types have direct influence on education services. Visual literacy skills are highly influential on teaching materials designed by teachers, timely and effective use of teaching materials with appropriate methods and techniques, and being able to organize the message to be conveyed to the student visually (for example being able to make simple schemes and drawings).

Visual mathematics literacy has indispensable meaning and importance in both daily life and mathematics education. This is because mathematics visualized in constructs makes is easier to understand the relationships among objects (Ayguner, 2016). For some students, visual representations are obligatory to learn mathematics. The reason for some students to fail in mathematics is the inadequacy of visual elements and helpers used in mathematics education. Visual regulations for such students will increase the success of the students, and therefore, their interest in the subject and participation in the classroom (Tutkun, Erdogan & Ozturk, 2014). According to Tekin and Tekin (2004), individuals who are visually literate in mathematics have the quality of recognizing and analyzing experiences based on shapes, space, time and movement using all senses, and the representations of the concepts.

Spatial Thinking

The concept of dimension, which constitutes the basis for understanding space geometry, has a significant role in the development of mathematical thinking (Manin, 2006). Dimension concept has taken place in the primary and secondary school curricula in recent years. In order to understand the subjects in geometry, the concepts of point, line, plane and volume should be understood well. Dimensions of these concepts should be mentioned when they are introduced. In this regard, it is possible to say that geometry may be shaped by the concept of dimension. On the other hand, spatial thinking is the ability of creating the objects and status quo in one's mind. Drawing explanatory figures when solving problems, being able to transfer verbal problems into tables and graphics, and being able to understand the relationships among geometrical figures show that this skill can be improved (Turgut, 2010). Subject in the learning field of geometry play an important role in improving students' skills of objective and critical thinking, establishment of causality relationships, and numerical thinking (Oral & Ilhan, 2012). Therefore, the use of spatial skills through visualizing geometrical concepts and spatial situations constitutes the basis for spatial thinking. Improving spatial skill will also make it possible to understand mathematical and geometric subjects better. However, many studies are concerned with finding the methods that improve spatial skill the most. NCTM states that geometry will lead students to analyze the properties of geometric shapes by using spatial thinking and geometrical modelling in solutions of problems, and express geometric relationships mathematically. Accordingly, it was emphasized that the reasoning and justification skills of students will be improved by geometry, which is a natural field in mathematics (NCTM, 2000).

Visual Math Literacy

Visual literacy emerged as a result of the fact that human beings used visuals on the walls of caves, visuals are integrated into the education process, and visuals concepts increase the permanence of mental processes by materializing them. Visual literacy was described by Hortin (1980) as "readership of visual elements, the ability to think and learn with visual elements, and interpretation capacity that is, thinking visually." In this regard, the concept of visual mathematics literacy perception was described as "being able to perceive, express, interpret, assess and use the problems faced in daily life as visual and spatial, and similarly, being able to perceive, express, interpret, assess and use visual and spatial information in a mathematical sense" (Duran & Bekdemir, 2012).

According to NCTM, one of the primary goals of the process of learning geometry is the visual awareness of students as visually literate individuals. Many studies were conducted on this concept in the USA and Israeli, and even technology assisted computer programs were developed. The Visual Mathematics Institute was found in California, USA in 1975 to conduct studies on visual mathematics literacy perception. This institute, which aims to popularize mathematics among individuals, continues to carry on its duty at the University of California for visual mathematics research involving computer graphics and interactive environments (Marcolin & Abraham,

2006). On the other hand, visual mathematics literacy perception is also a research subject of the Education Technologies Department of the University of Haifa in Israel. One of the teams conducting studies in this department developed a computer program named Visual Math in the beginning of 1990s (Devraj, Butler, Gupchup & Poirier, 2010). When the software steps were examined, wide range of features of the software representing mathematical aspects of contextual problems drew attention (Yerushalmy, 2006). Students studying in 7th-12th grades improved their knowledge in geometry with a critical approach thanks to this software, which was based on geometry designs. The primary goal of Visual Math is to help students improve their algebra skills and ensure that they learn graphical reading techniques (Devraj, Butler, Gupchup & Poirier, 2010).

Visualization is the organization of data using visual elements (pictures, graphs, etc.) that can be easily perceived by the sense of seeing (Sevimli, Yildiz **&** Delice, 2008). As this concept is a different pathway in mathematical thinking than the thought of language and traditional algebra, it may be a strong and alternate resource for mathematical analyses of students (Konyalioglu, 2003). Usage of visualization in mathematics classes leads students to approach abstract concepts and constructs with different points of view. The mathematical relationships among visual objects are understood more easily with the help of visual mathematics literacy (Tutkun, Erdogan, & Ozturk, 2014). Considering the studies in the field of mathematics and geometry on visualization, it is seen that visualization is used intensively in geometrical processes (Nemirovsky & Noble, 1997). Visual literacy term is defined as the power of giving meaning to visual messages and composing message in a similar way (Alpan, 2008). In another definition for visual literacy is "the learned ability to interpret visual messages exactly and accurately, and to create such messages (Heinich, Molenda, Russell, & Smaldino, 1999, p. 64). Visual literacy is a required proficiency in education that requires teachers and instructors to be able to arrange, manipulate, and utilize graphics for the purpose of learning. By using their visual competency, they can enhance their students' learning and help them achieve academic success (Aisami, 2015).

There are a limited number of studies in the literature on visual mathematics literacy. Some of these studies were studies on developing a visual mathematics literacy scale. In this field, Bekdemir and Duran (2012) developed a visual mathematics scale for second grade students in elementary school education, and investigated the relationship between visual mathematics literacy and visual success in mathematics. Likewise, Authors (2016) developed a scale for determining the visual mathematics literacy of prospective teachers. Other studies focused on whether visual mathematics literacy levels and perception of self-efficacy changed based on variables such as perceived mathematics success, sex, and class level, and the relationships in this context. In a study by Tutkun, Erdogan and Ozturk, (2014), it was reported that the mathematics literacy levels and self-efficacy rates middle-school students were high, and these levels chanced based on sex, level of education and mathematics success. Ozdemir, Duran, and Kaptan (2016) detected a low-level and significant correlation between visual mathematics literacy and perception of self-efficacy, and reported that

the visual mathematics literacy and self-efficacy perception of students differed significantly in favor of female students. Similarly, Duran, and Bekdemir (2013) found a positive, medium-level and significant relationship between visual mathematics literacy and perception of self-efficacy. They also determined that perception of self-efficacy in visual mathematics was a significant indicator of visual mathematics success.

It is thought that visual mathematics literacy perception of prospective teachers is an important factor in mathematics education as well as the visual mathematics literacy perception of students. In addition to these, as many visuals are included in the field of geometry, it is assumed that the success in this field is associated with visual mathematics literacy perception. Based on the literature review, it is considered to be appropriate in this study to examine the relationship between success in geometry and visual mathematics literacy perception of prospective elementary school mathematics teachers. In this regard, the present study aims at addressing the following questions:

RQ_1: What is the relationship level between geometry success and visual mathematics literacy perception and its sub-dimensions?

RQ_2: What is the predictive power of visual mathematics literacy perception and its sub-dimensions for geometry success?

Method

Research Design

A descriptive method was employed in this study. Descriptive studies attempt to describe a given situation comprehensively and in detail. Descriptive screening studies are commonly conducted on educational issues. This is because researchers conduct descriptive studies in order to summarize the characteristics of individuals, groups or physical environments (Buyukozturk, Kilic Cakmak, Akgun, Karadeniz & Demirel, 2012, p. 22). In this study, relational screening model was preferred in order to obtain the data indicating the relationship between visual mathematics literacy perception and geometry success. The relational screening model is a research model that aims at determining the coexistence or degree of relationship between two or more variables.

Research Sample

The population of the study consisted of prospective teachers in the field of Mathematics Education at the, Faculty of Education, Firat University in the fall semester of 2015-2016 academic year. The sample consisted of 232 volunteer prospective teachers, 97 (41.8%) were male and 135 (58.2%) were female selected through simple random sampling method. The number of students in this sample group were 72.5% of all students. The arithmetic mean of prospective teachers' age was 26.23. The reason for choosing the method of simple random sampling was that,

in this method, all units constituting the population have the same possibility of being included in the sample (Can, 2013, p. 26).

Research Instruments and Procedures

The study used the Visual Mathematics Literacy Perception Scale and the Geometry Success Test (GST), which were developed by researchers as data collection tools. The primary assessment tool, the Visual Mathematics Literacy Perception Scale, was developed by Ilhan (2015) in order to determine the visual mathematics literacy perception levels of prospective teachers, and is a 5-point Likert-type scale consisting of 37 items. This scale consisted of five sub-dimensions: visual perception, geometric field, spatial intelligence, concretion, and create a pattern. Some items in the scale are as follows; "I can draw a three-dimensional shape from front to top view", "By breaking a three-dimensional object, I can obtain new three-dimensional objects", "I can find the general term from a shaped pattern with steps 3 and 5", "Modeling a decimal number to create the problem, I can solve it", "I can make geometric proof of Pythagoras" and "I can make geometric modeling of exponential numbers". Factor analysis studies were conducted to investigate the construct validity of the scale. Factor analysis methods used in the research were exploratory and confirmatory factor analyzes. With the exploratory factor analysis, it was tried to determine what factors were related to the scale. In order to ensure the validity of the scope, a pool of 60 items, which aimed to measure scale in specific dimensions, was formed as a result of the opinions of teacher candidates and the related literature review. The factor load of the scale ranged from 0.410 to 0.716. The internal consistency reliability coefficients of the scale were 0.820, 0.740, 0.891, 0.763 and 0.852 for the five sub-dimensions respectively, and the test-retest reliability coefficients were 0.764, 0.833, 0.856, 0.823 and 0.841 for the five sub-dimensions respectively. The Cronbach's Alpha reliability coefficient of the scale was 0.904. The Cronbach's Alpha reliability coefficient of the scale employed in this study was calculated as 0.889. If this coefficient is greater than 0.7, it is considered that the mean scores of the scale is reliable (Buyukozturk, 2018, p. 183).

The second assessment tool, the GST, was developed by the researchers in order to determine the geometry success levels of the students. The geometry outcomes of the elementary school mathematics curriculum were examined before preparing the success test. A draft test form was prepared, consisting of 26 multiple-choice questions previously used in the Academic Personnel and Postgraduate Education Entrance Exams between the years of 2010-2015, which included these outcomes. These questions were submitted to two faculty members, experts in the field of mathematics education, for their opinions in terms of content. In line with the feedbacks of these expert faculty members, 6 questions were taken out and the final test was created, consisting of 20 questions. The highest score was determined as 20 (each of the question is 1 point) and the lowest score was determined as 1 for this specific test. The reliability coefficients of this test were examined using KR-20 reliability coefficients and split-half reliability. KR-20 reliability coefficient for a study examines whether all the questions in a scale constitutes a homogeneous structure or not (Kalayci, 2010, p. 405). In split-half reliability, the test is split into two pairs as singular-plural, first half-second half and neutral. Based on the relationships between the two halves, the

correlation coefficient was calculated for the whole test by using Spearman Brown formula (Buyukozturk, 2018, p. 183). As a result of the analyses conducted, the KR-20 reliability coefficient of the GST was calculated as 0.794 and split-half reliability was calculated as 0.693. In this study, item difficulty and distinctiveness indices were calculated for each question. These indices are given in Table 1.

Table 1

Item Difficulty Index and Item Distinctiveness Index of GST

Item No	*Item Difficulty Index*	*Item Distinctiveness Index*
1	0.431	0.592
2	0.520	0.560
3	0.602	0.571
4	0.493	0.613
5	0.391	0.530
6	0.511	0.661
7	0.590	0.582
8	0.432	0.520
9	0.510	0.581
10	0.673	0.773
11	0.621	0.520
12	0.460	0.552
13	0.510	0.470
14	0.572	0.691
15	0.531	0.633
16	0.545	0.550
17	0.343	0.591
18	0.304	0.542
19	0.371	0.661
20	0.330	0.550
Total	0.493	0.589

The item difficulty index values of the substances found in a test vary between 0-1 and the items with difficulty indexes between 0.30 and 0.70 are considered as substances with an average difficulty level. In addition, substances with a substance difficulty index of less than 0.3 are found to be difficult and substances with a substance difficulty index of over 0.7 are considered to be difficult substances (Tekin, 1997; Yilmaz, 1998). The difficulties in the GST test (pj) ranged from 0.271 to 0.673. Therefore, it can be said that the majority of the substances in the test have a moderate difficulty. The overall difficulty index of the test was 0.493. In other words, it is possible to say that the test has moderate difficulty in general. The discriminant index (rjx) values ranged between -1 and +1, and a value of 0.40 and above indicated that the substances were very well distinguished. If this value is in the range of 0.30 to 0.39, the best substance should be corrected if it is in the range of 0.20 and 0.29, and if it is 0.19 and lower, it can be considered as the item to be removed from the test (Buyukozturk, 2018). When the discriminant index of GSAYB test items were examined, it was seen that the discrimination indexes of all items were greater than 0.40. That is to say, all items were suitable for testing in terms of discrimination.

During data collection, teacher candidates were informed about the application process. Forms were reproduced by foreseeing the number of samples. Pencil and eraser needs of teacher candidates were met. The implementation period lasted for two weeks. Visual mathematics literacy perception scale was administered in the first week, and in the second week, geometry achievement test was applied. In the application process, the classes of the related university were used. During the applications, attention was paid to keep the environment quiet and to offer individual solutions for the problems of the teacher candidates. Thus, the implementation process was completed without any problems.

Data Analysis

The relationship between the visual mathematics literacy perception and geometry success of prospective teachers was calculated using the Pearson Moments Multiplication Correlation Coefficient method, and the Multiple Regression Analysis method was used to examine whether visual mathematics literacy perception and its sub-dimensions were meaningful predictors of geometry success or not. Additionally, the mean and standard deviation values for visual mathematics literacy perception and sub-dimensions, its sub-dimensions and geometry success were also calculated. The visual mathematics literacy perception scale had 37 items and the items were scored in 5-point Likert-type. In the geometry achievement test, there were 20 items, and the items were scored as 1 for the correct answers and 0 for the wrong answers. In this direction, the visual mathematics literacy perception of students was taken as the independent (predictive) variable and geometry success was taken as the dependent (predicted) variable. Multiple Regression Analysis is a method related to the prediction of the dependent variable on the basis of two or more independent variables associated with the dependent variable (Buyukozturk, 2018, p. 98). The data were analyzed using SPSS (Statistical Package for Social Sciences) version 21.0.

Results

In this section of the study, we will discuss the findings related to the relationship between visual mathematics literacy perception, its sub-dimensions and geometry success of the prospective teachers; and then, findings will be presented in relation to determining the predictive power of visual mathematics literacy perceptions and, its sub-dimensions in terms of geometry success.

The relationship between visual mathematics literacy perception, its sub-dimensions and geometry success of the prospective teachers was calculated using correlation analysis. The results of the data analysis are shown in Table 2.

Table 2

The Relationship Levels Between Visual Mathematics Literacy Perceptions and Sub-Dimensions with Geometry Success

Variables		*Mean*	*S.S.*	*G1*	*G2*	*G3*	*G4*	*G5*	*G6*
G1	Geometry Success	9.957	4.179	1					
G2	Visual Perception	2.967	0.739	0.076*	1				
G3	Geometric Field	4.044	0.653	0.175*	0.435*	1			
G4	Spatial Intelligence	3.366	0.313	0.259*	0.501*	0.479*	1		
G5	Concretion	4.359	0.680	0.196*	0.344*	0.421*	0.402*	1	
G6	Pattern	3.826	0.792	0.107*	0.485*	0.596*	0.434*	0.375*	1
G7	Visual Math Literacy	3.712	0.635	0.191*	0.062*	0.150*	0.249*	0.172*	0.077*

*p<0.01; n=232

When the data in Table 2 were examined, the mathematics teacher candidates' Visual Math Literacy Perception sub-dimension's highest mean belonged to Concretion dimension's ($\overline{X}$=4.359), and this dimension appeared to follow Geometric Field dimension's. Lowest mean belonged to Visual Perception dimension's ($\overline{X}$=2.957). Prospective teachers' geometric achievement scores were below mean ($\overline{X}$=9.957). All of the correlations between the sub-dimensions appeared to be significant. The relationship between Visual Math Literacy Perception with all sub-dimensions was meaningful at the level of p<0.01. Also, when the relation of geometry success to sub-dimensions was examined, Spatial Intelligence (Correlate=0.259) dimension had the highest mean. This was followed by Concretion dimension's mean (Correlate=0.196). Geometry success' correlation with Visual Math Literacy Perception was seen 0.191.

It has been assumed in the study that visual mathematical literacy and sub-dimensions significantly predict the geometric success. Regression analysis was performed in this direction, and the findings are given in Table 3.

Table 3

The Results of the Multiple Regression Analysis Regarding the Predictive Power of Visual Mathematics Literacy Perception and its Sub-Dimensions

Predictive Variables	*R*	*R²*	*Change (R²)*	*Std.β*	*t*	*F*	*p*
1. Visual Perception-Visual Perception	0.760	0.577	0.001	0,031 0.030	7.644 6.654	1.337	0.000
2. Visual Perception-Geometric Field	0.175	0.031	0.031	0.002 0.174	3.086 2.425	3.624	0.028
3. Visual Perception-Spatial Intelligence	0.268	0.072	0.064	- 0.034 - 0.084	4.980 4.040	8.873	0.000

Table 3 Continue

4. Visual Perception-Concretion	0.196	0.038	0.011	0.005	3.348	4.582	0.011
				0.028	2.790		
5. Geometric Field-Geometric Field	0.175	0.031	0.026	0.112	3.192	7.279	0.007
				0.042	2.698		
6. Geometric Field-Spatial Intelligence	0.265	0.070	0.067	0.043	2.585	8.658	0.000
				0.046	3.124		
7. Geometric Field-Concretion	0.219	0.048	0.110	0.070	2.029	5.773	0.004
				0.046	2.041		
8. Geometric Field-Pattern	0.175	0.031	0.175	0.112	3.114	3.624	0.028
				0.053	2.129		
9. Spatial Intelligence-Spatial Intelligence	0.259	0.067	0.029	0.064	5.041	16.475	0.000
				0.010	4.059		
10. Spatial Intelligence-Concretion	0.276	0.076	0.019	0.032	2.521	9.461	0.000
				0.016	3.068		
11. Spatial Intelligence-Pattern	0.259	0.067	0.020	0.032	3.903	8.203	0.000
				0.021	3.686		
12. Concretion-Concretion	0.196	0.038	0.099	0.017	3.709	9.176	0.003
				0.001	3.029		
13. Concretion-Pattern	0.200	0.040	0.077	0.222	2.906	4.778	0.009
				0.085	2.611		

When Table 3 was examined, the sub-dimensions of visual mathematics literacy perception such as visual perception, geometric field, spatial intelligence, concretion and pattern were used as the predictors of visual math literacy perceptions in the regression analysis (R=0.760, p=0.00<0.01). It was observed that the spatial intelligence and concretion sub-dimensions of visual mathematics literacy perception explained 7.6% of the data. The two best sub-dimensions of visual mathematics literacy perception were spatial intelligence (R=0.268, p=0.00<0.01) and concretion (R=0.196, p=0.00<0.01).

In this study, after predictive power of the sub-dimensions of the literacy perception was examined, predictive power of the geometric success of visual mathematics literacy perception sub-dimensions was investigated. The findings are given in Table 4.

Table 4

Predictive Power of the Geometric Success of the Visual Mathematics Literacy Perception Sub-Dimensions

Variables	***B***	***Standard Error***	***Beta***	***t***	***p***
Constant	4.975	1.712	-	2.906	0.004
Visual Perception	0.031	0.027	0.076	1.157	0.249
Geometric Field	0.112	0.042	0.175	2.698	0.007
Spatial Intelligence	0.260	0.064	0.259	4.059	0.000
Concretion	0.241	0.079	0.196	3.029	0.003
Pattern	0.189	0.115	0.107	1.635	0.103

When Table 4 was examined, it was seen that the highest ordering power was in the spatial intelligence sub-dimension with 0.259, followed by concretion dimension (0.196), geometric field (0.175), pattern (0.107); and these differences were statistically significant. On the other hand, the lowest ordering power was in the visual perception dimension with 0.076, but this was not significant. After determining the geometry success of the sub-dimensions, the overall power of the visual mathematical literacy perception was investigated in terms of eometric success. The findings are given in Table 5.

Table 5

Predictive Power of the Geometric Success of the Visual Mathematics Literacy Perception

Variance Source	***Sum of Squares***	***df***	***squares mean***	***F***	***p***	***R***
Regression	146.834	1	146.834	8.685	0.004	0.191
Error	388.735	230	16.908			
Total	4035.569	231				

When Table 5 was examined, it was seen that visual mathematics literacy perception had a significant effect on geometric success at the 0.01 level ($F(1,230)=8.685$; $p=0.004$). The regression coefficient between visual mathematical literacy perception and geometric success was calculated ($R=0.191$, $p=0.00<0.01$). It was observed that geometric success of visual mathematics literacy perception explained 1.91% of the data.

Discussion, Conclusion and Recommendations

Recently, visuals are gaining importance in all areas of the field of mathematics. Moreover, these visuals strengthen the relationship between mathematics and real life, and encourage permanent learning. While high visual mathematics literacy perception levels of students make it possible for them to understand the described mathematical concepts better, high visual mathematics literacy perception levels of teachers make it possible for them to create more useful and healthier visual materials, as well as facilitate permanent learning by creating more efficient teaching processes (MoNE, 2013).

In this study, the relationship levels of visual mathematics literacy perception and its sub-dimensions with each other and geometry success were examined, and then, the extent to which visual mathematics literacy perception and its sub-dimensions predicted geometry success was calculated. When the relationship of visual mathematics literacy perception and its sub-dimensions with geometry success was examined, it was determined that there was a statistically significant relationship with its sub-dimensions. All of the correlations between the sub-dimensions appeared significant. The relationship between Visual Math Literacy Perception with all its sub-dimensions was meaningful. Also, when the relation of geometry success to its sub-dimensions was examined, spatial intelligence dimension had the highest mean value, and the concretion dimension appeared to follow because visual mathematics literacy perception was closely related to spatial intelligence. In a study conducted by Bal

(2012), it was observed that prospective teachers had different geometric thinking levels, their attitudes towards geometry were high and there was a statistically significant, but a low-level of relationship, only in the dimension of 'anxiety' related to, geometric thinking levels and attitudes. Similarly, Gellert (2004) expressed that there was a statistically significant relationship between the concept of mathematics literacy perception and mathematics lessons where instructive materials were used. On the other hand, based on the literature review, it was seen that there were studies that found the relationship between visual mathematics literacy and geometry success on a medium or high level. In a study conducted by Kocaarslan and Celikturk (2013), it was determined that the visual literacy competence levels of the students in the faculty of education were generally high and there was a statistically significant and positive relationship between visual literacy levels and academic success. Another finding of this study was that geometric field was the most significant source that predicted geometry success. In other words, students who were more successful in geometry had higher visual mathematics perceptions based on the geometric field variable. Literature review showed that there are other studies supporting this finding. In parallel with these results, Duran and Bekdemir (2013) found a medium level, positive and statistically significant relationship between visual mathematics literacy self-efficacy perception and visual mathematics success.

Geometry success of prospective elementary school mathematics teachers in exams prepared by the Student Selection and Placement Center (SSPC), their performances in different selection and placement exams and their performances related to the field of geometry in their daily life constitute personal experiences in visual mathematics literacy perceptions. As a result of the positive and negative outcomes of these experiences, mathematical success of the students was associated with visual mathematical literacy. In Turkey, students answered 5.1 mathematics and geometry questions correct out of 40 in the Transition to Higher Education Examination in 2017; they answered 4.22 geometry questions correct out of 30 in the Undergraduate Placement Exam in 2016 (SSPC, 2017). According to studies in the literature, relationships among the sources were identified similarly, and their effects on geometry success was medium or low. In parallel to the findings of this study, Garderen determined in a study conducted in 2006 that students with visual-spatial impairment had deficiencies in understanding and differentiating information. Rapp (2009) expressed that students with visual-spatial intelligence were held back when teaching techniques were not supported by visuals in the classroom. Roblyer and Bennett (2001) propose that teachers will require skills that enable them to select appropriate materials for meaningful instruction, support the effective production of materials, design and teaching of specific activities that will facilitate deep learning as well as the ability to evaluate the level of student's visual literacy. Success in higher education is becoming more and more dependent on visual literacy skills (Nalinci & Yapici, 2015).

The findings of this study showed that there was a significant relationship between visual perception and geometry success. The sub-dimensions of visual mathematics literacy perception such as geometric field, spatial intelligence, concretion and pattern

were used as the predictors of visual math literacy perceptions in the regression analysis. One of the sub-dimensions of visual mathematical literacy perception was visual perception. It is not possible to ignore the relationship of this source with visual math literacy perceptions. Literature review showed that visual perceptions are directly or indirectly associated with geometry success (Bekdemir & Duran, 2012; Hortin, 1980; Karunaratne, 2000). Again, as a result of the findings, it was determined that there was a low-level relationship between pattern sub-dimension, which was a source of visual mathematical literacy, and geometry success. According to the literature review, pattern sub-dimension is directly or indirectly associated with geometry success. In another study, Tanisli and Kose (2011) adopted a numerical approach in which the prospective teachers transformed a linear shape pattern into a visual and shape pattern with a focus on the shape structure when determining the rule of pattern and sustaining the pattern to a close/distant step, and they used 26 strategies in these approaches.

Visual perception, geometric field and spatial intelligence sources were more effective on geometry success when compared to other sources. Each source explained approximately 7% of geometry success. Considering other sources, it was seen that the effect of spatial intelligence source was higher. Spatial intelligence, visual perception and geometric field sources should be improved in order to improve visual mathematical literacy of students. Furthermore, it is important to help students gain accurate, complete and successful experiences. Lastly, concretions and pattern sources were the fourth and fifth important sub-dimensions that affected geometry success. Each of these sub-dimensions explained approximately 4% of geometry success. When compared to other sub-dimensions, it was seen that the effect of concretions sub-dimension was lower. In line with the findings obtained as a result of this study, it is concluded that it is required to examine direct and indirect relationships between the sub-dimensions of visual mathematics literacy perception and geometry success. Duran and Bekdemir (2013) stated that visual mathematics literacy self–efficacy perception is a significant predictor of visual mathematics success. Bal (2012) claimed that teacher candidates they were located in different geometric levels, between a high level of attitude towards geometry and a low level of attitude with geometric thinking levels, but it was observed that the relationship was at a low level. Ozgen and Bindak (2011) determined in their study that the success scores in mathematics courses and the importance given to mathematics courses are significant predictors of mathematical literacy self – efficacy perceptions. In addition to these studies, there are other studies in literature that determined the factors predicting mathematics success (Dogan & Baris, 2010; Kayagil, 2010; Ozdemir, 2010; Ozer & Anil, 2011; Uredi & Uredi, 2005; Yilmaz, 2006). The finding that the predictive powers of the sub-dimensions of visual mathematics literacy perception are statistically meaningful brings the importance of this relationship to the forefront. Furthermore, it was also found that visual mathematics literacy perception had a significant effect on geometric success. The reason for this is that the research was conducted only at a university, or the number of samples was low. Another reason is that teacher candidates could not reflect their perception levels to their achievements. It is also assumed that the sample

in the study reflected all the information on the test and the scale. The work to be done in the future in accordance with these limitations can include;

1. research on larger samples related to the subject,

2. investigation of the relationship between visual mathematical literacy perceptions and geometric successes of different parts,

3. similar practices on existing mathematics teachers,

4. trainings on these concepts in the direction of the results,

5. investigating the concept of perception related to other learning areas of mathematics,

6. conducting experimental studies on the perception of visual mathematics literacy,

7. conducting projects to gain visual perception of visual mathematics literacy.

References

Aisami, R.S. (2015). Learning styles and visual literacy for learning and performance. *Procedia - Social and Behavioral Sciences*, 176, 538 – 545

Alpan, G. (2008). Visual literacy and instructional technology. *Yuzuncu Yil University Journal of Education Faculty, 5*(1), *74-102.*

İlhan, A. & Celik, H.C. (2016). Development of visual mathematics literacy scale for elementary education mathematics teacher candidates. *US-China Education Review, 6*(7), 391- 401.

Ayguner, E. (2016). *A comparison of eight grade students' self-efficacy perception of visual mathematics literacy and their actual performance* (Unpublished master's thesis). Eskisehir Osmangazi University, Institute of Educational Sciences, Eskisehir.

Bal, A.P. (2012). Teacher candidates' geometric thinking levels and attitudes to geometry. *Journal of Educational Sciences Research, 2*(1), 17-34.

Beauchamp, D.G., Braden, R.A., & Baca, J.C. (1994). Visual literacy in the digital age: Selected readings from the annual conference of the international visual literacy association. *Proceeding of 25th Annual Conference of the International Visual Literacy Association, 13-17 October 1994,* Blacksburg, Virginia.

Bekdemir, M., & Duran, M. (2012). Development of a visual math literacy self-efficacy perception scale (VMLSEPS) for elementary students. *On Dokuz Mayis University Faculty of Education Journal*, *31*(1), 89-115.

Buyukozturk, S. (2018). *Data analysis handbook for social sciences (24th edition).* Ankara: Pegem Akademi Publishing.

Buyukozturk, S., Kilic Cakmak, E., Akgun, O.E., Karadeniz, S., & Demirel, F. (2012*). Scientific research methods (11th edition).* Ankara: Pegem Akademi Publishing

Can, A. (2013). *Quantitative data analysis in the scientific research process with SPSS*. Ankara: Pegem Academy.

Cappelli, A. (2015). Implementing literacy strategies and activities to help math students in geometry. *Education Masters,* p. 318. Retrieved from http://fisherpub.sjfc.edu/education_ETD_masters/318

Colak, S. (2006). *An experimental study on the impact of material use on mathematical literary in the context of geometrical concepts in sixth grade students* (Unpublished master's thesis). Gazi University, Institute of Sciences, Ankara.

Demir, F., & Altun, M. (2018). Development of mathematical literacy question writing process and skills. *Education and Science, 43*(194), 19-41

Devraj, R., Butler, L.M., Gupchup, G.V., & Poirier, T.I. (2010). Active-learning strategies to develop health literacy knowledge and skills. *American journal of pharmaceutical education, 74*(8), 137.

Dogan, N., & Baris, F. (2010). The levels of attitude, value and self-efficacy of students' mathematical achievement levels in the TIMSS-1999 and TIMSS-2007 exams. *Journal of Measurement and Evaluation in Education and Psychology, 1*(1), 44-50.

Duran, M. (2012). *Relationship between visual math literacy self-efficacy perceptions with visual mathematics achievements of elementary 7th grade students* (Unpublished master's thesis). Erzincan University, Institute of Sciences, Erzincan.

Duran, M., & Bekdemir, M. (2013). Evaluation of visual math literacy self-efficacy perception with visual mathematics accomplishment. *Pegem Journal of Education & Instruction,* 3(3), 27-40

Dwyer, F. M. (1978). Strategies for improving visual learning. State College, PA: Learning Services.

Erden, M., & Akman, Y. (1995). *Educational psychology: Teaching developmental learning*. Ankara: Arkadas Publications.

Gatabi, A.R., Stacey, K., & Gooya, Z. (2012). Investigating grade nine textbook problems for characteristics related to mathematical literacy. *Mathematics Education Research Journal, 24*(4), 403–421. doi:10.1007/s13394-012-0052-5

Gellert, U. (2004). Didactic material confronted with the concept of mathematical literacy. *Educational Studies in Mathematics, 55*(1), 163-179.

Heinich, R., Molenda, M., Russell, J.D., & Smaldino, S. E. (1999). *Instructional media and technologies for learning* (6th ed.). Upper Saddle River, NJ: Prentice-Hall.

Hortin, J. A. (1980). *Visual literacy and visual thinking*. New York: America publication.

Ilhan, A. (2015). *Developing the visual mathematics literacy scale for primary mathematics teacher candidates and examining the relationship between visual mathematics literacy and geometry achievements* (Unpublished master's thesis). Firat University Institute of Educational Sciences, Elazig.

Kavale, K. (1982). Meta-analysis of the relationship between visual perceptual skills and reading achievement. *Journal of Learning Disabilities, 15*(1), 42-51.

Kayagil, S. (2010). *Prediction of mathematics achievement by critical thinking skills in seventh grade students* (Unpublished master's thesis). Selcuk University, Institute of Educational Sciences, Konya.

Karunaratne, W. (2000). *Case for adult literacy in South East Asia with special reference to Sri Lanka.* Paper presented at the Australian Council for Adult Literacy Conference, 21-23 September, Perth, Western Australia: ACAL.

Kalayci, S (Ed.). (2010). *SPSS applied multivariate statistical techniques.* Ankara: Asil Publication.

Karasar, N. (2011). *Scientific research method.* Ankara: Nobel Publication Distribution.

Kocaarslan, M., & Celikturk, Z. (2013). Determination of visual literacy competences of education faculty students. *Bartin University Education Faculty Journal, 2*(2), 344-362.

Konyaligolu, A. C. (2003). *Investigation of effectiveness of visualization approach on understanding of concepts in vector spaces at the university level* (Unpublished doctoral dissertation), Ataturk University, Institute of Sciences, Erzurum.

Levie, W. H. (1987). Research on pictures: A guide to the literature. In D. M. Willows & H. A. Houghton (Eds.), *The psychology of illustration: Vol I. Basic research* (pp. 1-50). New York: Springer-Verlag.

Marcolin, S., & Abraham, A. (2006). *Financial literacy research: Current literature and future opportunities.* In P. Basu, G. O' Neill & A. Travaglione (Eds.), Proceedings of the 3rd International Conference on Contemporary Business, Leura NSW, 21-22 September 2006. Australia: Faculty of Commerce, Charles Stuart University

Manin, Y. I. (2006). The notion of dimension in geometry and algebra. Bull. *Amer. Math. Soc. (N.S.), 43* (2), 139–161. Retrieved from http://www.ams.org/journals/bull/2006-43-02/S0273-0979-06-01081-0/ S0273-0979-06-01081-0.pdf

Turkish Ministry of National Education. (2013). *Primary mathematics lesson 6-8. Grade curriculum*. Ankara: Board of Education and Training Board.

Nalinci, G. Z., & Yapici, M. (2015). A discussion of the practicality of the visual literacy competency standards for higher education. *Anadolu Journal of Educational Sciences International, Art Education Special Issue, (1),* 225-241.

National Educational Psychological Service. (2015). V*isual perception.* Retrieved from https://www.education.ie/en/Schools-Colleges/Services/National-Educational-Psychological-Service-NEPS-/NEPS Guides/Cognitive-Skills/Cognitive-Skills-Visual-Perception.pdf

National Council of Teachers of Mathematics. (2000). *Principal and standards for school mathematics*. Reston: NCTM Publications.

Nemirovsky, R., & Noble, T. (1997). On mathematical visualization and the place where we live. *Educational Studies in Mathematics, 33*(2), 99–131

Olkun, S., Altun, A., & Deryakulu, D. (2009). Development and evaluation of a case-based digital learning tool about children's mathematical thinking for elementary school teachers (L-TEST). *European Journal of Teacher Education, 32*(2), 151-165.

Ontario Ministry of Education. (2004). *Leading math success: Mathematical literacy, grades 7–12. The Report of the Expert Panel on Student Success in Ontario*. Retrieved from http://www.edu.gov.on.ca/eng/document/reports/numeracy/numeracyreport.pdf

Oral, B., & Ilhan, M. (2012). Analysis of geometric thinking levels of candidate mathematics teachers of primary and secondary schools in terms of various variables. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education, 6*(1), 201-219.

Ozdemir, F. (2010). *Analysis of factors which effects the mathematics achievements in PISA 2003 high school students and Kanuni High School Students* (Unpublished master's thesis). Hacettepe University, Institute of Social Sciences, Ankara.

Ozdemir, F., Duran, M., & Kaplan, A. (2016). Investigation of middle school students' self-efficacy perceptions of visual mathematics literacy and perceptions of problem-solving skill. *Journal of Theoretical Educational Science, 9*(4), 532-554

Ozer, Y., & Anil, D. (2011). Examining the factors affecting students' science and Mathematics achievement with structural equation modeling. *Hacettepe University Journal of Education, 41*(1), 313-324.

Ozgen, K., & Bindak, R. (2011). Determination of self-efficacy beliefs of high school students towards math literacy. *Educational Sciences: Theory & Practice*, *11*(2), 1073-1089.

Rapp, W. H. (2009). Avoiding math taboos: Effective math strategies for visual-spatial learners. *Teaching Exceptional Children Plus, 6*(2), 1-12.

Roblyer, M. D., & Bennett, E.K. (2001). The fifth literacy: Research to support a mandate for technology-based visual literacy in pre-service teacher education. *Journal of Computing in Teacher Education, 17*(2), 8-15.

Sevimli, E., Yildiz, C., & Delice, A. (2008). *An overview of the visualization process in geometry questions: Where do I draw?* Paper presented at the 8th National Science and Mathematics Education Congress, Abant Izzet Baysal University, Bolu.

Sigmundsson, H., Anholt, S., & Talcott, J. B. (2010). Are poor mathematics skills associated with visual deficits in temporal processing? *Neurosci. Lett.* 469, 248–250.

Student Selection and Placement Center. (2017). *Numerical information for 2017-YGS and 2016 LYS exams.* Retrieved from http://www.osym.gov.tr/

Tambychik, T., Meerah, T. S. M., & Aziz, Z. (2010). Mathematics skills difficulties: A mixture of intricacies. *Procedia-Social and Behavioral Sciences, 7*(1), 171-180.

Tanisli, D., & Kose, N. Y. (2011). Generalization strategies about linear figural patterns: effect of figural and numerical clues. *Education and Science, 36*(160), 184-198.

Tat, E. T. (2018). Prospective mathematics teachers' perceived self-efficacy in mathematical literacy. *Elementary Education Online, 17*(2), 489-499

Tekin, B., & Tekin, S. (2004). A research on mathematical literacy levels of preservice mathematics teachers'. Retrieved from http://matder.org.tr

Tekin, H. (1997). *Measurement and evaluation in education*. Ankara: Mars Printing House.

Tutkun, O. E., Erdogan, D. G., & Ozturk, B. (2014). Levels of visual mathematics literacy self-efficacy perception of the secondary school students. *Middle Eastern & African Journal of Educational Research, 8*, 19-27.

Turgut, M. (2010). *The effect of technology assisted linear algebra instruction on pre-service primary mathematics teachers' spatial ability* (Unpublished doctoral dissertation). Dokuz Eylul University, Institute of Educational Sciences, Izmir.

Uredi, I., & Uredi, L. (2005). The predictive power of self-regulation strategies and motivational beliefs on mathematics achievement of primary school 8th grade students. *Mersin University Journal of Education Faculty, 1*(2), 250-260.

Van Garderen, D. (2006). Spatial visualization, visual imagery, and mathematical problem solving of students with varying abilities. *Journal of Learning Disabilities, 39*(6), 496-506.

Yerushalmy, M. (2006). *Challenging known transitions: Research of technology supported long-term learning.* Paper presented at the conference of the Seventeenth International Commission on Mathematical Instruction Study, Hanoi University of Technology, Hanoi.

Yilmaz, E. T. (2006). *The factors that effect the mathematics achievement of the students of Turkey in programme for international student assessment (PISA)* (Unpublished master's thesis). Hacettepe University, Institute of Social Sciences, Ankara.

Yilmaz, H. (1998). *Measurement and evaluation in education (3rd bs.).* Ankara: Micro Publishing.

Görsel Matematik Okuryazarlığı Alt Boyutlarının Geometri Başarısını Yordama Gücü Nedir?

Atıf:
Ilhan, A., Tutak, T., & Celik, H. C. (2019). What is the predictive power of visual mathematics literacy perception and its sub-dimensions for geometry success? *Eurasian Journal of Educational Research, 80,* 1-24, DOI: 10.14689/ejer.2019.80.1

Özet

Problem Durumu: Okuryazarlık kavramı günümüzde şekillenen bir kavram olarak bilinmekle birlikte tarihsel sürecinin de çok eski olmadığını araştırmalar neticesinde görmek mümkündür. Ayrıca okuryazarlık kavramı çok genel bir kavramdır ve bu kavramla ilişkili yeni okuryazarlık türleri literatürde her geçen gün oluşturulmuş ve oluşturulmaktadır. İnsanların ilk çağlardan beri mağara duvarlarında görselleri kullanması, görsellerin eğitim sürecinin neredeyse tamamına entegre edilmesi ve görsel kavramların zihinsel süreçte somutlaştırmayı sağlayarak kalıcılığı artırması görsel okuryazarlık kavramını doğurmuştur. Görsel okuryazarlık, Hortin (1980) tarafından, "görsel elemanları okuma ve yorumlama kapasitesiyle beraber görsel öğeler ile düşünme ve öğrenme becerisi, yani görsel olarak düşünebilme" şeklinde tanımlanmıştır. Eğitim alanında farklı okuryazarlıklara dair ortak yanların bütünleşmesinden doğan sanatsal matematik veya görsel matematik okuryazarlığı gibi okuryazarlıkların tanımlanması ön plana çıkmıştır. Bu noktadan hareketle görsel matematik okuryazarlığı "günlük hayatta karşılaşılan problemleri görsel veya uzamsal, tersine görsel veya uzamsal bilgileri de matematiksel olarak algılayabilme, ifade edebilme, yorumlayabilme, değerlendirme ve kullanabilme yeterliğidir" (Duran ve Bekdemir, 2013). Görsel matematik okuryazarlığının günlük yaşamda ve matematik eğitiminde vazgeçilmez bir anlamı ve önemi vardır. Görsel okuryazarlık ve görsel matematik okuryazarlığı kavramları, geometri öğrenme alanı ile bilişsel okuryazarlık becerileri arasındaki ilişki sonucunda ortaya çıkmıştır. Geometri matematiksel öğrenmenin temel alanlarından biri olmasına rağmen, geometriyle ilgili akademik başarının altında yatan bilişsel süreçler detaylı bir şekilde incelenmiştir. Öğretmen adaylarının görsel matematik okuryazarlığı algılarının, matematik eğitiminde önemli bir faktör olduğu düşünülmektedir. Bunlara ek olarak, birçok görselin geometri alanına dâhil edilmesi nedeniyle, bu alandaki başarının görsel matematik okuryazarlığı algısı ile ilişkili olduğu düşünülmektedir. Dolayısıyla görsel matematik okuryazarlık algısı ve alt boyutları ile geometri başarısı arasındaki ilişkinin incelenmesinin alan yazına yararlı olacağı öngörülmektedir.

Araştırmanın Amacı: Bu araştırmanın amacı, öğretmen adaylarının görsel matematik okuryazarlığı ile geometri başarıları arasındaki ilişkiyi incelemektir.

Araştırmanın Yöntemi: Mevcut araştırma, betimsel nitelikli bir çalışmadır. Betimsel çalışmalar, verilen bir durumu olabildiğince tam ve dikkatli bir şekilde tanımlamaya çalışır. Eğitim alanındaki araştırmalarda, yaygın olarak betimsel yöntem tarama çalışmaları yapılmaktadır. Mevcut araştırmada görsel matematik okuryazarlığı ile geometri başarısı arasındaki ilişkiyi belirten verilerin elde edilmesi noktasında ilişkisel tarama modeli tercih edilmiştir. Araştırmanın çalışma grubunu, 2015-2016 öğretim yılı

güz döneminde, Fırat Üniversitesi, Eğitim Fakültesi, matematik eğitimi anabilim dalında okuyan matematik öğretmen adayları oluşturmaktadır. Örneklemde ise bu öğrencilerden basit rastgele örnekleme yöntemiyle seçilen 232 (97 erkek, 135 kadın) öğrenci yer almıştır.

Araştırmada, veri toplama araçları olarak araştırmacı tarafından geliştirilen ve öğretmen adaylarının görsel matematik okuryazarlık algı düzeylerini belirlemek amacıyla Görsel Matematik Okuryazarlığı Ölçeği ve geometri başarı düzeylerini incelemek için Geometri Başarı Testi kullanılmıştır. Araştırmada öğretmen adaylarının görsel matematik okuryazarlıkları ile geometri başarıları arasındaki ilişkinin hesaplanmasında Pearson Momentler Çarpım Korelâsyon Katsayısı yöntemi, görsel matematik okuryazarlığı algısı ve buna ait alt boyutların geometri başarısının yordayıcısı olup olmadığı Çoklu Regresyon Analizi yöntemiyle incelenmiştir.

Araştırmanın Bulguları: Çalışma bulguları dikkate alındığında görsel algı kaynağının geometri başarısı ile arasında düşük düzeyde ilişki tespit edilmiştir. Nitekim görsel matematik okuryazarlığının veri kaynaklarından biride görsel algıdır. Bu kaynağın geometri başarısı ile ilişkisini göz ardı etmek mümkün değildir. Literatür taraması yapıldığında görsel algıların geometri başarısı ile doğrudan veya dolaylı yönden ilişkili olduğu görülmektedir (Bekdemir ve Duran, 2012; Karunaratne, 2000). Yine bulgular neticesinde görsel matematik okuryazarlığının veri kaynaklarından biri olan örüntü oluşturma boyutunun geometri başarısı ile arasında düşük düzeyde ilişki tespit edilmiştir. Literatür taraması yapıldığında da örüntü boyutunun geometri başarısı ile doğrudan veya dolaylı yönden ilişkili olduğu görülmektedir. Tanışlı ve Köse (2011)'nin yapmış olduğu araştırmada, öğretmen adayları lineer şekil örüntüsünü yakın/uzak bir adıma devam ettirmede ve örüntünün kuralını belirlemede sadece şeklin yapısına odaklanılan görsel ve şekil örüntüsünün sayı örüntüsüne dönüştürüldüğü sayısal yaklaşımı benimsemişler, bu yaklaşımlar altında da toplam 26 strateji kullanmışlardır.

Araştırmanın Sonuçları ve Önerileri: Görsel algı, geometrik bilgi ve uzamsal zekâ kaynakları geometri başarısı üzerinde diğer kaynaklara göre daha fazla öneme sahiptir. Bu kaynaklar geometri başarısının yaklaşık %7'sini açıklamaktadır. Özellikle uzamsal zekâ kaynağının geometri başarısı üzerindeki etkisi diğer kaynaklar göz önünde bulundurulduğunda daha yüksek çıkmıştır. Öğrencilerin görsel matematik okuryazarlıklarını geliştirmek için öncelikle uzamsal zekâ kaynaklarıyla beraber görsel algı ve geometrik bilgi kaynakları geliştirilmelidir. Son olarak, somutlaştırma ve örüntü oluşturma kaynakları geometri başarısına etki eden dördüncü ve beşinci önemli boyutlardır. Bu boyutların her biri geometri başarısının yaklaşık %4'ünü açıklamaktadır. Özellikle somutlaştırma boyutunun geometri başarısı üzerindeki etkisi diğer boyutlar göz önünde bulundurulduğunda daha düşük çıkmıştır. Elde edilen bulgular doğrultusunda doğrudan ve dolaylı olarak görsel matematik okuryazarlığı algısı boyutlarının geometri başarısı ile arasındaki ilişkiyi incelemenin gerekli olduğu sonucuna ulaşılmıştır. Duran ve Bekdemir (2013), görsel matematik okuryazarlığı öz-yeterlik algısının, görsel matematik başarısının anlamlı bir yordayıcısı olduğunu söylemişlerdir. Özgen ve Bindak (2011) çalışmalarında matematik dersi başarı puanının ve matematik dersine verilen önemin, matematik

okuryazarlığı öz-yeterlik inancına yönelik anlamlı birer yordayıcı olduklarını tespit etmişlerdir. Bu çalışmaların yanında matematik başarısını yordayan faktörlerin neler olduğunu belirleyen çalışmalar da literatürde mevcuttur (Doğan ve Barış, 2010; Kayagil, 2010). Genel anlamda görsel matematik okuryazarlığı alt boyutlarının geometri başarısını yordama gücü düşük çıksa da anlamlı olması ilişkinin önemini ön plana çıkarmaktadır. Sonucun böyle çıkmasının sebebi, araştırmanın yalnız bir üniversitede yapılmış olması veya örneklem sayısının düşük olması olabilir. Ayrıca araştırmada örneklemin teste ve ölçeğe tüm bilgilerini yansıttıkları varsayılmıştır. Bu sınırlılıklar doğrultusunda ileride yapılacak olan çalışmalara, konuyla ilgili daha büyük örneklemler üzerinde araştırmaların yapılması, farklı sayısal bölümlerin görsel matematik okuryazarlık algılarıyla geometri başarıları arasındaki ilişkilerinin incelenmesi hâlihazırda bulunan matematik öğretmenleri üzerinde benzer uygulamalar yapılması ve sonuçlar doğrultusunda bu kavramlar hakkında eğitimlerin verilmesi önerilmiştir.

Eurasian Journal of Educational Research 80 (2019) 25-38

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

The Opinions of Primary School Teachers on their Creative Thinking Skills*

Ali Riza ERDEM[1], Duygu Cag ADIGUZEL[2]

ARTICLE INFO

Article History:
Received: 22 Feb. 2017
Received in revised form: 16 Oct. 2018
Accepted: 8 Mar. 2019
DOI: 10.14689/ejer.2019.80.2

Keywords
Productive, innovative, thinking, education, opinion

ABSTRACT

Purpose: This study aimed to identify the views of primary school teachers on their creative thinking skills. The following research questions were addressed based on the aim of the study: (1) What are the primary school teachers' views on their critical thinking skills? (2) Do primary school teachers' views on their creative thinking skills differ according to their gender, seniority in the profession, and educational level?

Research Methods: In sampling distribution, the proportional sampling method was used and 421 classroom teachers took part in the study. The data were obtained through "*How Creative Are You?*" scale. In the analysis process, percentage and arithmetic means were used as parametric test, while Man Whitney U and Kruskal Wallis were used as non-parametric tests.

Findings As a result of the data analysis obtained from the scales which were administered to the primary school teachers, the level of creativity of the primary school teachers were below average. Elementary school teachers' views on their creativity levels were examined based on the variables of gender, professional seniority and educational level. The analyses revealed no statistically significant differences between primary school teachers' views on their creativity levels according to their gender, professional seniority and educational level.

Conclusions: According to the results of the analysis, there was no statistically significant difference among the views on the level of creativity in terms of gender, seniority and educational level.

* This study is a part of a master's thesis titled "The relationship between classroom teachers 'creative thinking skills and teacher behaviors' contribution to the development of creative thinking skills of students" under the supervision of Prof Dr Ali Rıza ERDEM.
31 May-3 June 2016: An extended version of oral proclamation held in III. Conference of International Euroasia Education Research organized by Muğla Sıtkı Koçman University Faculty of Education and Anı Publishing.

[1] Corresponding Author, Adnan Menderes University, TURKEY, E-mail: arerdem@adu.edu.tr, ORCID: 0000-0001-9704-9529
[2] MEB, TURKEY, E-mail: duygucag@gmail.com, ORCID: 0000-0002-3360-6776

Introduction

Thinking is an active, purposive and organized cognitive process that individuals apply to understand the situation that they are in. According to De Bono (1978), thinking is about discovering an experience intentionally with purposes such as understanding, planning, decision making, problem solving, judging and acting. Individuals' learning to think means thinking in different ways. The type of thinking in which individuals feel themselves more comfortable among others is creative thinking. Vernon (1989) defines creativity as the ability to produce a new idea or view on a topic and make an invention. Creativity is not only about revealing a new product, but also synthesizing based on all known information, and then discovering different solutions or thinking the functions of objects in an extraordinary way. For Weisberg (2006), creative thinking is the process in which the creative product comes out. As for Maclure (1991), one of the most important objectives of contemporary education is to improve students' thinking skills. In a study titled "*Curricula for Problem Solving and Creative Thinking*", Isaksen and Sidney (1985) examined the attitudes, knowledge and behaviors of 152 curriculum developers, and compared traditional learning and creative learning. The results showed that 87% of the participants intentionally planned the development of creative thinking and problem-solving skills. Moreover, 65% produced their own curricula. D. J. Treffinger's model revealed that most of them preferred creative thinking techniques. In their study titled "*Creative Potential and Socio-emotional Relationships Beyond Academic Assessment in Preschool Children*", Diener, Wright, Brehl and Siyah (2016) focused on social behaviors of children, and examined the connections of creative potential in the preschool period. However, school age did not shed light on childhood. Great duties fall to particularly primary school teachers to develop creative thinking skills in children. Developing children's creativity has an important place among the learning objectives. Teachers need to fully know what creativity is and how it can be developed (Baysal, Carikci & Yasar, 2018; Cellek, 2002; Demirci, 2007; Doganay, 2017; Emir, Ates, Aydin, Bahar, Durmus, Polat & Yaman, 2004; Erktin, 2002; Karatas & Ozcan, 2010; Oncu, 2003; Ozden, 1999; Ozerbas, 2011; Ozgenel & Cetin, 2017; Temizkan, 2011; Tican, 2013; Ucan, Tascı & Ovayolu, 2008; Ulger, 2014; Yildirim & Turk, 2018).

In the Turkish Teaching program, which has been renewed, it is important to enrich vital experience, via a way that lets the analytic and creative thinking improve, as well as to know the historical accumulation and to reach the ways to re-produce it. Reaching the aims in the entrepreneurship perfection and taking initiative, which are among the basic skills in the Turkish Program, include creativity, taking risks, innovation as well as planning and carrying out projects (MNE, 2017).

For primary school teachers, actively using creativity in the classroom is a facilitative factor in teaching children. In this way, children easily build up the knowledge they need to acquire. Teachers play a guiding role in building knowledge. Teachers who have the creative thinking skills try to find different solutions when they encounter a problem instead of avoiding this problem (Aslan &

Cansever, 2009). In this way, primary school teachers also support the development of creative thinking in children while leading to form a more persistent educational environment by using different instructional methods and techniques suitable for every class and topic. With regard to raising creative individuals as targeted in the primary curriculum and developing creative thinking skills in students, this study was needed to identify primary school teachers' levels of the creative thinking skills.

This study aimed to identify views of primary school teachers on their creative thinking skills. The study is thought to provide guidance to teacher training institutions and in-service trainings to teachers. The following research questions were addressed based on the aim of the study: (1) What are the primary school teachers' views on their creative thinking skills?, (2) Do primary school teachers' views on their creative thinking skills differ according to their gender, seniority in the profession, and educational level?

Method

Research Design

This study focusing on creative thinking skills based on primary school teachers' views was a descriptive one adopting a survey model. According to Karasar (2007, p. 77), the screening model is an approach used to describe a fact of the past and the present without changing it. Indeed, such a study attempts to define a fact, person, or object of research in its own condition. The subject is not exposed to any change or effect.

Participants

The population of the study consisted of primary school teachers (Grades 1, 2, 3 and 4) working in primary schools in Merkezefendi and Pamukkale districts of Denizli province in the 2015-2016 school-year. Since it was not possible to reach the whole population, a sample was selected to represent this population. This sample was selected based on the significance level of 0.5. The lower limit in the necessary sample was calculated as 306 with ±5% sampling error at the confidence interval of 95%. "*Proportional cluster sampling*" method was used, and 421 primary school teachers were selected as the sample. Within the sample, 65.1% of the individuals were male, and 39.9% were female. As for seniority in the profession, 25.4% had 1-10 years of experience, 37.5% had 11-20 years, and 37.1% had 21 years of experience and above. Regarding educational level, 77.9% of the participants graduated from an education faculty, while 22.1% of them graduated from other teacher training institutions such as a teacher's training school or a faculty of science and humanities.

Data Collection and Application

In the data gathering process, the scale "How Creative Are You?" was used. This test was developed by Eugene Raudsepp, and translated into Turkish by Sabire Coban based on the original form. The validity and reliability statistics of the test

were also calculated by Sabire Coban, and the Cronbach's Alpha coefficient was found as 0.95 for creativity (Coban, 1999). The data obtained through the "*How Creative Are You?*" scale were coded as "*(5) Strongly Disagree*", "*(4) Disagree*", "*(3) Neutral*", "*(2) Agree*" and "*(1) Strongly Agree*". The answer options that were discontinuous were turned into "continuous" to be able to interpret the results yielded in statistical procedures. The interval of four in the scale "*How Creative Are You?*" was divided into five options (4:5=0.80), the resulting value was added to the lowest number representing the options, and the results were interpreted as not creative for 1.00 – 1.80, creativity level being below the average for 1.81 – 2.60, moderate creativity level for 2.61 – 3.40, creativity level being above the average for 3.41 – 4.20, and high creativity level for 4.21-5.00.

Validity refers to the suitability of an instrument for its purpose of employment. As for reliability, it is the extent to which an instrument measures the data accurately. The most important criterion that determines the quality of scientific works is the validity and reliability of the instruments used (Uzgoren, 2012). The validity and reliability studies of the scale "How Creative Are You?" were conducted by Sabire Coban to be used in her doctoral dissertation, and the Cronbach's Alpha coefficient was 0.95 for creativity (Coban, 1999). In addition, the scale was administered to 421 teachers, and the reliability analyses were conducted again. Its Cronbach's Alpha coefficient was found as 0.88. Accordingly, it can be argued that the scale was valid and reliable.

Data Analysis

Kolmogorov- Smirnov test was used to determine whether the data distribution was normal. According to the results of scale, data obtained through "*How Creative Are You?*" did not show normal distribution (K-s)-z =1,538 p=0.018). In the analysis process, percentage and arithmetic mean were used as parametric tests, while Mann Whitney U and Kruskal Wallis were used as non-parametric tests.

Results

This section presents the findings revealed through the analysis of the data to answer the research problems.

Findings for the First Research Question and Interpretations

The first research question addressed in the study was "*What are the primary school teachers' views on their creative thinking skills?*".

Table 1

Primary School Teachers' Views on Their Creative Thinking Skills

N	*Minimum*	*Maximum*	$\overline{x}$	*SD*
421	1,00	4,60	2,4533	0,58966

As is seen in Table 1, the results showed that the primary school teachers' creative thinking levels were below the average (x=2.45). This overlaps with the findings of other studies. In a study titled "*Creativity, professional burnout and life satisfaction in primary school teachers*", Sahin (2010) reported that 80.6% of the participants' creativity levels were identified as not creative, 16.6% were as moderately creative, 2.5% were as creative above the moderate level, and 0.2% of them were identified as creative.

Findings for the Second Research Question and Interpretations

The second research question of the study was "*Do primary school teachers' views on their creative thinking skills differ according to their (a) gender, (b) seniority in the profession, and (c) educational level?*".

Table 2

Results of the Mann-Whitney Test for Primary School Teachers' Views on Their Creative Thinking Skills according to Gender

Gender	*N*	$\overline{x}$	*U*	*p*	*Difference*
Female	274	216,17			
Male	147	201,37	1.872	0.234	Not Significant
Total	421				

p>0.05

As can be seen in Table 2, the results revealed no statistically significant difference between creative thinking skills based on the primary school teachers' views. This is consistent with the findings of other studies. In a study titled "*Examining the relationship between preschool teaching students' creativity and problem solving levels*", Zeytun (2010) did not reveal any statistically significant difference according to gender. In another study titled "*Examining the relationship between teachers' adjustment in marriage and their creativity*", Gulererli (2014) did not find a significant difference between the arithmetic means of the groups as a result of the independent samples t-test performed to determine whether there was a significant difference in the teachers' scores of the creativity scale according to the gender variable. However, there are also research findings that contradict with the findings of the current study. In a study titled "*The relationship between primary school teachers' creativity and organizational commitment*", Altin (2010) reported that the teachers' perceptions of their creativity levels were different according to their gender, and this difference was in favor of the female teachers. With regard to the mean scores of creativity levels, the female teachers perceived themselves more creative compared to the male teachers.

Table 3

Results of the Kruskal-Wallis Test for Primary School Teachers' Views on Their Creative Thinking Skills according to Seniority in the Profession

Seniority	*N*	$\overline{x}$	*K*	*P*	*Difference*
1-10 years	107	218,94			
11-20 years	158	202,18	1.414	0.493	Not Significant
21+ years	156	214,48			
Total	421				

As is seen in Table 3, when the primary school teachers' views of their creative thinking skills were examined according to the seniority variable through the Kruskal-Wallis Test, no statistically significant difference was found between their creativity levels. This is consistent with the findings of other studies. In a study titled "*Creativity, Professional Burnout and Life Satisfaction in Primary School Teachers*", Sahin (2010) revealed a significant difference in one-way ANOVA performed to determine whether there was a difference between the participants' creativity levels according to their seniority.

Table 4

Results of the Mann-Whitney Test for Primary School Teachers' Views on Their Creative Thinking Skills according to Level of Education?

Educational Level	*n*	$\overline{x}$	*U*	*P*	*Difference*
Faculty of Education	328	210,86			
Other (e.g. Teacher's Training School, Faculty of Letters)	93	211,49	1.521	0.965	Not Significant

p> 0.05

As is seen in Table 4, primary school teachers' creative thinking skills did not show a significant difference according to their educational level. However, there are research findings that contradict with this finding of the current study. In a study titled "*The relationship between primary school teachers' organizational commitment and creativity*", Altin (2010) reported that the teachers' creativity levels differed according to their undergraduate education, and this difference was in favor of those graduated from an education faculty ($t=3.08$, $p<.05$). When the mean scores regarding the creativity levels were examined, it was seen that the teachers who graduated from an education faculty perceived themselves more creative compared to those graduated from other faculties.

Discussion, Conclusion and Recommendations

The results of the analyses on the first research question " *What are the primary school teachers' views on their creative thinking skills* " showed that primary school teachers' creative thinking levels were below the average (x=2,45). In a study titled "*Creativity, professional burnout and life satisfaction in primary school teachers*", Sahin (2010) reported that 80.6% of the participants' creativity levels were identified as not creative, 16.6% were as moderately creative, 2.5% were as creative above the moderate level, and 0.2% of them were identified as creative. The result of this study is consistent with the current study, which brings up the question why teachers' creativity is low. This study is thought to be guiding in examining the reasons why teachers' creativity is low.

According to the results of the analyses regarding the second research question "Do primary school teachers' views on creative thinking skills differ according to their (a) gender, (b) seniority in the profession, and (c) educational level?";

There was no statistically significant difference between creative thinking skills based on the primary school teachers' views. This is consistent with the findings of other studies. In a study titled "*Examining the relationship between preschool teaching student's' creativity and problem solving levels*", Zeytun (2010) did not reveal any statistically significant difference according to gender. In another study titled "*Examining the relationship between teachers' adjustment in marriage and their creativity*", Gulererli (2014) did not find a significant difference between the arithmetic means of the groups as a result of the independent samples t-test performed to determine whether there was a significant difference in teachers' scores in the creativity scale according to the gender variable. However, there are also research findings that contradict with the findings of the current study. In a study titled "*The relationship between primary school teachers' creativity and organizational commitment*", Altin (2010) reported that the teachers' perceptions of their creativity levels were different according to gender, and this difference was in favor of the female teachers. With regard to the mean scores of creativity levels, female teachers perceived themselves more creative compared to male teachers.

There was no significant difference between the primary school teachers' creativity levels according to the seniority variable. This is consistent with the findings of other studies. In a study titled "*Creativity, professional burnout and life satisfaction in primary school teachers*", Sahin (2010) revealed a significant difference in the one-way ANOVA performed to determine whether there was a difference between the participants' creativity levels according to their seniority in the profession.

There was no statistically significant difference between primary school teachers' creative thinking skills based on their level of education. However, there are also research findings that contradict with the findings of the current study. In a study titled "*The relationship between primary school teachers' organizational commitment and creativity*", Altin (2010) reported that teachers' creativity levels differed according to their undergraduate education, and this difference was in favor of those graduated

from an education faculty (t=3.08, p<.05). When the mean scores regarding the creativity levels were examined, it was seen that the teachers who graduated from an education faculty perceived themselves more creative compared to those graduated from other faculties.

As a result of the analyses based on the data gathered through the scale, primary school teachers' creativity level was found to be below the average. Elementary school teachers' creativity levels being below the average can be interpreted as that they do not use their creativity fully or could not use it. Primary school teachers can be enabled to improve their creativity through pre-service and in-service trainings. This is particularly important for them to teach "*creative thinking skills*" to primary school students. Elementary school teachers' views on their creativity levels were examined according to the variables of gender, professional seniority and educational level. The analyses revealed no statistically significant differences between primary school teachers' views on their creativity levels according to their gender, professional seniority and educational level.

The following suggestions can be offered based on the research findings: (1) According to the results of the analyses regarding the research question " *What are the primary school teachers' views on their creative thinking skills* ", primary school teachers' creativity levels were below the average. For this reason, primary school teachers should be exposed to seminars to develop their creativity levels. (2) Considering that primary school teachers' creativity levels are low, studies can be conducted to reveal why their creativity levels are low, and how this can be improved.

References

Altın, B. (2010). *Ilkogretimde gorevli ögretmenlerin orgutsel baglılıklarıyla yaratıcılıkları ilişkisi [Relation of creativity with organizational commitment of primary school teachers]* (Yayımlanmamıs yuksek lisans tezi). Maltepe Universitesi, Sosyal Bilimler Enstitusu, Istanbul.

Aslan, N., & Cansever, B.A. (2009). Egitimde yaratıcılıgın kullanımına iliskin ogretmen tutumları [The primary school teachers' attitudes for creativity in education]. *Tubav Bilim Dergisi*, *2*(3), 333-340.

Baysal, Z. N., Carıkcı, S., & Yaşar, B. (2018). Ogretim elemanı/uyelerinin düsünme becerileri ogretimine iliskin görüslerinin incelenmesi [Analysis of academics' views on teaching thinking skills]. *Inönü Universitesi Egitim Fakultesi Dergisi, 19*(2), 174-188.

Cellek, T. (2002). Yaratıcılık ve egitim sistemimizdeki boyutu [Creativity and the size of our education system]. *Universite ve Toplum Bilim, Egitim ve Dusunce Dergisi, 2*(1), 02-04.

Coban, S. (1999). *Yoneticilerin yaratıcılık duzeyleri ile liderlik tarzları arasındaki iliski [Relationship between managerial creativity levels and leadership styles]*

(Yayımlanmamıs doktora tezi). Istanbul Universitesi, Sosyal Bilimler Enstitusu, İstanbul.

Demirci, C. (2007). Fen bilgisi ogretiminde yaratıcılıgın erisi ve tutuma etkisi [The effects of the creative approach ıi the science teaching on achievement and attitude]. *Hacettepe Universitesi Egitim Fakultesi Dergisi, 32*, 65-75.

Diener, M.L., Wright, C., Brehl, B., & Siyah, T. (2016). Okul oncesi cocuklarında akademik degerlendirme otesinde yaratıcı potansiyeli ve sosyo-duygusal iliskileri [Academic beyond assessment creative potential and socio-emotional relationships in preschool children]. Retrieved from http://dx.doi.org/10.1080/10400419.2016.1229975

Doganay, A. (2017). Ust duzey dusunme becerilerinin ogretimi [Teaching of higher level thinking skills]. In A. Doğanay (Ed.), *Ogretim ilke ve yontemleri* (pp. 328-384). Ankara: Pegem Akademi Yayıncılık.

Emir, S., Ates, S., Aydın, F., Bahar, M., Durmus, S., Polat, M., & Yaman, H. (2004). Ogretmen adaylarının yaratıcılık düzeyleri [The creativity level of teachers cantidates]. *Abant Izzet Baysal Universitesi, 2*(9), 105-116.

Erktekin, E. (2002). Ilkogretimde düsünme becerilerinin gelistirilmesi [Developing thinking skills in primary education]. *M.Ü. Ataturk Egitim Fakultesi Egitim Bilimleri Dergisi, 16*, 61-70.

Gulererli, N. (2014). *Ogretmenlerin evlilik uyumları ile yaratıcılıkları arasındaki ilişkinin incelenmesi [Investigation of the relationship between teachers' marriage harmony and creativity]* (Yayımlanmamıs yuksek lisans tezi). Marmara Universitesi, İstanbul.

Isaksen, S. G., & Parnes, S. J. (1985). Curriculum planning for creative thinking and problem solving. *The Journal of Creative Behavior*, *19*(1), 1-2.

Karatas, S., & Ozcan, S. (2010). Yaratıcı düsünme etkinliklerinin ogrencilerin yaratıcı düsünmelerine ve proje gelistirmelerine etkisi [The effects of creative thinking activities on learners' creative thinking and project development skills]. *Ahi Evran Universitesi Egitim Fakultesi Dergisi, 11(*1), 225-243.

Karasar, N. (2007). *Bilimsel arastırma yöntemi[Scientific research method].* Ankara: Nobel Yayınevi.

MEB. (2017). Ilköğretim programı [Primary education program]. Retrieved from http://mufredat.meb.gov.tr/ProgramDetay.aspx?PID=222.

Oncu, T. (2003). Torrance yaratıcı dusunme testleri-şekil testi aracılıgıyla 12-14 yasları arasındaki cocukların yaratıcılık düzeylerinin yas ve cinsiyete gore karsılastırılması [Comparison of creativity levels of children ages 12-14 by age and sex by means of Torrance creative thinking test-shape test]. *Ankara Universitesi Dil ve Tarih Cografya Fakultesi Dergisi, 43*(1), 221-237.

Ozden, Y. (1999). *Ogrenme ve ogretme [Learning and teaching].* Ankara: Pegem Yayıncılık.

Ozerbaş, M. A. (2011). Yaratıcı dusunme ogrenme ortamının akademik basarı ve bilgilerin kalıcılığa etkisi [The effect of creative thinking teaching environment on academic achievement and retention of knowledge]. *GÜ, Gazi Egitim Fakultesi Dergisi, 31*(3), 675-705.

Ozgenel, M., & Cetin, M. (2017). Marmara yaratıcı dusunme egilimleri olceginin gelistirilmesi: Gecerlik ve guvenirlik calısması [Development of the Marmara creative thinking dispositions scale: Validity and reliability analysis]. *Marmara Universitesi Atatürk Egitim Fakultesi Egitim Bilimleri Dergisi*, *46*, 113-132.

Sahin, E. (2010). *Ilkogretim ogretmenlerinde yaratıcılık, mesleki tukenmislik ve yasam doyumu [Creativity, occupational burnout and life satisfaction in primary school teachers]* (Yayımlanmamıs yuksek lisans tezi). Sakarya Universitesi, Sakarya.

Temizkan, M. (2011). Turkçe ogretiminde yaratıcı düsünmeyi gelistirme bakımından Nasreddin Hoca fikralari [Nasreddin Hodja stories ın terms of ımproving creative thinking ın turkish language teaching]. *Mustafa Kemal Universitesi Sosyal Bilimler Enstitusu Dergisi, 8*(16), 195-223.

Tican, C. (2013). *Yansıtıcı düsünmeye dayalı ogretim etkinliklerinin ogretmen adaylarının yansıtıcı dusunme becerilerine, elestirel dusunme becerilerine, demokratik tutumlarına ve akademik basarılarına etkisi [The effects of reflective thinking-based teaching activities on pre-service teachers' reflective thinking skills, critical thinking skills, democratic attitudes and academic achievement]* (Yayımlanmamış doktora tezi). Gazi Universitesi, Egitim Bilimleri Enstitusu, Ankara.

Ucan, O., Tasci, S., & Ovayolu, N. (2008). Eleştirel düşünme ve hemşirelik [Critical thinking and nursing]. *Fırat Sağlık Hizmetleri Dergisi, 3*(7), 17-27.

Uzgoren, N. (2012). *Bilimsel arastırmalarda kullanılan temel istatistiksel yöntemler ve SPSS uygulamaları [Basic statistical methods used in scientific researches and SPSS applications].* Bursa: Ekin Basın Yayın Dagıtım.

Ulger, K. (2014). Ogrencilerin yaratıcı dusunme gelisimlerinin incelenmesi [The ınvestigation of the students' creative thinking development]. *Egitim ve Bilim*, *39*(175), 275-284.

Weisberg, R. W. (2006). Expertise and reason in creative thinking: evidence from case studies and the laboratory. In J. C. Kaufman & J. Baer (Ed), *Creativity and reason in cognitive development* (pp. 7-42). New York, NY, US: Cambridge University Press.

Yildirim, B., & Turk, C. (2018). Sınıf ogretmeni adaylarının STEM egitimine yönelik gorusleri: Uygulamalı bir çalısma [Pre-service primary school teachers' views about STEM education: An applied study]. *Trakya Universitesi Egitim Fakultesi Dergisi*, *8*(2), 195-213.

Zeytun, S. (2010). *Okul oncesi ogretmenligi ogrencilerinin yaratıcılık ve problem çözme duzeyleri arasındaki iliskinin incelenmesi [A study of the relationship between creativity and problem solving levels of preschool teacher's student]* (Yayımlanmamış yuksek lisans tezi). Dokuz Eylul Universitesi, İzmir.

Yaratıcı Düşünme Becerileri Hakkında Sınıf Öğretmenlerinin Görüşleri

Atıf:

Erdem, A. R., & Adiguzel, D. C. (2019). The opinions of primary school teachers on their creative thinking skills. *Eurasian Journal of Educational Research, 80*, 25-38, DOI: 10.14689/ejer.2019.80.2

Özet

Problem Durumu: Yaratıcı düşünme süreci dinamik, üretici ve özgürdür. Problemlere her açıdan bakabilmek ve farklı çözüm yolları bulabilmek gerekir. Çevresine yaratıcı gözlerle bakabilmeli, tüm kaynakların farkına varabilmeli ve gerektiğinde yararlanabilmelidir. "*Yaratıcılık*" bugüne kadar birçok yazar tarafından farklı yönleri vurgulanarak ele alınmış, çok yönlü bir kavram olarak algılanmış ve tanımlanmıştır. Örneğin bazı yazarlar düşünme süreçleri üzerinde dururken, bazı yazarlar ortaya çıkan ürüne odaklanmışlardır. Kimi yazarlar ise yaratıcı kişiliğin sahip olduğu özelliklere vurgu yapmışlardır. Farklı görüş ve bakış açılarına karşın fikir birliğine varılan nokta; yeni bir ürünün ortaya çıkış sürecidir. Yaratıcılık; sadece yeni bir ürün ortaya koymak değil, bilinen tüm bilgilerden sentez yapabilme ve sonrasında farklı çözüm yollarını keşfedebilme ya da nesnelerin işlevlerini alışılmışın dışında düşünebilmektir. Yaratıcı düşünme, problemlere eleştirel açıdan bakabilmek, daha önce aralarında ilişki kurulmamış nesneler yâda düşünceler arasında ilişki kurabilmek ve yeni önermeler de bulunmaktır. Alışılmışın dışında, özgün, farklı çözüm yollarından giderek yeni sonuçlar geliştirebilmektir. Maclure'e (1991) göre çağdaş eğitimin en önemli hedeflerinden biri öğrencilerin düşünme becerilerini geliştirmektir. Öğrencilerde yaratıcı düşünme becerisini geliştirebilmek için özellikle sınıf öğretmenlerine çok büyük görevler düşmektedir.

Sınıf öğretmeni için sınıf ortamında yaratıcılığı aktif olarak kullanmak, çocuğa öğretilmek isteneni kolaylaştırıcı bir etkendir. Öğrenciler bu yolla öğrenmesi gereken bilgiyi kolaylıkla inşa edebilmektedirler. Öğretmen bilginin inşa edilme aşamasında yönlendirici rol üstlenir. Yaratıcı düşünme becerisine sahip bir öğretmen bir sorun ile karşı karşıya kaldığı zaman bu sorundan kaçmak yerine o soruna farklı çözümler bulmaya çalışır (Aslan, Cansever, 2009). Sınıf öğretmeni her derste her konuya uygun farklı öğretim yöntem tekniklerini kullanarak daha kalıcı bir eğitim ortamının oluşmasına öncülük etmesinin yanı sıra öğrencilerde de yaratıcı düşünmenin gelişimine destek sağlamış olur. İlköğretim programının hedeflediği yaratıcı bireylerin yetişebilmesi için, öğrencilerde yaratıcı düşünme becerisi geliştirilebilmesi

için sınıf öğretmenlerinin yaratıcı düşünme beceri düzeyleri belirlemek amacıyla bu araştırmanın yapılmasına gerek duyulmuştur.

Araştırmanın Amacı: Bu araştırmanın amacı ilkokullarda görev yapan sınıf öğretmenlerinin yaratıcı düşünme becerileri hakkındaki görüşlerini belirlemektir. Araştırmanın amacı doğrultusunda alt problemler şu şekilde belirlenmiştir: (1) Sınıf öğretmenlerinin görüşlerine göre yaratıcı düşünme becerileri nedir? (2)Sınıf öğretmenlerinin görüşlerine göre yaratıcı düşünme becerileri cinsiyete, meslekteki kıdeme, eğitim durumuna göre farklılık göstermekte midir?

Araştırmanın Yöntemi: Sınıf öğretmenlerinin görüşlerine göre yaratıcı düşünme becerilerini belirlemeye yönelik yapılan bu araştırma tarama modelinde betimsel bir çalışmadır. Araştırmanın evrenini, 2015-2016 eğitim-öğretim yılında Denizli ili, Merkezefendi ve Pamukkale ilçelerinde görev yapmakta olan ilkokullardaki sınıf öğretmenleri (1, 2, 3, 4. Sınıf öğretmenleri) oluşturmaktadır. Örneklem almada "*oranlı küme örnekleme*" yöntemi kullanılarak 421 sınıf öğretmeni örnekleme alınmıştır. Veriler "*Ne kadar yaratıcısınız?*" ölçeği ile toplanmıştır. Verilerin çözümlenmesinde SPSS paket programı kullanılmıştır. "*Ne kadar yaratıcısınız?*" ölçeği ile elde edilen veriler kesinlikle katılmıyorum seçeneği 5, katılıyorum seçeneği 4, kararsızım seçeneği 3, katılıyorum seçeneği 2, kesinlikle katılıyorum seçeneği 1 olarak kodlanmıştır. "*Ne kadar yaratıcısınız?*" ölçeğindeki 4 aralık 5 seçeneğe bölünmüş (4: 5 = 0.80); bulunan sayı seçenekleri temsil eden en alt sayıdan itibaren ilave edilerek: 1.00 – 1.80 yaratıcı değil, 1.81 – 2.60 yaratıcılık düzeyi ortalamanın altında, 2.61 – 3.40 yaratıcılık düzeyi orta, 3.41 – 4.20 yaratıcılık düzeyi ortalamanın üstünde, 4.21-5.00 yaratıcılık düzeyi yüksek şeklinde yorumlanmıştır. Cronbach alpha iç tutarlık katsayısı "*Ne kadar yaratıcısınız?*" ölçeği için 0.88 bulunmuştur. Dağılımın normal olup olmadığını belirlemek için Kolmogorov Simirnov testi uygulanmıştır. "*Ne kadar yaratıcısınız?*" ölçeği ile elde edilen sonuçlara (K-s)-z =1,538 p=0.018) göre verilerin normal dağılım göstermediği belirlenmiştir. Çözümlemelerde de parametrik testlerden yüzde, aritmetik ortalama; non-parametrik testlerden ise Man Whitney U ve Kruskal Wallis kullanılmıştır.

Araştırma Bulguları: Sınıf öğretmenlerine uygulanan ölçekten elde edilen verilerin analizi sonucunda sınıf öğretmenlerinin yaratıcılık düzeyi ortalamanın altında çıkmıştır. Sınıf öğretmenlerinin görüşlerine göre yaratıcılık düzeyi cinsiyet, mesleki kıdem ve mezuniyet durumu değişkenleri açısından da incelenmiştir. Yapılan analiz sonuçlarına göre sınıf öğretmenlerinin yaratıcılık düzeylerine ilişkin görüşleri arasında cinsiyet, mesleki kıdem ve mezuniyet durumu değişkenleri açısından anlamlı bir farklılık yoktur.

Araştırmanın Sonuçları ve Önerileri: Sınıf öğretmenlerine uygulanan ölçekten elde edilen verilerin analizi sonucunda sınıf öğretmenlerinin yaratıcılık düzeyi ortalamanın altında çıkmıştır. Sınıf öğretmenlerinin görüşlerine göre sınıf öğretmenlerinin yaratıcılık düzeylerinin ortalamanın altında çıkması, sınıf öğretmenlerinin yaratıcıklarını tam anlamıyla kullanmadıkları veya kullanamadıkları şeklinde yorumlanabilir. Sınıf öğretmenlerinin yaratıcılık düzeylerine ilişkin görüşleri arasında cinsiyet, mesleki kıdem ve mezuniyet durumu değişkenleri açısından anlamlı bir farklılık yoktur. Araştırma bulgularına

dayanılarak şunlar önerilebilir: (1) "Sınıf öğretmenlerinin görüşlerine göre yaratıcı düşünme becerileri nedir?" analiz sonuçlarına göre Sınıf öğretmenlerinin yaratıcılık düzeyleri ortalamanın altında çıkmıştır. Bu nedenle; Sınıf öğretmenleri yaratıcılık düzeylerinin geliştirilmesi konusunda seminerlere tabii tutulmalıdır. (2) Sınıf Öğretmenlerinin yaratıcılık düzeylerinin düşük olduğu sonucu göz önünde bulundurularak Sınıf öğretmenlerinin yaratıcılık düzeylerinin neden düşük olduğu ve nasıl yükseltilebileceği ile ilgili çalışmalar yapılabilir.

Anahtar Sözcükler: Üretken, yenilikçi, düşünme biçimi, eğitim, görüş.

Eurasian Journal of Educational Research 80 (2019) 39-58

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

Impacts of Urban Institutions of Higher Education on the Community: A Social Approach

Bui Phu HUNG*

ARTICLE INFO

Article History:
Received: 21 Jan. 2017
Received in revised form: 23 Dec. 2017
Accepted: 17 Aug. 2018
DOI: 10.14689/ejer.2019.80.3

Keywords
Social sustainability, higher education, impacts, community, educational institutions

ABSTRACT

Purpose: This study aimed to investigate the effects of urban institutions of higher education on social sustainability of the community. Three urban colleges in Vietnam were involved in this study. **Methods:** Questionnaires and interviews were administered to 120 local residents and 1470 students and college employees (office and teaching staff) living in the nearby communities. In the interviews, the researcher made efforts to explore in-depth information by providing prompts in order to confirm, interpret and supplement the quantitative data collected from the questionnaires. **Findings:** All participants responded that the colleges contributed greatly to the social equity through rentals, food consumption, and other services available. However, though some local residents took a positive view of social diversity, others were conservative. The social cohesions took place in the community where its members were willing to welcome new cultures. Most local residents were not confident about community security when there were many new members. Noise and traffic also caused much concern to local residents. **Implications for Research and Practice:** It is suggested that local residents should distribute equital gains to physically and financially challenged people in the community. Other implications were given to researchers, policy makers, community members, and student tenants living around college. Further studies can take a holistic approach to sustainable development as a result of the effects of educational institutions. Researchers may also take into account more educational institutions of higher education in other contexts.

* Faculty of Foreign Languages - Van Hien University, VIETNAM, buiphuhung@yahoo.com
ORCID: https://orcid.org/0000-0003-3468-4837

Introduction

The Rationale

Sustainable development has been a concern for discussions in recent decades as it deals with many areas in the world. It has motivated research on how to develop a place without causing harms to any others. However, studies on sustainable developement mainly focus on the associations between this concern and environmental and/or economic issues (e.g. Johansson, Segerstedt & Jakobsson, 2016; Reed & Wilkinson, 2005; Yu, Tong, Tang, Yuan & Chen, 2018). Researchers have found that society has also had some interactions with other fields as environmental and/or economic changes have some influences on humans and vice versa (Cook & Esuna, 2014). Recent studies have given models for sustainable development with three main domains: economy, environment and society (Rogers, Jalal & Boyd, 2007). Social sustainability should be a concern to researchers.

As built-in constructions, instituions of higher education may have certain impacts on the community. Dempsey, Bramley, Power and Brown (2011) and Dave (2011) believe that the construction and existence of an organization may also affect the neighborhood to a certain extent as built environments play a role in the social sustainability. In Vietnam, many colleges and universities have off-campus accommodations for students. Several private colleges do not have any dormitories (Sheridan, 2010). Off-campus students may have positive and negative effects on the community (Omar, Abdullah, Yusof, Hamdan, Nasrudin & Abullah, 2011). Also, in the past 20 years, education in Vietnam has been developing and accepting transformation (Hayden & Thiep, 2007); re-location and construction of new campuses have taken place (Nha & Tu, 2015). Higher education institutions in urban areas in Vietnam have potential impacts on the community.

Previous studies show a lack of interest in exploring impacts of institutions of higher education on communities, particularly from the perspective of social sustainability. Previous research (e.g. Nieuwenhuis, Hooimeijer, Dorsselaer & Vollebergh, 2013; Nieuwenhuis & Hooimeijer, 2015) mainly focused on the effects of the neighbourhoods on students' academic achievements at nearby schools. This paper investigated the impacts of three colleges located in the south of Vietnam on social sustainability of the neighborhoods by administering questionnaires to local residents and their staff and students living around the colleges. It attempted to answer the following questions:

1) What impacts of institutions of higher education on the community are perceived by the local residents?
2) What impacts of insitutions of higher education on the community are perceived by college students and staff?

Theoretical framework

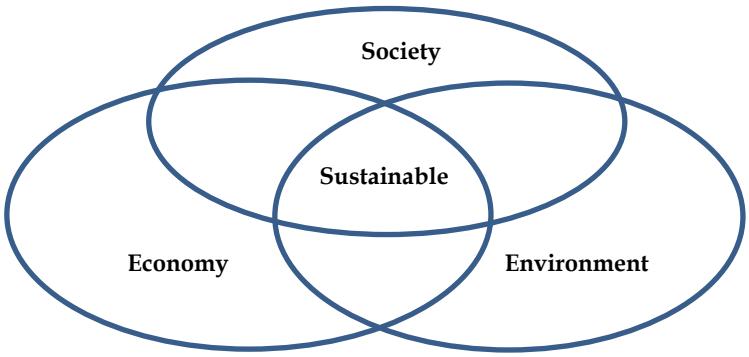


Figure 1. *Three Components in Sustainable Development*

(Adapted from Adams, 2006)

There are three main components in sustainability (Figure 1) in that society, economy and environment have equally significant roles in sustainable development (Adams, 2006). This model was applied in a study by Vallance, Perkins and Dixon (2011). However, Cato (2009) argues that economy plays a central role in sustainable development, an environment can restrict the social development and economic development (Figure 2). No matter what model is suggested, it is clear that sustainable development is influenced by the social, economic and environmental factors. According to a contemporary and currently applied approach to social sustainability, all these three domains are interrelated in that the other two domains are embedded within the environment. Therefore, social sustainability is comprised of all human activities that have impacts on sustainable development.

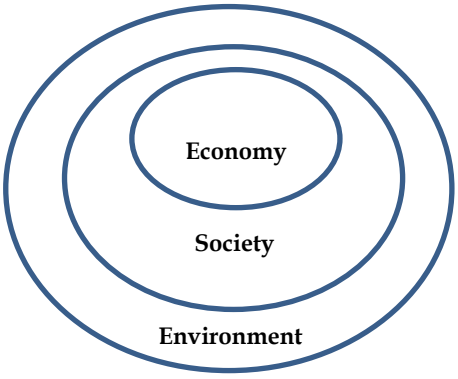


Figure 2. *Three Pillars of Sustainable Development (Adapted from Cato, 2009)*

A framework for social sustainability has been established by researchers. Woodscraft, Bacon, Caistor-Arendar and Hackett (2011) introduces the framework for

social sustainability. Accordingly, social sustainability has four dimensions: amenities and infrastructure, social and cultural life, voice and influence and space to grow. Vallance et al. (2011) describe the three main components of social sustainability (Figure 3) as *development, maintenance* and *bridge.* The development aspect may be tangible or intangible. It is "about meeting basic needs, inter- and intra-generational equity". Maintenance is understood as what maintained in a society. Bridge refers to the change of behavior to achieve objectives which can be ecological or physical regarding the environment. Sen (2000) outlines a more detailed framework for social sustainability with six indicators: equity, diversity, social cohesions, quality of life, democracy and governance, and maturity. Equity sustainability takes place when the monetary or financial values of the community is secured. The community should also create opportunities for diversity. Social cohesions or social interactions take place when formal and/or informal interactions between members inside and outside the community take place. In other words, all community members should think that they belong to a unit. Quality of life is also crucial in that all he community members' basic needs, such as well-being, housing, education, employment and security, are satisfied. The social sustainability is also measured by the democratic and governing policies. Maturity is defined as the community members' development through their communication styles, behavior patterns and informal education. All these dimensions of the paradigm are interrelated and equally contribute to social sustainability. Polese and Stren (2000, pp. 16-17) make a brief definition of social sustainability as "development (and/or growth) that is compatible with the harmonious evolution of civil society, fostering an environment conducive to the compatible cohabitation of culturally and socially diverse groups while at the same time encourging social integration, with improvements in the quality of life for all segments of the population."

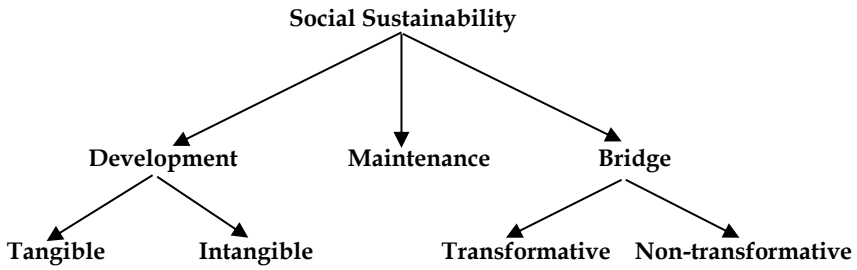


Figure 3. *Dimensions of Social Sustainability*

(Adapted from Sen, 2000)

The construction and existence or re-location of any large organization, especially schools and shopping malls, usually have some impacts on the place where they are located as these institutions are examples of long-term investments and have interactions with the neighborhood (Tanriogen, 2018; Wey, 2018). The neighborhood may have some influences on the success of educational institutions (Nieuwenhuis et

al., 2013), and educational institutions also have effects on the neighborhood. There is an interaction between these two communities, but the effect of the neighboorhood on school is not really direct (Nieuwenhuis & Hooimeijer, 2015). However, the influence of educational institutions, especially in higher education, on the neighborhood can be clear. This influence is especially straighforward in case of off-campus students in Asian contexts. The interactions between off-campus college students in Asia and the community where the college is located are usually strong. Therefore, the influence can be positive or negative (Omar et al., 2011). Accordingly, many college students are from other places, and when they live off campus, they choose to reside in the vicinity of the school to avoid or reduce transit. Some students who do not dwell relatively close to school may also take place of the catering service in the neighborhood. Local restaurants also cater the academic and office staff in the school. Their expenditures on rental, food, entertainment and necessities in turn contribute to the business development of the nearby community to a certain extent. They may cause traffic congestion. Local people can also get employed in the neighboring institutions. The interactions between a school (teachers, office staff and students) and local people may also lead to behavior exchange. Lawhan (2009) believes that existence of an educational instituion may make the members of the neighborhood feel a sense of community. In summary, there should be research on the impacts of urban insitutions of higher education on the community from the perspective of social sustainability.

Research Methods

Research Approach and Design

The qualitative and quantitative approaches have their own strengths and weaknesses, and purposes. Regarding the circumstance that this topic was not an interest of research at the time of the implementation of this study, a mixed-methods approach was adopted (Hung & Van, 2018; McMillan & Schumacher, 2001, pp. 372-373). The use of this approach was considered to triangulate the data for reliability (McKim, 2017, p. 203; Tashakkori & Creswell, 2007). To avoid the contradiction between these different sources of data, the interviews to collect qualitative data were conducted within two days after the administration of the questionnaires which were used to collect quantitative data. The time interval was used to analyze the participants' responses to the questionnaires. The data from the interviews conducted within 48 hours after the occurence could be 95% reliable (Gass & Mackey, 2000, pp. 14-17). The participants' responses from the questionnaires were used to confirm, interpret, and supplement their responses in the interviews. This model was applied by Hung (2017) and Hung, Truong and Nguyen (2018). The data triangulation was believed to increase the resarch reliability and validity.

However, considering research ethics, all the participants were neither forced to answer the questions nor to criticize their contradictions in the collected data. The administration of questionnaires and interviews had no room for any type of discrimination. All the participants were treated with dignity and respect. Their identities and answers were kept confidential. The names of the insitutions were also

kept anonymous. In this study, the researcher played the roles as question writer, questionnaire administrator, interviewer, and data analyst.

Research Sample

1470 college employees and junior and senior students living as tenants close to three colleges located in Ho Chi Minh City, Vietnam and 120 local residents living close these colleges were randomly selected to answer the questionnaires. All the participants had dwelled in the neaby communities five years by the implementation of this study. After the questionnaires were administered, 65 students and 45 people in the vicinity were interviewed. The selection of the people for interviews was based on their answers to the questionnaires. They were informed of the importance of their answers in the study. As the communities around these three colleges were considered the target participants to be impacted by the colleges, their responses were collected prior to those from the students and college staff.

Instruments and Procedure

The questionnaires were designed for participants' ratings in a Likert-scale of 1-5. It applied the framework by Sen (2000). However, regarding the practical purpose of this study on the effects of the schools in the nearby households as well as the student participants involved in this study, This framework was adapted for the participants to reflect what they had experienced in places. It had four main clusters: equity, diversity, social interactions and quality of life. The interview questions were open-ended to collect qualitative data and also applied this framework. Table 1 gives detailed descriptions of this framework.

Table 1

Question Framework

Category	Description
Equity	The influence of the college, including staff and students, to the incomes of the local community
Diversity	The extent to which the college contributes to the social diversity of the local community
Social cohesions	The effects of the interactions between the local residents and the college staff and students on the local residents' life
Quality of life	The improvement or decline in the quality of life, such as noise and traffic congestion, as a result of the existence of the college

Validity and Reliability

The questionnaires were first pilot-tested with 20 students and 20 local residents. These participants were not involved in the main study, and the data collected was used for revising the questionnaires. After the pilot study, some items in the clusters were removed as their Cronbach Alpha values were relatively low, $\alpha < .7$. Some other items were linguistically simplified to facilitate the participants' answers. Data triangulation was also believed to increase research validity and reliability. Finally, the comparison of local residents', students' and staff's responses was considered to contribute to the reliability and validity of the findings.

Data Analysis

The quantitative data from the participants' responses to the questionnaires were input into SPSS 22 for statistial analysis. Mean scores (M), standard deviations (SD), significance values (p), and Cronbach Alpha values (α) were first achieved. Correlation between the clusters and descriptive statistics were also used and discussed in the findings. The qualitative data from the interviews, which were audio-recorded, were coded and then theme-analyzed and transcribed. The participants' responses in the interviews were classified into four main themes: equity, diversity, social interactions, and quality of life. The data from observations before and after the constructions of the colleges were coded and analyzed into the aforementioned themes.

Results

The Impacts of the Colleges on the Local Community

Table 2

Local Residents' Opinions of the Effects of the Three Colleges

n=120	**College A** (n=35)				**College B** (n=45)				**College C** (n=40)			
	M	SD	*p*	*α*	M	SD	*p*	*α*	M	SD	*p*	*α*
Equity	4.18	.52	.000	.82	4.43	.61	.000	.83	4.30	.54	.000	.80
Diversity	3.67	.61	.003	.83	3.93	.77	.003	.80	3.86	.72	.003	.82
Social cohesions	4.09	1.32	.000	.82	4.16	1.21	.000	.82	3.66	1.62	.000	.84
Quality of life	3.92	1.21	.000	.81	4.02	1.36	.000	.82	3.68	1.26	.000	.81

The local residents gave answers to the questionnaires about the effects of the three colleges surveyed. Overall, they believed the three colleges had moderate and strong effects on their communities. The strongest impact was thought to generally be on the quality of life, and the weakest impact was on the diversity of the community at all three colleges. More specifically, College A had less impact on all the categories asked

about than the others. College B had strongest impact on the equity of the neighborhood than the others, but it was assumed to have as much effect on the diversity of the community as College C College A had less effect on this concern. Regarding social interactions and quality of life, College C had slightly stronger effect. Table 2 also shows that the impacts of the colleges were significant, $p < .005$, and Cronbach Alpha coefficients were assured, $\alpha > .8$. the data achieved from the interviews generally confirmed the local residents' responses to the questionnaires. The local residents' responses provided additional information about the effects as a contribution to social sustainability. Details can be found in each sub-category below.

Equity

The financial contributions which the colleges made to the equity of the communities around these institutions were significant. In general, they made the greatest contribution to the rents. The mean scores about the contributions to rents which College A, College B and College C made were 4.45, 4.72 and 4.47 respectively. In the interviews, most local residents responded that many of the students took advantage of the homestay service offered in the neighborhoods, but relatively few employees did not live in the homestay in the neighborhoods. The local residents around College A revealed that almost no employees from this college used this service. However, the office staff from College B were considered to contribute greatly in terms of rental. Homestay owners added that the rents from the office staff were higher than those from students. Students usually shared the room with their friends; however, the office staff in the homestay lived with their spouse and children. Most of them rent comfortable rooms and paid higher than students. This may have explained the reason why the contribution of rental to the neighborhood around College B was higher.

Table 3

Local Residents' Responses about the Effects of the Colleges on the Community Equity

n=120	**College A** (n=35)				**College B** (n=45)				**College C** (n=40)			
	M	SD	*p*	α	M	SD	*P*	α	M	SD	*p*	α
Rental	4.45	.42	.000	.80	4.72	.41	.000	.81	4.47	.47	.000	.82
Food	3.71	.56	.002	.81	3.61	.72	.003	.80	3.78	.63	.002	.80
Entertainment	4.23	.72	.000	.80	4.74	.81	.000	.83	4.45	.73	.000	.81
Other services	4.31	.65	.000	.82	4.63	.94	.000	.82	4.51	.74	.000	.83

Regarding the college staff's and students' use of the catering service, the mean scores for this influence was less than those for rental, entertainment, and other services. The local residents around College B said that many people (college employees and students) did not order food from the nearby restaurants or food

suppliers. From the local residents' responses to the questionnaires, the community around College B got marginally less benefited than College A and College C in terms of their expense on food. Their responses in the interviews provided the reasons. Most married tenants around College B preferred to cook for themselves as married couples wanted to have a sense of family during the dining time, but most students living in the homestay around College A and College C often ordered food from the neighboring communities.

The academic staff and students living as tenants around the three colleges also made great contributions to the community equity regarding their use of entertainment. Their contributions were significant. The staff and students from College B were thought to spend more on entertainment than those from College A and College C. The data from the interviews explained that the places where they often went to for recreation were typically karaoke shops, coffee shops, and cinemas. The colleges were also thought to use other services available in the communities. Their expenditures were often on basic needs, such as washing powder, shampoo, and household items. Although these items were not really costly, they were the category which people from the colleges spent much on.

Although the colleges made some contributions to the equital development of the communities, the development was not really sustainable. Table 3 illustrates that the college staff and students made significant contributions in terms of rental, food, entertainment and other services. However, there was not evidence of social sustainability in these communities. In the interviews, many local residents responded that the financially and physically challenged people and the elderly in the neighborhoods did not receive direct benefits from this equital development. These disadvantaged people got some benefits from the governments, but the benefits were insufficient for them to make the ends meet because they rested against the policy set by the central government. The communities did not have any funds for helping these people. Honestly, they mainly depended on the the financial support from their families. They occasionally received financial aids from charity offered by non-government associations located outside the communities.

The local residents' responses to the questionnaires about the effects of the colleges on the diversity of the communites (Table 4) show that there had been much change to the recipes and foods in the places. This category had been affected the most. The local residents also provided some reasons for this. Some foreign teachers from College B and College C lived in the community close to the institutions, and they did not cook, but ate out or ordered foods from the nearby restaurants. In addition, most students living as tenants here were immigrants from other places in Vietnam and Laos. The catering services in the neighborhood had to customize their recipes to meet the customers' needs. The dishes were added to their menus, and they also updated the recipes to get more customers. However, College A did not have any foreign teachers and students living around the campus, so the foods in this place was less diverse. The food diversity in this community was mostly affected by native students living around campus. In the interviews, most local residents revealed that they resisted to change. Although all the participants accepted that food diversity was inevitable, 45

participants (37.5%) responded that they only cooked traditional foods at home and did not encourge food diversity. Whenever they ate out, the middle-aged and the elderly only ordered traditional foods. Interestingly, young people seemed to be more adaptable; they welcomed a wide variety of dishes.

Diversity

Table 4

Local Residents' Opinions of the Effects of the Colleges on Community Diversity

n=120	**College A** (n=35)				**College B** (n=45)				**College C** (n=40)			
	M	SD	*p*	α	M	SD	*p*	α	M	SD	*p*	α
Food	3.95	.71	.002	.78	4.61	.82	.000	.77	4.56	.94	.000	.80
Thoughts and beliefs	3.26	.52	.004	.83	3.45	.76	.004	.81	3.31	.86	.004	.81
Lifestyles	3.59	.62	.003	.80	3.92	.63	.003	.82	4.15	.72	.000	.79
Spare time activities	3.87	.81	.001	.81	3.74	.54	.002	.80	3.40	.63	.004	.82

The second most influenced category in the community around College B and College C was the lifestyles, but it was about interests in the community around College A. According to the local residents, the lifestyles of the students and the academic staff living as tenants influenced the houseowners a great deal. For instance, as students usually stayed up late, most local residents, except the retired people, now went to bed at around midnight instead of 8 or 9 o'clock as they used to. Young people picked it up first, and then spead it out to other family members. All the traditional families who had dinner, watched television, and communicated together were affected. This constituted 56.5%. Nevertheless, those families whose members lived independently were not really affected. In this type of family, most young people went to bed around midnight, and the middle-aged went to bed at around 10 o'clock. Other clear-cut examples of lifestyle diversity included preferrences of fashion, habits, and behaviors. Local residents also responded that young people were eager to change and adapt to the diversity. On the one hand, many teenagers picked up Korean and Japanese fashion styles, but resisted Korean and Japanese sets of etiquette because they were thought to be complicated. On the other hand, they picked American style of communication. In traditional communication etiquette of Vietnam, teenagers were advised to bow slightly or to shake with two hands in business to show their respect to elder people, but they did not do so. Instead, they preferred to smile and greet with Vietnamese equivalent of "hello" and shake with only one hand even when greeting with older people. Local residents attributed this change to the impact of foreigners' lifestyles. An examination into the the demographic features showed that 4 Korean teachers, 2 Japanese teachers, and 5 American teachers from College B lived in the

nearby community for 3.5 years, and 5 Korean teachers, 3 Japanese teachers, and 5 American teachers from College C dwelled in the community around it for about 4 years. Also, these foreigners communicated with young people more often than the elderly in the communities. The demographic features of these foreigners were in line with the local residents' opinions about their influence on the young people's lifestyles in the communities.

Thoughts and beliefs were considered to be influenced the least in terms of the community diversity. The differences in the residents' opinions about the effects of the colleges on the communities were marginal, within the range of slightly over 3.25 to 3.45. Statistical analysis showed that these effects were significant, $p < 0.05$. The residents' responses in the interviews confirmed this. In general, thoughts and beliefs in Vietnam were of relative uniformity and solidity. The students and Vietnamese academic staff living as tenants softened some family traditions. For example, parents' words were not law any longer. That is, families were democratic. However, most families still practiced ancestral worship, a long-lasting tradition in Vietnamese culture. Considering the effects of foreigners, residents responded that the 3-4 year period was too short to change the common thoughts and beliefs. In addition, although most of these foreigners practiced Christianity, the communities around the three schools still remained buddhists or pagoda-goers. None of them changed their religions or appeared to have an intent to change their religions.

The reasons for the influence on spare time activities originated from their contact with the people from the colleges. College A had sports teams and clubs, and local residents around it were allowed to join them. However, the local residents around College B revealed that its influence was from the students and staff. Previously, people had preferred to read books and watch television in their spare time. Nonetheless, as the tenants wanted to relieve their homesickness, they set up sports teams. The most popular ones were soccer and badminton because they were easy to play and did not require any special instruments. Critically, some residents said that College B neither often organize sporting events for its students nor did it have a stadium. Foreigners here also interacted with the community much. That was the reason why the physical activities done in this community were also diverse. Interestingly, College C had a stadium, but it only served its students and staff. Students and staff living in the community around the college did not participate regularly in the events held by the community.

Social Cohesions

Statistical analysis (Table 5) demonstrates that the colleges had significant effects of the cohesions of the communities. College B was considered to influence the community the most. The community close to College C moderately influenced, with a mean score of slightly above the average. College A, according to the residents' responses, had considerable effects on its community activities and equality. In the interviews, the residents responded that students and staff (including foreigners) of College B living as tenants were sociable. They participated in most of the community activities, such as sporting events and celebrations. The students and Vietnamese staff

also initiated a few leisure activities for community cohesions. As a result of the existence of young people in the community, parents seemed to become more tolerant. As a result, equality was obtained. However, all the students and staff from College A living as tenants were Vietnamese. Their participation in community events was easy, but they usually returned to their hometown on holidays and in the summer. That did not mean they had no effects on the community cohesions. Their interactions with the local residents in daily life activities were the source of change. Although College C shared similarities with College B in that its Vietnamese and foreign staff lived off campus, they did not interact a great deal with the local residents. Their influences were less direct.

Table 5

Local Residents' Opinions of the Effects of the Colleges on the Community Cohesions

n=120	**College A** (n=35)				**College B** (n=45)				**College C** (n=40)			
	M	SD	*p*	α	M	SD	*p*	α	M	SD	*p*	α
Social event	4.15	1.26	.001	.85	4.27	1.09	.000	.87	3.72	1.07	.000	.78
Interactions	3.92	1.34	.002	.81	4.08	1.13	.001	.83	3.31	1.23	.004	.76
Equality	4.22	.85	.000	.82	4.13	1.04	.001	.81	3.97	.97	.002	.80

Quality of Life

The local residents' responses revealed the three colleges had different influence patterns on the quality of life of local residents (Table 6). College A had most influence on the mental health of the community, while College B and College C had most influences on sense of safety and living conditions. A closer look shows that College B had more influence on its nearby community than did the other two colleges. The local residents explained, in the interview, that the Vietnamese and foreign staff from College B and College C helped to improve the locals' living condition. The local residents added interesting information that College B seemed to be from wealthier families, and their expenditures added materialistic values to homestay owners.

The statistical analysis also showed that the local residents were concerned about sense of safety as an effect of the colleges. Regarding this concern, a female resident named Dan around College A said "They have different backgrounds. We do not know much about their families. We felt unsafe." Another resident said "Traffic has been becoming worse, especially in peak hours. It is always crowded with students here now." A landlord around College B said "Although most tenants here looked wealthy, we are not worried about theft, but violence. It may influence our children's development." Another resident responded "Traffic is terrible. Streets are congested with students in the daytime. Traffic accidents, as a result, are more frequent now." A resident around College C was also concerned about violence. He said "The staff from the college were nice, but the students sometimes drink alcohol. They gather and drink

beer or wine. Their friends outside the community also come over and join them. Thus, violence occurs between the students and their friends or between them and the local young people. However, they apologize afterwards." Another local resident living close to College C said "Thefts sometimes happened, but landowners did not call the police because they did not want the police to look into their business." She added "We do not care much about those tenants who have been living here for a year or more. We know them well. First-year students living as tenants often caused trouble. Sometimes they are asked to move." In a word, the local residents were concerned about traffic problems, theft, violence, and first-year students' adaptability.

Table 6

Local Residents' Opinions of the Effects of the Collges on the Community Life Quality

n=120	**College A** (n=35)				**College B** (n=45)				**College C** (n=40)			
	M	SD	*p*	α	M	SD	*p*	α	M	SD	*p*	α
Physical health	3.25	1.01	.005	.76	3.21	1.52	.005	.74	3.09	1.34	.005	.74
Mental health	4.22	1.35	.001	.84	4.19	1.47	.001	.83	3.32	1.28	.004	.78
Sense of safety	4.15	1.16	.001	.83	4.32	1.39	.001	.87	4.21	1.32	.001	.82
Living condition	4.03	1.51	.001	.85	4.37	1.09	.000	.86	4.10	1.21	.001	.83

The quantitative data also showed that the colleges had the least effect on public physical health. In the interviews, the local residents confirmed that these influences were positive. As College A allowed the local people to join its institutional activities, young people in the community joined the school league. Public health had been improving, as a result. College B, although did not welcome the participants of the local residents in the school activities, was considered to have positive effect on public health. A local resident said "Since the college began to operate, sports clubs have been constructed in the community to serve both local residents and tenants." However, College C was believed to have the least effect on public health of the nearby community.

Regarding mental health, most local residents around College C believed that the influence was caused mostly by the students. They were noisy. College A and College B were supposed to have more influence on mental health of the nearby communities. Most of the responses criticized public noise caused by the students. A landowner revealed that "They laugh and speak very loud. I warned them many times. Sometimes they made loud noise at night." A female resident said "They make noise even in public in the daytime. They speak loud along the street." Another resident

responded "Because the community did not have a soccer court, some local young people together with the tenants played soccer in the streets late in the afternoon. It was very noisy." However, one landlady in the community around College B provided positive feedback that she was less worried because her living condition had been improving. She added "I have learned about cultures of Japan, United States of America, and Korea from the foreigners from the college."

Effects of Institutions of Higher Education on the Community Perceived by College Students and Staffs

Table 7

College employees' and students' opinions of the effects of the colleges on community

n=1,470	**College A** (n=423)				**College B** (n=520)				**College C** (n=527)			
	M	SD	*p*	α	M	SD	*p*	*α*	M	SD	*p*	α
Equity	4.03	1.24	.001	.85	4.12	1.31	.001	.82	4.15	1.05	.001	.83
Diversity	3.03	1.36	.005	.80	3.21	1.42	.004	.81	3.06	1.34	.005	.78
Social cohesions	3.97	1.43	.003	.83	4.32	1.25	.000	.86	3.23	1.56	.004	.79
Quality of life	3.78	1.02	.002	.83	4.11	1.14	.000	.84	3.91	1.42	.002	.80

The employees and students of the three colleges living off campus responded that their impacts on all areas of concerns about the communities were at least moderately significant (Table 7). In general, their responses confirmed the local residents' opinions about the effects of the colleges. However, the employees and students believed that they had less influence than did the local residents.

Regarding the equital contributions, the respondents around College B and College C were supposed to have more effect than College A. They explained that college staffs and students living off campus contributed greatly to developments of the communities by their expenses, especially overheads and food. They further explained that the electricity and water supply they used were overcharged by the landowners. They were mostly double-priced. Regarding social equity, most of them revealed that many disadvantaged people did not receive sufficient care. Most elderly and disabled people mainly depended on their families. Communal aids were rare, but some people personally received financial aids or charity from generous people. One student said "Some people cannot even afford health care." He added "Health care for disabled and old people should be free of charge." A Vietnamese employee said "Health care is partially included in the health insurance for working people, but unemployed people had to pay the whole bill." A foreign employee also responded "Health care for physically and financially challenged people should be free of charge.", "There should be a community fund for this if health care benefit is not provided by the government.", he added.

Diversity was considered to be the least concerned area, with a mean score slightly above the average. The diversity of the community around College B was considered to be the most influenced. In the interviews, the students and employees revealed that harmony was always a problem. Young people could integrate easily, but new lifestyles, thoughts, and interests were not really welcomed by the middle-aged and elderly. They seemed to be closer to those people born in the community than to people from other places. A student further explained "I do not think it is an example of generation gap. They even criticize lifestyles of foreigners here. I sometimes overheard about it." A Vietnamese employee detailed that "They even made jokes of tenants' accents." A foreigner said "I do not care much about their discrimination, although I know it exists." She added "The problem is that I find it hard to communicate with the local people sometimes. You know, communications between people in the same community are inevitable." Another foreigner said "I usually ask my students or school staff to assist me when I am in need of help." It can be seen from the tenants' responses that the local residents did not really encourage diversity. However, the effects of the colleges on the communities were unwanted and unavoidable.

The influences of the colleges on social cohesions were also considered by the employees and students to be significant. College B was assumed to have the most effect in this respect. In the interviews, the respondents said that unimportant social events or community activities provided circumstances for their interactions with the community members. They also said that interactions between tenants and landowners were rare. For example, student Thanh said "Do I talk to them only when it is really necessary." A further student said "When there are important celebrations, such as national holidays, tenants here often return home. Foreigners travel instead." A foreigner revealed interesting information "Language barriers often make interactions between local people and foreign tenants impossible." Another foreigner added "Culture shocks make communications unfrequent." He recounted his personal experience that the local people often asked him personal questions. He explained "It seemed that I was being investigated. That made me uncomfortable." However, most students and employees as tenants around College B said that people were open-minded and understanding. Many of them could speak English, although not very proficiently. They respected personal styles and privacy. In summary, social cohesions in the community around College B were considered to be modified by its employees and students living around the campus as a result of their interactions. However, the unfrequent interactions between the employees and students as tenants around College A and College C made effects on the communities less direct and solid.

Finally, the employees' and students' opinions about their effects on quality of life were generally positive. All of them responded that their expenses improved the living conditions of the communities, which in turn greatly contributed to the well-being of the communities. One student said "Landowners do not worry about income now." They accepted that some tenants were very noisy. "But noise is usually caused by new comers.", another student said. They also expressed that "Traffic is unavoidable, but students should not be blamed. Traffic congestion is a concern of macro-management." However, they admitted that some students rode motorbikes

carelessly, and accidents, therefore, happened. One student added "Local young people caused more traffic accidents than tenants. They behaved carelessly because they thought they were indigeous." The Vietnamese and foreign employees confirmed their words by the students. Concerning sense of safety, the respondents around College A also revealed that they did not feel a sense of security; however, the respondents from those living around College B and College C were mostly positive.

Discussion, Conclusions, and Recommendations

The domains of sustainable development may be hard to be separated. A conclusion if a place has sustainably developed required examinations into different indicators (Davidson, 2010). This study investigated only social sustainability as an indicator of sustainable development. From the results, it can be said that the communities around College A, College B and College C had been developing. However, social sustainability was still quite far. Effects of built institutions on the nearby community were both possitive and negative. From the results, the effect on social equity was positive and highly appreciated by the community. However, the impacts on diversity, social cohesions, and quality of life were both positive and negative. From the positive view, quality of life developed as a result of economic development. However, the local residents, especially the middle-aged and elderly, took the detrimental effects serious.

The findings also illustrated that the re-location or construction of college or university may cause effects on equity, diversity, social cohesions, and quality of life. An educational institution with most students and employees, including teachers, living off campus is usually of the most influence. The construction of dormitories for these groups may be ideal. However, in contexts where this is impossible, school administrators are advised to establish a professional or business network with the landowners in the neighborhood to assist their students' and employees' lodging.

The employees from the colleges living off campus were of more benefit than harm to the community. The local residents did not really complain much about them. They were considered to add much materialistic values to the community. Foreign teachers from the colleges were assumed not to cause any concern of safety either. However, these people did not integrate much with the community in which they lived. From the view of social sustainability, this was not really positive. All the community members should have had social interactions or cohesions instead. Their cohesions could lead to strong development of the whole community through meetings.

Cultural differences were generally unwelcome. This made the expatriate teachers feel uncomfortable in communication with local residents. Open-mindedness may be helpful. The use of your own culture to judge people from other cultures cannot bridge the gaps between different ethnic groups (Hoa & Vien, 2018). Also, new comers to a particular community, in case of college and university, can be students and non-native teachers who find the local culture and practices different from theirs. They should find some ways to understand and adapt to the local norms. Adaptation can

give them a sense of well-being. Mutual respect should be an issue in contemporary life. In the age of globalization, it is common that people living in the same community are from different places, and mutual respect is appreciated, which in turn contributes to social sustainability (Bagceli Kahraman & Onur Sezer, 2017). In summary, accepting the existence of different cultures may help to make the social cohesions in this case better.

In addition, apart from contributions of new community members, such as equity, infrastructure should be an issue of concern. When an educational institution of higher education is constructed, there should be a concern of transportation. The governmental authority of the community should improve transport systems. Schools should also have mass transits for their own students. The equital gains from the colleges should be partly distributed as a return to the community as financial aids to disadvantaged people or funds for infrastructure development and maintenance. Community planning is essential in that it should predict and foresee both positive and negative effects so that necessary preparations can be done.

It is also significant to educate people, including local residents and tenants about the importance of social sustainability. In particular, education can be implemented cognitively, which may help carve knowledge and result in retention of knowledge (Hung, Vien & Vu, 2018). Alternatively, education of social sustainability can be undertaken from a social approach in that people and media play a crucial role in transmitting knowledge and arouse awareness of social sustainability.

What's more, sense of security is essential. Noise and violence may make people unproductive and may cause detrimental influence on health. In celebrations, young people may make uncontrolled and unexpected noise and violence. This probably accumulates a bad public image in a long run. Interactions in community are essential so that sympathy is grasped. In interactions between groups of people in a community, people understand what they should do and what they are expected to do. Thus, they can make adjustments.

As this study was a type of primary research on the effects of educational institutions on community from the view of social sustainability, it was difficult to conclude if these communities had been improved over time. In a broad scale, because social sustainability was not widely researched in the world, it may be early to construct a model or framework about the effects of institutions of higher education on the neighborhood. However, the findings could generally confirm the model by Cato (2009) that economic development could be in the center of sustainable development. The respondents accepted that their living conditions, well-being and health improved tremendously from the tenants' expenditures. Finally, the impacts of the institutions of higher education involved in this study were mainly based on the participants' responses. Although a number of measures were taken to increase the research validity and reliability, this could not describe the whole picture of the issue.

Further studies can take a holistic approach to sustainable development as a result of the effects of educational institutions. They may also include observations of the

changes of a community over time. Researchers may also take into account more educational institutions of higher education in other contexts.

References

Adams, W. M. (2006). The future of sustainability: Rethinking environment and development in the twenty-first century. *Report of the IUCN Renowned Thinkers Meeting.*

Bagceli Kahraman, P. & Onur Sezer, G. (2017). Relationship between attitudes of multicultural education and perceptions regarding cultural effect of globalization. *Eurasian Journal of Educational Research, 67,* 233-249. Doi: 10.14689/ejer.2017.67.14.

Cato, S. M. (2009). *Green economics*. Earthscan Publications: London.

Cook, S. & Esuna, D. (2014) Rethinking social development for post-2015 world. *Development*, *57*(1), 30-35.

Dave, S. (2011). Neighborhood density and social sustainability in cities and urban countries. *Sustainable Development, 19,* 189-206. Doi: 10.1002/sd.433.

Davidson, M. (2010). Social sustainability and the city. *Geography Compass, 4*(7), 872-880. Doi: 10.1111/j.1749-8198.2010.00339.x.

Dempsey, N., Bramley, G., Power, S. & Brown, C. (2011). The social dimension of sustainable development: Defining urban social sustainability. *Sustainable Development, 19*(5), 289-300. Doi: 10.1002/sd.417

Gass, S. M. & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hayden, M. & Thiep, L. Q. (2007). Institutional autonomy for higher education in Vietnam. *Higher Education Research and Development*, *26*(1), 73-85. Doi: 10.1080/07294360601166828.

Hoa, C. T. H. & Vien, T. (2018). The integration of intercultural education into teaching English: What Vietnamese teachers do and say. *International Journal of Instruction, 12*(1), 441-456. Doi: 10.29333/iji.2019.12129a.

Hung, B. P. (2017). Vietnamese students learning the semantics of English prepositions. *GEMA Online Journal of Language Studies, 17*(4). 146-158. Doi: 10.17576/gema-2017-1704-10.

Hung, B. P., Truong, V. & Nguyen, N. V. (2018). Students' responses to CL-based teaching of English prepositions. *Eurasian Journal of Educational Research, 73*(3), 41-58. Doi: 10.14689/ejer.2018.73.3.

Hung, B. P. & Van, L. (2018). Depicting and outlining as pre-writing strategies: Experimental results and learners' opinions. *International Journal of Instruction, 11*(2), 451-464. Doi: 10.12973/iji.2018.11231a.

Hung, B. P., Vien, T. & Vu, N. N. (2018). Applying cognitive linguistics to teaching English prepositions: A quasi-experimental study. *International Journal of Instruction, 11*(3), 327-346. Doi: 10.12973/iji.2018.11323a.

Johansson, T.; Segerstedt, E.; Olofsson, T. & Jakobsson, M. (2016). Revealing social values by 3D city visualization in city transformations. *Sustainability, 8*(2), 195. Doi: 10.3390/su8020195.

Lawhan, L. L. (2009). The neighborhood unit: Physical design or physical determinism? *Journal of Planning and History, 8*(2), 11-132. Doi: 10.1177/1538513208327072.

McKim, C. A. (2017). The values of mixed methods research: A mixed methods study. *Journal of Mixed Methods Research, 11*(2), 202-222. Doi: 10.1177/1558689815607096.

McMillan, J. H. & Schumacher, S. (2001). *Research in education: A conceptual introduction* (5th ed.); Longman: New York.

Nha, N. V. & Tu, V. N. (2015). Higher education in Vietnam: Current situation, challenges and solutions. *VNU Journal of Science, 31*(4), 85-97.

Nieuwenhuis, J. & Hooimeijer, P. (2015). The association between neighborhoods and educational achievement, a system review and meta analysis. *Journal of Housing and the Built Environment, 31*(2), 321-347. Doi: 10.1007/s10901-015-9460-7.

Nieuwenhuis, J., Hooimeijer, P. Dorsselaer, S. V. & Vollebergh, W. (2013). Neighborhood effects on school achievement: The mediating effect of parenting and problematic behavior, *Environment and Planning A, 45*(9), 2135-2153.

Omar, D. Bt., Abdullah, F., Yusof, F., Hamdan, H., Nasrudin, N. & Abullah. (2011). The impacts of off-campus students on local neighborhood in Malaysia, *International Journal of Social, Human Science and Engineering, 5*(10), 24-30.

Polese, M. & Stren, R. (2000). *The social sustainability of cities: Diversity and the management of change;* University of Tokyo Press: Tokyo.

Reed, R. G. & Wilkinson, S. J. (2005). The increasing importance of sustainability for building ownership. *Journal of Corporate Real Estate, 7*(4), 339-350. Doi: 10.1108/14630010510700831.

Rogers, P. P., Jalal, K. F. & Boyd, J. A. (2007). *An introduction to sustainable development;* Earthscan Publications: London.

Sen, A.K. (2000). The ends and means of sustainability: Keynote address. *The International Conference on Transition to sustainability*, Tokyo.

Sheridan, G. (2010). *Vietnam: Preparing the higher education sector development project (HESDP) – Project number: 4207901.* Accessed from https://www.adb.org/sites/default/files/project-document/63092/42079-01-vie-tacr-03.pdf on September 3rd, 2018.

Tanriogen, Z. M. (2018). The possible effects of the 4th industrial revolution on Turkish educational system. *Eurasian Journal of Educational Research, 77,* 163-184. Doi: 10.14689/ejer.2018.77.9.

Tashakkori, A. & Creswell, J. W. (2007). Developing publishable mixed methods manuscripts. *Journal of Mixed Methods Research, 1*, 107-111. Doi: 10.1177/1558689806298644.

Vallance, S., Perkins, H. C. & Dixon, J. E. (2011). What is social sustainability? A clarification of concepts. *Geoforum*, *42(*3), 341-348. Doi: 10.1016/j.geoforum.2011.01.002.

Yu, Y.; Tong, Y.; Tang, W.; Yuan, Y. & Chen, Y. (2018). Identifying spatiotemporal interactions between urbanization and eco-environment in urban agglomeration in the middle reaches of the Yangtze River, China. *Sustainability, 10*(1), 290. doi:10.3390/su10010290.

Wey, Wann-Ming (2018). A commentary on sustainably built environments and urban growth management. *Sustainability, 10(*11), 1-5. Doi: 10.3390/su10113898.

Woodscraft, S., Bacon, N., Caistor-Arendar, L. & Hackett, T. (2011). *Design for social sustainability: A framework for creating thriving new communities*. Social Life: London.

Eurasian Journal of Educational Research 80 (2019) 59-88

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

Student and School Level Variables related to Elementary School Students' Attitudes towards Science*

Esme HACIEMINOGLU[1]

ARTICLE INFO

Article History:
Received: 5 Aug. 2018
Received in revised form: 15 Oct. 2018
Accepted: 17 Mar. 2019
DOI: 10.14689/ejer.2019.80.4

Keywords
Attitude, Hierarchical linear modeling (HLM), Educational resources, cross-sectional survey

ABSTRACT

Purpose: In the literature, there are lots of quantitative studies regarding students' attitudes, and results of the studies revealed that students have negative attitudes towards science. Therefore, it is necessary to study the factors that predict the attitudes of the students towards science taking into account the factors at both the school and the student level. The purpose of this study was to investigate what school- and student-level factors are associated with student' attitudes towards science.

Research Methods: The overall design of this study is mainly a cross-sectional survey and correlational. The convenience sampling method was used in this study and 2975 elementary students in different schools and cities of Turkey constituted the sample of this study. The Test of Science Related Attitude, Learning Approach Questionnaire, Achievement Motivation Questionnaire, School Background Questionnaire were used as data collection tools. Hierarchical Linear Modeling (HLM) was selected as a modeling technique for data analyses.

Findings: This study provides a general overview about students' attitudes towards science. The quality of school's educational resources, learning and motivational factors, factors related to student feelings and outside activities, and some student characteristics significantly contributed to the students' attitudes towards science.

Implications for Research and Practice: Enriching science lessons with materials, increasing students' curiosity, making them learn by discovering and making them do experiments have potentials to contribute students' attitudes, and experimental research can be used to explore these effects. In order to examine the reasons why students like science or not in detail, qualitative studies should also be conducted.

*This study was partly presented at the 5th International Eurasian Educational Research Congress in Antalya, 02 May – 05 May, 2018

[1] Corresponding Author, Akdeniz University, TURKEY, e-mail: ehacieminoglu@gmail.com, ORCID: https://orcid.org/0000-0001-9318-3931

Introduction

According to Petty and Cacioppo (1996), attitude is a tendency of conducting behavior and is not observable. Attitude can be defined as positive or negative feelings about a human, object or subject (as cited in Genc, 2015). Students' attitudes towards science have been studied in different research contexts for years. In the literature there are lots of quantitative studies regarding students' attitudes, and results of the studies revealed that students have negative attitudes towards science. This negative attitude prevents many students from exhibiting positive attitudes towards scientific research and from continuing scientific inquiry (Arisoy, 2007; Azizoglu & Cetin, 2009; Hacieminoglu, 2016). Researchers tried to develop students' attitudes toward science using different teaching methods. One of these studies is conducted by Akamca and Hamurcu (2005). It aimed to investigate the effects of science lessons designed according to multiple intelligence theory on fifth grade students' achievement, permanence of knowledge and attitudes towards science lessons. At the end of the study, conducted through Pre-test – post-test control group quasi-experimental design, It was found that instruction based on multiple intelligences had a significant effect on achievement and permanence but no significant relationship was found between the instruction method and students' attitudes towards science lessons. In another study, (Celik, 2006) examined the effect of web-based science lesson instruction on 9th grade students' problem solving skills and attitudes towards science lesson. In the experimental group, web-based research and discussion environment was created to support the instruction process. Results suggested that web-based instruction did not have a significant effect on students' attitudes towards science lessons.

Sasmaz Oren and Tezcan (2009) conducted a quasi-experimental study to examine the effect of learning cycle approach on 7th grade students' attitudes towards science lessons. In the experimental group lessons were taught through learning cycle method for five weeks. Results suggested that attitudes of students in the experimental group increased significantly compared to the attitudes of students in the control group in which the instruction is done through traditional methods. Similarly, Cibik (2009) aimed to investigate the effect of project-based learning approach on 7th grade students' attitudes towards science lessons. Pre-test – post-test control group quasi-experimental design was used in the study. Project-based learning approach was explained to teachers and students in the experimental group. The implementation process lasted six weeks and every week, students presented their studies. Results suggested that project-based learning approach significantly increased students' attitudes towards science lessons. Koc and Bocek (2012) examined the effect of activities done with simple materials on the 7th grade students' attitudes towards science lessons. Pre-test – post-test control group quasi-experimental design was used in the study. In the experimental group, the force and movement unit was taught through methods such as direct instruction, question-answer and discussion, and along with these methods students conducted experiments with the materials they can find in their immediate environment. At the end of the study it was found that the

experiments done with simple materials had a significant effect on their attitudes towards science lessons.

There are also studies assessing technology integrated classroom on students' attitudes towards science. One of these studies conducted by Gomleksiz and Fidan (2013) aimed to determine the effect of computer-supported mind map (CSMM) technique on 7th grade students' achievement, attitudes towards science and technology lesson and permanence of knowledge. They also aimed to find out views of teachers and students regarding the activities. Results revealed that CSMM technique was more effective in teaching science and technology lesson than the traditional teaching methods. This shows that CSMM technique is more beneficial for improving academic success in teaching of science and technology lesson. As a result of the activities done through CSMM technique students' achievement and attitudes towards science and technology lesson improved more than the students taught through traditional methods. Another study is the one conducted by Sakiz, Ozden, Aksu and Simsek (2014). Researchers aimed to examine the effects of smart board use on the fourth grade student's achievement and attitudes towards science lesson. Pre-test – post-test control group quasi-experimental design was used in the study. The sample group was composed of 78 fourth grade students. Results revealed that smart board use had a positive effect on students' achievement and attitudes towards science lesson.

There are also some similar experimental studies regarding the effects of new teaching strategies (such as, STEM, discovery, out of school learning environment) entering the new science program on students' attitudes towards science. Firstly, Yamak, Bulut and Dundar (2014) aimed to examine the effect of STEM approach on fifth grade students' attitudes towards science lesson and scientific process skills. The sample of the study was composed of 20 elementary school students studying in Ankara. In this study, which was conducted through single group pre-test – post-test design, three different STEM activities were applied. In the implementation process, teacher explained the task and gave materials. Students discussed the process, tested and evaluated their models, and lastly presented them. Results suggested that STEM education positively affected students' attitudes towards science lesson and their scientific process skills. Secondly, one of the studies examining the effects of discovery method was conducted by Unal and Ergin (2006). It aimed to investigate the effect of activities prepared in accordance with the constructivist approaches for the teaching of liquid and gas pressure unit on 7th grade students' achievement and attitudes towards science lesson. Results revealed that activities done through worksheets prepared in line with discovery approach improved students' achievements considerably but it was also found that these activities improved students' attitudes towards science in a low level. Similarly, Kucuk (2014) aimed to determine the effects of simulation method on the 7th grade students' achievement in light unit and their attitudes towards science lesson. Pre-test – post-test control group design was used in the study. In experimental group, lessons were taught through simulation method, and constructivist approach was used in control group. Results showed no significant differences between two groups. Another study on light and sound unit was

conducted by Teker, Kurt and Karamustafaoglu (2017). It aimed to examine the effect of using discovery learning method in light and sound unit on fifth grade students' achievement and attitudes towards science lesson. Pre-test – post-test control group quasi-experimental design was used in the study. Discovery learning approach was used while teaching the experimental group. In control group lessons were taught according to the curriculum and students' book prepared and offered by Ministry of Education. Results suggested that the discovery learning approach had a significant effect on students' achievement and attitudes towards science lesson. Thirdly, one of the studies conducted about science festivals, an out-of-school learning environment, belong to Yildirim and Sensoy (2016). It aimed to investigate the effects of science festivals on 6th grade students' attitudes towards science lesson. Lessons were taught according to the 6th grade science lesson curriculum in both experimental and control groups. However, students in experimental group spent an additional 15-minute of one lesson in every week on the science festival preparations. Some information regarding science festivals, scientific methods, research steps and project diaries were given to the students in experimental group. It was seen from the results that attitudes of the students in the experimental group increased significantly and this increase was preserved for three months after the study. Lastly, Keles and Oner (2016) aimed to examine the effect of elective science applications course on seventh grade students' attitudes towards science lesson. The sample of this quasi-experimental study was composed of 212 seventh grade students studying in Agrı. The implementation process lasted 33 weeks. Activities were done with the students taking this course. Post-test was applied to all the students whether they took this course or not. Results revealed that elective science applications course had a significant effect on 7th grade students' attitudes towards science lesson.

In the literature there are also some descriptive and correlational studies regarding students' attitudes towards science. One of these studies conducted by Bozdogan and Yalcin (2005) aimed to determine 6th, 7th and 8th grade students' attitudes towards physics topics. Descriptive survey model was used in the study. The sample group composed of 337 students (172 male, 165 female). Data were collected through an attitude scale prepared for 6th, 7th and 8th grades separately. It had 33 items in total. Results revealed that while the class level increases students' attitudes towards physics experiments decrease. When the educational opportunities were taken into consideration, attitudes of students studying in second type of schools (better educational opportunities and more teachers) were the highest. After that came the third type of schools (educational opportunities were the best, had the most teachers, were in the city center and were mostly preferred) and first type of schools (limited educational opportunities, less teachers and located in the countryside). Another study examining primary school students' attitudes towards science lesson was conducted by Kozcu-Cakir, Senler and Gocmen-Taskin (2007). It aimed to determine the relationship between primary school students' attitudes towards science lesson and various variables. These variables were as follows: class level, gender, the residential area they live in, educational level of their parents, socio-economic status of their parents, presence of a study room of their own and frequency of using laboratories in science lessons. Descriptive correlational design was used in the study. 440 primary

school students studying in central districts of Mugla participated in this study. Results revealed that there was a significant difference between 6th grade students' and 7th and 8th grade students' attitudes towards science in favor of 6th graders. However no significant difference was found between the attitudes of 7th and 8th graders. It was found that as the class level increased the attitude scores decreased. No significant difference was found between students' attitudes towards science lesson and their gender, residential area they live in, parents' education status and family's socio-economic status. It was also found that having a private room positively affected their attitudes towards science lesson. It can be concluded that because the students having a private room had a more comfortable study environment, they were more successful in their lessons and their attitudes increased accordingly. According to results of the one-way analysis of variance, a significant difference was found between students' science lesson achievement and their attitudes. The students successful in science lessons had more positive attitudes towards science. Moreover, when the laboratory use was examined, a significant difference was found between the groups who use science laboratories sometimes, usually and always and the group using it never. The difference was in favor of the laboratory using group.

The other one is the study of Karacam, Mirza and Elitok (2013) that aimed to compare the 6th grade students' attitudes towards science lesson to their frequency of watching documentaries and their gender. 157 students studying at the sixth grade of an elementary school in Ankara participated in the study. Results revealed that students watching documentaries about science topics more often have more positive attitudes towards science lesson than students watching them rarely.

There are also some recent studies related to students' attitudes towards science and related variables. One of these studies belongs to Kapici and Akcay (2016). The purpose of this study was to investigate middle school students' attitudes towards science. Sample of this study constituted of 2063 fourth to eight grade students from all regions of Turkey. Results revealed that girls have more positive attitudes than boys, and while students' grade level increases, attitudes toward science decrease. The other study conducted by Keles and Aydin (2017) examined elementary school students' attitudes towards science in terms of class level. 649 students (161 fifth grade, 174 sixth grade, 152 seventh grade and 162 eighth grade) participated in the study. Results revealed that students' attitude scores only decrease as they get through from fifth to sixth grade, and that attitude scores stay almost the same for the other class levels.

In educational research data generally have a nested structure. Each student might be nested within some schools or classrooms. Beside this, these schools or classrooms might be nested within any other location such as a district, province, region, or country. If these hierarchical data are analyzed with traditional linear model, some of the basic assumptions especially the independence of observation is violated. The students in the same group (a classroom or a school) are more similar than the students in different groups. All the students of a school or a class are affected by the school or class atmosphere in the same manner. Additionally, the factors that affect the students in the same group (a school or a classroom) have the same effects on only if all the

students are in the same group. For example; the educational resources of school A affect all the students of school A in the same way. On the other hand, they do not have any effects on the students of school B. Therefore, students in different groups can be independent; however the students in the same group like same classroom or same school have the same value on some classroom or school factors. If these hierarchical data are analyzed with the traditional linear model, some of the basic assumptions, especially the independence of the observation, are violated. If the independence of observation assumption is violated, estimating the coefficients can be biased, and the estimates of standard errors can be smaller than they should be. Each group is represented by its own sub-models when the data are analyzed using the hierarchical linear modeling technique (HLM). These sub-models not only show relationships among the variables at the same level, but also reveal the effect of the variables at this level on the other level. For this reason, HLM is a more reliable statistical technique for identifying relationships in studies for hierarchical structural data (Raudenbush & Bryk, 2002). In recent years student, classroom and school level predictors have been investigated using large scale data such as TIMMS and PISA. While some studies in the literature (Huffman, Lawrenz, and Minger, 1997; Le Mare and Sohbat, 2002) support the conclusion that students with female teachers feel more comfortable and safer, Gilmartin, Denson, Li, Bryant and Aschbacher, (2007) found no relation between the number of female teachers in schools and their interest in science courses (as cited in Hacieminoglu, Ertepinar, Yilmaz-Tuzun & Cakir, 2015). Meta-analysis studies of Fuller (1987), Hanushek (1997) and Burtless (1996) revealed that most of the studies they investigated supported the idea that teaching materials, and instructional resources such as school library, availability of science laboratories in science teaching were positively related to students' achievement. Students' attitudes towards science is strongly associated with students' achievement. Therefore, it is necessary to study the factors that predict the attitudes of the students towards science taking into account the factors at both the school and the student level.

The purpose of this study was to investigate what school- and student-level factors are associated with students' attitudes towards science. The specific research questions were:

1. Are there any differences in students' attitudes towards science among schools?

2. Are proportion of female science teachers, ability grouping between science classes, quality of school's physical infrastructure and quality of school's educational resources associated with students' attitudes towards science?

3. Are students' background characteristics, factors related to students characteristics, factors related to student feelings and outside activities and earning and motivational factors associated with students' attitudes towards science?

Method

Research Design

The overall design of this study is mainly a cross-sectional survey and correlational. Fraenkel and Wallen (2003) stated that the survey type of research is used to describe the characteristics of a population through asking a set of questions. Moreover, correlational type of research is used to determine the relationships among two or more factors without any manipulation.

Research Sample

The convenience sampling method was used in this study and 2975 sixth, seventh and eight grade elementary students in different schools and cities of Turkey constituted the sample of this study. The distribution of the students' demographic characteristics was presented in Table 1.

Table 1

Demographic and Sociodemographic Characteristics of Participants

Demographic Characteristics		Number	% Percent
Region	Marmara	430	14.4
	Black Sea	382	12.9
	Central Anatolia	1004	33.7
	Aegean	633	21.2
	Mediterranean	392	8.5
	Eastern Anatolia	114	3.9
	Southeastern Anatolia	20	0.7
Gender	Female	1531	51.5
	Male	1444	48.5
Income	Low	721	24.2
	Medium	1520	51.1
	High	734	24.7

Research Instruments and Procedures

The Test of Science Related Attitude, Learning Approach Questionnaire, Achievement Motivation Questionnaire, School Background Questionnaire were used as data collection tools for this study. The information about data collection instruments was summarized in Table 2.

Table 2
Information about Data Collection Tools

	The Test of Science Related Attitude	**Learning Approach Questionnaire**	**Achievement Motivation Questionnaire**	**School Background Questionnaire**
Scale type	5-point Likert-type scale	4-point Likert-type scale	4-point Likert-type scale	
Developed by	Fraser (1978)			OECD Publications (2004, p.316)
Translated and adapted into Turkish	Arisoy (2007)	Caliskan, (2004)	Caliskan, (2004)	
Used by	Hacieminoglu (2016)	Bou Joude (1992) Cavallo and Schafer (1994)	Bou Joude (1992) and Cavallo and Schafer (1994)	Hacieminoglu, Ertepinar,Yilmaz-Tüzün & Cakir (2015)
Items	40 items	22 items	14 items	**Information obtained from the principals of each school**
Dimentions and reported cronbach alpha reliability index	Adaptation of scientific attitudes,	Rote learning α .81	Learning goals α .94	**School SES**
	Enjoyment of science lessons,		Performance goals α .82	**Ability grouping between science classrooms,**
	Leisure interest in science,	Meaningful learning α.76	Self-efficacy α .89	**Proportion of female science teachers,**
	Career interest in science			**Quality of school's physical infrastructure,**
	Total Scale α .78 by Fraser (1978)	Cavallo, Rozman and Potter (2004)	Cavallo, Rozman and Potter (2004)	"school buildings and grounds", "heating/cooling and lighting systems, "instructional space (eg., classrooms)" (not at all=1, very little=2, to some extent=3, and a lot=4)
Dimentions and cronbach alpha reliability index	Adaptation of scientific attitudes α .68	Rote learning α .77	Learning goals α .83	
		Meaningful learning α.71		
			Performance goals α .73	**Quality of school's educational resources**
	Enjoyment of science lessons α .83		Self-efficacy α .75	"Instructional materials (eg. textbooks)", "science laboratory equipment and material", "computers for instruction", "library materials", "audio-visual resources" (not at all=1, very little=2, to some extent=3, and a lot=4)
	Leisure interest in science α .84			
	Career interest in science α .80			

Factors

In this study factors considered are labeled as outcome factors, student level factors (Level-1) and school level factors (Level-2), and presented in Table 3.

Outcome Factors

The outcome factors of this study were students' attitudes towards science.

Student Level (Level-1) Factors

Student Level (Level-1) Factors were students' background characteristics, factors related to students characteristics, factors related to student feelings and outside activities, learning and motivational factors.

Students' Background Characteristics

Students' background characteristics were their socio-economic status, parents' education level and parents' occupational status.

Factors related to Student Characteristics

Grade level, Science achievement (science grades) and gender were factors related to student characteristics. Science grades refer to the achievement scores obtained from trial high school exam test, which is a standardized test applied by the Ministry of National Education for all elementary schools.

Factors related to Student Feelings and Outside Activities

The course they like most, whether they read articles or books regarding science, whether they benefit from the internet sites regarding science, whether they watch documentaries and whether they share their ideas about science subject with their families were the factors related to student feelings and outside activities.

Learning and Motivational Factors

Performance goal orientation, learning goal orientation, self-efficacy, rote learning approach and meaningful learning approach were the factors regarding learning and motivation.

School Level (Level-2) Factors

School Level (Level-2) factors were proportion of female science teachers, ability grouping between science classes, quality of school's physical infrastructure and quality of school's educational resources.

Table 3

Outcome, Student and School level Factors

Outcome Factors	***Student Attitude Toward Science***		
Student Level (Level-1) Factors	Students' background characteristics	Socio economic status	INCOMEME INCOMEHI
		Parents' education level	DUMMYCOL DUMMYGRA
		Parents' occupational status	PARENTOC
	Factors related to student characteristics	Grade level	GRADE7 GRADE8
		Science achievement	SCIENGRA
		Gender	GENDER
	Factors related to Student Feelings and Outside Activities	The course they like most	DUMMYLIK
		Whether they read articles or books regarding science	READINGB
		Whether they benefit from internet sites regarding science	INTERNET
		Whether they watch documentaries	DOCUMENT
		Whether they share their ideas about science subjects with their families	SHARINGI
	Learning and Motivational Factors	Performance goal orientation	PERFGOAL
		Learning goal orientation	LEARNGOA
		Self-efficacy	SELFEFFI
		Meaningful learning approach	MEANINGF
		Rote learning approach	ROTELEAR
School Level (Level-2) Factors	Factors related with school characteristics	Proportion of female science teachers	FEMALESC
		Ability grouping between science classes	ABILITYG
		Quality of school's physical infrastructure	PHYSICAL
		Quality of school's educational resources	QUALITYE

Data Analyses

Hierarchical Linear Modeling (HLM) was selected as a modeling technique in order to investigate how the school level factors and student level factors were related to students' attitudes towards science because of the nested structure of the data sets that means students nested within schools. Models were developed by using HLM 6.0 in order to examine the relations between school level and students level factors.

Results

Results of Research Question I (One-Way ANOVA with Random Effects)

With respect to attitudes towards science in order to answer the first research question regarding if there were any differences in students' attitude toward science among schools, one-way ANOVA with random effects model was conducted.

For i= 1, …, nj students in school j, and j = 1, …, 23 schools, equations at two levels are:

Level 1 (Students level) Model:

$Y_{ij} = \beta_{0j} + r_{ij}$

Level 2 (School level) Model:

$\beta_{0j} = \gamma_{00} + u_{0j}$

where

Y_{ij} = the endogenous factor, attitude toward science for i^{th} students in j^{th} school

β_{0j} = the intercept (the mean attitude toward science for the j^{th} school)

r_{ij} = the student level error

γ_{00} = the grand mean

u_{0j} = the random effect associated with unit j (school)

The final estimation of fixed effects obtained from analysis of variance model of attitudes toward science is represented in the Table 4.

Table 4

Final Estimation of Fixed Effects for One-Way ANOVA with Random Effects for Attitudes towards Science

Fixed Effect	*Coefficient*	*Standard Error*	*t-ratio*	**p-value**
Average school mean, γ_{00}	3.546715	0.028411	124.836	0.000

The analysis of variance indicates that average school mean of attitudes towards science, the grand-mean of attitudes towards science (γ_{00}), was statistically different from zero. That means there were significant differences among schools.

The grand-mean of attitudes towards science was 3.546 with a standard error of 0.028, indicating a 95% confidence interval of:

Confidence Interval = 3.546 ± 1.96 (0.028) = (3.491, 3.600)

Table 5

Final Estimation of Variance Components for One-Way ANOVA with Random Effects for Attitudes towards Science

Random Effect	*Variance Component*	*df*	*Chi-square*	**p-value**
School mean, u_{0j}	0.01327	22	89.69055	0.000
Level-1 Effect, r_{ij}	0.47801			

The final estimation of variance components obtained from the one-way ANOVA with random effects model is represented in Table 5.

The findings also indicated significant (p< .005) variation does exist among schools in their attitudes towards science (χ^2 = 89.69055, df = 22). The result also revealed that school level factors might account for the differences in the students' attitudes toward science.

At the student level Var(r_{ij}) = σ^2 = 0. 47801. At the school level, τ_{00} is the variance of the true school means, β_{0j}, around the grand-mean, γ_{00}. Var(u_{0j})= τ_{00} = 0.01327.

The intraclass correlation (ICC), which represents proportion of variance in Y (attitude toward science) among schools, is

ICC = τ_{00} / (τ_{00} + σ^2) = 0.01327/ (0.01327+ 0. 47801) = 0.027

indicating that about 2.7% of the variance in attitudes toward science is among schools.

HLM also provides an estimate of the reliability of the sample mean in any school. The reliability is an estimate of the true school mean and is affected by the sample size within each school. The overall estimate of reliability is the average of the school reliabilities. ρ = .715 indicating that the sample means tend to be a reliable indicator of true school means. The equation for determining reliability of the mean attitudes towards science within each school is:

ρ = τ_{00} / [τ_{00} + (σ^2 / nj)]. Therefore, the reliability of the sample mean varies from school to school because the sample size, nj, varies.

In the following models, additional level 1 (student level) factors will be tried to reduce the variation within schools (σ2) and additional level 2 (school level) factors will be tried to explain between school differences (τ00).

Results of Research Question II (Means as Outcomes Model)

In order to answer the second research question regarding which of the school level factors are associated with students' attitudes toward science, means-as-outcome model was applied.

Equations at two levels are:

Level 1 (Students level) Model:

$Y_{ij} = \beta_{0j} + r_{ij}$

Level 2 (School level) Model:

$\beta_{0j} = \gamma_{00} + \gamma_{01}$ (QUALITYE) + γ_{02} (FEMALESC) + γ_{03} (ABILITYG) + γ_{04} (PHYSICAL) + u_{0j}

for j = 1, 2, …, n schools

where

β_{0j} = the school mean on students' attitudes toward science

γ_{00} = the intercept (the grand mean for students' attitude toward science, that is the average of the school means on students' attitude toward science scores across the population of schools)

γ_{01} = the differentiating effect of quality of school's educational resources on the school mean on students' attitude toward science.

γ_{02} = the differentiating effect of proportion of female science teachers on the school mean on students' attitude toward science.

γ_{03} = the differentiating effect of ability grouping between science classes on the school mean on students' attitude toward science.

γ_{04} = the differentiating effect of quality of school's physical infrastructure on the school mean on students' attitude toward science.

τ_{00} = the conditional variance or school level variance in β_{0j} after controlling for these school level factors.

u_{0j} = the residual

The model was first run with all four factors, but *Proportion of female science teachers, Ability grouping between science classes, Quality of school's physical infrastructure* were not significant and were removed from the final analysis. The final estimation of fixed effects obtained from means as outcomes model of students' attitude toward science is represented in Table 6.

The results revealed significant and positive relationship between quality of school's educational resources and students' attitudes toward science (γ_{01}= 0.003, se=0.001).

Table 6

Final Estimation of Fixed Effects for Means as Outcomes Model for Attitudes toward Science

Fixed Effect	*Coefficient*	*Standard Error*	*t-ratio*	**p-value**
Model for School Means[1]				
Intercepts, γ_{00}	3.546	0.025	137.868	0.000
QUALITYE, γ_{01}	0.003	0.001	2.125	0.045

[1] The student level factors were Grand Mean Centered before analysis.

The final estimation of variance components obtained from means as outcomes model is represented in Table 7. The degrees of freedom for this model (Means as Outcomes Model) is based on the number of schools with sufficient data, and the number of school level factors included in the model.

Degrees of Freedom = J – Q – 1, where

J = the number of schools with sufficient data

Q = number of school level factors included in the model

Thus, all schools were used in this analysis and degrees of freedom for this model is:

df = J – Q – 1 = 23 – 1 – 1 = 21

Table 7

Final Estimation of Variance Components for Means as Outcomes Model for Attitudes toward Science

Random Effect	*Variance Component*	*df*	*Chi-square* χ^2	**p-value**
School mean, u_{0j}	0.01005	21	68.16538	0.000
Level-1 Effect, r_{ij}	0.47819			

The residual variance between schools (τ_{00} = 0.01005) is substantially smaller than the original variance (τ_{00} = 0.01327) resulting from the analysis of variance model. This reduction is due to the inclusion of school level factors.

Proportion of variance explained at

$$\text{level 1} = \frac{\tau_{00}(\text{ANOVA}) - \tau_{00}(\text{Means as Outcomes})}{\tau_{00}(\text{ANOVA})}$$

Proportion of variance explained at level 1= $\frac{0.01327 - 0.01005}{0.01327} = 0.242$

This result indicated that 24.2% of the true between school variance in students' attitude toward science was accounted for *Quality of school's educational resources.*

Finally, in order to examine whether the school attitude toward science means vary significantly when quality of school's educational resources is controlled chi-square statistics was conducted. Chi-square statistic χ^2 was found as 68.1653 (df=21, p= .00). This finding indicated that this school level factor namely quality of school's educational resources did not account for all the variation in the intercepts. However, even after controlling for quality of school's educational resources, schools still varied significantly in their average attitude toward science views.

Results of Research Question III (Random Coefficient Model)

In order to answer the fifth research question regarding which of the student level factors help to explain the difference in understanding the students' attitude toward science Random Coefficient Model was conducted.

The equations to answer this question are:

Level 1(Students level):

$Y_{ij} = \beta_{0j} + \beta_{1j}$(GRADE7) + β_{2j}(GRADE8) + β_{3j}(SCIENGRA) + β_{4j}(GENDER) + β_{5j} (INCOMEME) + β_{6j}(INCOMEHI) + β_{7j}(DUMMYCOL) + β_{8j}(DUMMYGRA) + β_{9j} (PARENTOC) + β_{10j}(DUMMYLIK) + β_{11j}(READINGB) + β_{12j}(INTERNET) + β_{13j}(DOCUMENT) + β_{14j}(SHARINGI) + β_{15j}(PERFGOAL) + β_{16j}(LEARNGOA) + β_{17j}(SELFEFFI) + β_{18j}(MEANINGF) + β_{19j}(ROTELEAR) + r_{ij}

Level 2(School level):

$\beta_{0j} = \gamma_{00} + u_{0j}$

$\beta qj = \gamma_{q0} + u_{qj,}$

where

Y_{ij} = Attitude toward Science of student *i* in class *j*

β_{0j} = the school mean on attitude toward science

β_{1j} = the differentiating effect of 7th grade level in school j

β_{2j} = the differentiating effect of 8th grade level in school j

β_{3j} = the differentiating effect of science achievement in school j

β_{4j} = the differentiating effect of gender in school j

β_{5j} = the differentiating effect of medium level income in school j

β_{6j} = the differentiating effect of high-level income in school j

β_{7j} = the differentiating effect of college education level as a highest educational level of parents in school j

β_{8j} = the differentiating effect of graduate education level as a highest educational level of parents in school j

β_{9j} = the differentiating effect of highest parental occupational status in school j

β_{10j} = the differentiating effect of the course student like most in school j

β_{11j} = the differentiating effect of if students read articles or books regarding science in school j

β_{12j} = the differentiating effect of if students benefit from internet sites regarding science in school j

β_{13j} = the differentiating effect of if students watch documentary film in school j

β_{14j} = the differentiating effect of if students share their ideas about science subjects with their families in school j

β_{15j} = the differentiating effect of students' performance goal orientation in school j

β_{16j} = the differentiating effect of students' learning goal orientation in school j

β_{17j} = the differentiating effect of students' self-efficacy in school j

β_{18j} = the differentiating effect of students' meaningful learning approach in school j

β_{19j} = the differentiating effect of students' rote learning approach in school j

βqj = the coefficient for factor *q* for class j after accounting for other factors

γ_{00} = the average of school mean on attitude toward science across the population of schools

γ_{q0} = the average *q* factor- attitude toward science slope across those schools

u_{0j} = the unique increment to the intercept associated with school *j*

u_{qj} = the unique increment to the slope associated with school *j*

The final random coefficient model included the factors significantly related to attitudes toward science, and the factors both significantly related to attitudes toward science and randomly varying. The final estimation of fixed effects obtained from random coefficient model of is displayed in the Table 8.

Table 8

Final Estimation of Fixed Effects for Random Coefficient Model for Attitude toward Science

Fixed Effect	*Coefficient*	*Standard Error*	*t-ratio*	**p-value**
Overall mean attitude toward science[1], γ_{00}	3.213	0.038	83.380	0.000
GRADE7, γ_{10}	-0.045	0.021	-2.096	0.036
GRADE8, γ_{20}	-0.132	0.047	-2.814	0.005
SCIENGRA, γ_{30}	0.020	0.009	2.044	0.041
READINGB, γ_{40}	0.219	0.025	8.764	0.000
INTERNET, γ_{50}	0.126	0.023	5.392	0.000
SHARINGI, γ_{60}	0.166	0.024	6.903	0.000
SELFEFFI, γ_{60}	0.166	0.020	7.946	0.000
MEANINGF, γ_{60}	0.437	0.022	19.035	0.000
ROTELEAR, γ_{60}	-0.310	0.021	-14.557	0.000

[1] The student level factors were Group Mean Centered before analysis.

The Grade-Attitudes toward science *slope* coefficients indicates that students from different grades had significantly different attitudes toward science. Students from seventh grades (γ_{10}= -0.045, se= .021) and eighth grades (γ_{20}= -0.132, se=.047) performed significantly lower than the students from sixth grades on students' attitudes toward science.

The Science grade- attitudes toward science slope coefficients (γ_{30}= .020, se= .009) indicates that students' science achievement is significantly and positively related to students' attitudes toward science. Students having higher achievement had better attitudes toward science than the other students.

The reading books - attitudes toward science slope coefficients (γ_{40}= .219, se= .025) indicates that the more the students read articles or books regarding science, the better attitude toward science the students had.

The Internet- attitudes toward science slope coefficients (γ_{50}= .126, se= .023) indicates that whether they benefited from internet sites regarding science is significantly and positively related to students' attitudes toward science.

The Students' sharing their ideas - attitudes toward science slope coefficients (γ_{80}= .166, se= .024) indicates that students' sharing their ideas about science subject with

their families was significantly and positively related to students' attitudes toward science.

The Self efficacy- attitudes toward science slope coefficients (γ_{90}= .166, se= .020) indicates that students' self-efficacy was significantly and positively related to students' attitudes toward science. Students having high self-efficacy had better attitudes toward science.

While the rote learning approach - attitudes toward science slope coefficients (γ_{90}= -.310, se= .021) indicates that students' rote learning approach was significantly and negatively related to students' attitudes toward science, the meaningful learning approach - attitudes toward science slope coefficients (γ_{80}= .437, se= .022) indicates that students' meaningful learning approach was significantly and positively related to students' attitudes toward science, Students having rote learning approach had lower attitudes toward science. Students having meaningful learning approach had higher attitudes toward science.

The final estimation of variance components obtained from random coefficient model is displayed in Table 9.

Table 9

Final Estimation of Variance Components for Random Coefficient Model for Attitudes toward Science

Random Effect	*Variance Component*	*df*	*Chi-square* χ^2	**p-value**
School mean, u_{0j}	0.10861	22	125.38959	0.000
Level-1 Effect, r_{ij}	0.54394			

Variance among the school means τ_{00}= 0.108 with a chi-square statistic of 125.389 is found to be statistically significant (p=.000).

The variances in the Analysis of Variances Model and Random Coefficient Model will be compared to calculate the variance explained at the student level. It can be compared by creating an index of the proportion of reduction in variance at the student level by comparing the σ^2 estimates from these two models.

Proportion of variance explained at

$$\text{level 1} = \frac{\sigma^2(\text{ANOVA}) - \sigma^2(\text{Random Coefficient})}{\sigma^2(\text{ANOVA})}$$

$$\text{Proportion of variance explained at level 1} = \frac{0.47801 - 0.29587}{0.47801} = 0.3810$$

By including these student level factors (seventh grade level, eight grade level, science achievement, whether they read articles or books regarding science, whether they watch documentaries, whether they benefit from internet sites regarding science, whether they share their ideas about science subject with their families, students' self-efficacy, students' meaningful learning and rote learning approach) as predictors of students' attitudes toward science within school variance was reduced by 38.1%. Therefore, these factors account for about 38% of the student level variance in attitudes toward science.

Discussion, Conclusion and Recommendation

This study provides a general overview about students' attitudes toward science and the predictive variables related to their attitude. Results of the One-Way ANOVA with random effects in HLM analysis revealed that there are significant differences in students' attitudes toward science among schools. These differences among schools in terms of students' attitudes could be resulted from many factors such as schools, classrooms, teachers, and students. In the present study factors regarding schools and students were examined and results were discussed. George (2000) found that 7th grade students in suburban schools had more positive attitudes toward science than students in metropolitan and rural schools. Also, Bozdogan and Yalcin (2005) indicated that attitudes of students studying in second type of schools (better educational opportunities, more teachers) were the highest. After that came the third type of schools (educational opportunities were the best, had the most teachers, were in the city center and were mostly preferred) and first type of schools (limited educational opportunities, less teachers and located in the countryside).

When the results of the study regarding school level variable were examined, it was found that while proportion of female science teachers, ability grouping between science classes, and quality of school's physical infrastructure were not significantly contributed to the model, the quality of school's educational resources significantly contributed to the students' attitudes toward science. In this study, instructional materials (eg. textbooks), science laboratory equipment and materials, computers for instruction, library materials, and audio-visual resources were regarded as quality of school's educational resources. Results revealed that the more educational resources the schools had, the more positive attitudes students developed towards science. When the students' attitudes were evaluated in terms of socio-economic status and parents' educational level it was found that parents' socio-economic status and their educational level did not significantly contribute to students' attitudes towards science. The results of the present study supported that independent from parents' socio-economic status and their educational level, the more quality of school's educational resources was, the more positive attitudes students developed towards science. Although the parents' socio-economic status is not so good, if the schools

improve their instructional materials (eg. textbooks), science laboratory equipment and materials, computers for instruction, library materials, and audio-visual resources, that is to say, if they create a rich and active learning environment, students could develop positive attitudes towards science. Hacieminoglu(2016) and Perera, Bomhoff and Lee (2014) supported the idea that, even if the students have a low socio-economic level, they can be encouraged by their parents to perform well if their families see the value of science and display a positive attitude toward science. Kozcu-Cakir, Şenler and Gocmen-Taskin (2007), Sinan, Sardag, Salifoglu, Cakir and Karabacak (2012), Sulun, Ekiz and Sulun (2009) stated that while students' attitudes towards science is positively related to frequency of having science lessons in the laboratory.

When the results of the study were examined in terms of students' characteristics variables, it was found that while gender do not significantly contribute to the model, grade level and science achievement significantly contribute to students' attitudes toward science. Results reflected that gender had no significant contribution to the model, namely there was no significant difference between boys and girls in their attitudes towards science. Our findings are similar to those of Dhindsa and Chung (2003), Kaya and Boyuk (2011), Kozcu Cakir, et al. (2007), Miller, Lietz and Kotte (2002), Sinan, et al. (2012) and Smist, Archambault and Owen (1994). On the other hand, our findings contradict the findings of Catsambis (1995), Hacieminoglu (2016), Jones, Howe, and Rua (2000), Piburn and Baker (1993) and Greenfield (1996) since these studies revealed that, male students' attitudes toward science was higher than that of the female students. Also, Gurkan and Gokce (2000) found that female students' attitudes towards science lesson were higher than male students. On the other hand, while there is a positive correlation between students' attitudes towards science and their science achievement, their attitudes are negatively correlated with their grade level. Namely, as the students' science achievement increases, their attitudes increase too. The findings of the present study are similar to those obtained by Hacieminoglu, (2016), Beaton, Mullis, Martin, Gonzalez, Kelly & Smith (1996); Kozcu Cakir et al. (2007); Sulun, Ekiz and Sulun (2009); Weinburgh, (1995). Similarly, Gurkan and Gokce (2000) found that students who were successful in science lessons had a positive attitude towards science. With respect to the grade level; as the students' class level increases, their attitudes towards science decreases. This could be resulted from the exams which students take to attend a prestigious high school. Students prefer to memorize the information while studying these multiple-choice exams and this might affect their attitudes towards science negatively. Related supporting literature, (Bozdogan and Yalcin, 2005; George, 2000; Kapici and Akcay , 2016; Kozcu Cakir, et al., 2007; Sinan, et al., 2012) often reveals that the attitude scores of students are higher in primary school, and that as the students get older their attitudes decrease relatively. On the other hand, in the study of Keles and Aydin (2017) results revealed that students' attitude scores only decrease as they get through from fifth to sixth grade, and that attitude scores stay almost the same for the other class levels. Also, Kaya and Boyuk (2011) stated that students' attitude scores towards science lesson differ in favor of 8th graders and that no difference is found for the other class levels.

Regarding to the factors related to student feelings and outside activities following variables contributed positively to students' attitudes towards science; whether they read articles or books regarding science, whether they benefited from the internet sites regarding science, whether they shared their ideas about science subject with their families. Similarly, Christidou (2011) suggested that web sites and TV programs about science had an effect on students' attitudes and interests towards science. Researcher stated that identifying the negative elements of books, magazines and TV programs about science by examining their contents positively affected students' attitudes towards science lesson (as cited in; Karacam, Mirza and Elitok, 2013). Also, Long and Steinke (1996) stated that TV programs about science affected students' attitudes towards science lesson (as cited in; Karacam, Mirza and Elitok, 2013). George and Kaplan (1998) examined the effect of parent and teacher variables on students' attitudes towards science. Results revealed that parents directly or indirectly affected their children's attitudes towards science. Researchers stated that parents directly affect their children's attitudes by talking to them about their lessons and they indirectly affect the attitudes by encouraging them to join activities or going museums and libraries (as cited in; Karacam, Mirza and Elitok, 2013). On the other hand, in this study watching documentaries had no significant contribution to the students' attitudes toward science. Karacam, Mirza and Elitok (2013) supported that students watching documentaries about science topics more often have more positive attitudes towards science lesson than students watching them rarely.

Learning and motivational factors were investigated in this study. They were performance and learning goal orientation, self-efficacy, meaningful learning approach and rote learning approach. From these variables while self-efficacy and meaningful learning approach contributed positively, rote learning approach contributed negatively to the students' attitudes towards science. Performance and learning goal orientation had no significant contribution to the model as supported by Hacieminoglu (2016). Similarly, Christidou (2011), George (2000) and Sinan, et al. (2012) also found that there was a positive correlation between students' attitudes towards science lesson and their self-efficacy perceptions. Students who believed in their abilities to perform well in science were more likely to have higher attitudes towards science than students who were less efficacious.

Recommendation

Since hierarchical data were used in relational screening model to get more reliable results, techniques which do not violate independence of observation hypothesis such as HLM must be used. As is seen in this study, different features related to different schools, classes and teachers could affect attitudes towards science differently. In this study, students' attitudes towards science differed among schools. When the school variable was examined, it was found that quality of school's educational resources significantly predicted students' attitudes towards science. Thus, improving instructional materials, science laboratory equipment and materials, computers for instruction, library materials, and audio-visual resources in all schools will have a

potential to contribute students' attitudes towards science. So, although families' socio-economic status is below average, students' attitudes towards science will be positively affected since the school will provide the materials that their families could not. Moreover, another finding supporting this one was that the following variables contributed positively to students' attitudes towards science; whether they read articles or books regarding science, whether they benefit from the internet sites regarding science, whether they share their ideas about science subject with their families. If the quality of school's resources are improved students can read articles and books regarding science and can do research by taking advantage of the internet sites regarding science in school's computer laboratory. Teachers should assign projects that lead students to do research regarding science topics. Teachers should raise the awareness of families about what they can do with their children, and with the help of parent-teacher association students' attitudes towards science should be improved. Teachers should increase self-confidence of students by valuing the projects and research done by them. They have to make students feel successful in the classroom environment and thus, by improving their students' self-efficacy, they can help their students have positive attitudes towards science. As the students' class level increases, their attitudes towards science decreases. This could be resulted from the exams which students take to attend a prestigious high school. Students prefer to memorize the information while studying these multiple-choice exams and this affects their attitudes towards science negatively. Therefore, teachers should arrange teaching learning environments in order to provide meaningful learning experiences, and they should prefer assessment and evaluation methods based on performance evaluation rather than multiple-choice exams. If teachers ask questions requiring memorization, students prefer to learn by memorizing. Therefore, teachers should ask questions to make students think about how to use knowledge both in the exams and in the teaching-learning process. So, students prefer meaningful learning to memorizing. There are many quantitative studies examining students' attitudes towards science. Enriching science lesson with materials, increasing students' curiosity, making them learn by discovering and making them do experimental observations affect their attitudes positively. In order to examine the reasons why students like science or not in detail, qualitative studies should also be conducted.

References

Akamca, G.Ö., & Hamurcu, H. (2005). Çoklu zeka kuramı tabanlı öğretimin öğrencilerin fen başarısı , tutumları ve hatırda tutma üzerindeki etkileri[The effects of multiple intelligence theory based teaching on students' science achievement, attitudes and retention]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 28,* 178-187.

Arisoy, N. (2007). Examining 8th grade students' perception of learning environment of science classrooms inrelation to motivational beliefs and attitudes. Unpublished Theses in Middle East Technical University, Ankara, Turkey

Azizoglu, N., & Cetin, G. (2009). The effect of learning style on middle schools students' motivation and attitudes towards science, and the relationships among these variables. *Kastamonu Education Journal, 17*(1), 171-182.

Beaton, A.E., Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., Kelly, D.L. & Smith, T.A. (1996). *Mathematics Achievement in the Middle School Years, Centre for the Study of Testing, Evaluation and Educational Policy.* Boston College, Boston.

BouJaoude, S. B. (1992). The relationship between high school students' learning strategies and the change in their misunderstandings during a high school chemistry course. *Journal of Research in Science Teaching, 29*(7), 687-699.

Bozdogan, A. E. & Yalcin, N. (2005). İlköğretim 6., 7. ve 8. sınıf öğrencilerinin fen bilgisi dersindeki fizik konularına karşı tutumları [Attitudes of 6th, 7th and 8th grade students towards physics subjects in science lesson]. *Gazi Üniversitesi Kırşehir Eğitim Fakültesi Dergisi, 6*(1), 241-247.

Burtless, T. G. (1996). Does money matter? The effect of school resources on student achievement and adult success. Washington, DC: Brookings Institution Press.

Catsambis, S. (1995). Gender, race, ethnicity, and science education in the middle grades. *Journal of Research in Science Teaching, 32*(3), 243-257.

Cavallo, A. M. L., & Schafer, L. E. (1994). Relationships between student's meaningful learning orientation and their understanding of genetic topics. *Journal of Research in Science Teaching, 31*(4), 393-418.

Cavallo, A. M. L., Rozman, M., & Potter, W. H. (2004). Gender differences in learning constructs, shifts in learning constructs, and their relationship to course achievement in a structured inquiry, yearlong collage physic course for life science majors. *School Science and Mathematics, 104*(6), 288-301.

Caliskan, I. S. (2004). *The effect of inquiry-based chemistry course on students' understanding of atom concept, learning approaches, motivation, self-efficacy, and epistemological beliefs.* Unpuplished master thesis, The Middle East Technical University, Ankara.

Christidou, V. (2011). Interest, attitudes and images related to science: Combining students' voices with the voices of school science, teachers, and popular science. *International Journal of Environmental & Science Education, 6*(2), 141-159

Celik, O. (2006). *Ağ tabanlı fen öğretiminin öğrencilerin problem çözme becerilerine ve fene yönelik tutumlarına etkisi[The effect of network based science teaching on students' problem solving skills and attitudes towards science].* Yayımlanmamış yüksek lisans tezi, Dokuz Eylül Üniversitesi, İzmir.

Cibik, A. S. (2009). Proje tabanlı öğrenme yaklaşımının öğrencilerin fen bilgisi dersine yönelik tutumlarına etkisi [The effect of project-based learning approach on students' attitudes towards science lesson]. *Elementary Education Online 8*(1), 36-47.

Dhindsa, H. S. & Chung, G. (2003). Attitudes and achievement of Bruneian science students. *Internalational Journal of Science Education, 25*, 907-922.

Fraenkel, J. R., & Wallen, N. E. (2003). How to design and evaluate research in education (5th ed.). Boston: McGraw Hill.

Fraser, J. B. (1978). Development of a test of science-related attitudes, *Science Education, 62*(4), 509-515.

Fuller, B. (1987). What school factors raise achievement in the third world? *Review of Educational Research, 57* (3), 255–292.

Genc, M. (2015). The effect of scientific studies on students' scientific literacy and attitude. *Ondokuz Mayis University Journal of Faculty of Education, 34*(1), 141-152.

George, R. (2000), Measuring change in students' attitudes toward science over time: an application of latent variable growth modeling. *Journal of Science Education and Technology*, (9), 213-225.

Gilmartin, S., Denson, N., Li, E., Bryant, A. & Aschbacher, A. (2007). Gender ratios in high school science departments: The effectof percent female faculty on multiple dimensions of students' science identities. *Journal of Research in Science Teaching, 44* (7), 980–1009.

Gomleksiz, M. N., & Fidan, E. K. (2013). Fen ve teknoloji dersinde bilgisayar destekli zihin haritası tekniğinin öğrencilerin akademik başarısına, tutumlarına ve kalıcılığa etkisi. *Gaziantep University Journal of Social Sciences, 12*(3), 403-426.

Greenfield, T. A. (1996). Gender, etnicity, science achievement and attitudes. *Journal of Research in Science Teaching, 33*, 901-933.

Gurkan,T., & Gokce, E. (2000). İlköğretim öğrencilerinin fen bilgisi dersine yönelik tutumları[Attitudes of primary school students towards science lesson]. *IV. Fen Bilimleri Eğitimi Kongresi Bildirileri*, 6-8 Eylül 2000, Hacettepe Üniveristesi. Ankara, ss. 188-192.

Hacieminoglu, E., Ertepinar, E., Yilmaz-Tuzun, O., & Cakir, H. (2015). Students and school characteristics related to elementary school students' nature of science views. Education 3-13: *International Journal of Primary, Elementary and Early Years Education,43*(6), 700-721. doi:10.1080/03004279.2013.865655

Hacieminoglu, E. (2016). Elementary school students' attitude toward science and related variables. *International Journal of Environmental & Science Education, 11*(2), 35-52.

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis,19* (2), 141–164.

Jones, G.M., Howe, A., & Rua, M. J. (2000). Gender differences in students' experiences, interests and attitude toward science and scientists. *Science Education, 84*, 180-192.

Kapici, H. O., & Akcay, H. (2016). Middle school students' attitudes toward science, scientists, science teachers and classes. *Asia- Pasific Forum on Science Learning and Teaching, 17*(1), 10.

Karacam, S., Mirza, Y., & Elitok, S. (2013). Fen konularına ilişkin belgesel izleme sıklığı ve cinsiyetin fen ve teknoloji dersine yönelik tutumlar üzerine etkisi[The effect of the frequency of watching documentary related to science subjects and gender on attitudes towards science and technology course]. *Journal of Duzce University Institute of Social Sciences, 3*(1),62-85.

Kaya, H. & Boyuk, U. (2011). İlköğretim II. kademe öğrencilerinin fen ve teknoloji dersine ve fen deneylerine karşı tutumları [Elementary school students' attitudes towards science and technology and science experiments]. *TÜBAV Bilim Dergisi, 4*(2), 120-130.

Keles P.U. & Oner A. (2016). Seçmeli bilim uygulamaları dersinin 7. sınıf öğrencilerinin fen tutumuna etkisi [The effect of elective course "science applications" on 7th grade students on attitudes towards science]. *Current Research in Education, 3* (1), 13-22.

Keles P.U. & Aydin S. (2017). The latitudinal analysis of secondary school students' attitudes to science course. *International Journal of Social Sciences and Education Research, 3*(3), 711-719.

Koc, A. & Bocek, U. (2012). Basit malzemelerle yapılan deneylerin fene yönelik tutuma etkisi [The effect of experimenting with simple materials on students' attitudes towards science]. *Türk Fen Eğitimi Dergisi 9*(4), 102-118.

Kozcu-Cakir, N., Senler, B. & Gocmen-Taskin, B. (2007).İlköğretim II.kademe öğrencilerinin fen bilgisi dersine yönelik tutumlarının belirlenmesi [Investigating the attitudes of elementary school students towards science lesson]. *Türk Eğitim Bilimleri Dergisi ,5*(4),637-655.

Kucuk, T. (2014) Işık ünitesinde simülasyon yönteminin kullanılmasının öğrencilerin fen başarısına ve fen tutumlarına etkisi [The effect of using simulation in light unit on students' science achievement and attitude toward science]. Yayımlanmamış yüksek lisans tezi, Çanakkale Onsekiz Mart Üniversitesi, Çanakkale.

Miller, L., Lietz, P., & Kotte, D. (2002). On decreasing gender differences and attitudinal changes. Factors influencing Australian and English pupils' choice of a career in science. *Psyhology, Evolution, and Gender, 4*, 69- 92.

Organisation for Economic Co-Operation and Development (2004). *Learning for tomorrow's world: First results from PISA 2003*. Paris: OECD Publications.

Piburn, M. D., & Baker, D.R. (1993). If I were the teacher…qualitative study of attitude toward science. *Science Education 77,* 393-406.

Perera, L. D. H., Bomhoff, E. J. & Lee, G. H.Y. (2014). *Parents' attitudes towards science and their children's science achievement.* Discussion Paper, ISSN 1441-5429, Monash University, Business and Economics, Department of Economics.

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical Linear Models: Applications and Data Analysis Methods. Thousand Oaks, CA: Sage.

Sakiz, G., Ozden, B., Aksu, D., & Simsek , Ö. (2014). Fen ve teknoloji dersinde akıllı tahta kullanımının öğrenci başarısına ve dersin işlenişine yönelik tutuma etkisi [The effect of use of smart boards in science and technology course on student achievement and attitude towards lesson]. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 18* (3), 257-274.

Sasmaz Oren, F., & Tezcan, R. (2009). İlköğretim 7. sınıf fen bilgisi dersinde öğrenme halkası yaklaşımının öğrencilerin tutumları üzerine etkisi [The effect of using learning cycle approach in 7th grade science course on students' attitudes]. *İlköğretim Online, 8*(1), 103-118.

Sinan, O., Sardag, M., Salifoglu, A., Cakir, C. & Karabacak, U. (2012). İlköğretim 5-8. sınıf öğrencilerinin fen tutumları ve öz yeterliliklerinin incelenmesi [Investigation on elementary 5th -8th grade students' attitudes towards science and self-efficacy]. *X. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi*, 27-30.06.2012, Niğde Üniversitesi, Niğde.

Smist, J. M., Archambault, F. X., & Owens, S. V. (1994). Gender difference in attitude toward science. Paper presented at the annual meeting of the American Educational Research Association (New Orleans, LA, USA).

Sulun, Y., Ekiz, S. O. & Sulun, A. (2009). Proje yarışmasının öğrencilerin fen ve teknoloji dersine olan tutumlarına etkisi ve öğretmen görüşleri [The effect of the project competition on students' attitudes towards science and technology course and teachers views]. *Erzincan Eğitim Fakültesi Dergisi, 11*(1), 75-94.

Teker, S., Kurt, M., & Karamustafaoglu, O. (2017). Işığın ve sesin yayılması ünitesini buluş yoluyla öğrenmenin öğrenci başarısı ve tutuma etkisi [The effect of inquiry based learning of light and sound unit on student achievement and attitude]. *Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 2017*(27), 835-863.

Weinburgh, M. (1995). Gender differences in student attitudes toward science: a metaanalysis of the literature from 1970 to 1991. *Journal of Research in Science Teaching, 32*, 387-398.

YildirimH. I., & Sensoy, Ö. (2016). Bilim şenliklerinin 6.sınıf öğrencilerinin fen bilimleri dersine yönelik tutumlarına etkisi [The effects of science festivities on 6th grade students' attitudes towards science course]. *Türk Eğitim Bilimleri Dergisi 14*(1), 23-40.

Unal, G., & Ergin, O. (2006). Buluş yoluyla fen öğretiminin öğrencilerin akademik başarılarına, öğrenme yaklaşımlarına ve tutumlarına etkisi [The effect of

inquiry based science teaching on students' academic achievement, learning approaches and attitudes]. *Türk Fen Eğitimi Dergisi, 3* (1), 36-52.

Yamak, H., Bulut , N., & Dundar, S., (2014). 5. sınıf öğrencilerinin bilimsel süreç becerileri ile fene karşı tutumlarına fetemm etkinliklerinin etkisi [The effect of fetemm activities on 5th grade students' attitudes towards science with their scientific process skills]. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, *34*(2), 249-265.

İlköğretim Öğrencilerinin Fen Bilimine Yönelik Tutumlarına İlişkin Öğrenci ve Okul Düzeyi Değişkenleri

Atıf:

Hacieminoglu, E. (2019). Student and school level variables related to elementary school students' attitudes towards science. *Eurasian Journal of Educational Research, 80, 59-88*, DOI: 10.14689/ejer.2019.80.4

Özet

Problem Durumu: Literatürde, öğrencilerin tutumları konusunda çok sayıda nicel çalışmalar vardır ve çalışmaların sonuçları, öğrencilerin önemli ve artan oranda fen bilimleri ile ilgilenmediğini ve fen bilimlerine yönelik negatif tutuma sahip olduğunu ortaya koymuştur. Pek çok öğrencinin, bilimsel araştırmaya devam etmelerini engelleyen, fen bilime karşı olumsuz duygu ve tutumları vardır. Eğitim ile ilgili araştırma verilerinde genellikle iç içe gruplanmış bir yapı vardır. Her öğrenci bazı okullarda veya sınıflarda gruplanmış olabilir. Bunun yanı sıra, bu okullar ya da sınıflar, il, bölge ya da ülke gibi başka herhangi bir lokasyonda gruplanmış olabilir. Bu hiyerarşik (aşamalı) veriler geleneksel doğrusal modelle analiz edilirse, özellikle gözlemin bağımsızlığı gibi temel varsayımların bir kısmı ihlal edilir. Veri aşamalı doğrusal modelleme yöntemi (HLM) kullanılarak analiz edildiğinde her grup kendi alt modelleri ile temsil edilmektedir. Bu alt modeller aynı seviyedeki değişkenler arasındaki ilişkilerle beraber, bu seviyedeki değişkenlerin diğer seviyedekilere olan etkisini de ortaya koymaktadır. Bu nedenle HLM, araştırmacılar için hiyerarşik yapıya sahip verilerdeki ilişkileri tanımlamak için daha güvenilir bir istatistiksel tekniktir. Öğrencilerin hem okul hem de öğrenci düzeyindeki özelliklerini göz önünde bulundurarak fen bilimine yönelik tutumlarını yordayan faktörler ile ilgili sınırlı sayıda çalışmalar bulunmaktadır.

Araştırmanın Amacı: Bu çalışmanın amacı öğrencilerin fen bilimlerine yönelik tutumlarını yordayan okul ve öğrenci düzeyindeki faktörleri belirlemektir.

Araştırmanın Yöntemi: Bu çalışmanın deseni temel olarak kesitsel ve korelasyonel bir araştırmadır. Bu çalışmada uygun örnekleme yöntemi kullanılmış ve bu çalışmanın

örneklemini Türkiye'nin farklı okul ve şehirlerindeki 2975 ilköğretim öğrencisi oluşturmuştur. Veri toplama araçları olarak fen bilimleri ile ilgili tutum ölçeği, öğrenme yaklaşımı ölçeği, başarı motivasyon ölçeği, ve okul özellikleri anketi kullanılmıştır. Veri analizinde aşamalı doğrusal modelleme(HLM) kullanılmıştır.

Araştırmanın Bulguları: Bu çalışma, öğrencilerin fen bilimlerine yönelik tutumları ve ilgili yordayıcı değişkenler hakkında genel bir bakış sunmaktadır. Aşamalı doğrusal modelleme sonuçları farklı okullardaki öğrencilerin fen bilimlerine yönelik tutumları anlamlı düzeyde farklılıklar gösterdiğini ortaya koymuştur. Öğrencilerin tutumları bakımından okullar arasında olan bu farkın okulla ilgili faktörler, sınıfla ilgili faktörler, öğretmenle ilgili faktörler, öğrencilerle ilgili faktörler gibi bir çok yordayıcısı olabilir. Bu çalışmada okul ve öğrenci düzeyindeki faktörler incelenmiş ve sonuçlar tartışılmaya çalışılmıştır. Çalışmanın sonuçları okul özellikleri ile ilgili değişkenler açısından incelendiğinde kadın fen öğretmenlerinin oranı, fen sınıfları arasında beceri gruplaması, okulun fiziki altyapısının kalitesinin modele anlamlı bir katkısı olmamakla birlikte, okulun eğitim kaynaklarının kalitesi, öğrencilerin fen bilimine yönelik tutumlarına anlamlı bir katkıda bulunmuştur.Bu çalışmada okuldaki eğitim kaynaklarının kalitesi olarak, öğretim materyalleri, fen laboratuar ekipmanı ve materyalleri, eğitim amaçlı bilgisayarlar, kütüphane materyalleri, görsel-işitsel kaynaklar ele alınmıştır. Çalışmanın sonuçları okulların eğitsel kaynakları ne kadar yeterli ise, öğrencilerin fen bilimine yönelik tutumlarının o kadar olumlu olduğunu göstermektedir.Öğrencilerin tutumları ailelerin sosyoekonomik durumu ve eğitim düzeyi, açısından değerlendirildiğinde ailelerin sosyoekonomik durumu ve ailelerin eğitim düzeyi, öğrencilerin fen bilimime yönelik tutumlarına anlamlı ölçüde katkıda bulunmamıştır. Çalışmanın sonuçları öğrenci özellikleri ile ilgili değişkenler bakımından incelendiğinde ise cinsiyetin modele anlamlı bir katkısının olmadığı, sınıf düzeyi ve fen başarısının ise öğrencilerin fen bilimlerine yönelik tutumlarına anlamlı düzeyde katkı sağladığı belirlenmiştir. Fen başarısı öğrencilerin fen bilimlerine yönelik tutumlarına pozitif düzeyde anlamlı katkı sağlarken, sınıf düzeyi değişkeni öğrencilerin fen bilimlerine yönelik tutumlarına negatif duzeyde anlamlı katkı sağlamıştır. Yani öğrenci başarısı arttıkça öğrenciler fen bilimlerine yönelik daha olumlu tutum sergilerken, sınıf düzeyi arttıkça öğrenciler fen bilimlerine yönelik negatif tutum göstermektedirler. Bunun yanısıra öğrenci duyguları ve okul dışı faaliyetlerle ilgili faktörler göz önünde bulundurulduğunda öğrencilerin fen bilimleri ile ilgili makaleler veya kitaplar okuyup okumadıkları, fen bilimleri ile ilgili internet sitelerinden yararlanıp yararlanmadıkları, fen bilimleri ile ilgili fikirlerini aileleri ile paylaşıp paylaşmadıkları.gibi değişkenler öğrencilerin fen bilimlerine karşı tutumlarına olumlu katkılar sağlanmıştır. Öğrencilerin öğrenme ve motivasyonla ilgili özellikleri incelendiğinde bu değişkenlerden, öz-yeterlik ve anlamlı öğrenme yaklaşımı öğrencilerin fen bilimlerine karşı tutumlarına olumlu düzeyde katkı sağlarken, ezbere öğrenme yaklaşımı ise öğrencilerin fen bilimlerine karşı tutumlarını negatif düzeyde yordamıştır.

Sonuç ve Öneriler

İlişkisel çalışmalarda daha güvenilir sonuçlar elde etmek için hiyerarşik yapıda veriler kullanıldığında HLM gibi gözlemlerin bağımsızlığı varsayımını ihlal etmeyen

teknikler kullanılmalıdır. Bu çalışmadada görüldüğü gibi farklı okullar sınıflar ve öğretmenler ile ilgili farklı özellikler öğrencilerin fen bilimlerine yönelik tutumunu farklı şekillerde etkileyebilir. Bu çalışmada öğrencilerin fen bilimlerine yönelik tutumları okullar arasında farklılık göstermektedir. Okul ve öğrenci özellikleri ile ilgili değişkenler öğrencilerin fen bilimlerine yönelik tutumlarını farklı düzeylerde yordadığı sonucuna ulaşılmıştır. Bu çalışmanın sonuçları göstermektedirki, öğrencilerin ailelerinin sosyoekonomik düzeyi ve eğitim düzeyinden bağımsız olarak okulların eğitim kaynaklarının kalitesi ne kadar gelişmiş olursa öğrenciler fen bilimlerine yönelik o kadar olumlu tutum geliştirmektedirler. Ailenin sosyoekonomik durumu çok iyi olmasa bile okullar öğretim materyalleri (ör. Ders kitapları), fen laboratuar ekipmanı ve materyalleri, eğitim amaçlı bilgisayarlar, kütüphane materyalleri, görsel-işitsel kaynaklarını arttırırlarsa ve geliştirirlerse, yani zengin ve aktif bir öğrenme ortamı oluştururlarsa öğrenciler fen bilimlerine yönelik olumlu tutum geliştirebilirler. Öğretmenler öğrencileri bilgisayar labaratuvarlarında fen bilimleri ile ilgili internet sitelerinden yararlanarak araştırmalarını yapabilmeleri konusunda cesaretlendirmeliler ve onlara araştırma yapmaya yönlendirecek projeler vermelidirler. Öğretmenler öğrencilerin kendilerinin yaptıkları araştırmalara ve projelere değer vererek onların özgüvenlerini arttırmalılar. Öğrencilerin sınıf seviyesi yükseldikçe fen e yönelik olumsuz tutuma sahip olmaktadırlar. Öğrenciler çoktan seçmeli soru tipinde olan liselere giris sınavlarına çalışırken bilgileri ezberleme yolunu tercih etmektedirler, bu da fen e yönelik tutumu olumsuz bir şekilde etkilemektedir. Fen Bilimleri programımız araştırma sorgulama temelli öğretimi temel almaktadır. Bu nedenle öğretmenlerimiz öğrenciyi anlamlı öğrenmeye sevk eden, ezbere öğrenme den uzaklaştıracak şekilde öğretim tasarlamaya özen göstermeli, ve çoktan seçmeli sınavlardan ziyade performans değerlendirmeyi temel alan ölçme değerlendirme yöntemlerini tercih etmelidirler. Öğretmen hem sınavlarda hemde öğretim sürecindeki sorularında bilgiyi bilmek yerine bilgiyi nasıl kullanacağını bileceği mantıksal düşünme yeteneğini geliştirecek şekilde sorular sormalıki hem öğrenci anlamlı öğrenmeyi tercih etsin, hemde ezberleyerek öğrenmekten uzaklaşsın. Öğrencilerin fen bilimlerine yönelik tutumlarını inceleyen çok sayıda nicel araştırma mevcuttur, öğrencilerin feni sevmelerinin veya sevmemelerinin nedenlerini daha detaylı inceleyebilmek için bu konuda nitel araştırmalar da yapılmalıdır.

Eurasian Journal of Educational Research 80 (2019) 89-112

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

Effect of Cooperative Learning Supported by Reflective Thinking Activities on Students' Critical Thinking Skills*

Fatma ERDOGAN[1]

ARTICLE INFO

Article History:
Received: 11 Nov. 2018
Received in revised form: 17 Mar. 2019
Accepted: 23 Mar. 2019
DOI: 10.14689/ejer.2019.80.5

Keywords
Mathematics education, higher order thinking skills, jigsaw technic, middle school.

ABSTRACT

Purpose: It is assumed that cooperative work and critical thinking skills will come into prominence in 2020. In this context, the aim of this study is to examine the effect of cooperative learning supported by reflective thinking activities on seventh grade students' critical thinking skills during mathematics course.

Method: In this study, a quasi-experimental model with pretest-posttest control group was applied. In the experimental group, cooperative learning method supported by reflective thinking activities was applied. In the control group, mathematics teaching was carried out in accordance with the curriculum of the mathematics course. The study group was composed of 70 seventh grade students. Cornell Critical Thinking Test, Level X was used as the data collection tool. Dependent and independent samples t-tests were used in data analysis, and ANCOVA was applied to determine the difference between the post-tests scores of the groups.

Findings: In the study, when the pre-test scores of the experimental and control groups were checked, a significant difference was found between the corrected CCT-X post-test mean scores. This difference was found to be in favor of the experimental group. Based on this finding, cooperative learning supported by reflective thinking activities can be said to have a positive effect on students' critical thinking skills.

Implications for Research and Practice: In future research, the effects of different reflective thinking strategies on critical thinking skills can be examined in cooperative learning environment, and their advantages and disadvantages can be discussed. Student's critical thinking skills can be analyzed by qualitative methods.

* This study was partly presented at the 5th International Eurasian Educational Research Congress in Antalya, 02-05 May, 2018.

[1] Fırat University, TURKEY, e-mail: f.erdogan@firat.edu.tr, ORCID: https://orcid.org/0000-0002-4498-8634

Introduction

Today, individuals are expected to have strong interpersonal skills (Organisation for Economic Co-operation and Development [OECD), 2017). In today's conditions, cooperative learning (CL) plays a role in helping individuals acquire the desired skills. The CL can be defined as a learning method in which students with a common purpose work together in small groups, where each group member is responsible for the learning of other members (Johnson, Johnson & Smith, 2014). CL is one of the most important applications in mathematics education (Parveen, Yousuf & Mustafa, 2017). According to the research results, CL has a positive effect on students' mathematics achievement (Asha & Al Hawi, 2016; Cumhur & Elmas-Baydar, 2017; Parveen et al., 2017), communication and social skills (Pandya, 2017), mathematical understanding, and logical inference skills (Asha & Al Hawi, 2016). CL has many benefits in the field of mathematics education.

It can be said that the skills expected from the students have changed according to today's conditions in CL and problem solving process in mathematics education. In this context, the OECD (2017) has combined its ability to monitore and reflect with cooperative problem-solving competencies. Dewey (1933) emphasizes the importance of cooperation in the reflection process. Therefore, reflective thinking (RT) is a concept that needs to be addressed in the CL process.

Reflection is the ability of the student to present a subject or a problem state, and to present his / her own thoughts, attitudes, knowledge and abilities (Schön, 1987). Dewey (1933) states that learning consists of reflections on experiences. RT is an important thinking skill in terms of mathematics education (Kramarski, Weiss & Sharon, 2013; McNaught, 2010). Mathematical learning process requires to build up interrelation between concepts, strategy selection and reflection (Kramarski et al., 2013). The results of the research show that RT supports meaningful learning in mathematics (Inoue & Buczynski, 2011; McNaught, 2010) and provides students to reach correct and logical solutions (Agustan, Juniati & Siswono, 2016).

In addition to the importance of RT in mathematics education, some strategies are important in the process of providing students to reflect. One of the strategies used in the development of RT skills is "Writing". Writing is an effective tool for students to express, explore, organize, and reflect ideas about mathematical content and process (Freeman, Higgins & Horney, 2016; Inoue & Buczynski, 2011; Suhaimi, Shahrill, Tengah & Abbas, 2016). With "Journal writing", which is one of the writing activities, students can reflect on their experiences, points they are strong or weak, important points of an event, how they deal with a power situation, and their feelings (Farrah, 2012; Guce, 2017, 2018; Mitchell & Coltrinari, 2001). Students are able to discuss with their group members and evaluate their own activities with the "Group discussion" strategy that transforms reflective activities into a social activity (Aldahmash, Alshmrani & Almufti, 2017; Kohen & Kramarski, 2012). Students focus on their own thinking processes, processes, activities through questions that guide peer interaction and reflection in the "Reflective dialogue"

strategy (Kohen & Kramarski, 2012; Wille, 2017). Furthermore, if a student is guided to be present in the "Self-evaluation", he/she asks himself/herself questions, is trying to understand how he/she learns, and can think about these solutions by being aware of his/her strengths and weaknesses. (Agustan et al., 2016). The students evaluate their mental processes and reflect on their thoughts audibly at the end of the evaluation with "Thinking aloud" strategy (Taggart & Wilson, 2005).

When the literature is examined, it is seen that there are studies in which different strategies are applied in order to improve RT skills of students in the field of mathematics education (Agustan et al., 2016; Freeman et al., 2016; Guce, 2017; Kramarski et al., 2013; McNaught, 2010; Suhaimi et al., 2016). However, in mathematics education in Turkey, there is a limited number of studies on improving the skills of RT (Kizilkaya, 2009). Therefore, in Turkey, studies on the RT skills are inadequate, it is necessary and important for the literature to carry out studies in which RT activities are applied in mathematics education.

One of the concepts related to both CL and TR is critical thinking. Critical thinking is reflective and logical thinking, focused on deciding what to do and what to believe (Ennis, 1996). In addition, critical thinking is the judgmental problem-solving process aimed at improving knowledge (Tiruneh, Verburgh & Elen, 2014). Critical thinking related to mathematical skills such as problem solving, questioning, analysis is an important part of mathematics education (Palinussa, 2013; Sumarna, Wahyudin & Herman, 2017; Su, Ricci & Mnatsakanian, 2016). Due to the nature of critical thinking, critical thinking requires reflection and sociability (Choy & Oo, 2012).

In the studies, it has been reported that CL is a method that develops students' critical thinking skills (Garcha & Kumar, 2015; Loes & Pascarella, 2017; Quines, 2017; Tiruneh et al., 2014; Silva et al., 2016). Critical thinking is also associated with RT (Ghanizadeh, 2017). In this context, a reflective thinker also has critical thinking skills (Evin-Gencel & Guzel-Candan, 2014; Tican & Taspinar, 2015). Studies show that students' activities to improve their skills of RT have positive effects on critical thinking skills (Aryani, Rais & Wirawan, 2017; Farrah, 2012). Considering research carried out in Turkey, few studies have been found to examine reflective and critical thinking levels of pre-service teachers and examine the effect of RT activities on critical thinking skills, (Demir, 2015; Evin-Gencel & Guzel-Candan, 2014; Tican & Taspinar, 2015)

With the adoption of the constructivist approach in Turkey, skills such as cooperation, reflective thinking, and critical thinking are included in mathematics curriculum. However, according to results in PISA 2015, Turkey ranks last among OECD countries in cooperative problem-solving skills (OECD, 2017). Therefore, in order to reach the objectives of the mathematics curriculum, it is necessary to conduct more comprehensive and different activities related to CL in the field of mathematics education. It is thought that the present research will contribute to the

literature in terms of presenting the RT activities that can be applied in CL environments.

In today's world, changes continue without slowing down. It is assumed that cooperative work and critical thinking skills will come into prominence in 2020 in the knowledge of the World Economic Form (Gray, 2016). However, effectiveness of CL on learning outcomes is discussed. A number of studies have shown that CL does not bring significant cognitive, social, and affective gains for students (Berkun & Ada, 2017; Souvignier & Kronenberger, 2007). In this context, research has been conducted to examine CL under different forms. In some studies, CL was supported by multiple intelligence (Isik & Tarim, 2009), metacognitive strategies (Mevarech & Amrany, 2008), inquiry strategy (Souvignier & Kronenberger, 2007) or problem-solving strategies (Yazlik & Erdogan, 2016). In these previous studies, unstructured CL techniques were used, and the cooperative group structures were not discussed in detail. In this context, it can be said that more research is needed on the effectiveness of CL.

In recent years, it can be seen that CL and RT are the concepts discussed together. RT skills are combined with cooperative problem-solving competencies (OECD, 2017). In studies, it was stated that as a result of students' reflections in CL environments, their skills such as problem solving, questioning, linking old and new information, making plans and strategies, and self-regulation skills were developed (Applefield, Huber & Moallem, 2000; Bransford, Brown & Cocking, 2000; Gagnon & Collay, 2006; King, Goodson & Rohani, 2013; Lan, 2007; Kramarski & Kohen, 2017). In the studies that analyzed the interrelationship of CL and RT skills mentioned above, the importance of reflection in CL environments was emphasized while the cooperative group structure was not explained in detail. In these studies, the use of structured CL techniques and the integration and implementation of different strategies to improve skills of RT have been ignored. Despite CL and work together to address the RT concept abroad, the number of research carried out in this field in Turkey (Guvenc, 2011) is quite limited. It is considered important to investigate the outcomes of CL's support in the teaching environment with RT strategies.

Considering the research conducted in Turkey in the field of mathematics education; It will be seen that the number of studies on CL, RT and critical thinking concepts is quite low. This research is specific for the use of a structured CL technique, for describing the use of different RT strategies in the CL environment, and for the detailed presentation of the materials used in the implementation of strategies. This study differs from previous studies in terms of CL structure of previous studies, and in terms of integrating RT strategies. Therefore, this study is thought to provide a different perspective on the effectiveness of CL. In accordance with the stated reasons, the aim of this study is to examine the effect of CL supported by RT activities on seventh grade students' critical thinking skills during mathematics course.

Method

Research Design

In this study, a quasi-experimental model with pretest-posttest control group was applied. In the quasi-experimental model, due to the difficulty of artificially forming groups, paired groups are randomly assigned as experimental groups (Fraenkel & Wallen, 2006). In this study, there is an experiment and a control group. In the experimental group, CL method supported by RT activities was applied. In the control group, no special teaching method was applied. An appropriate mathematics instruction was conducted in accordance with the current mathematics curriculum.

Study Group

This study was carried out during the academic year 2016-2017, with a total of 70 students in the seventh grade in a secondary school located in Turkey's Eastern Anatolia province. A random method was adopted to determine the experimental and control groups. There were 36 students in the experimental group and 34 in the control group. While 21 of the students in the experimental group were females (58%) and 15 of them were males (42%), 20 of the students in the control group were females (59%) and 14 of them were males (41%). The number of female students in both groups was higher than that of male students. It can be stated that the number of students in two study groups is quite close to each other.

Data Collection Tool

Cornell Critical Thinking Test, Level X (CCT-X) was used as the data collection tool. This test was developed by Ennis and Millman (1985) and adapted to Turkish by Mecit (2006). CCT-X, which is one of the most widely used tests for measuring critical thinking skills at elementary level all over the world, is a three-choice multiple-choice measurement tool consisting of 72 items in total. In the test with the "Yes, No, Maybe" options, each question has only one correct answer. A maximum of 72 points can be obtained from the test. The general Cronbach Alpha coefficient of the scale was calculated as .75 by Mecit (2006), and it was calculated as .77 in this study.

Procedures

The experimental phase of the study lasted 25 class hours. Considering the awareness program and the pretest-posttest implementation periods for the recognition of strategies and materials, this study was completed within a total of 35 class hours. The intervention in the experimental and control groups were carried out by the researcher. The same problems were studied in the groups and the studies in the groups were started and completed in parallel time periods.

The research stages were carried out by determining the strategies of RT and material preparation, pilot application, awareness program, and implementation of RT activities in CL groups. In the process of designing RT activities, firstly literature review in the field of CL, RT, and mathematics education was done.

Strategies that are found successful in previous studies are examined (Colley, Bilics & Lerch, 2012; Kramarski et al., 2013; Lan, 2007; Taggart & Wilson, 2005). In the present study, writing, journal writing, group discussion, reflective dialogue, self-evaluation and thinking aloud strategies, which are frequently used in mathematics education, directing students to thinking, discussing and questioning, developing problem solving skills were used.

After the determination of the strategies, materials were designed by the researcher in order to construct the implementation of these strategies. It was aimed to successfully manage the CL process and group dynamics through structured materials. The results of successful research were examined in the editing of RT materials (Brockbank & McGill, 2006; Kohen & Kramarski, 2012; Lan, 2007; Mevarech & Kramarski, 2014; Michalsky & Kramarski; 2015; Mitchell & Coltrinari, 2001; Taggart & Wilson, 2005; Wilson & Jan, 1993). In this direction, journal, group discussion form, reflective dialogue form, self-evaluation form materials, which would be used in the implementation process of RT activities were designed. In order to discuss RT strategies and materials, opinions of the experts and mathematics teachers who took mathematics education were taken, and necessary arrangements were made. The RT strategies, activities and materials applied in the experimental group are described below.

Writing. In the awareness program, students were told that they had to write everything they had learned and everything passed through their minds. During the experimental implementation process, students were directed to write on the study papers about each situation they thought and did during the activities. Materials such as journal, group discussion form and self-evaluation form, which were developed to provide reflection of students inside and outside the classroom, were structured. Thus, it was aimed to develop critical thinking skills of students with writing activities in CL environment. Writing strategy is a general strategy used throughout the entire experimental implementation.

Journal writing. Journal writing strategy was used in order to make students remember what they had learned in the course, to review their experiences and to make reflections by self-evaluation. When writing the jornal, steps to be followed were based on the work of Mitchell and Coltrinari (2001) and "journal" material was designed (Appendix, 1). At the end of each math course, journals written at home were examined by the teacher, and the students were provided with feedback.

Group discussion. With this strategy, it was aimed to create new thoughts with the interaction of the community and to reflect these thoughts in the learning process. It is important for group discussions to be carried out with well-structured activities and the questions that lead students to the reflection process. In this context, students were directed to make reflective discussions as a group, to answer questions of what and why. The "group discussion form" (Appendix, 2) was prepared based on research using questions that led students to reflective inquires (Kohen & Kramarski, 2012; Lan, 2007; Mevarech & Kramarski, 2014). At the end of

the discussions, students were asked to write their common ideas as a group on the group discussion form.

Reflective dialogue. With this strategy, students were required to review and question what they have learned as a result of their dialogue with their peers. In the study, a structured "reflective dialogue form" including CL principles was developed based on the study of Brockbank and McGill (2006) (Appendix, 3). Reflective dialogue form was used as a guide to ask students task-oriented questions. Paired students asked questions in the form to each other and answered them aloud.

Self-evaluation. With self-evaluation strategy, it was aimed to improve students' ability to criticize the learning process with a critical approach, to raise awareness of deficiencies and mistakes, to correct them if there are deficiencies and errors, and to improve their monitoring skills. In this direction, "self-evaluation form" (Appendix, 4) was developed based on CL principles and type of reflective prompt (generic, judgment or modification) that Michalsky and Kramarski (2015) stated in their research.

Thinking aloud. In this study, students were asked to think aloud during reflective dialogue process or during group discussions with their friends. Students were directed to interact with their friends openly for their explanations.

The contents of the activity worksheets were based on the seventh grade ratio and proportion sub-learning area attainments. The study included 19 eighth grade students, who were not included in the scope of the study group, who had the previous year's ratio and proportion sub-learning attainments. Pilot training was conducted with these students during 15 class hours. Students were given RT materials, the meaning of questions and expressions, and how to use the materials were explained. Students were informed about the Jigsaw-I technique and the strategies they would use. Students actively participated in the RT strategies and used materials. After the pilot implementation, the statements were revised in terms of functionality, student feedback, and clarity of the materials and activities.

Following the application of the pretest, six hours of awareness program was conducted in the experimental group. In the awareness program, firstly information about the implementation of CL, the Jigsaw-I technique, teacher and student roles were presented. Next, the strategies and materials to be implemented were explained in detail. In this context, information about the purpose and duration of the study was given and the questions of the students were answered by introducing the materials.

Following the awareness program, CL method, which was supported by RT activities, was applied in the experimental group. Throughout the experimental process, the teacher played a guiding role in the research. In the study, Jigsaw-I technique, one of the CL techniques, was applied by taking into consideration the steps mentioned in the literatüre (Aronson & Patnoe, 1997; Souvignier &

Kronenberger, 2007; Un-Acikgoz, 2011). The application stages of the Jigsaw-I technique and the strategies and materials used in these stages were announced.

Creation of main groups. In the first phase, students were divided into heterogeneous groups taking into account the gender and the average of the previous year's mathematics course. Each group was assigned a letter (A, B, C, D, E, F), consisting of six groups of six students. The group names of these main groups were determined by asking the students. Ratio and proportion sub-learning attainments were grouped according to the number of group members. The teacher asked the group members to share headings from each group. Students in the group were coded according to the topics they took (eg, students in group A, A1, A2, A3, A4, A5, A6). The subjects were distributed to all students in the same coded group with the same topic (For example, students with A1, B1, C1, D1, E1, F1 studied the same subject). The students were given worksheets on the subject and were asked to get prepared for the topic. The students were asked to write their journal at home.

Expertise. In the following process, in each group, groups of experts formed by gathering the same subject and having the same code were determined. Students in the expert groups were allowed to work together with their friends on their specialization topics. First of all, students in expert groups worked in the CL environment without an RT strategy. Then, the process was supported by RT strategies to enable students to collaborate, interact and reflect. With the reflective dialogue strategy, students were able to answer the questions in a reflective dialogue form. Thus, students were directed to share their thoughts, to reflect their ideas and to help each other.

Also, students were asked to work with group discussion strategy. Students discussed the questions in the group discussion form by thinking aloud, and they wrote and prepared a joint report on the topic to be told to their friends in the main groups. Thus, all groups were taught in the same way according to a common report and missing learning was prevented. In the next course hour, students were asked to complete self-evaluation form to evaluate their own learning. At the end of each mathematics course, students were asked to write their journals at home.

Consolidation. In the expert groups, students working on their subjects returned to their main groups and told their group members in order. At this stage, students asked questions to their fellow students about the topic, and they discussed by thinking aloud. At the end of each topic, students working with the group discussion strategy completed the group discussion form and the individual self-evaluation form. Also, at the end of each math course, students wrote their journals at home.

Completion and evaluation. At this stage, a different sub-topic was presented to each of the main groups in order to integrate learning. After the group presentations, whole topic was summarized by the teacher. In the group presentations, the first three successful groups were given a certificate of achievement.

In the control group; mathematics teaching was conducted under the guidance of current mathematics curriculum. Students in the control group solved the same problems with the experimental group. The teacher gave a presentation with the guidance of the mathematics curriculum. During the course, the same problems used in the experimental group were solved by the teacher, or the teacher directed these problems to the students by question-answer method. At the end of the course, students were asked questions, concepts or problems they could not solve, in order to complete the missing learning of the students. Difficulties were solved on the board by the teacher or another student. At the end of the topic, topic summary was made by the teacher or by the students. Therefore, it can be said that the control group had a different learning environment than the experimental group.

Data Analysis

In the data analysis, the Shapiro-Wilks test was applied to analyze the normal distribution. When analyzing the data, independent group t-test and dependent group t-test were used, depending on the data type. Covariance analysis (ANCOVA) was performed to examine the difference between the posttest scores of the groups. The Bonferroni test was used to determine the source of the difference between the adjusted scores. The significance level was taken as $p < .05$ in the research process.

Results

In this section, the findings obtained from the CCT-X pre and posttest scores of the experimental and control groups were presented. In the analysis of the tests, normal distribution of scores was analyzed by the Shapiro-Wilks test. The results were presented in Table 1.

Table 1

Shapiro-Wilks Normality Results Regarding CCT-X

Group	Pre-test		Post- test	
	Shapiro-Wilks	p	Shapiro-Wilks	p
Experimental	.96	.22	.97	.33
Control	.97	.59	.98	.79

In Table 1, Shapiro-Wilk Test normality results, CCT-X pre-test scores (w= .96, p= .22> .05) and CCT-X post-test scores (w= .97, p=. 33> .05) were given. Shapiro-Wilk Test normality results of the control group were presented for the pre-test scores of CCT-X; (w= .97, p= .59> .05) and for CCT-X post-test scores; (w= .98, p= .79> .05). Based on these findings, it was determined that the test scores showed normal distribution. Before RT activities in the experimental group, CCT-X pre-test scores of the study groups were analyzed by independent group t-test, and the results were shown in Table 2.

Table 2

Independent Group T-Test Results for Comparing CCT-X Pre-Test Scores

Group	N	$\bar{x}$	sd	df	t	p
Experimental	36	27.36	8.13	68	.02	.98
Control	34	27.32	6.05			

When Table 2 was analyzed, there was no statistically significant difference between the CCT-X pre-test scores of the experimental and control groups [$t_{(68)}$= .02, p> .05]. In the light of these findings, it can be said that critical thinking skills of the students in the experimental and control groups were equal before the experimental procedures. In the study, CCT-X pre- and post-test scores were analyzed with dependent group t-test after CL process supported by RT activities in the experimental group, and the results were given in Table 3.

Table 3

Dependent Group T-Test Results for the CCT-X Pre- and Post-Test Scores of Experimental Group

Test	N	$\bar{x}$	sd	df	t	p
Pre-test	36	27.36	8.13	35	-6.05	.00
Post-test	36	33.97	7.79			

According to Table 3, the CCT-X post-test mean scores of the experimental group was higher than the pre-test mean scores. As a result of the dependent group t-test analysis for the CCT-X pre- and post-test scores, there was a statistically significant difference [$t_{(35)}$=-6.05, p< .05]. The resulting significant difference was in favor of the post-test. After the teaching practices in the control group, the CCT-X pre- and post-test scores were analyzed by the dependent group t-test, and the results were presented in Table 4.

Table 4

Dependent Group T-Test Results for the CCT-X Pre- and Post-Test Scores of Control Group

Test	N	$\bar{x}$	sd	df	t	p
Pre-test	34	27.32	6.05	33	-1.44	.16
Post-test	34	28.41	6.47			

When Table 4 was examined, it was found that the CCT-X pre-test and post-test mean scores of the control group were quite close to each other. According to the results of the dependent group t-test for the CCT-X pre- and post-test scores of the control group, no statistically significant difference was found [$t_{(33)}$= -1.44, p> .05].

ANCOVA was applied to determine whether the CCT-X post-test mean scores differed for the experimental and control groups after the initial CCT-X pre-test

scores were controlled. Before applying ANCOVA, ANCOVA assumptions were examined. The first assumption was the normal distribution of the data with the Shapiro-Wilks test. The findings of these data were given in Table 1. As can be seen from Table 1, it was found that CCT-X pre-test and post-test scores had a normal distribution. In order to control the hypothesis of linear relationship between linearity, dependent variable (CCT-X post-test scores) and covariate (CCT-X pre-test scores), the overall distribution of scores was examined by scatter plot for each of the groups. For each group, a linear relationship was found between the dependent variable and the covariate. Then, the assumption of homogeneity of regression slopes, which was one of the main assumptions of ANCOVA, was investigated. These results are shown in Table 5.

Table 5

Regression Slope for CCT-X

Source	Sum of squares	df	Mean square	F	p
Corrected model	2252.15	3	750.72	27.65	.00
Intercept	531.65	1	531.65	19.58	.00
Group	103.97	1	103.97	3.83	.06
Pre-test	1641.93	1	1641.93	60.48	.00
Group*Pre-test	23.04	1	23.04	.85	.36
Error	1791.69	66	27.15		
Total	72497.00	70			
Corrected total	4043.84	69			

As can be seen from Table 5, it was determined that covariate and post-test scores did not show a statistically significant interaction [$F_{(1,66)}$= .85, p= .36> .05]. Based on this finding, it was seen that the assumption of homogeneity of the regression slopes was ensured. In addition, in order to examine the assumption of homogeneity of variances, Levene's Test was applied and it was determined that the variances between the groups were homogeneous (F=2.03, p= .16> .05). According to the findings, ANCOVA assumptions were obtained. Based on these results, to control the pre-test scores of the groups, ANCOVA was applied to determine whether there was a significant difference between the corrected post-test scores. These results are presented in Table 6.

Table 6

ANCOVA Results for Corrected CCT-X Post-Test Scores

Source of variance	Sum of squares	df	Mean square	F	p
Pre-test (Reg.)	1688.48	1	1688.48	62.34	.00
Group	535.58	1	535.58	19.77	.00
Error	1814.73	67	27.09		
Total	72497.00	70			

According to Table 6, the CCT-X pre-test scores of the experimental and control groups were checked and a significant difference was found between the corrected CCT-X post-test mean scores [$F_{(1,67)}$=19.77, p <.05]. This difference was found to be in favor of the experimental group. When the Bonferroni test results for the corrected CCT-X post-test mean scores were examined, critical thinking skills of students in the experimental group ($\bar{x}$= 33.96) were found to be higher than those of the students in the control group ($\bar{x}$= 28.43).

Discussion, Conclusion and Recommendations

According to the findings of the study, there was no statistically significant difference between the CCT-X pre-test mean scores of the students in the experimental and control groups. Based on this finding, it can be said that critical thinking skills of the students in the study groups were equal before the experimental applications in the experimental group.

In the study, after the experimental implementations, it was concluded that the CCT-X post-test mean score of the experimental group was significantly higher than the pre-test mean score. However, no significant difference was found between the CCT-X pre-test and post-test mean scores of the control group. Based on this finding, it can be said that the mathematics courses carried out in the control group did not contribute to the development of students' critical thinking skills. Based on the findings of the research, in addition to stressing that students have critical thinking skills in mathematics curriculum (Ministry of National Education [MoNE], 2018), it can be said that structured activities for the development of these skills should be applied in classroom environments.

Applefield et al. (2000) stated that, in the context of social constructivist vision, creating knowledge and skills required in students is based on interpersonal interaction. In cooperative environments based on social constructivism, students who have discussions and reflect on their learning develop their critical perspectives. Research findings of Palinussa (2013) also support this result. Palinussa (2013) stated that students' critical thinking skills were influenced by classroom learning environment. Therefore, it is stated that it is insufficient to emphasize these skills in the curricula only in order to ensure that students are

developed as critical thinkers. The need for a supportive atmosphere for the development of skills in the classroom is also indicated. Chukwuyenum (2013) showing a similar result to Palinussa (2013) emphasizes that critical thinking skills should be taught by reflecting in social interaction and by integrating them into mathematics course.

Another important finding in the research; when the pre-test scores of the experimental and control groups were checked, is that a significant difference was found between the corrected CCT-X post-test mean scores. This difference was found to be in favor of the experimental group. Based on this finding, CL supported by RT activities can be said to have a positive effect on students' critical thinking skills.

It supports Colley et al. (2012) who reported that reflective learning environments are created on the basis of educating individuals who think critically. This finding of the research is consistent with the results of the research examining the effects of providing cooperative environments on critical thinking skills (Adams, 2013; Gagnon & Collay, 2006; Gillies, 2006; Gorlewski & Greene, 2011; Guce, 2017; Lucena & San Jose, 2016; Mitchell & Coltrinari, 2001; Parsons & Stephenson, 2005; Silva et al., 2016; Webb & Farivar, 1994). Webb and Farivar (1994) reported that students' inadequately developed communication skills could adversely affect their use of CL, while Gillies (2006) emphasized that students who benefit most from CL are benefited from guiding support areas. It was determined that critical perspectives of the students who analyzed peers and gave feedback to each other, analyzed their thoughts, interacted with their peers, and made statements to their peers in CL groups were developed during the study. In a number of studies; it was stated that classroom discussion, CL, journal writing, dialogue, and reflection were influential factors in the development of critical thinking skills of teaching environments (Adams,2013; Gorlewski & Greene, 2011; Guce, 2017; Mitchell & Coltrinari, 2001). Gagnon and Collay (2006) stated that students have criticized the meaning they have formed together and their own thinking process by both group reflection and by making individual projections in CL groups. Considering the research findings in the literature, it is observed that students who develop interactive thinking and reflections in CL groups develop their critical thinking skills.

In the study, it was thought that writing and journal writing strategies were effective in the development of critical thinking skills of students in the CL group supported by RT activities more than the students in the control group. Previous studies have supported this conclusion. In the studies, it was stated that students explain their reasoning, correctness of solutions, thinking processes by writing and they made reflections about mathematics learning process in mathematics courses (Freeman et al., 2016; Inoue & Buczynski, 2011). Suhaimi et al. (2016) stated that students who write journal in mathematics courses organize their thoughts, thus journal writing improves communication and critical thinking skills of the students. In addition, it can be said that supporting CL with group discussion strategy is one of the factors that improve students' critical thinking skills. King et

al. (2013) and Gibson (2008) have obtained findings that support this view in their studies. They reported that well organized and managed group discussions developed students' critical thinking skills in their studies.

It can be said that reflective dialogue strategy used in this research contributed to the development of students' critical thinking skills. In research supporting this view, it was determined that students questioned their knowledge and focused on thinking about the activities they reflected critically on the learning process as a result of peer dialogues (Brockbank & McGill, 2006; Kohen & Kramarski, 2012; Wille, 2017). In addition, it is emphasized in researches that verbal language skills about explaining and asking questions (Souvignier & Kronenberger, 2007) and how to communicate in CL groups should be taught (Gillies, 2006). With the reflective dialogue strategy, it can be said that questioning and reasonable thinking skills developed within the scope of critical thinking,

In the study, critical thinking skills can be said to develop in a positive way as a result of self-evaluation. Previous research findings support this view (Mevarech & Kramarski, 2014; Michalsky & Kramarski, 2015; Wilson & Jan, 1993). At the end of the learning activities with self-evaluation strategy, students can learn to think critically (Michalsky & Kramarski, 2015; Wilson & Jan, 1993).

Used in research; writing, journal writing, reflective dialogue, group discussion, self-evaluation, and thinking aloud strategies were used alone in previous research (Agustan et al., 2016; Aldahmash et al., 2017; Guce, 2017, 2018; Kohen & Kramarski, 2012; Lan, 2007; Quines, 2017; Tican & Taspinar, 2015). In this study, integrating the use of RT strategies in a structured CL environment can be seen as one of the strengths of the research. In future research, the effects of different RT strategies on critical thinking skills can be examined in CL environment, their advantages and disadvantages can be discussed.

In mathematics curriculum (MoNE, 2018), students are expected to have high-level thinking skills. However, it can be said that strategies, activities and materials for developing these skills are not sufficiently involved in the curriculum. PISA results also show that students in Turkey are quite inadequate in cooperative problem-solving skills (OECD, 2017). Therefore, RT activities can be used as a useful tool for supporting CL and developing critical thinking skills.

First of all, teachers who will apply CL and RT activities in a classroom should have high-level thinking skills such as reflective thinking and critical thinking. Gagnon and Collay (2006) state that teachers should structure and manage the processes of projecting about their cooperative experiences. In teacher training programs, it is suggested that pre-service teachers should implement their experiences with CL and RT activities.

This research was limited to the seventh grade level. However, the independent applicability of the RT strategies and materials used in the CL process are considered to be one of the strengths of the research. In future research, the effects of CL supported by RT activities on mathematical achievement, metacognitive

skills, self-efficacy beliefs, attitude, et al. can be determined. Experimental research can be carried out in different subject areas and at different grade levels by using CL supported by RT activities. Students' critical thinking skills can be analyzed by qualitative methods and more detailed data can be gained.

In subsequent studies, supporting different CL techniques with RT activities can be investigated in terms of academic achievement, self-regulation, metacognitive skills and attitudes. In the classroom environment, student interaction processes can be analyzed in detail through qualitative research. It is also important that the physical state of the class is designed to allow student interaction for the implementation of CL and RT activities. In this context, not only for mathematics courses but only for all courses, it is suggested that the physical structure of the class should be arranged in such a way that teachers can easily apply methodological approaches.

References

Adams, J. W. (2013). *A case study: Using lesson study to understand factors that affect teaching creative and critical thinking in the elementary classroom* (Unpublished doctoral thesis). Drexel University, US.

Agustan, S., Juniati, D., & Siswono, T. Y. E. (2016). Reflective thinking in solving an algebra problem: A case study of field independent-prospective teacher. Retrieved May 12, 2017, from http://iopscience.iop.org/article/10.1088/1742-6596/893/1/012002/pdf.

Aldahmash, A. H., Alshmrani, S. M., & Almufti, A. N. (2017). Secondary school science teachers' views about their reflective practices. *Journal of Teacher Education for Sustainability, 19*(1), 43-53.

Applefield, J. M., Huber, R., & Moallem, M. (2000). Constructivism in theory and practice: Toward a better understanding. *The High School Journal, 84*(2), 35–53.

Aronson, E., & Patnoe, S. (1997). *The jigsaw classroom: Building cooperation in the classroom* (2nd ed.). Wokingham: Addison- Wesley Educational Publishers Inc.

Aryani, F., Rais, M., & Wirawan, H. (2017). Reflective learning model in improving student critical thinking skills. *Global Journal of Engineering Education, 19*(1), 19-23.

Asha, I. K., & Al Hawi, A. M. (2016). The impact of cooperative learning on developing the sixth-grade students decision-making skill and academic achievement. *Journal of Education and Practice, 7*(10), 60-70.

Berkun, D. N., & Ada, T. (2017). Permutasyon ve olasilik konusunun ogretiminde bilgi degisme tekniginin kullanilmasinin akademik basariya ve hatirda tutma duzeyine etkisinin incelenmesi [The effect of exchange of knowledge

method on the academic success and recall level in teachıng permutation and probability subject]. *Kastamonu Education Journal, 25*(6), 2295-2310.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school.* Washington, Dc: National Academy Press.

Brockbank, A., & McGill, I. (2006). *Facilitating reflective learning through mentoring and coaching*. London, Philadelphia, PA: Kogan Page.

Choy, S. C., & Oo, P. S. (2012). Reflective thinking and teaching practices: A precursor for incorporating critical thinking into the classroom. *International Journal of Instruction, 5*(1), 167-182.

Chukwuyenum, A. N. (2013). Impact of critical thinking on performance in mathematics among senior secondary school students in Lagos State. *IOSR Journal of Reasearch & Method in Education (IOSR-JRME), 3*(5), 18-25.

Colley, B. M., Bilics, A. R., & Lerch, C. M. (2012). Reflection: A key component to thinking critically. *The Canadian Journal for the Scholarship of Teaching and Learning, 3*(1), 1-19.

Cumhur, F., & Elmas-Baydar, H. (2017). İsbirlikli ögrenme yonteminin ebob-ekok konusu ogretimindeki etkililigi [The effect of cooperative learning method in the teaching of GCD-LCM]. *Kastamonu Education Journal, 25*(5), 1663-1680.

Demir, S. (2015). Evaluation of critical thinking and reflective thinking skills among science teacher candidates. *Journal of Education and Practice, 6*(18), 17-22.

Dewey, J. (1933). *How we think.* Boston: Heath & CO.

Ennis, R. H., & Millman, J. (1985). *Cornell critical thinking test (Level X)*. Pacific Grove, CA: Critical Thinking Press & Software.

Ennis, R. H. (1996). *Critical thinking*. Upper Saddle River, NJ: Prentice-Hall.

Evin-Gencel, İ., & Guzel-Candan, D. (2014). Ogretmen adaylarının elestirel dusunme egilimleri ve yansitici dusunme duzeylerinin incelenmesi [Investigation of critical thinking tendency and reflective thinking levels of teacher candidates]. *International Journal of Curriculum and Instructional Studies, 4*(8), 55-68.

Farrah, M. (2012). Reflective journal writing as an effective technique in the writing process. *An-Najah University Journal for Research – Humanities, 26*(4), 997-1025.

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

Freeman, B., Higgins, K. N., & Horney, M. (2016). How students communicate mathematical ideas: An examination of multimodal writing using digital technologies. *Contemporary Educational Technology, 7*(4), 281-313

Gagnon, G.W., & Collay, M. (2006). *Constructivist learning design: Key questions for teaching to standarts*. California: Corwin Press.

Garcha, P. S., & Kumar, K. (2015). Effectiveness of cooperative learning on critical thinking dispositions of secondary school students. *Issues and Ideasin Education, 3*(1), 55–62.

Ghanizadeh, A. (2017). The interplay between reflective thinking, critical thinking, self-monitoring, and academic achievement in higher education. *Higher Education, 74*(1), 101-114.

Gibson, S. A. (2008). An effective framework for primary-grade guided writing instruction. *The Reading Teacher*, *62*(4), 324–334.

Gillies, R. M. (2006). Teachers and students verbal behaviors during cooperative and small-group learning. *British Journal of Educational Psychology*, *76*(2), 271–287.

Gorlewski, J., & Greene, K. (2011). Research for the classroom: The power of reflective writing. *The English Journal, 100*(4), 90-93.

Gray, A. (2016). *The 10 skills you need to thrive in the fourth industrial revolution***.** Retrieved November 15, 2017, from https://www.weforum.org/agenda/2016/01/the-10-skills-you-need-to-thrive-in-the-fourth-industrial-revolution.

Guce, I. (2017). Investigating college students' views on mathematics learning through reflective journal writing. *International Journal of Evaluation and Research in Education, 6*(1), 38-44.

Guce, I. (2018). The role of journal writing in senior high school students' attitude toward mathematics. *International Journal of Evaluation and Research in Education, 7*(2), 132-141.

Guvenc, H. (2011). Yansitma materyalleriyle desteklenen isbirlikli ogrenmenin Turkce ogretmeni adaylarinin ozduzenlemeli ogrenmelerine etkileri [The effects of cooperative learning supported with reflective materials on Turkish teacher candidates self-regulation]. *Education and Science, 36*(159), 3-13.

Inoue, N., & Buczynski, S. (2011). You asked open-ended questions, now what? Understanding the nature of stumbling blocks in teaching inquiry lessons. *The Mathematics Educator, 20*(2), 10–23.

Isik, D., & Tarim, K. (2009). The effects of the cooperative learning method supported by multiple intelligence theory on Turkish elementary students mathematics achievement. *Asia Pacific Educ. Rev., 10*, 465-474.

Johnson, D. W., Johnson, R. T., & Smith, K. (2014). Cooperative learning: Improving university instruction by basing practice on validated theory. *Journal on Excellence in College Teaching, 25*(3-4), 85-118.

Kizilkaya, G. (2009). *Yansitici dusunme etkinlikleri ile desteklenmis web tabanli ogrenme ortamlarinin problem cozme uzerine etkisi [The effect of web-based learning environments supported with reflective thinking activities to problem solving]* (Unpublished doctoral dissertation). Hacettepe University, Ankara.

King, F. J., Goodson, L., & Rohani, F. (2013). *Higher order thinking skills. Center for advancement of learning and assessment.* Retrieved June 23, 2015 from http://www.cala.fsu.edu/files/higher_order_thinking_skills.pdf.

Kohen, Z., & Kramarski, B. (2012). Developing self-regulation by using reflective support in a video-digital microteaching environment. *Journal for Education Research International, 3,* 1-35.

Kramarski, B., & Kohen, Z. (2017). Promoting preservice teachers' dual self-regulation roles as learners and as teachers: Effects of generic vs. specific prompts. *Metacognition Learning, 12*(2), 157-191.

Kramarski, B., Weiss, I., & Sharon, S. (2013). Generic versus context-specific prompts for supporting selfregulation in mathematical problem solving among students with low or high prior knowledge. *Journal of Cognitive Education and Psychology, 12*(2), 197-214.

Lan, G. H. (2007). A cooperative learning programme to enhance mathematical problem solving performance among secondary three students. *The Mathematics Educator, 10*(1), 59-80.

Loes, C., & Pascarella, E. (2017). Collaborative learning and critical thinking: Testing the link. *Journal of Higher Education, 88*, 726-753.

Lucena, R. J., & San Jose, A. E. (2016). Co-operative learning in enhancing the speaking skills of students: A Phenomenological approach. *International Journal of Advanced Multidisciplinary Research*, *3*(2), 67-71.

McNaught, K. (2010). Reflective writing in mathematics education programmes. *Reflective Practice, 11*(3), 369–379.

Mecit, Ö. (2006). *The effect of 7E learning cycle model on the improvement of fifth grade students'critical thinking skills* (Unpublished doctoral dissertation). Middle East Technical Universiy, Ankara.

Mevarech Z. R., & Amrany C. (2008). Immediate and delayed effects of meta-cognitive instruction on regulation of cognition and mathematics achievement. *Metacognition Learning, 3*(2), 147–157.

Mevarech, Z., & Kramarski, B. (2014). *Critical maths for innovative societies: The role of metacognitive pedagogies*. Paris: OECD Publishing.

Michalsky, T., & Kramarski, B. (2015). Prompting reflections for integrating self-regulation into teacher technology education. *Teachers College Record, 117*(5), 1–38.

Ministry of National Education. (2018). *Matematik dersi ogretim programi (İlkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. siniflar)[Mathematics curriculum (Primary and secondary 1, 2, 3, 4, 5, 6, 7 and 8 grades)].* Ankara: MEB Yayinlari.

Mitchell, C., & Coltrinari, H. (2001). Journal writing for teachers and students. In T. Barer Stein, & M. Kompf (Eds.), *The craft of teaching adults* (pp. 21-28). (3rd ed.). Toronto: Irwin Pub.

OECD (2017). *PISA 2015 results (Volume V): Collaborative problem solving, PISA*. Paris: OECD Publishing.

Pandya, S. (2017). Understanding students' opinion on co-operative learning implementation in Mathematics. *European Journal of Education Studies, 3*(7), 132-153.

Parveen, Q., Yousuf, M. I., & Mustafa, S. (2017). An experimental study on the effect of cooperative learning on students' academic achievement and students' perceptions towards cooperative learning. *The Anthropologist, 27*(1-3), 69-76.

Palinussa, A. L. (2013). Students' critical mathematical thinking skills and character: Experiments for junior high school students through realistic mathematics education culture-based. *Journal on Mathematics Education, 4*(1), 75-94.

Parsons, M., & Stephenson, M. (2005). Developing reflective practice in student teachers: Collaboration and critical partnerships. *Teachers and Teaching: Theory and Practice, 11* (1), 95–116.

Quines, E. (2017) Effectiveness of cooperative learning approach in developing critical thinking skills of secondary students. In G. Teh, & S. Choy (Eds.), *Empowering 21st century learners through holistic and enterprising learning* (pp.115-123). Singaporep: Springer.

Schön, D. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.* San Francisco: Jossey Bass.

Silva, H., Lopes, J., Dominguez, C., Payan-Carreira, R., Morais, E., Nascimento, M.M., & Morais, F. (2016). Fostering critical thinking through peer-review between cooperative learning groups. *Revista Lusófona de Educação, 32,* 31-45.

Souvignier, E., & Kronenberger, J. (2007). Cooperative learning in third graders' jigsaw groups for mathematics and science with and without questioning training. *Br J Educ Psychol*, *77,* 755–771.

Su, H. F., Ricci, F. A., & Mnatsakanian, M. (2016). Mathematical teaching strategies: Pathways to critical thinking and metacognition. *Journal of Research in Education and Science (IJRES), 2*(1), 190-200.

Suhaimi, Z., Shahrill, M., Tengah, K. A., & Abbas, N. A. H. (2016). Incorporating the use of writing-to-learn strategy in grade 10 mathematics lessons: The students' perspectives. *Journal of Mathematics Education at Teachers College*, *7*(2), 11-20.

Sumarna, N., Wahyudin, & Herman, T. (2017). The increase of critical thinking skills through mathematical investigation approach. *IOP Conf. Series: Journal of Physics: Conf. Series 812*(2017) 012067, 1-8.

Taggart, G. L., & Wilson, A. P. (2005). *Promoting reflective thinking in teachers: 50 action strategies.* California: Corwin Press.

Tican, C., & Taspinar, M., (2015). The effects of reflective thinking-based teaching activities on pre-service teachers' reflective thinking skills, critical thinking skills, democratic attitudes, and academic achievement. *Anthropologist, 20*(1, 2), 111-120.

Tiruneh, D. T., Verburgh, A., & Elen, J. (2014). Effectiveness of critical thinking instruction in higher education: A systematic review of intervention studies. *Higher Education Studies*, *4*(1), 1-17.

Un-Acikgoz, K. (2011). *Aktif ogrenme [Active learning].* İzmir: Bilis Yayincilik.

Webb, N., & Farivar, S. (1994). Promoting helping behavior in cooperative small groups in middle school mathematics. *American Educational Research Journal, 31(2),* 369-395.

Wille, A. M. (2017). Imaginary dialogues in mathematics education. *Journal für Mathematik-Didaktik, 38*(1), 29-55.

Wilson, J., & Jan W. L. (1993). *Thinking for themselves developing strategies for reflective learning.* Australia: Eleanor Curtain Publishing.

Yazlik, D. O., & Erdogan, A. (2016). Isbirlikli ogrenme ile birlikte kullanilan problem cozme stratejilerinin ogrenci basarisi uzerine etkisi [The effects of problem solving strategies used in combinations with cooperative learning on learner achivement]. *Ahi Evran Universitesi Kirsehir Egitim Fakultesi Dergisi, 17*(3), 1-16.

Yansıtıcı Düşünme Etkinlikleriyle Desteklenen İşbirlikli Öğrenmenin Öğrencilerin Eleştirel Düşünme Becerileri Üzerine Etkisi

Atıf:

Erdogan, F. (2019). Effect of cooperative learning supported by reflective thinking activities on students' critical thinking skills. *Eurasian Journal of Educational Research, 80*, 89-112, DOI: 10.14689/ejer.2019.80.5

Özet

Problem Durumu: Dünya Ekonomik Forumu'nun bilgilendirmesinde 2020 yılında bireyin sahip olması gereken beceriler kapsamında işbirlikli çalışma ve eleştirel düşünme becerisinin ön plana çıkacağı varsayılmaktadır. Ayrıca, OECD izleme ve yansıtma becerisini işbirlikli problem çözme yetkinlikleri ile birleştirmiştir. Dolayısıyla, son yıllarda, işbirlikli öğrenme (İÖ) ve yansıtıcı düşünmenin (YD) birlikte ele alınan kavramlar olduğu görülmektedir. Araştırmalarda, öğrencilerin İÖ ortamlarında yansıtmalar yapmaları sonucunda, problem çözme, sorgulama, eski ve yeni bilgileri ilişkilendirme, plan yapma, stratejiler geliştirme ve öz-düzenleme becerilerinin geliştiği belirtilmektedir. Daha önce yapılan, İÖ ve YD becerilerinin karşılıklı ilişkisini analiz eden araştırmalarda, İÖ ortamlarında yansıtmanın önemi vurgulanırken, işbirlikli grup yapısı ayrıntılı açıklanmamıştır. Bu araştırmalarda, yapılandırılmış İÖ teknikleri kullanımı ve YD becerisini geliştirmeye yönelik farklı stratejilerin İÖ tekniklerine entegre edilerek uygulanması göz ardı edilmiştir. Yurtdışında İÖ ve YD kavramlarını birlikte ele alan çalışmalar yapılmasına rağmen Türkiye'de araştırmaların oldukça sınırlı olduğu belirlenmiştir. Öğretim ortamlarında İÖ'nün YD stratejileriyle desteklenmesi sonucu elde edilen çıktıların araştırılması önemli görülmektedir. Türkiye'de yapılan araştırmalara bakıldığında matematik eğitimi alanında İÖ, YD ve eleştirel düşünme kavramlarını inceleyen oldukça az sayıda araştırma olduğu saptanmıştır. Mevcut araştırma, yapılandırılmış bir İÖ tekniğinin kullanılması, İÖ ortamında farklı YD stratejilerin kullanımının betimlenmesi ve stratejilerin uygulanması sürecinde kullanılan materyallerin ayrıntılı sunulması, İÖ yapısına uygun YD stratejilerinin bütünleştirilmesi açısından daha önce yapılan araştırmalardan farklılaşmaktadır. Dolayısıyla, bu araştırmanın İÖ'nün etkililiği üzerine farklı bir bakış açısı sunacağı düşünülmektedir.

Araştırmanın Amacı: Bu araştırmanın amacı, matematik dersinde YD etkinlikleriyle desteklenen İÖ'nün yedinci sınıf öğrencilerinin eleştirel düşünme becerileri üzerindeki etkisini incelemektir.

Yöntem: Araştırmada nicel araştırma yaklaşımlarından ön test-son test kontrol gruplu yarı deneysel model uygulanmıştır. Bu araştırmada bir deney ve bir kontrol grubu yer almaktadır. Deney grubunda YD etkinlikleriyle desteklenen İÖ yöntemi uygulanmıştır. Mevcut araştırmada, matematik eğitimi alanında sıklıkla

kullanılan, öğrencileri düşünmeye, tartışmaya ve sorgulamaya yönelten, problem çözme becerisini geliştirdiği vurgulanan, İÖ ve grupla çalışma yapısına uygun olduğu görülen yazma, günlük yazma, yansıtıcı diyalog, grup tartışması, öz-değerlendirme ve sesli düşünme stratejileri kullanılmıştır. Bu doğrultuda YD etkinliklerini uygulama sürecinde kullanılacak olan günlük, grup tartışma formu, yansıtıcı diyalog formu, öz-değerlendirme formu materyalleri tasarlanmıştır. Kontrol grubuna ise özel bir öğretim yöntemi uygulanmamış, mevcut matematik dersi öğretim programına uygun bir matematik öğretimi gerçekleştirilmiştir.

Araştırma, 2016–2017 eğitim öğretim yılında Türkiye'nin Doğu Anadolu bölgesinde yer alan bir ilde bulunan bir ortaokulun yedinci sınıflarında öğrenim gören toplam 70 öğrenci ile yürütülmüştür. Araştırmada veri toplama aracı olarak Cornell Koşullu Sorgulama Testi-Form X (CCT-X) kullanılmıştır. Veriler çözümlenirken, veri türüne bağlı olarak, bağımsız grup t-testi ve bağımlı grup t-testi kullanılmıştır. Grupların son test puanları arasındaki farkı incelemek amacıyla tek faktörlü kovaryans analizi (ANCOVA) uygulanmıştır.

Bulgular: Araştırmada, deney ve kontrol gruplarının CCT-X ön test puanları arasında istatistiksel açıdan anlamlı bir farklılık olmadığı görülmüştür [$t_{(68)}$= .02, p> .05]. Deney grubunun CCT-X son test ortalamasının ön test ortalamasına göre yüksek olduğu belirlenmiştir. Deney grubunun CCT-X ön ve son test puanları için yapılan bağımlı grup t-testi analizi sonucunda istatistiksel olarak anlamlı bir farklılık olduğu saptanmıştır [$t_{(35)}$= -6.05, p< .05]. Ancak, kontrol grubunun CCT-X ön ve son test puanlarına yönelik uygulanan bağımlı grup t-testi sonucuna göre istatistiksel açıdan anlamlı bir farklılık bulunamamıştır [$t_{(33)}$= -1.44, p> .05]. Deney ve kontrol gruplarının CCT-X ön test puanları kontrol edilirken, düzeltilmiş CCT-X son test ortalama puanları arasında anlamlı bir farklılık tespit edilmiştir. Bu farklılığın deney grubu lehine olduğu belirlenmiştir.

Sonuç ve Öneriler: Araştırmada, deney ve kontrol grubundaki öğrencilerin CCT-X ön test puan ortalamaları arasında istatistiksel olarak anlamlı bir farklılık tespit edilmemiştir. Deneysel uygulamalar sonrasında, deney grubunun CCT-X son test puan ortalamasının, ön test puan ortalamasına göre anlamlı derecede yüksek olduğu sonucuna ulaşılmıştır. Ancak, kontrol grubuna ait CCT-X ön test ve son test puan ortalamaları arasında anlamlı bir farklılık tespit edilmemiştir. Araştırmada elde edilen diğer bir önemli bulgu olarak, deney ve kontrol gruplarının ön test puanları kontrol edilirken düzeltilmiş CCT-X son test ortalama puanları arasında anlamlı bir farklılık bulunmuştur [$F_{(1,67)}$= 19.77, p< .05]. Bu farklılığın deney grubu lehine olduğu belirlenmiştir. Bu bulguya dayanarak, YD etkinlikleriyle desteklenen İÖ'nün öğrencilerin eleştirel düşünme becerileri üzerinde pozitif etkisinin olduğu söylenebilir.

Bu araştırmada, YD stratejilerinin yapılandırılmış İÖ ortamında, bütünleştirilerek kullanılması araştırmanın güçlü yönlerinden biri olarak görülebilir. Yapılacak araştırmalarda, İÖ ortamlarında farklı YD stratejileri karşılaştırılarak eleştirel düşünme becerileri üzerindeki etkileri incelenebilir, avantaj ve dezavantajları irdelenebilir. YD etkinlikleri İÖ'nün desteklenmesi ve eleştirel düşünme

becerilerinin geliştirilmesi için faydalı bir araç olarak kullanılabilir. İÖ sürecinde kullanılan YD stratejilerinin ve materyallerin konudan bağımsız uygulanabilirliği araştırmanın güçlü yönlerinden bir diğeri olarak görülmektedir. YD etkinlikleriyle desteklenen İÖ kullanılarak farklı konu alanlarında ve farklı sınıf düzeylerinde deneysel araştırmalar yapılabilir. Yapılacak araştırmalarda, farklı İÖ tekniklerinin YD etkinlikleriyle desteklenmesinin akademik başarı, öz-düzenleme, üstbilişsel beceri ve tutum gibi değişkenler üzerindeki etkisi araştırılabilir. Sınıf ortamında öğrenci etkileşim süreçleri nitel araştırmalarla ayrıntılı olarak incelenebilir. Ayrıca, sınıfların fiziksel yapısının öğretmenlere metodolojik yaklaşımları kolaylıkla uygulayacakları şekilde düzenlenmesi önerilmektedir.

Anahtar Kelimeler: Matematik eğitimi, üst düzey düşünme becerileri, birleştirme tekniği, ortaokul.

Appendix

Appendix 1.

Journal

Dimensions	Expression in the journal
Descriptive	1) What was done in the lesson today?
Metacognitive	2) What are your feelings, beliefs, attitudes about the activities in the lesson?
Analytical	3) Why were the activities and practices in the class made?
Evaluator	4) What aspects did you find successful / unsuccessful in the activities? 5) What were the chapters in which you were successful or forced in group work?
Reconstructor	6) What kind of change can be made about the activities? Why? 7) What would you like to have in future activities? 8) What are your recommendations to your group friends to make the activities more successful?

Appendix 2.

Group Discussion Form

- What did you learn at the end of the activities? Please summarize.
- Is the subject fully understood? What are unclear points?
- What are the similarities and differences between the subjects learned and the previous subjects?
- Is it necessary to change the activities? Why? If yes, what is the change?
- What were the problem-solving strategies used throughout the activities?
- Do you have a different strategy proposal? If so, what are these strategies?
- What were the hard points during the events? Why?
- What have you been doing to deal with the points you are struggling with?

Appendix 3.

Reflective Dialogue Form

Question type	Question statement
Task-oriented	• What is the purpose of this activity? • Could you describe your role in this activity? • Can you tell me about your friend's duties at this activity?
Process-oriented	• How did you complete the objectives of this activity? • What is the most interesting situation in this activity? • Do you list the situations in which you are successful? • Can you list situations in which you have difficulty with this activity? • What was your friend's help with at this activity? • How did you help your friend in this activity? • If you didn't complete this activity, what were the things that prevented you? • What can help you to perform better at this activity? • What do you feel about this activity? • What is your advice to a friend?

Appendix 4.

Self-Evaluation Form

- I'm the best at the activity.....
- The worst things in the activity.....
- The most important knowledge / skills I gained at the activity.....
- My friend's contributions / damages ....
- What do you think can be done differently during the activity process?

Eurasian Journal of Educational Research 80 (2019) 113-132

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

University Education and Creativity: An Assessment From Students' Perspective*

Ruya EHTIYAR[1], Gozdegul BASER[2]

ARTICLE INFO

Article History:
Received: 11 Aug. 2018
Received in revised form: 27 Nov. 2018
Accepted: 26 Mar. 2019
DOI: 10.14689/ejer.2019.80.6

Keywords
Creativity, Creative thinking, University education

ABSTRACT

Problem Statement: This study consists of three main problems: first what university students understand from creativity, secondly what they conceptualize about creativity in university education, and finally their personal evaluations and recommendations about creativity in university education.

Purpose of Study: The study aimed to find out the perceptions, assessments, comments, experiences and suggestions of a group of fourth year university students related to university education and creativity.

Method: Designed as a qualitative research, phenomenological methodology was followed, data was collected using focus group interview and analyzed via thematic analysis.

Findings: The results indicated that students found their university education mostly as uncreative, and experienced with limited number of creative practices during their education, and they thought that university education did not contribute to their creativity potential in general.

Implications for Research and Practice: This study provided findings related to the meaning and evaluation of creativity as well as creative and uncreative practices, effect of university education on student's creativity potential, and recommendations for university education as it was perceived by university students. It is expected that this research will lead to a broader scale research with the basic data it supplies. The research related to university education and creativity is limited in Turkey, and this study supplies basic and important findings.

* This study was partly presented at the V[th] International Eurasian Educational Research Congress in Antalya, 2-5 May, 2018

[1] Akdeniz University, TURKEY, e-mail: ehtiyar@akdeniz.edu.tr, ORCID: https://orcid.org/0000-0003-2719-2156

[2] Antalya Bilim University, TURKEY, e-mail: gozdegul.baser@antalya.edu.tr ORCID: https://orcid.org/ORCID: 0000-0002-1450-191X

Introduction

In our today's world, proposing new ideas, new applications and practice have significant importance as the social, economic and technological environment encourage "innovation", "entrepreneurship", "differentiation", "customization", and "novelty" etc. These concepts basically emphasize creativity and creative thinking. Creativity and creative thinking have many personal, cognitive, behavioral, and cultural dimensions as they include a multidimensional phenomenon. Creativity is substantial for the young adults to be able to cope with ambiguous, complex and fast changing world awaiting them. Since the university education is crucial to shape their career, creative thinking ability is assumed to be effective for students' intellectual abilities and capabilities.

On the other side, universities question their education from the aspect of novelty, adaptability and technology whether they fit the needs of young adults for their future career. University education can be considered as the last step for young learners to be ready for their creative thinking capacities, potentials and abilities to cope with the volatile and uncertain future, where they are expected to act as "change agents" or "future-makers" to be able to survive in the future environment.

This research proposes that students are creative and possess creativity potential as many researches pointed out in various studies (Amabile et al. 2005; Craft et al. 2001; Lakota,2007). Yet, creativity potential is subject to environmental factors. It can be supported, encouraged and cultivated as well as weakened, suffocated and even killed (Robinson, 2006; Seeling, 2012). The education system may lead the students mainly to memorize or to think. Drucker (1969) argued that all a student could do is to repeat what somebody had already done which would not require creativity. As Scott (2000) stated higher education systems are powerful tools not only as "knowledge factories" certainly, but also as "open zones" in which social transformation and cultural creativity can occur. Higher education needs to prepare young adults for a fast-changing working environment.

This study focuses on a group of university students as an interactive social area. There has been many studies related to elementary schools and creativity in Turkey (Ucus, 2017). However, there is a lack of research among university students when it comes to creativity in general (Papaleontiou-Louca et al. 2014, p.138).

Creativity in university education concerns teaching for creativity as well as teaching creatively (Papaleontiou-Louca et al. 2014, p.138). Teaching for creativity is acknowledged as forms of teaching that are aimed to develop young students' own creative thinking or behavior, and teaching creatively accounts for "using imaginative approaches and applications to make learning more interesting and effective". Teaching for creativity must involve creative teaching techniques (Morris, 2006, p.4). However, this study does not aim to determine the difference or emphasize the effects of teaching for creativity or teaching creatively. This study aims to clarify the understanding of a group of students and how they assess their university education from the point of creativity. In this study, researchers aim to discuss the present situation related to university education and creativity at a Turkish state university

through a group of fourth year students and get their views and comments in detail. Fourth year students are considered under this research as they are more experienced compared to the previous year students.

Literature Review

Creativity is a multi-dimensional concept and it has been generally accepted that creativity is a complex concept for which there is an absence of a particular definition (Prentice, 2000). Various definitions of creativity point out new, original ideas, knowledge etc. which would result in a change or a social or technological value: "the achievement of something remarkable and new, something which transforms and changes a field of endeavor in a significant way…. The kinds of things that people do that change the world" (Feldman et al. 1994), and "a person's capacity to produce new or original ideas, insights, restructurings, inventions or artistic objects, which are accepted by experts as being of scientific, aesthetic, social, or technological value" (Vernon, 1984). This process is defined as an exceptional human capacity: "exceptional human capacity for thought and creation" (Ryhammar & Brolin, 1999). Costello (2000) argued that creativity involves problem solving, i.e. thinking "outside the box". All these efforts are stated to be "unique" and "original": "the ability to produce something novel, something that is unique and original" (Torrance, 1970). Plucker et al. (2004) came up with the following definition: "Creativity is the interaction among aptitude, process and environment by which an individual or group produces a perceptible product that is both novel and useful as defined within a social group". In addition, "Creativity is rated among the most important human mental attributes or human capital that is considered by researchers as the driving force behind economic development, technical advances, work place leadership, and life success" (Chew et al. 2017).

In education, the most common assessments of creativity are the Torrance Tests of Creative Thinking (TTCT) which are still the most popular creativity assessment tools in education settings. Torrance identified four components by which individual creativity could be assessed: "fluency: the ability to produce a large number of idea; flexibility: the ability to produce a large variety of ideas; elaboration: the ability to develop, embellish, or fill out an idea, and originality: the ability to produce ideas that are unusual, statistically infrequent, not banal or obvious". Having a psychodynamic approach, Torrance (1969) searched the place of creativity within education. He focused on the four P's of creativity as "the creative person, the creative product, the creative process and the creative press". He proposed that creative thinking be rewarded in schools because it allow students to understand how better to achieve their potentialities. However, more recently investigations focus on understanding and evaluating the creative mind in terms of intelligence (Gardner, 1995).

Besemer and Treffinger (1981) classified creativity into components as: "novelty of how new the product is in terms of techniques, processes, concepts, the capacity of a product to spark further creative products inspired by it; the potential of a product to

'transform', or create a radical shift in approach, resolution the extent to which a product meets a need, or resolves a situation, synthesis the extent to which a product combines elements which are unlike, into a coherent whole". In addition, Glăveanu (2018) proposed that educators should be much more reflexive when using definitions, theories or assessment tools for creativity, and notice which kind of creativity they recognize and which kind they ignore.

Creativity in education has received strong concern since 1950s basing on the idea that education needs to prioritize the development and encouragement of creativity (Papaleontiou-Louca, et al. 2014). The success of the Soviets to launch the first artificial satellite, Sputnik, is another development that has accelerated creativity efforts (Özaşkın & Bacanak, 2016). Mostly starting from 1950's, education professionals tried to develop many strategies about how to cover creativity in education (Craft, 2001).

Jackson et al. (2014) searched the views of academic teachers on the core features they associated with being creative in eight disciplinary areas and discovered certain features as: "being imaginative, being original, being curious with an enquiring disposition, being resourceful, being able to combine, connect, synthesize, being able to think critically and analytically, being able to represent ideas and communicate them to others".

Amabile (1983) proposed a simple model of creativity and determined three essential elements as expertise, the ability to think creatively about relevant problems and opportunities, and the will to engage. Jackson (2014) added context to this model as context gives the reasons for being creative. This model suggests that creativity requires a context to support creativity, e.g. cultural, technological, teaching environment, and it is an interaction of expertise, task motivation, and creativity skills.

Kaufman and Beghetto (2009) proposed that creativity lies on a continuum and follows continuous progress and change to be investigated. Their four category model of creativity tries to investigate the nature, scope and influence of individuals' creativity starting from mini-c to little-c, pro-c and big-c. "Pro-c" creativity is associated with creative acts of people who have mastered a field, including, but not only, people involved in Professional activity; "little-c" creativity is the everyday creative acts of individuals who are not particularly expert in a situation; and "mini-c" is the novel and personally meaningful interpretation of experiences, actions and events made by individuals. Both mini-c and little-c forms of creativity are relevant to higher education learning and curriculum designs, and teaching and learning strategies could usefully be used to encourage and facilitate them. They pointed out the fact that if students are not encouraged to be creative, they may stay on the mini-c, and if they are encouraged, they may go further on the continuum.

Torrance (1965) examined the attitudes of over 1000 teachers in five different countries and found out that teachers were rewarding students for being well mannered, doing work on time and being obedient, popular and willing to accept the judgements of teachers, but on the other hand punishing students who were good at guessing, questioning and who were daring in their opinions. This approach is still considered widely and prevails in many educational establishments of today

(Papaleontiou-Louca, et al. 2014). However, creativity needs change and change needs going out of what is standard. Teachers need to change their standard views if they want to have creative students. Developing contemporary education policies and strategies, and teaching creativity and innovation professionally in educational programs are not enough alone, and teachers who will apply them must implement contemporary approaches to creative behavior (Ozmusul, 2012).

Today's universities are supposed to be in parallel with Industry 4.0 which requires interconnected, digital services and a new view on teaching and learning. This requires the application of innovative procedures and approaches. It requires young adults with a strong sense of self-confidence and desire for being original, creative and able to cope with big data. If students are to become unique, autonomous individuals, they need to feel worthy, competent and trustworthy. However, the education system does not promote and welcome creative thinking well enough because in some cases creativity does not "go with" the curriculum, the education system has focused and promoted "parroting" which is the favored and "right/correct" way to learn although it may result in uncreative ways of teaching (Papaleontiou-Louca, et al. 2014). University education is expected to be far from creating similar "parrots", but rather it concentrates on achieving individuals who will be able to take risks and be innovative. University education needs to be far from "memorizing" and concentrate on knowledge production rather than knowledge adoption. Cachia et al. (2009) also mentioned that although students are viewed as the center of teaching and learning processes and procedures, they do not have an active role in general.

In high level education, teaching practices should focus on more than promoting the transmission of contents and routine information (Deverell & Moore, 2014), and they should train students to inquire and investigate, problematize, take risk, think, evaluate, and act critically with high self-confidence. They should also include a diversity of approaches, enthusiasm for teaching and the promotion of curiosity, self-regulation and intrinsic motivation for the progress of creativity (Hargreaves, 2008; Sternberg, 2004). In addition, assessment of the students and the criteria of success will need to be changed (Boud & Dochy, 2010). The success criteria will need to include more than grades, and will need to be based on some outputs like projects, thesis, systems or ideas proposed.

Students sometimes learn to repress or hide their talent of creativity because they might not get a "good grade". Although they are expected to be creative, creativity is seldom a clear objective of the assessment procedures. Overall student grades usually consist of quizzes, assignments and participation, and these usually form the main method of assessment. Many students from different fields of study differ in their perceptions of creativity (Glück et al.,2002). According to the results of a study involving 264 students at a foundation university operating in Istanbul, a positively significant relationship was found between innovation tendency and entrepreneurial potential that was linked to creativity potential (Ensari and Alay, 2017).

In addition, a study using CREA (Creativity tool to measure) applied to 17 students found no differences between sexes (Carrasco,2017). Creating classroom environments

in which creativity is highlighted and used is important in terms of increasing the quality of education (Karaca & Koray, 2017). Artistic endeavors like painting, music, handicraft, and dramatic arts like literature, cinema, and theatre are important in promoting students' life and creativity (Wales, 2017).

Method

Research Design

A qualitative approach was selected to find out the perceptions, assessments, comments, experiences and suggestions of a group of fourth year university students because qualitative research is more involved with understanding individuals' perceptions of the world, and investigating insights rather than statistical analysis (Silverman, 2005). Designed as a qualitative research, a phenomenological methodology was followed. Phenomenological methodology aims to understand the experiences of the individuals about a phenomenon, defines what an individual is experiencing, and describes the essence of an individual's experiences (Saban & Ersoy, 2017).

Research Sample

The research was conducted in a faculty in a state university, in the city centre of Antalya in Turkey. A non-probability sample was preferred as the sample derives from the researcher targeting a particular group, in the full knowledge that it does not represent the wider population, it simply represents itself. This is the frequent case for qualitative research such as action, ethnography or case (Cohen et al. 2007). The participants in this study were tourism faculty fourth year students who had an overall GPA higher than the class average. The students were invited to the study by one of the researchers. There were seven female and three male students who accepted to participate in the study. Eight of the participants were 21 years old, one of them was 22, and one of them was 20 years old.

Research Instruments and Procedures

Data was collected via a focus group interview. Focus group interviews provide rich and high variety information which quantitative research may not supply as well as providing in-depth data and preventing misunderstandings (Çokluk et al. 2011). Focus group is a form of qualitative research involving interviews in which a group of people are asked about their perceptions, opinions, beliefs, and attitudes about a concept or topic. A focus group is formed by people from similar backgrounds or experiences to discuss a specific topic of interest, guided by a moderator who introduces topics and helps the group to participate in a lively and natural discussion atmosphere. Focus group is a qualitative analysis method investigating the recent context and its content (Creswell, 2016). It usually consists of eight people (Baş & Akturan,2008). The number of people may change between four to 15 people (Çokluk et al. 2011). Phases of focus group interview is planning and organizing the focus

group, group composition, conducting the focus group, recording the responses, analyzing data, and reporting the findings (Dilshad & Latif, 2013).

Data were collected in May 2018. This included a 120 minutes recorded focus group interview with the participants. The answers of the participants were recorded upon their permission. Data were collected in the meeting room in the Faculty, participants sat in a round form, and there were the two researchers and the participants. One of the researchers acted as the moderator and the other acted as an observer and took notes during the session.

Data Analysis

Thematic analysis was used to analyze data. Thematic analysis is a "method for identifying, analyzing and reporting patterns (themes) within data" and "it is a flexible and useful research tool which provides a rich and detailed, yet complex, account of the data" (Braun & Clarke, 2006). Data analysis began with readings of interview transcripts from focus group interview conversations with fourth year university students. The purpose was to determine the essence of the phenomenon and experiences of students amidst factors affecting their creativity potential to understand the perceptions, assessments, comments, experiences and suggestions of a group of fourth year university students.

During thematic analysis, data were organized categorically, reviewed repeatedly and coded continually. Interview transcripts were reviewed. The recorded data were listened for three times, converted to written form; data were grouped into themes and sub-themes by two different researchers, and the groupings formed by the researchers were found to be matching which proved the validity of the study. Quotations were listed upon each person's relevant sayings as P1, P2, P3, P4, P5, P6, P7, P8, P9, P10 demonstrating participant 1,2, etc.

The mapping of interview questions was carried out in three stages: general questions about creativity, questions about creativity and university education, and questions related to their personal recommendations. Firstly, they were asked about the meaning of creativity. Secondly, they were asked to evaluate the university education they had from the point of creativity. Next, they were asked how they would define the creative and uncreative practices in their university education. After they were asked how the university education affected their creativity potential, they were asked about the factors that influenced creativity and whether there was an increase in their potential. Finally, they were asked for their recommendations for creativity in university education.

Results

The findings of the research can be grouped under six main themes as the meaning of creativity, creativity in university education, creative and uncreative practices in university education, effect of university education on student's creativity potential, factors influencing creativity, and students' recommendations.

Meaning of Creativity

Meaning of creativity is defined as a kind of imagination, thinking differently, completing what is lacking, and being different than what is common. Theme, subthemes, codes and frequencies are listed on Table 1.

Table 1

The Ideas Of Participants Related To The Meaning Of Creativity And Their Frequency

Theme	Subthemes	f	%	Codes	f
Meaning of Creativity	Differentiation	14	56	Different actions	6
				Different ideas	5
				Different results	3
	Novelty	11	44	Imagination abilities	4
				New abilities	2
				New environment	2
				New products	3

According to P1, "Creativity is a kind of imagination, I can say new methods, practices". P2 said creativity was about new ideas and putting them in action. P3 said creativity was about new methods which were different than all the present applications. P9 said "I think creativity is to present a product which did not exist before, and to create differentiation in this sense". P6 said it was establishing novelty, however this novelty was like the realization of something uncommon. P8 described creativity as the application of things that were not seen before and the ones who could achieve this were creative. P7 explained creativity as completing something that was lacking in some way.

Creativity in University Education

Creativity in university education reminds diversity in education, creative lessons, different applications, entrepreneurship and intellectual encouragement.

Table 2

The Ideas of The Participants Related to Creativity in University Education and Their Frequency

Theme	Subthemes	f	%	Codes	f
Creativity in university education	Differentiation	11	38	Different lessons	8
				Different seminars	3
	Entrepreneurship	10	34	Entrepreneurship Lessons	3
				Practice	4
				Motivation	3
	Teaching methods	8	28	Creative methods	5
				Using arts in teaching	3

P1 said that creativity in education was to teach different things that could be applied which would not be customary. P9 said that "When we talk about creativity in university education, lessons that will increase creativity come to my mind. These can be different practices in lessons, seminars, projects etc." P2 talked about different methods to be applied in teaching. P3 said that university education should include more different things and should be appropriate for practice. P1 said that university education is limiting their creativity as a result of the standard methods of teaching: "We see slides in lessons and we are restricted. We do not disseminate our own ideas; this is how university education is". When talking about creativity in university education, entrepreneurship came to the mind of P7. P4 said that there was too much theory but no practice.

Creative and Uncreative Practices in University Education

Participants were asked about their comments related to the creative and uncreative practices in university education. Table 3 shows their comments and their frequency.

Table 3

Creative and Uncreative Practices in University Education and Their Frequency (N=10)

Theme	**Subthemes**	**f**	**%**	**Codes**	**f**
Creative and uncreative practices	Teaching methods	17	74	Memorization	8
				Traditional Teaching	5
				No creativity	4
	Creative applications	6	26	Sector communication	3
				Watching films	2
				Documentary	1

P10 said that he had not seen any kind of creativity in the first three years since university gave only preliminary information. P9 mentioned that the first two years were completely memorization, and the examinations were tests which in a way prevented students to think. P7 told that there was no system to make students think differently. P6 said that creativity could be with practice, and some lessons could be given related to creativity. P4 said that except for a few lessons, they memorized completely. P3 said that there were lessons just to fill the curriculum, and added "We saw the alternative tourism topics in introduction to tourism. I wish there were beneficial lessons in the first term instead of unnecessary lessons, though they may be theoretical, and I wish we went for internships in the second term, this would be more logical. If we went to internship directly and to places that would add something to us, this would have been better". P5 said that especially in the first two years, they didn't need to think.

Effect of University Education on Student's Creativity Potential

Effect of university education on student's creativity potential was perceived as neutral by 50 %, and 40% of the students believed that their creativity potential decreased, and 10 % thought that university education increased their creativity potential. Table 4 shows the theme, subtheme, codes and frequencies.

Table 4

Participants' Thoughts Related to The Effect of University Education on Student's Creativity Potential and Their Frequency (N=10)

Theme	**Subthemes**	**f**	**%**	**Codes**	**f**
Effect of University Education on Creativity Potential	Positive effect	7	32	Self confidence	2
				Different thinking	3
				Network	2
	Negative effect	10	45	Loss of time	6
				Loss of energy	4
	Neutral effect	*5*		No change	5

According to P9"…university education increased my creativity, however not on the level that I imagined, but I also don't think that it decreased". P8 said that it increased by the efforts of some of the professors. P7 said that it was on the average. P6 said "In my opinion, university education decreases creativity, they don't show us different points of view, if we focus on the topic, we can reach all the topics that we have been taught from any book. It is a loss of time". P5 said that she didn't take anything from the lessons in the university: "I think it decreases creativity. I came here from İstanbul. When I was there, my opinions were brighter, here I am only molded somehow. I also didn't take anything from the lessons at the university, I only got something when I made internship". P10 said that university education increased his creativity potential, and the reason why they were at the university was to know the sector and develop themselves.

Factors influencing creativity

Factors influencing creativity are declared to be Professors, Environment, Society, Lifestyle, Families, Friends, Traditional way of life, Trial, Books, Films, Activity groups, Economic conditions, Place of birth, Different places and people, Learning, Observation, Fashion, Growing up in a small town, and Social media. Table 5 demonstrates participants' ideas related to the factors influencing creativity.

Table 5

Participants' Ideas Related To Factors Influencing Creativity and Their Frequency (N=10)

Theme	Subthemes	f	%	Codes	f
Factors Influencing Creativity	People influencing our education	10	29	Professors	6
				Families	4
	Person by himself	11	31	Self-interest	7
				Perception	4
	Environment	14	40	Place born	2
				School	3
				Books	5
				Social environment	2
				Economic environment	2

P1 said that people around him, his professors and himself could influence his creativity. P2 said that creativity would be more influenced by the environment and mostly by the ideas of professors. P3 declared that the society, lifestyle, socio-economic conditions, family and friends would influence creativity. P4 also said that family and traditional way of life would influence creativity. P5 said that observation and trials would influence creativity. P6 said "I think creativity would be affected by books, films, activity groups that we belong to, also by economic conditions we live in. If you are not trapped in a cage, you will be creative but if you are trapped in a small place since you were born, you will not be creative, it comes from your family." P7 said that different places, another country or society would influence creativity. P9 said that observation skill might have an affect as well as fashion and the place a person grew up.

Recommendations

The recommendations of the students for a more creative education are listed as meeting with professionals from the sector, lessons to be more interesting, no attendance obligation to lessons, attending hobby activity clubs, being let free, no memorization, more practice, more contact with professors, and encouragement by professors. Table 6 lists the theme, subthemes, codes and frequencies.

Table 6

Recommendations of Participants Related to Creativity in University Education and Their Frequency (N=10)

Theme	Subthemes	f	%	Codes	f
Recommendations for university education creativity	Freedom	9	29	Attendance obligation	5
				Choice of interest	4
	Method of teaching	22	71	Free environment	7
				No memorization	9
				More practice	6

P1 said that he wished he could meet with more professionals so as to have a different vision. P2 said "… the lessons should be more attractive and there should not be an obligation for attendance to lessons. Since it is obligatory, I do not want to listen to the lesson". P7 said that the attendance obligation did not make her feel well, and she played with her phone during the lessons. P6 recommended that the newcomers could attend the social clubs, traveling club, etc. to increase their creativity. P5 said that students should be left free to develop creative thinking. P4 said that professors could follow more creative and efficient way of teaching to initiate the creativity potential of students.

Discussion, Conclusion and Recommendations

The findings of the study regarding the meaning of creativity are consistent with the findings of many studies (Dacey et al.1998; Feldman et al. 1994; Rhyammer and Brolin, 1999; Torrance, 1970;). Some students concerned creativity to be related with entrepreneurship as it is also mentioned by Ensari and Alay (2017).

The evaluations of a group of university students related to creativity in their education indicated that they found creativity affecting their intellectual skills. They considered creativity to be an important issue to influence their personal development as well as their future career. The learning environment of universities seems to have an influence on the creative performance (Barron & Harrington, 1981; Oldham & Cumming, 1996; Scott & Bruice, 1994). However, almost half of the students thought that there were no changes in their creativity potential, and more than half of them thought that university education even decreased their creativity potential. Therefore, university education system, curricula, teaching techniques, as well as assessment techniques need to be revised as universities may have a considerable role in enhancing creativity which is supposed to influence students' future career and life. Universities have a significant role as preparing students for future challenges and opportunities, by supporting their flexibility and creativity, in order to have students of the future "with skills to manage life" (Sternberg, 2004).

Creativity is influenced by intrinsic and extrinsic factors. The findings of this study indicated that young adults found their environment, society, life style, family, friends, fashion etc. influencing their creative abilities. What is interesting from this study was their belief about the effect of their relationship with their professors at the university. They thought that their relationship might have an influence on their creativity potential. This points out the influence of professors on young adults' creativity.

Amabile (1983) and Jackson's (2014) model of creativity points out three main areas and context. The results of this study support the effect of context as relationships with professors, curricula, university and education culture to support creativity; and secondly teaching staff as expertise in creativity and the motivation and skills of the students. The motivation of the students was observed to be high and they expressed that they were willing to participate any novel application and research.

One of the main interesting points as a barrier for creativity is the memorization pressure. Students found this situation as threatening their creativity potential. This view supports some of the researchers criticizing university education to create "parrots" through rote memorization skills (Papaleontiou-Louca, et al. 2014), and students not having an active role and creativity not playing a central role (Cachia et al. 2009).

The research results point out critical information to consider about creativity for the research group. First, creativity in university education is underestimated and not given considerable attention in general. The students do not feel to establish or develop creative skills, practices, experiences and applications. Only one of the students think that university education contributed to her creativity potential. This main result of the study is contradicting the vision of raising highly qualified human power who will be ready for Industry 4.0 age in a fastly changing, competitive, innovative and challenging environment as they can see no progress in their creative abilities and critical thinking process.

University education needs to be taught creatively and creativity should replace the pressure of memorization. Besides, students want to feel free and want to express themselves. University teachers can try to be "information guides" instead of being "information exigents". Thirdly, Turkish students find a strong correlation between entrepreneurship and creativity, therefore university curricula can involve more entrepreneurship lessons or applications.

In summary, the mission of university as to contribute to the intellectual potential of the people of future from the point of creativity needs to be reconsidered. Systematic concern can recover the curricula, research abilities, coordination with industry, less memorization pressure, and freedom to produce new ideas and projects. The students need less pressure to memorize and have the opportunity to investigate and create their own ideas based on observation, knowledge, and experience.

University academicians and instructors should be aware of creativity and be ready for it in the context of their education. It is a serious fact to consider creative teaching and evaluate the results. Otherwise, all efforts will be wasted, and the creative

and innovative thinking model of the individual will be a dream. The importance of valuing creativity in teacher education should be emphasized.

Students usually have a willingness to learn more and improve their thinking and creativity skills because of the increasingly volatile, uncertain, complex and ambiguous world awaiting them, they need to be ready for the future. They need to know their creative abilities, their potential, and they need creative thinking abilities as the "change agents" of the future. They need more imagination and less pressure to memorize. Imagination and creativity have a power that keeps us apart from everything in the world, and that is what makes a difference (Robinson, 2015). They need to learn about risk taking and failures and how to learn from their failures.

Cultural differences in every society has an impact on teaching systems. It is a fact that creative teaching and teaching for creativity cannot be standardized as well as the education in general. Therefore, each country should establish its own model for creativity especially in teacher education, and secondly in education in general. Because a model successfully applied by a country cannot guarantee the same results if applied in other countries in the same way (Özmusul, 2012).

The contribution of this study for future studies is that the results give important clues to revise and examine the university education. Young learners need more potential for innovative thinking, self-confidence, imagination, and divergent thinking (Craft, 2001).

Universities all over the world are in the era of transferring into fourth generation universities. They aim to form links and projects between government and industry through academic consultancy, research and development centers, programs, entrepreneurship projects, and student-industry collaboration (Papaleontiou-Louca, et al. 2014). This new era requires innovation and creative thinking abilities, risk taking, problem solving, being "change agents" and being "future oriented". For this purpose, universities need new tools like digital simulations, games, project-based lessons, research and development centers, and students need more practice rather than memorizing what is already known. The findings of this study indicate that students want to get close to the professionals, and they want to be more experienced before they are graduated from the university.

The university of the future will have its main focus as "improved thinking skills and creativity, it will expand its reach to untraditional areas, change the mix of its offerings, broaden its student base, and develop more creative delivery of learning ways" (Papaleontiou-Louca, et al. 2014). People of the future will need to think creatively, develop new ideas, products and services, new jobs, new processes and methods, new ways of thinking and living, new enterprises, new sectors, new business models, and new social models. Increasingly, innovation and creativity spring not from individuals thinking and working alone, but through cooperation and collaboration with others to draw on existing knowledge to create new knowledge (OECD,2018).

Finally, creativity is a multi-dimensional concept and it needs a systematic view as stated by Kaufman and Beghetto (2009). It starts with as a mini-c and evolves to pro-c. Also, it needs a framework (Dewulf & Baillie, 1999) as CASE. The study findings indicate that students think that their creativity is influenced by many factors like professors, environment, family, friends, society as mentioned above. Creativity in education should not be limited to university education, in reverse it should be considered in the whole body of the education system. "It seems that teaching for creativity will not be explored unless it adds value to the learning process, the individual and to the university, government, industry and the community stakeholders" (Papaleontiou-Louca, et al. 2014). Therefore, it will be beneficial to search for creativity in the future studies from many aspects.

References

Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. Journal of Personality and Social Psychology, 45(2), 357-376.

Amabile, T. M., Barsade, S. G., Mueller, J. S. & Staw, B. M. (2005). Affect and creativity at work. Administrative science quarterly, 50(3), 367-403. http://www.erevistas.csic.es/ficha_articulo.php?url

Barron, F. & Harrington, D. M. (1981). Creativity, intelligence, and personality. Annual review of psychology, 32(1), 439-476.

Baş, T. & Akturan, U. (2008). Nitel araştırma yöntemleri NVivo 7.0 ile nitel veri analizi. .Ankara: Seçkin Yayıncılık.

Besemer, S. P.& Treffinger, D. J. (1981). Analysis of creative products: Review and synthesis. The Journal of Creative Behavior, 15(3), 158-178.

Boud, D.& Dochy, F. (2010). Assessment 2020. Seven propositions for assessment reform in higher education. http://www.iml.uts.edu.au/assessment-futures/

Braun V., Clarke V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3,77-101.

Brolin, C. (1992). Kreativitet och kritiskt tandande. Redsckap for framtidsberedskap'[Creativity and critical thinking. Tools for preparedness for the future. Krut, 53, 64-71.

Cachia, R., Ferrari, A., Kearney, C., Punie, Y., Van den Berghe, W.& Wastiau, P. (2009). Creativity in schools in Europe: A survey of teachers. European Commission-Joint Research Center-Institute for Prospective Technological Studies, Seville. http://ftp.jrc.es/EURdoc/JRC55645.

Carrasco, AMV.(2017).Assessing creativity to boost the inclusion in musical education.ARTSEDUCA, 16,32-48.

Chew, E.S.,Abd Hamid M.A. & Madar, A.R. (2017), Conceptual Framework for Designing and Developing a Creativity Enhancement Module in Education Incorporating Indigenous Perspectives. Pertanika Journal of Social Science and Humanities, 25,SI,67-81.

Craft, A. (2001). An analysis of research and literature on creativity in education. Qualifications and Curriculum Authority, 1-37.

Craft, A., Jeffrey, B.& Leibling, M. (Eds.). (2001). Creativity in education. A&C Black. London.

Creswell, J. W. (2016). Nitel araştırma yöntemleri: Beş yaklaşıma göre nitel araştırma ve araştırma deseni. Siyasal Kitabevi.

Cohen, L., Mannion, L. & Morrison K. (2007).Research Methods in Education. Routledge, Taylor & Francis Group.

Costello, P. J. (2000). Thinking skills and early childhood education. London: David Fulton Publishers Ltd.

Çokluk, Ö., Yılmaz, K.& Oğuz, E. (2011). Nitel bir görüşme yöntemi: Odak grup görüşmesi. Kuramsal Eğitimbilim Dergisi, 4(1), 95-107.

Dacey, J. S., Lennon, K.& Fiore, L. B. (1998). Understanding creativity: The interplay of biological, psychological, and social factors. San Francisco, Jossey-Bass, 1st edi.

Deverell, A.& Moore, S. (2014). Releasing creativity in teaching and learning: the potential role of organisational legitimacy and increased dialogue. Innovations in Education and Teaching International, 51(2), 164-174.

Dewulf, S.& Baillie, C. (1999). CASE: Creativity in Art, Science and Engineering: How to foster creativity. Great Britain Department for Education and Employment, UK.

Dilshad, R. M.& Latif, M. I. (2013). Focus Group Interview as a Tool for Qualitative Research: An Analysis. Pakistan Journal of Social Sciences (PJSS), 33(1).

Drucker, P. F. (1969). Management's new role. Harvard Business Review, 49-54.

Ensari, M. Ş.& Alay, H. K. (2017). Üniversite Öğrencilerinin Yenilikçilik Eğilimi ile Girişimcilik Potansiyelleri Arasındaki İlişkiye Ailelerin Girişimcilik Öyküsünün Aracı Etkisinin İncelenmesi Üzerine Bir Araştırma.

Feldman, D. H., Csikszentmihalyi, M. & Gardner, H. (1994). Changing the world. A framework for the study of creativity. Westport, CT: Praeger.

Gardner, H. (1995). Multiple intelligences:" myths and messages". The International Schools Journal, 15(2), 8.

Glăveanu, V. P. (2018). Educating which creativity? Thinking Skills and Creativity, 27, 25-32.

Glück, J., Ernst, R.& Unger, F. (2002). How creatives define creativity: Definitions reflect different types of creativity. Communication Research Journal, 14(1), 55-67.

Hargreaves, J. (2008). Risk: the ethics of a creative curriculum. Innovations in Education and Teaching International, 45(3), 227-234.

Jackson, N. (2014). Developing students' creativity through a higher education. In International Symposium on 'The Cultivation of Creativity in University Students. http://www.normanjackson.co.uk/creativity.html

Jackson, M. O., Shaw M. & Wisdom J. (ed) (2014), Developing Creativity in Higher Education.Abingdon: Routledge.

Kaufman, J. C.& Beghetto, R. A. (2009). Beyond big and little: The four c model of creativity. Review of general psychology, 13(1), 1.

Karaca, T. & Koray,O. (2017). The Effect of Using the Creative Reversal Act in Science Education on Middle School Students' Creativity Levels. Eurasian Journal of Educational Research, 67,199-214.

Lakota, A.B. (2007). Presentation at the meeting on High Level Group on Educational Policies, Ljublijana.

Papaleontiou-Louca, E., Varnava-Marouchou, D., Mihai, S., & Konis, E. (2014). Teaching for creativity in universities. Journal of Education and Human Development, 3(4), 131-154.

Morris, W. (2006). Creativity–its place in education. New Zealand.

OECD (2018). The future of education and skills: Education 2030.https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf, retrieved on June 25th,2018.

Oldham, G. R. & Cummings, A. (1996). Employee creativity: Personal and contextual factors at work. Academy of Management Journal, 39(3), 607-634.

Özaşkın, A.G. & Bacanak, A., (2016). Eğitimde Yaratıcılık Çalışmaları: Neler Biliyoruz? Eğitim ve Öğretim Araştırmaları Dergisi, 5(25), 212-226.

Özmusul, M. (2012). Öğretmen eğitiminde yaratıcılık ve inovasyon. Kastamonu Eğitim Dergisi, 20(3), 731-746.

Plucker, J. A., Beghetto, R. A.& Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. Educational psychologist, 39(2), 83-96.

Prentice, R. (2000). Creativity: a reaffirmation of its place in early childhood education. Curriculum Journal, 11(2), 145-158.

Ryhammar, L. & Brolin, C. (1999). Creativity research: Historical considerations and main lines of development. Scandinavian Journal of Educational Research, 43(3), 259-273.

Robinson, K. (2006). Do schools kill creativity? TED.www://ted.com/talks/lang/eng/ken_robinson.

Robinson, K. (2015). Yaratıcı Öğrenciler. Sola Yayınları, İstanbul.

Saban, A. & Ersoy A. (2017). Eğitimde Nitel Araştırma Desenleri, 1. Baskı.

Scott, A. J. (2000). The cultural economy of cities: essays on the geography of image-producing industries. Sage Publication, London.

Scott, S. G. & Bruce, R. A. (1994). Determinants of innovative behavior: A path model of individual innovation in the workplace. Academy of Management Journal, 37(3), 580-607.

Seeling, T. (2012). A crash course on Creativity. www.youtube.com/watch?V=Dle_GvFIbqV.

Silverman, D. (2005). Doing Qualitative Research. London: Sage Publications.

Sternberg, R. J. (2004). Teaching College Students that Creativity Is a Decision. Guidance & Counselling, 19(4), 196-200.

Torrance, E. P. (1965). Rewarding Creative Behavior; Experiments in Classroom Creativity.

Torrance, E. P. (1969). Prediction of adult creative achievement among high school seniors. Gifted Child Quarterly, 13(4), 223-229.

Torrance, E. P. (1970). Encouraging creativity in the classroom. WCB/McGraw-Hill.

Ucus, S. (2017). Exploring creativity in social studies education for elementary grades: Teachers' opinions and interpretations. Journal of Education and Learning, 7(2), 111.

Vernon, P.E. (1984). Intelligence, Cognitive styles, and brain lateralization. International Journal of Psychology, 19 (1-4), 435-455.

Wales, P. (2017). Creativity and Democrasy in Education: Practices and Politics of Learning through the Arts. Pedagogies, 12 (4),413-415.

Üniversite Eğitimi ve Yaratıcılık: Öğrencilerin Bakış Açısından Bir Değerlendirme

Atıf:
Ehtiyar, R., & Baser, G. (2019). University education and creativity: An assessment from students' perspective. *Eurasian Journal of Educational Research, 80,* 113-132, DOI: 10.14689/ejer.2019.80.6

Özet

Problem Durumu: Günümüz dünyasında sosyal, ekonomik ve teknolojik ortam hızlı bir değişim içindedir. Bu değişime uyum sağlayabilmek için, farklı ve yenilikçi bakış açılarına ihtiyaç duyulmakta; belirsiz ve karmaşık ortamlarda yeni fikirler, yeni uygulamalar ve yenilikçi yaklaşımlar gerekmektedir. Buradan hareketle, yaratıcılık özellikle genç insanların kariyerlerini ve yaşama bakış açılarını şekillendirmede temel olabilecek önemli konulardan birisi olarak görülmektedir. Üniversite eğitimi, gençlerin profesyonel yaşamlarına hazırlandıkları ve hızlı değişen global ortama katılmaya hazırlandıkları eğitim aşaması olup gençlerin eğitiminde, oldukça önemli bir yere sahiptir. Bu araştırma, üç ana problemi içermektedir: öncelikle üniversite öğrencilerinin yaratıcılıktan ne anladıkları, ikinci olarak, üniversite eğitiminde yaratıcılığı nasıl kavramlaştırdıkları ve son olarak üniversite eğitiminde yaratıcılık hakkındaki kişisel değerlendirmeleri ve önerileri.

Araştırmanın Amacı: Araştırma, bir grup dördüncü sınıf öğrencisinin üniversite eğitimi ve yaratıcılıkla ilgili algılarını, değerlendirmelerini, yorumlarını, deneyimlerini ve önerilerini bulmayı amaçlamaktadır. Araştırma kapsamında nitel araştırma yaklaşımı benimsenerek, öğrencilerin algı, yorum, tecrübe ve önerileri değerlendirilmiştir.

Araştırmanın Yöntemi: Nitel bir araştırma olarak tasarlanmış olup, fenomenolojik metodoloji izlenmiştir. Veriler odak grup görüşmesi kullanılarak toplanmış olup, tematik analizle incelenmiştir. Fenomenolojik metodoloji, bireylerin bir kavram ile ilgili tecrübelerini anlamayı hedefler ve bireyin tecrübelerinin özünü tanımlar (Saban and Ersoy, 2017). Odak grup görüşmesi ile bir devlet üniversitesinde okumakta olan 10 tane dördüncü sınıf öğrencisi ile araştırma yapılmıştır. Çalışma Mayıs 2018'de gerçekleştirilmiştir.

Araştırmanın Bulguları: Bulgular, öğrencilerin üniversite eğitimini çoğunlukla yaratıcı bulmadıklarını ve eğitimleri esnasında sınırlı şekilde yaratıcı uygulama deneyimlediklerini ve üniversite eğitiminin genel olarak yaratıcılık potansiyellerine katkıda bulunmadığını düşündüklerini göstermektedir. Yaratıcılığın anlamı farklılaşma ve yenilik olarak iki ana temada ifade edilmiş ve bu temaların altında ise yeni eylem, fikir, sonuç, hayal etme, ürün ve çevre kavramları yer almıştır. Üniversite eğitiminde yaratıcılık kavramı altında öğrenciler, farklılaşma, girişimcilik ve öğretme metotları gibi ana temalardan bahsetmişlerdir. Farklılaşma teması altında farklı dersler ve seminerler olmasının eğitimde yaratıcılığı destekleyeceğini; üniversite eğitiminde girişimciliğin onlara yaratıcılığı anımsattığını ve bu amaçla girişimcilik ile

ilgili derslerin, uygulamaların ve girişimciliğin motive edilmesi gibi konuların gündeme gelmesini düşündüklerini belirtmişlerdir. Üniversite eğiminde yaratıcılık denildiğinde öğretme metotlarının yaratıcı olmadığını ancak yaratıcı olması gerektiğini belirterek, olabildiğince sanatsal metot kullanımının yaratıcılığı destekleyeceğini düşündüklerini belirtmişlerdir. Üniversite eğitiminde yaratıcı ve yaratıcı olmayan uygulamalara örnek vermeleri istenildiğinde öğretme metotları yaratıcı olmayan uygulamalar olarak ifade edilmiştir. Öğretme metotlarının yaratıcı olmamasının nedenleri arasında ezbere dayanması, geleneksel anlayışta olması ve yaratıcılık içermemesi belirtilmiştir. Yaratıcı uygulamalara örnek olarak sektörle iletişim halinde olmak, konularla ilgili film, belgesel vb. farklı kaynaklardan bilgi edinmek gibi konular açıklanmıştır. Öğrencilere, üniversite eğitiminin yaratıcılık potansiyellerini nasıl etkilediği konusundaki düşünceleri sorulmuştur. Ağırlıklı bölümü üniversite eğitiminin yaratıcılık üzerinde olumsuz bir etkisinin olduğunu düşündüklerini belirtmiş olup, zaman ve enerji kaybı olarak ifade etmişlerdir. Olumlu etkisi olduğunu düşünen öğrenciler ise özgüven, farklı düşünme yeteneklerini gelişmesi ve iletişim ağlarının gelişmiş olması gibi konulardan söz etmişlerdir. Olumlu ya da olumsuz etkisi olmadığını düşünen öğrenciler de olmuştur. Yaratıcılığı etkileyen faktörlerin ne olduğu konusundaki fikirleri sorulduğunda ise öğrenciler eğitimlerini etkileyen insanlar (eğitimciler ve aileler), kişinin kendisinin (ilgi, algı) ve çevrenin yaratılcılık üzerinde etkili olduğunu belirtmişlerdir. Çevre içinde ise kişinin doğup büyüdüğü yer, gittiği okullar, ilgi duyduğu kitaplar ve yetiştiği sosyal çevre yer almaktadır. Son olarak, üniversite eğitiminde yaratıclık konusunda önerileri sorulduğunda, öğrenciler özgürlük ve öğretme metotları olarak iki ana tema üzerinde görüş ifade etmişlerdir. Üniversite eğitiminde daha özgür bir ortam olmasını istediklerini ve bu amaçla devam zorunluluğu ve ilgi alanı dışındaki derslere devam etme zorunluluğu gibi konuların yaratıcılığı olumsuz etkilediğini düşündüklerini söylemişlerdir. Ayrıca öğretme metotlarında özgür bir ortama ihtiyaç olduğunu, ezbere dayanan bir eğitim istemediklerini ve daha çok uygulama görmek istediklerini ifade etmişlerdir.

Araştırmanın Sonuçları ve Önerileri: Araştırma sonuçları, üniversite öğrencilerinin üniversite eğitimini genel olarak yaratıcı bulmadıklarını ve aldıkları eğitimin yaratıcılık potansiyellerini olumlu yönde geliştirmediğini; üniversite eğitiminde öğretme metotlarının daha yenilikçi olmasını istediklerini, ezber yöntemi yerine daha özgürlükçü ve uygulamaya dönük çalışmaları tercih etmek istediklerini göstermektedir. Çalışma, yaratıcı ve yaratıcı olmayan uygulamaların yanı sıra yaratıcılığın anlamı ve değerlendirilmesi ile ilgili, üniversite eğitiminin öğrencinin yaratıcılık potansiyeli üzerindeki etkisi ve üniversite öğrencileri tarafından algılandığı şekilde üniversite eğitimi ile ilgili bazı öneriler sunmaktadır. Araştırmanın sınırlı bir örneklemi kapsadığı unutulmamalıdır. Bununla birlikte, bu araştırmanın sağladığı temel verilerle daha geniş kapsamlı bir araştırmaya yol açması beklenebilir. Türkiye'de üniversite eğitimi ve yaratıcılıkla ilgili araştırmalar sınırlıdır ve araştırma bulguları üniversite eğitiminde yaratıcılık konusuna dikkat çekmektedir.

Anahtar Kelimeler: Yaratıcılık, Yaratıcı düşünme, Üniversite eğitimi.

Eurasian Journal of Educational Research 80 (2019) 133-150

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

Reliability of Essay Ratings: A Study on Generalizability Theory

Hakan ATILGAN[1]

ARTICLE INFO

Article History:
Received: 2 Aug. 2018
Received in revised form: 8 Jan. 2019
Accepted: 8 Mar. 2019
DOI: 10.14689/ejer.2019.80.7

Keywords
Generalizability Theory, generalizability, reliability, essay rating, essay rater reliability, writing ratings

ABSTRACT

Purpose: This study intended to examine the generalizability and reliability of essay ratings within the scope of the generalizability (G) theory. Specifically, the effect of raters on the generalizability and reliability of students' essay ratings was examined. Furthermore, variations of the generalizability and reliability coefficients with respect to the number of raters and optimal number of raters for obtaining optimal reliability of the rating of the writing ability of a student, which is considered to be an implicit trait as a whole and in its sub-dimensions of wording/writing, paragraph construction, and title selection, were determined.

Research Methods: The student sample of the study comprised 443 students who were selected via random cluster sampling, and rater sample of this study comprised four Turkish teachers. All the essays written by the students in the sample were independently rated on a writing skill scale (WSS), which is an ordinal scale comprising 20 items, by four trained teachers. In this study, data analysis was performed using the multivariate $p^{\bullet} x\, i^{\circ} x\, r^{\bullet}$ design of the G theory.

Finding: In the G studies that were performed, variances of the rater (*r*) as well as item and rater (*ixr*) were low in all sub-dimensions; however, variance of the object of measurement and rater (*pxr*) was relatively high. The presence of trained raters increased the reliability of the ratings.

Implications for Research and Practice: In the decision (D) study analyses of the original study conducted using four raters, the G and Phi coefficients for the combined measurement were observed to be .95 and .94, respectively. Further, the G and Phi coefficients were .91 and .90, respectively, for the alternative D studies that were conducted by two trained raters. Thus, rating of essays by two trained raters may be considered to be satisfactory.

[1] Ege University, TURKEY, E-mail: hakan.atilgan@ege.edu.tr, ORCID: https://orcid.org/0000-0002-5562-3446

Introduction

Different tools are used depending on the feature of education that is to be measured. One of these measurement tools is essay-type examinations, which are appropriate for measuring high-level skills, including writing, self-expression in a native or foreign language, problem solving, creative thinking, critical thinking and synthesis step behaviours (Atilgan, Kan & Aydin, 2017; Turgut & Baykul, 2010). Cohen, Swerdlik and Philipps (1996) also emphasised that essay-type examinations require organisation, planning and writing skills. Writing is a critical skill (Graham, Harris & Hebert, 2011); therefore, writing and writing-based essay-type examinations constitute a primary mechanism by which students can display their knowledge (Graham, 2006).

Furthermore, essay-type examinations are tests by which students are expected to display their academic content knowledge (Bereiter, 2003). Generally, a student writing an essay must gather his/her thoughts about a given subject, create an idea, and organise his/her thoughts. Essay-type examinations are more recognised compared to other types of examinations for measuring writing ability of a student (Atilgan, Kan & Aydin, 2017; Schoonen, 2005). From this viewpoint, essay-type examinations are considered to be essential measurement tools in the field of education. However, even though essay-type examinations exhibit various advantages while measuring writing ability of a student, it exhibits various disadvantages, such as the creation of errors, because of the complexity and versatility of essay-type examinations (Shavelson, Baxter & Gao, 1993).

Because there are differences between writing abilities of various students, students are not expected to achieve identical ratings in essay-type examinations. Furthermore, ratings will vary from one student to another, thereby reflecting differences between their writing abilities. However, a student's rating is affected by several extraneous factors. With respect to writing, which is a complex ability, these extraneous factors include several variance sources such as the task, type of task, rater, rating tool, essay topic, student's interest in the topic, essay type (such as descriptive, analytical, narrative or argumentative), time constraint, rating process, interaction, and other such factors (Schoonen R., 2005; Sudweeks, Reeve & Bradshaw, 2005). Moreover, changes in ratings that are obtained based on this variance are considered to be measurement errors.

Similar to all ratings, main objective of measurements in essay-type examinations is to accurately evaluate the measured feature of students (Kim, Schatschneider, Wanzek, Gatlin & Otaiba, 2017; Nitko & Brookhart, 2011; Nunnally & Bernstein, 1994). However, as mentioned previously, apart from a student's writing efficiency, measurement errors arising from the sources of variance, such as raters, tasks and measurement tools, also affect measurement results (Schoonen, 2012). Presence of errors from such sources of variance while measuring writing abilities complicates the determination of reliability (Bouwer, Beguin, Sanders & van den Berg, 2015).

Nitko and Brookhart (2011, p. 219) indicate that intra-rater reliability is low because of the nature of essay-type questions. In particular, rater is the source of variance that

affects the reliability of essay-type examinations. Because the same rater may rate the same essay differently at different times (Block, 1985; Cooper, 1984), the same essay may achieve inconsistent ratings when multiple raters rate it independently from each other (Baykul, 2000; Tugut, 1995). Furthermore, scoring reliability can be increased when the raters are provided with a high level of training (Weilgle, 1994; Weigle, 1998). However, raters may interpret the rating criteria differently and rate differently despite their high level of training (Gebril, 2009; Schoonen R., 2005; Swartz, et. al., 1999). Several studies have shown that raters differ in their implementation of the rating criteria in terms of rigidity and generosity (Atilgan, 2008; Cumming, Kantor & Powers, 2002; Eckes, 2008; Kan, 2007; Kondo-Brown, 2002).

Measurement errors that are caused by this differentiation among raters result in inconsistency in rating and decrease in reliability. Furthermore, determination of the accuracy of rating obtained via essay-type examinations depends on the measurement errors that arise from the sources of variance. Simultaneously, to minimise the interference of such errors with the measurement results, sources of these errors should be accurately understood; moreover, measurement conditions should be designed accordingly. The generalizability theory (G theory) is an appropriate methodology for designing measurement tools by determining the errors arising from multiple sources of variance.

Generalizability Theory

While determining reliability, the classical test theory considers only the errors that are obtained from a single source of variance such as items, raters and time (Crocker & Algina, 1986; Lord & Novick, 1968; Miller, Linn & Gronlund, 2009; Thordike, 1971). For example, in case of test–retest reliability, source of variance (error) is considered to be time, whereas, in case of Cronbach's alpha reliability coefficient, source of variance (error) is items. However, in some measurements, multiple sources of variance can exist. For example, in several multifaceted measurements, items that are rated using multiple raters, items and raters as well as their interactions are considered to be sources of potential variance. The G theory, which can simultaneously consider all the sources of potential variances and their interactions (Atilgan, 2008; Brennan, 2001a; Crocker & Algina, 1986; Cronbach, 1984; Nunnally & Bernstein, 1994; Shavelson & Webb, 1991), has been proposed by Cronbach et. al., (Cronbach, Rajaratnam & Gleser, 1963; Cronbach, Gleser, Nanda & Rajaratnam, 1972) as an expansion of the classical test theory for overcoming its limitations.

In a measurement scenario, a G study is conducted for determining the effects of error sources by analysing all error sources together and for defining the universe of admissible observation. The G theory can divide observed ratings into facets, interaction of facets and random errors. For example, the most prevalent G theory is a completely crossed design (p × i × r), where performances of the objects of measurement (p) are rated by multiple raters (r) using multiple items (Atilgan, 2008). In this design, p, i and r are referred to as facets. The p × i × r design of the G theory contains seven variances $\left(\sigma_p^2,\ \sigma_i^2,\ \sigma_r^2,\ \sigma_{pi}^2,\ \sigma_{pr}^2,\ \sigma_{ir}^2,\ \sigma_{pir,e}^2\right)$ comprising three main and four interaction effect variances (Atilgan, 2008; Brennan, 2001a; Shavelson & Webb,

1991). In the G study, these variances can be estimated using analytic variance techniques. Furthermore, relative error variance (δ) is defined, as presented in Equation 1, using the variances of interaction between estimated components of variance, including the objects of measurement and other facets.

$$\delta = \frac{\sigma_{pi}^2}{n_i} + \frac{\sigma_{pr}^2}{n_r} + \frac{\sigma_{pir,e}^2}{n_i n_r} \quad (1)$$

Furthermore, absolute error variance (Δ), as presented in Equation 2, is defined using the main effects of facets (except for the objects of measurement) and interaction variances among all the facets.

$$\Delta = \frac{\sigma_i^2}{n_i} + \frac{\sigma_r^2}{n_r} + \frac{\sigma_{pi}^2}{n_i} + \frac{\sigma_{pr}^2}{n_r} + \frac{\sigma_{ir}^2}{n_i n_r} + \frac{\sigma_{pir,e}^2}{n_i n_r} \quad (2)$$

The generalizability coefficient ($E\rho^2$) is defined for performing relative measurements, as presented in Equation 3, using relative variance (δ). Furthermore, reliability (Phi) coefficient (Φ) is defined for performing absolute measurements, as presented in Equation 4, using absolute error variances (Δ) (Atilgan, 2008; Brennan, 2001a; Shavelson & Webb, 1991).

$$E\rho^2 = \frac{\sigma_p^2}{\sigma_p^2 + \delta} \quad (3)$$

$$\Phi = \frac{\sigma_p^2}{\sigma_p^2 + \Delta} \quad (4)$$

The decision study (D) is conducted for determining the optimum conditions of facets, including the number of items and raters, using variances obtained from the G study for minimising the errors in a measurement design (Brennan, 2001a; Crocker & Algina, 1986; Shavelson & Webb, 1991). Furthermore, change in measurement error and reliability can be estimated by increasing or decreasing the number of each facet, such as item and rater, using the D study. Thus, measurement designs can be determined in which the conditions of facets may be considered to be optimal for achieving the desired level of reliability.

Several studies have been conducted based on G theory from the viewpoint of rating writing abilities and reliability of ratings. In some of these studies, rater and task (Kim, Schatschneider, Wanzek, Gatlin & Otaiba, 2017), rater and occasion (Sudweeks, Reeve & Bradshaw, 2005) and rater's years of experience (Dogan & Uluman, 2017) are examined as facets. In some studies related to the reliability of the writing ratings, certain traits, such as the topic of writing task, content or use of language, whether rating is analytic or holistic (Schoonen, 2005), whether rating guidance is used (Kan, 2007), the number of essay samples (Graham, Hebert, Sandbank & Harris, 2016), essay type (such as argumentative, narrative) (Bouwer, Beguin, Sanders & van den Ber, 2015) and different task types (Gebril, 2009), are considered to be the facets. Although several studies have determined the intra-rater reliability, only a few generalizability studies have studied the ratings using trained raters. Studies related to scoring and

generalizability of writing skills have mostly focused on writing skills in foreign languages, and the G theory analyses have been conducted using univariate patterns having sample widths of lower than 200. It is assumed that this study, which is conducted using the multivariate G theory pattern, will contribute to the literature with a large sample, where raters have been trained to rate writing skills in their native language.

Herein, the generalizability and reliability of the essay ratings, which measure writing abilities of the objects of measurement in their native Turkish language, have been examined in the context of multivariate G theory. In this context, the effects of raters who have been trained on the subject of rating are considered to be effective with respect to generalizability and reliability of essay ratings. This study has attempted to denote the manner in which the coefficients of generalizability and reliability change according to the number of raters while rating writing ability, which is an implicit trait, as a whole and in its sub-dimensions of title selection, paragraph construction and wording/writing along with a suitable number of raters for ensuring optimal reliability. Thus, this study intended to broaden our knowledge related to assessment of essay writing skills and to create a reference for obtaining a sufficiently reliable rating of essays.

Method

Research Design

The present study aimed to investigate generalizability and reliability of the essay ratings. The following sections describe the research sample, data collection procedure, tool and research data, and data analysis.

Research Sample

Atilgan (2013) indicates that a sample size of 400 is sufficient for performing an accurate and reliable estimation of the G and Phi coefficients. Therefore, size of the student sample of the study is targeted to be greater than 400. Therefore, three districts, namely Bayrakli, Bornova and Karsiyaka, in the provincial centre of İzmir, Turkey, and one school from each of the three districts have been selected to constitute a random cluster sample. All the 8th-grade students of these three schools constituted student sample of the study. Student sample size comprised a total of 443 students and contained 75, 165 and 204 students from each school according to the school sizes. A student sample size of 443 was sufficient for performing the G theory analyses. Because the selection of raters who are experts in the field will increase rating reliability (Schoonen, Vergeer, & Eiting, 1997), rater sample comprised four instructors chosen among Turkish instructors who are experts in their field.

Data Collection Tool and Research Data

All the students who constituted the sample were asked to write an essay. The topic of the essay was selected from the topics provided by three Turkish teachers and two

experts of educational measurement and assessment. Furthermore, instructions on the essay topic were given as follows:

> Success is not a gift that can be obtained because of coincidence but is a product of a certain amount of hard work. It is a victory that is achieved because of planned and determined work. The key to being successful is not to work for several hours but to work in a planned manner. Those who hold this key have no alternative but to succeed. Based on this explanation, write an essay explaining the importance of planned work

According to the abovementioned instructions, students wrote their essays in their own schools during Turkish class in one period (45 min) in a writing area that did not exceed the standard writing area, which can be defined as 70 lines and approximately one and a half pages of an A4-sized paper.

Furthermore, the Writing Skill Scale (WSS) (Dogan, 2015) was used for rating students' essays. This scale, which is an ordinal scale, comprised 20 items. Each item is rated on a quaternary-scale (none=0, insufficient=1, partially sufficient=2 and sufficient = 3). Because of the application of exploratory factor analysis for determining factorial construct validity, three factors with eigenvalues of greater than one were obtained. These three factors explained 82.82% of the total variance. Because of Varimax rotation, factor loads were observed to be between .74 and .87 in 14 items of the first sub-dimension, between .84 and .89 in 3 items of the second sub-dimension, and between .87 and .97 in 3 items of the third sub-dimension. These sub-dimensions were examined by experts, and the first sub-dimension was named as wording/writing (14 items), the second sub-dimension as paragraph construction (3 items), and the third sub-dimension as title selection (3 items).

Training raters with respect to rating can increase rating reliability (Weilgle, 1994; Weigle, 1998). Moreover, a good knowledge of the rating criteria affects the reliability of the ratings (Schoonen, 2005). Therefore, training was provided to four selected Turkish lesson teachers for understanding how to rate and how to use the scoring scale. Furthermore, essays to be rated were divided into four and distributed to the raters. Raters were requested to write their ratings in a separate electronic tablet that was reserved for each rater. Essays that were obtained from the raters who finished rating the essays provided to them were given to other raters. Thus, it was ensured that every rater rated all essays and that they were completely independent of each other in rating. A data matrix containing 443 × 20 dimensions was obtained because all students' papers were rated by each rater using a 20-item ordinal scale with three sub-dimensions. Furthermore, data matrices of four teachers were combined and prepared for analysis.

Data Analysis

The 20-item WSS used for rating comprised three sub-dimensions with a different number of items. Thus, sub-dimensions will be fixed facets and items will be nested in these facets. When sub-dimensions are crossed with 's', 'x' and symbolised as nested in ':', the design becomes a univariate G theory design that can be symbolised as $p \times (i:s) \times r$ because all objects of measurement (p) are rated by all raters (r) on all items (i) in each sub-dimension (s). Brennan (2001a) refers to such designs as the 'table

of specifications' designs that comprise a sub-dimension (or tests) and items in a sub-dimension. Such a design is considered to be balanced when the number of items in each sub-dimension is equal; otherwise, it is considered to be unbalanced. This study used an unbalanced design because the number of items in sub-dimensions was different. Brennan (2001a, p. 86) states in G theory that the usage of multivariate G theory analysis instead of univariate analysis in unbalanced designs, as in this study, is a more convenient and powerful methodology. Furthermore, univariate analysis creates uncertainty and complexity in estimates and designs with unequal number of items in sub-dimensions, whereas multivariate analysis ensures separate estimation of variance and covariance components in each fixed facet sub-dimension (Brennan, 2001a, p. 276); therefore, herein, a multivariate $p^{\bullet} x\, i^{\circ} x\, r^{\bullet}$ design of G theory is used. In this design, superscripted and filled circle '•' denotes that the facet is crossed with fixed multivariate data, and unfilled circle '◦' denotes that the facet is nested in multivariate data (Brennan, 2001a; Brennan, 2001b).

Variance components are estimated for sub-dimensions in G study conducted using the multivariate design $p^{\bullet} x\, i^{\circ} x\, r^{\bullet}$of the G theory. Herein, the generalizability coefficient ($E\rho^2$) was calculated for performing relative measurements, and reliability coefficient Φ was calculated for performing the absolute, sub-dimension and compound measurements. In the alternative D study, the $E\rho^2$ and Φ coefficients were calculated with an increased and decreased number of rater scenarios for sub-dimensions and compound measurements. All the G theory analyses were conducted using the mGENOVA 2.1 PC (Brennan, 2001b) version software.

Results

The findings are presented below respectively in two stages which are labelled as multivariate generalizability study and multivariate decision study.

Multivariate Generalizability Study

In generalizability (G) study using the multivariate design $p^{\bullet} x\, i^{\circ} x\, r^{\bullet}$ of the G theory, three main (*p*, *i* and *r*) and four interaction effect variances (*pxi*, *pxr*, *ixr* and *pxixr,e*) were estimated. These variances, which were separately estimated for the sub-dimensions and their percentages in the total variance, are presented in Table 1.

Table 1

Variances and Percentages for the Sub-dimensions Estimated using G study

	Title selection		Paragraph Construction		Wording/Writing	
Source*	Variance	%	Variance	%	Variance	%
P	.90326	73.90	1.08388	75.51	.32054	49.88
I	.06262	5.12	.00141	.10	.03129	4.87
R	.03361	2.75	.02942	2.05	.0102	1.59
p x i	.05269	4.31	.00116	.08	.0311	4.84
p x r	.08986	7.35	.24923	17.36	.08243	12.83
i x r	.00871	.71	.00192	.14	.02315	3.60
P x i x r, e	.07161	5.86	.06831	4.76	.14391	22.39
Total	1.22236	100.00	1.43533	100.00	.64262	100.00

: P: object of measurement, i: item, r: rater, e: error

Title Selection Sub-dimension. The percentage of the object of measurement (*p*) variance, which is also referred to as the universe variance, in the total variance is expected to be greater than the remaining main and interaction variances for an optimal measurement (Brennan, 2001a; Shavelson & Webb, 1991). Thus, the object of measurement variance (*p*) having the greatest variance (79.90%) in the total variance of this sub-dimension has denoted individuals' diversity with respect to 'title selection' abilities in the essays that they have written. Item (*i*) variance constituted 5.12% of the total variance. Relatively large percentage of variance associated with the items can be interpreted as differentiation of items in the "title finding" sub-dimension. The fact that another main effect variance, rater (*r*) variance, which was the focal point of this study, constituted a relatively small fraction (2.75%) of the total variance showed that there is little discrepancy among raters' ratings in the 'title selection' sub-dimension. Thus, only a few differences presumably existed in terms of generosity or rigidity in this sub-dimension with regard to ratings for all objects of measurement by four raters. The fact that variance of the interaction effect between the object of measurement and item (*pxi*), which was estimated as 4.31% of the total variance, was relatively high denoted that relative conditions of the objects of measurement differed between various items in the 'title selection' sub-dimension. Variance of interaction effect between the object of measurement and rater (*pxr*) constituted 7.35% of the total variance. This observation denotes that certain raters rated certain objects of measurement rigidly or generously in this 'title selection' sub-dimension, i.e. the relative rankings of certain objects of measurement differed for certain raters. Variance of the interaction effect between item and rater (*ixr*) constituted .71% of the total variance. The fact that the share of this variance in total variance was close to zero denoted that the raters rated students from one item to another in a consistent manner. The final variance, i.e. residual variance, comprised trilateral interaction occurring among the object of measurement, rater and item as well as error variance (*pxrxi,e* or residual). It has been concluded that relative rankings of the objects of measurement in this sub-dimension constituted 5.86% of the total variance of trilateral interaction variance of the objects of measurement, rater and item along with remaining error sources that were not taken into consideration during the G study.

Paragraph Construction Sub-dimension. Variance estimated for the object of measurement (*p*) main effect was the greatest constituting 75.51% of the total variance, denoting the diversity of the ability of 'paragraph construction' in submitted essays. Item (*i*) variance constituted .10%, which was a considerably small fraction, of the total variance. This denoted that items in the paragraph construction sub-dimension only exhibited a minor variation. The fact that rater (*r*) variance constituted a small fraction of the total variance with 2.05% denoted that there was a minor discrepancy between the ratings of the raters in the sub-dimension. Percentages of bilateral variance of the interaction effect between the object of measurement and item (*pxi*) and between item and rater (*ixr*) were .08% and .14%, respectively, and were observed to be close to zero. Thus, relative conditions of the objects of measurement among the items of this sub-dimension differed slightly, and raters rated the objects of measurement from one item to another in a consistent manner. On the contrary, variance of interaction effect between the object of measurement and rater (*pxr*) constituted 17.36% of the total

variance as the greatest variance after object of measurement (*p*) variance, which is the universal rating variance. This denoted that certain raters rated certain objects of measurement either rigidly or generously. The residual variance, which is the trilateral interaction among the object of measurement, rater and item as well as error variance (*pxrxi,e* or residual) and the variance of the relative rankings of the object of measurement, rater and trilateral interaction of item and other error sources that were not considered in the G study, were found to constitute 4.76% of the total variance.

Wording/Writing Sub-dimension. Object of measurement (*p*) variance, which is the universal rating variance, constituted a smaller percentage, 49.88%, of the total variance when compared with other sub-dimensions. However, for the object of measurement (*p*) main effect variance, which was the greatest variance in the total variance, students' 'wording/writing' ability diversity has been put forth in essays, although in a lesser degree in comparison with other sub-dimensions. Item (i) variance, which constituted 4.87% of the total variance, had a relatively higher percentage and showed differentiation of items in this sub-dimension. The fact that rater (r) variance, whichwas the focal point of this study, constituted a relatively small portion of the total variance with 1.59% and denoted that the ratings of the raters in this sub-dimension showed little discrepancy or that there were few differences in terms of generosity or rigidity. The fact that the bilateral variance of the interaction effect between the object of measurement and item (*pxi*), which constituted 4.31% of the total variance,was relatively large denoted that the relative conditions of the objects of measurement differed in this sub-dimension. The fact that bilateral variance of interaction effect between the item and rater was 3.60% of the total variance denoted that rating stability of the raters while rating the objects of measurements between various items was lower when compared to that observed in other sub-dimensions. Bilateral variance of interaction effect between the object of measurement and rater (*pxr*) constituted 12.83% of the total variance. This denoted that certain raters rated certain objects of measurement either rigidly or generously. The residual variance, which occurred because of trilateral interaction among the object of measurement, rater and item as well as error variance (*pxrxi,e* or residual), exhibited the second greatest variance percentage, i.e. 22.39% of the total variance. This indicated that relative rankings of the objects of measurement exhibited a great variance of trilateral interaction among the object of measurement, rater and item, which was larger than the remaining error sources that were not considered in the G study.

Multivariate Decision Study

In the decision study (D) with a multivariate design $p^{\bullet} x\, i^{\circ} x\, r^{\bullet}$ of the G theory, the G and Phi coefficients were calculated for four raters of the original study and for higher and lower number of raters as alternatives in each sub-dimension and in the compound measurement. Different number of raters and the G and Phi coefficient estimates for sub-dimensions and compound ratings are presented in Table 2.

Table 2

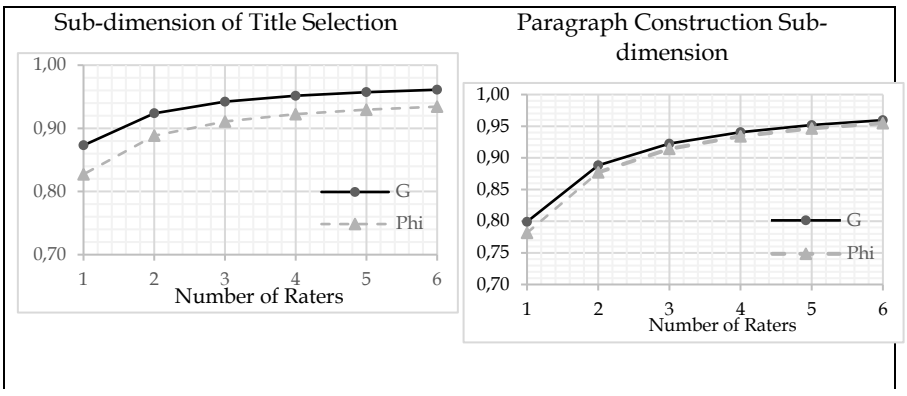
G and Phi Estimates for Different Number of Raters

Number of Raters	Sub-dimensions						Compound Measurement	
	Title Selection		Paragraph Construction		Wording/ Writing			
	G	Phi	G	Phi	G	Phi	G	Phi
6	.96	.93	.96	.96	.95	.94	.97	.96
5	.96	.93	.95	.95	.94	.93	.96	.95
4	*.95*	*.92*	*.94*	*.93*	*.93*	*.91*	*.95*	*.94*
3	.94	.91	.92	.91	.91	.89	.94	.93
2	.92	.89	.89	.88	.87	.85	.91	.90
1	.87	.83	.80	.78	.77	.75	.84	.82

Note: The italicised figures are the original number of raters.

The G coefficient ($E\rho^2$), which is calculated for the norm-referenced measurements, was obtained for the four raters as .95, .94 and .93 for 'title selection', 'paragraph construction' and 'wording/writing', respectively, and as .95 for compound measurement. The Φ coefficient, which measures the reliability of absolute (criterion-referenced) measurements, was calculated for the four raters who provided the ratings in the study as .92, .93 and .91 for the sub-dimensions of 'title selection', 'paragraph construction' and 'wording/writing', respectively, and as .94 for compound measurement.

The D study was conducted using different number of raters to determine the effect of the number of raters on the generalizability and reliability (dependability) of essay ratings, to determine the manner in which variances of the number of raters changed the G and Phi coefficients and to determine the optimal number of raters with the G theory perspective by considering manpower, time and economy without compromising psychometric quality. The effect of the number of raters obtained in the D study on the G and Phi coefficients for the sub-dimensions and compound measurement are presented in Figure 1.



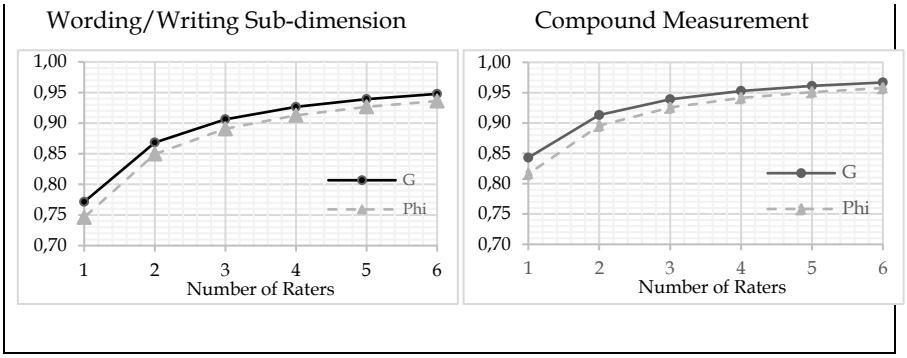


Figure 1. Sub-dimension and Compound Measurement G and Phi Coefficients for Different Number of Raters

Herein, four randomly selected Turkish course instructors were trained on how to rate students' essays. As presented in Table 2, all the G and Phi coefficients for sub-dimensions and compound measurement were greater than .90 with respect to the ratings of these four raters. As depicted in Figure 1, when the number of raters was increased from four to five, there was little gain in the G and Phi coefficients for sub-dimensions and compound measurement; when the number of raters was reduced to three, there was very little loss, and all the coefficients remain greater than .89. However, when the number of raters was reduced to two, there was some increase in the loss of the G and Phi coefficients for sub-dimensions and compound measurement. At the same time, in case two raters provided the rating, obtained G coefficients were .92, .89 and .87 and the Phi coefficients were .89, .88 and .85 for the sub-dimensions of 'title selection', 'paragraph construction' and 'wording/writing', respectively. As can be observed from Figure 1, when three raters instead of two provided the rating, the gain obtained decreased in the sub-dimensions and, particularly, in the compound measurement.

Discussion, Conclusion and Recommendations

One of the aims of this study was to determine the effect of raters on reliability. Therefore, ratings of raters, who were experts in their fields and who were trained on how to rate the essays and how to use the scale of rating, were analysed. In the G study, although wording/writing sub-dimension was smaller than title selection and paragraph construction sub-dimensions, the calculated variance of the object of measurement exhibited the highest share. The main effect variances of the raters were observed to be relatively small in the sub-dimensions, and this observation showed that ratings given for all the objects of measurement by the trained raters were consistent with each other. This result is similar to the findings of several previously conducted studies (Kim, Schatschneider, Wanzek, Gatlin & Otaiba, 2017; Schoonen R., 2005; Sudweeks, Reeve & Bradshaw, 2005) with respect to rating of writing abilities in the literature and shows that rater variance is small and that raters provide ratings

consistently with each other. Simultaneously, the fact that the percentage of the variance of interaction effect between item and rater (*ixr*) was small in all sub-dimensions can be attributed to raters being consistent in rating the items. Furthermore, people who will provide ratings should be chosen from relevant experts (Schoonen, Vergeer & Eiting, 1997) and should be trained; in these trainings (Weilgle, 1994; Weigle, 1998), they should be taught how to rate, and should also understand that provision of rating criteria affects the reliability of ratings (Schoonen, 2005). However, high percentage of variances of interaction effect between the object of measurement and rater (*pxr*) shows that certain raters were either rigid or generous in rating certain objects of measurement in all sub-dimensions. These results indicated that trained raters, who can provide consistent ratings for all objects of measurement and items, may rate a certain object of measurement more rigidly or generously and may not show the same level of consistency with respect to relative rankings of the objects of measurement. This situation (Schoonen, 2005) supports the view that even trained raters often cannot come to an agreement on rating. In this context, considering this topic while training experts will be appropriate to reduce variance between the object of measurement and rater (*pxr*) and to prevent differences between the ratings of certain raters. Moreover, with an increase in the experience of trained raters, this problem will decrease.

Another objective of this study was to establish a reference for providing future essay ratings by determining the optimal number of raters with respect to manpower, time and economy without compromising on the psychometric quality. In the analyses of the K study using the multivariate design $p^{\bullet} x\, i^{\circ} x\, r^{\bullet}$ of the G theory, the G and Phi coefficients were observed to be .95 and .94, respectively, for compound measurement among which the original coefficients were obtained using four raters, and these coefficients were observed to be greater than .90 and high in all the sub-dimensions. An increase in the number of raters with alternative K studies provided little gain in the coefficients that were obtained with four trained raters and that were observed to be already high. At the same time, when the number of trained raters was one, the G and Phi coefficients of compound measurement were obtained as .84 and .82, respectively; furthermore, when the number of trained raters was two, the G and Phi coefficients of the compound measurement were obtained as .91 and .90, respectively. This result is consistent with the finding of Kim, Schatschneider, Wanzek, Gatlin and Otaiba (2017), who suggested that one rater and several tasks are required to achieve a reliability of .80 and that two raters and several tasks are required to achieve a reliability of .90.

The results of this study suggest that two raters who are trained on the subject of rating will ensure that the G and Phi coefficients are greater than .90 while rating essay writing abilities of students. In this study, a crossed design was used. However, because a significant amount of time is required for all raters to rate all the persons, particularly in large scale tests, further research should be conducted using different designs as an alternative to crossed designs, such as nested design, by allowing some raters to rate some persons.

References

Atilgan, H. (2008). Using generalizability theory to assess the score reliability of the special ability selection examinations for music education programs in higher education. *International Journal of Research & Method in Education*, *31*(1), 63-76. https://doi.org/10.1080/17437270801919925.

Atılgan, H., Kan, A., & Aydin, B. (2017). *Egitimde olcme ve degerlendirme [Measurement and evaluation in education].* Ankara: Anı Yayıncılık.

Baykul, Y. (2000). *Egitimde ve psikolojide olcme: Klasik Test Teorisi ve uygulamasi [Measurement in education and psychology: Classical Test Theory and application].* Ankara: OSYM.

Bereiter, C. (2003). Foreword. In M. D. Shermis, & J. C. Burstein (Ed.), *Automated essay* (pp. 7-9). NJ: LEA: Mahwah.

Blok, H. (1985). Estimating the reliability, validity, and invalidity of essay ratings. *Journal of Educational Measurement*, *22*, 41-52. https://doi.org/10.1111/j.1745-3984.1985.tb01048.x.

Bouwer, R., Beguin, A., Sanders, T., & van den Berg, H. (2015). Effect of genre on the generalizability of writing scores. *Language Testing*, *32*(1) 83-100. https://doi.org/10.1177/0265532214542994.

Brennan, R. L. (2001a). *Generalizability Theory.* New York: Springer-Verlag.

Brennan, R. L. (2001b). *Manual for mGENOVA (version2.1). Iowa Testing Programmes, Occasional Papers Number 50.* Iowa City, IA: University of Iowa.

Cohen, R. J., Swerdlik, M. E., & Philips, S. M. (1996). *Pyschological testing and assessment: An inroduct on to test and measurement* (3th Edition). California: Mayfield Publishing Company.

Cooper, P. L. (1984). *The assessment of writing ability: A review of research.* Princeton, NJ: Educational Testing Service. GRE Board Research Report GREB No. 82-15R=ETS Research Report 84-12.

Crocker, L., & Algina, J. (1986). *Introduction to Classical and Modern Test Theory.* New York: Holt, Rinehart, and Winston.

Cronbach, J. L. (1984). *Essentials of psychological testing.* New York: Happers&Row Publishers.

Cronbach, L. J., Gleser, G., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability of scores and profiles.* New York, NY: Wiley.

Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. (1963). Theory of Generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology*, *16*, 137-163. https://doi.org/10.1111/j.2044-8317.1963.tb00206.x.

Cumming, A., Kantor, R., & Powers, D. E. (2002). Decision making while rating ESL/EFL writing tasks: A descriptive framework. *The Modern Language Journal*, *86*, 67–96. https://doi.org/10.1111/1540-4781.00137.

Doğan, N. (2015). Yazılı yoklamalar [Written examinations]. In H. Atılgan (Ed.), *Eğitimde ölçme ve değerlendirme [Measurement and evaluation in education]* (pp. 145-168). Ankara: Anı Yayıncılık.

Doğan, C. D., & Uluman, M. (2017). A comparison of rubrics and graded category rating scales with various methods regarding raters' reliability. *Educational Sciences: Theory & Practice*, *7*, 631–651. http://dx.doi.org/10.12738/estp.2017.2.0321.

Eckes, T. (2008). Rater types in writing performance assessments: A classification approach to rater. *Language Testing*, *25*, 155–185. https://doi.org/10.1177/0265532207086780.

Gebril, A. (2009). Score generalizability of academic writing tasks: Does one test method fit it all? *Language Testing*, *26*, 507–531. https://doi.org/10.1177/0265532209340188.

Graham, S. (2006). Writing. In P. Alexander, & P. Winne (Ed.), *Handbook of educational psychology* (pp. 457-478). NJ: Erlbaum: Mahwah.

Graham, S., Harris, K., & Hebert, M. (2011). *Informing writing: The benefits of formative assessment. A carnegie Corporation Time to Act report.* Washington, DC: Alliance for Excellent Education.

Kan, A. (2007). Effects of using a scoring guide on essay scores: Generalizability theory. *Perceptual and Motor Skills*, *105*, 891-905. https://doi.org/10.2466/pms.105.3.891-905.

Kim, Y.-S. G., Schatschneider, C., Wanzek, J., Gatlin, B., & Otaiba, S. A. (2017). Writing evaluation: rater and task effects on the reliability of writing scores for children in Grades 3 and 4. *Read Writ*, *30,* 1287-1310.

Kondo-Brown, K. (2002). A facets analysis of rater bias in measuring Japanese second language writing. *Language Testing*, *19,* 3–31. https://doi.org/10.1191/0265532202lt218oa.

Lord, F., & Novick, M. (1968). *Statistical Theory of mental test score.* California: Addison-Wesley Publishing Company.

Miller, D. M., Linn, R. L., & Gronlund, N. E. (2009). *Measurement assessment in teaching.* New Jersey: Pearson Education Inc.

Nitko, A. J., & Brookhart, S. M. (2011). *Educational assessment of student.* Boston, MA: Pearson Education.

Nunnally, J., & Bernstein, I. (1994). *Psychometric Theory (3rd Edition).* New York: McGraw-Hill, Inc..

Schoonen, R. (2005). Generalizability of writing scores: an application of structural equation modeling. *Language Testing*, *22*(1) 1-30. https://doi.org/10.1191/0265532205lt295oa.

Schoonen, R. (2012). The validity and generalizability of writing scores: The effect of rater, task and language. In E. Van Steendam, M. Tillema, G. Rijlaarsdam, & H. van den Berg (Ed.), *Measuring writing: Recent insights into theory, methodology and practice* (pp. 1-22). Leiden, The Netherlands: Brill.

Schoonen, R., Vergeer, M., & Eiting, M. (1997). The assessment of writing ability: Expert readers versus lay readers. *Language Testing*, *14*, 157-84 https://doi.org/10.1177/026553229701400203.

Shavelson, R. J., & Webb, M. N. (1991). *Generalizability Theory Aprime.* California: SAge Publication.

Shavelson, R. J., Baxter, G. P., & Gao, X. (1993). Sampling variability of performance assessments. *Journal of Educational Measurement*, *30*, 215-232.

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*, 239-261. https://doi.org/10.1016/j.asw.2004.11.001.

Swartz, C. W., Hooper, S. R., Montgomery, J. W., Wakely, M. B., de Kruif, R. E., Reed, M., Brown, T. T., Levine, M. D., & White, K. P. (1999). Using generalizability theory to estimate the reliability of writing scores derived from holistic and analytical scoring methods. *Education and Psychological Measurement*, *59*, 492–506. https://doi.org/10.1177/00131649921970008.

Thordike, L. R. (1971). *Educational measurement (2nd. Edition).* Washington: American Council on Education.

Tugut, F. (1995). *Egitimde olcme ve degerlendirme metodları [Measurement and evaluation methods in education].* Ankara: Nüve Matbaası.

Turgut, M., & Baykul, Y. (2010). *Egitimde olcme ve degerlendirme [Measurement and evaluation in education].* Ankara: Pegem Akademi.

Weigle, S. C. (1998). Using FACETS to model rater training effects. *Language Testing*, *15,* 263-287. https://doi.org/10.1177/026553229801500205.

Weilgle, S. C. (1994). Effects of training on raters of ESL compositions. *Language Testing*, 197-223. https://doi.org/10.1177/026553229401100206.

Kompozisyon Puanlarının Güvenirliği: Genellenebilirlik Kuramı Çalışması

Atıf:

Atilgan, H. (2019). Reliability of essay ratings: A study on generalizability Theory. *Eurasian Journal of Educational Research*, *80*, 133-150, DOI: 10.14689/ejer.2019.80.7

Özet

Problem Durumu: Kompozisyonların puanlanmasında puanlayıcılar arasındaki bu farklılaşmaların ölçme hatalarına neden olması puanların tutarsızlığı ve güvenirliğin düşmesi ile sonuçlanır. Kompozisyon tipi sınavlarla ölçülen becerilerin ne derece doğrulukla puanlanabildiğinin belirlenmesi varyans kaynaklarından gelen ölçme hatalarının ortaya konulmasına bağlıdır. Aynı zamanda ölçme sonuçlarına karışan bu tür ölçme hatalarının azaltılması için de bu hata kaynaklarının doğru şekilde bilinmesi ve ölçme durumunun ona göre desenlenmesi gerekir.

Araştırmanın Amacı: Bu Çalışmada çok değişkenli G Kuramı kapsamında bireylerin Türkçe anadilde yazma becerilerin ölçüldüğü kompozisyon puanlarının genellenebilirliği ve güvenirliği incelenmiştir. Bu bağlamda kompozisyon puanlarının genellenebilirliği ve güvenirliği üzerine yukarıda belirtildiği gibi daha etkili olduğu bilinen puanlama konusunda eğitilmiş puanlayıcıların etkisi üzerine odaklanılmıştır. Örtük özellik olan yazma becerisinin tümü ve alt boyutları olarak başlık bulma, paragraf oluşturma, anlatım-yazma boyutlarında puanlamada puanlayıcı sayısına göre genellenebilirlik ve güvenirlik katsayılarının nasıl değiştiği ve optimal bir güvenirlik için en uygun puanlayıcı sayısının ne olabileceği ortaya konulmaya çalışılmıştır. Böylece kompozisyon yazma becerilerinin değerlendirilmesi konusunda bilgimizi genişletmek ve kompozisyonların yeterince güvenilir puanlanması için referans oluşturmak amaçlanmıştır.

Araştırmanın Yöntemi: Çalışmada kullanılan okul örneklemi; Türkiye'de İzmir il merkezinden önce üç ilçe, sonra bu üç ilçenin her birinden birer okul yansız küme örneklem olarak seçilmiştir. Örnekleme seçilen okulların 8. sınıf öğrencilerinin tamamı öğrenci örneklemini oluşturmuştur. Öğrenci örneklemi 443 öğrenciden oluşmaktadır. Puanlayıcı örneklemi ise konusunda uzman olan Türkçe dersi öğretmenleri arasından seçilen dört öğretmenden oluşturulmuştur. Öğrencilerin kompozisyonlarını puanlamak için Yazma Becerileri Ölçeği (YBÖ) kullanılmıştır. Dereceleme ölçeği olan bu ölçekte 20 madde bulunmaktadır. Her bir madde dörtlü dereceleme ölçeği şeklinde puanlanmaktadır. Dört puanlayıcının kompozisyonların tümünü birbirlerinden bağımsız puanlamaları sağlanmıştır. Araştırmada G Kuramının çok değişkenli $p^{\bullet} x\, i^{\circ} x\, r^{\bullet}$ deseni kullanılmıştır. G Kuramının $p^{\bullet} x\, i^{\circ} x\, r^{\bullet}$ çok değişkenli deseniyle uygulanan G çalışmasında varyans bileşenleri alt boyutlar için kestirilmiştir. Araştırmada bağıl ölçmeler için Genellenebilirlik katsayısı ($E\rho^2$), mutlak ölçmeler için güvenirlik katsayısı (Φ) alt boyutlar ve birleşik ölçme için hesaplanmıştır. Alternatif D

çalışması ile Eρ^2 and Φ katsayıları puanlayıcı sayısının artırılması ve azaltılması senaryoları ile alt boyutlar ve birleşik ölçme için hesaplanmıştır.

Araştırmanın Bulguları: G Kuramının $p^{\bullet}x\, i^{\circ}x\, r^{\bullet}$ çok değişkenli deseni Genellenebilirlik (G) çalışması ile her bir alt boyut için üç ana (*p*, *i*, *r*) ve dört ortak etki varyansı (*pxi*, *pxr*, *ixr*, *pxixr,e*) kestirilmiştir. Başlık bulma, paragraf oluşturma ve anlatım/yazma alt boyutunda birey (*p*) sırasıyla %73.90, %75.51 ve %49.88 olarak hesaplanan varyanslar toplam varyanslar içindeki en büyük varyansa sahiptir. Bu sonuç bireylerin yazdıkları kompozisyonlarda "başlık bulma" beceri farklılıklarının ortaya konulabildiğini göstermektedir. Başlık bulma, paragraf oluşturma ve anlatım/yazma alt boyutlarının madde (*i*) varyansı toplam varyansları sırasıyla %5.12, %0.10 ve %4.87 olarak bulunmuştur. Paragraf oluşturma alt boyutu dışında nispeten büyük olan bu varyans yüzdesi; başlık bulma ve anlatım/yazma alt boyutunda maddelerin farklılaştığı biçiminde yorumlanabilir. Bu çalışmanın odak noktası olan puanlayıcı (*r*) varyansı başlık bulma, paragraf oluşturma ve anlatım/yazma alt boyutunda toplam varyansın sırasıyla %2.75, %2.05 ve %1.59 olarak hesaplanmıştır. Toplam varyansların nispeten küçük bir kısmını oluşturan puanlayıcı varyansları; puanlayıcıların alt boyutunda puanlamaları arasında tutarsızlıklarının az olduğunu göstermektedir. Başlık bulma, paragraf oluşturma ve anlatım/yazma alt boyutları için kestirilen birey ve madde (*pxi*) ortak etkisi toplam varyansların sırasıyla %4.31, %.08 ve %4.84'üdür. Başlık bulma, paragraf oluşturma ve anlatım/yazma alt boyutlarında varyansların nispeten büyük oluşu, bireylerin bu alt boyutunda maddeler arasında bağıl durumlarının farklılaştığını göstermektedir. Birey ve puanlayıcı (*pxr*) arasındaki ortak etkisi varyansı başlık bulma, paragraf oluşturma ve anlatım/yazma alt boyutlarında toplam varyansın sırasıyla %7.35, %17.36 ve %12.83'ünü oluşturmaktadır. Bu sonuç alt boyutlara belli puanlayıcıların belli bireyler için daha katı ya da daha cömert puanlama yaptıklarını göstermektedir. Madde ve puanlayıcı (*ixr*) arasındaki ortak etki varyansı başlık bulma, paragraf oluşturma ve anlatım/yazma alt boyutlarında toplam varyansın %.71, %.14 ve %3.60'ı olarak hesaplanmıştır. Başlık bulma ve paragraf oluşturma alt boyutlarında bu varyansların toplam varyansları içindeki payının sıfıra yakın olması, puanlayıcıların öğrencileri bir maddeden diğerine kararlı puanladıkları biçiminde yorumlanabilirken, anlatım/yama alt boyutunda aynı kararlığın olmadığını göstermektedir. Birey, puanlayıcı, madde arasında üç yönlü ortak etki ile hata varyansları (*pxrxi,e*) başlık bulma, paragraf oluşturma ve anlatım/yazma alt boyutlarında toplam varyansların %5.86, %4.76 ve %22.39'u olarak kestirilmiştir. Alt boyutlarda, özellikle anlatım/yazma alt boyutunda büyük olan bu varyanslar bireylerin bağıl konumlarının; birey, puanlayıcı, madde üç yönlü ortak etki varyansının G çalışmasında hesaba katılmayan diğer hata kaynaklarının büyüklüğünü göstermektedir.

G Kuramının $p^{\bullet}x\, i^{\circ}x\, r^{\bullet}$ çok değişkenli deseni Karar (D) Çalışması ile her bir alt boyut ve bütün ölçek için G ve Phi katsayıları çalışmanın orijinalinde puanlama yapan dört puanlayıcı için ve alternatif olarak daha az ve daha çok puanlayıcı sayıları için hesaplanmıştır. Bağıl ölçmeler için hesaplanan G katsayısı (Eρ^2) çalışmada puanlama yapan dört puanlayıcı için "başlık bulma", "paragraf oluşturma" ve "anlatım/yazma" alt boyutları için sırasıyla .95, .94, .93 birleşik ölçme için ise .95 olarak elde edilmiştir.

Mutlak ölçmeler için puanların güvenirliğinin bir ölçüsü olan Phi (Φ) katsayısı çalışmada puanlama yapan dört puanlayıcı için "başlık bulma", "paragraf oluşturma" ve "anlatım/yazma" alt boyutları için sırasıyla .92, .93, .91 ve birleşik ölçme için ise .94 olarak hesaplanmıştır. Puanlayıcı sayısının beş puanlayıcıya çıkarılması alt boyutlar ve birleşik ölçme için G ve Phi katsayılarında çok az kazanç sağladığı gibi, üç puanlayıcıya indirildiğinde ise kayıp çok az olmakta ve tüm katsayılar .89 ve üzerinde olmaktadır. Puanlayıcı sayısı ikiye indirildiğinden alt boyutlar ve birleşik ölçme için G ve Phi katsayılarında kayıp biraz daha artmakta ancak başlık bulma, paragraf oluşturma ve anlatım/yazma alt boyutları için sırasıyla G katsayıları .92, .89, .87; Phi katsayıları .89, .88, .85 ve birleşik ölçme için G katsayısı .91, Phi katsayısı .90 olarak elde edilmektedir.

Araştırmanın Sonuç ve Önerileri: Yapılan G çalışmalarında başlık bulma, paragraf oluşturma anlatım/yazma alt boyutlarında hesaplanan birey varyansı da en büyük paya sahiptir. Puanlayıcı ana etkisi varyansları alt boyutlarda göreli olarak küçük bulunmuştur. Bu sonuç literatürde yazma becerilerinin puanlanmasına ilişkin pek çok çalışmada puanlayıcı varyansının küçük ve puanlayıcıların birbirleri ile tutarlı puanlamalar yaptıkları bulguları ile benzerdir. Madde ve puanlayıcı (*ixr*) arasındaki ortak etki varyansı yüzdesinin tüm alt boyutlarda küçük olması puanlayıcıların maddeleri puanlamada tutalı oldukları şeklinde yorumlanabilir. Elde edilen bu sonuçlar puanlama yapacak kişilerin puanlama yapacakları konunun uzmanlarından seçilmesi, eğitilmesi ve bu eğitimlerde neyin nasıl puanlanması gerektiği, puanlama kriterlerinin verilmesi durumunda puanların güvenirliğinin yüksek olacağını göstermiştir. Ancak birey ve puanlayıcı (*pxr*) ortak etki varyansı yüzdelerinin tüm alt boyutlarda yüksek oluşu belli puanlayıcıların belli bireyleri puanlamalarında daha katı ya da cömert olduklarını göstermektedir. Bu bağlamda birey ve puanlayıcı (*pxr*) arasındaki ortak etki varyansının küçültülebilmesi ve böylece belli puanlayıcıların belli bireyleri puanlamalarında katılık ya da cömertlik bakımından farklılıkların olmaması için kompozisyon puanlayacak uzmanların eğitiminde bu konunun dikkate alınması yerinde olacaktır. Ayrıca puanlama yapacak uzman ve eğitimli puanlayıcıların puanlama deneyimlerinin artması ile bu sorunun da azalacağı düşünülebilir.

K çalışması analizlerinde, orijinali dört puanlayıcıyla yürütülen çalışmada birleşik ölçme için G katsayısının .95 ve Phi katsayısının .94 olduğu, tüm alt ölçeklerde bu katsayıların .90'ın üzerinde ve oldukça yüksek olduğu görülmüştür. Alternatif K çalışmaları ile puanlayıcı sayısının artırılması uzman ve eğitimli dört puanlayıcı ile elde edilen katsayılarda çok az kazanç sağlamıştır. Bununla birlikte uzman ve eğitimli puanlayıcı sayısının iki olması durumunda ise G katsayısı .91, Phi katsayısı .90 olarak elde edilmiştir. Bu sonuç .90 üzerinde bir güvenirliğe ulaşmak için iki puanlayıcının yeterli olduğunu göstermiştir.

Anahtar Kelimeler: Genellenebilirlik Kuramı, genellenebilirlik, güvenirlik, kompozisyon puanlama, kompozisyon puanlama güvenirliği, puanlayıcı güvenirliği, yazma puanlaması.

Eurasian Journal of Educational Research 80 (2019) 151-170

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

Using Inquiry-Based Laboratory Instruction to Improve Critical Thinking and Scientific Process Skills among Preservice Elementary Teachers

IRWANTO[1], Anip Dwi SAPUTRO[2], Eli ROHAETI[3], Anti Kolonial PRODJOSANTOSO[4]

ARTICLE INFO

Article History:
Received: 9 Jul. 2018
Received in revised form: 13 Feb. 2019
Accepted: 8 Mar. 2019
DOI: 10.14689/ejer.2019.80.8

Keywords
critical thinking skills, inquiry-based laboratory instruction, preservice elementary teachers, scientific process skills

ABSTRACT

Purpose: Although critical thinking skills (CTS) and scientific process skills (SPS) are the beneficial skills needed in the 21st century learning, the fact is that the acquisition of these two skills is still low. Research with a focus on improving these skills by using inquiry-based approach is also limited. Therefore, this quasi-experimental study aimed to enhance preservice elementary teachers' CTS and SPS by using Inquiry-Based Laboratory Instruction (IBLI).
Method: A pretest-posttest control group design was executed. A total of 43 students who attended Teaching Science in Elementary School-II Laboratory Course at the Muhammadiyah University of Ponorogo were divided into two groups using cluster random sampling. The experimental group (n=21) was taught by using IBLI, while the control group (n=22) was taught by using traditional laboratory method. The Oliver-Hoyo Rubric for Critical Thinking (OHRCT) and the Observation Checklist for SPS (OCSPS) were administered. The data were then analyzed by using normalized gain score and Mann-Whitney U test at significance level .05.
Findings: There was a significant difference in terms of CTS and SPS between control and experimental groups in favor of experimental group students. It was found out that gained CTS score of control and experimental group students was .58 and .80, while gained SPS score was .60 and .81, respectively. It can be highlighted that IBLI had a significant effect on preservice elementary teachers' performance compared to the conventional group.
Implications for Research and Practice: The findings suggest that IBLI is considered as the effective method to foster CTS and SPS of preservice elementary teachers. According to results, it is recommended that preservice teachers need to be given opportunities to develop hands-on and minds-on experiences in the science laboratory activities. The lecturers should utilize IBLI to develop students' various lifelong learning skills.

[1] Corresponding Author, Department of Educational Sciences, Yogyakarta State University, Yogyakarta, INDONESIA, e-mail: dr_irwanto@yahoo.com, irwanto.2016@student.uny.ac.id, ORCID: http://orcid.org/0000-0001-5842-5931
[2] Muhammadiyah University of Ponorogo, INDONESIA, ORCID: https://orcid.org/0000-0002-4544-2051
[3] Yogyakarta State University, Yogyakarta, INDONESIA, ORCID: http://orcid.org/0000-0002-0930-732X
[4] Yogyakarta State University, Yogyakarta, INDONESIA, ORCID: http://orcid.org/0000-0001-7603-5857

Introduction

Laboratory work is related to scientific skills that should be acquired by preservice teachers in order to improve scientific investigation, laboratory skills, and problem-solving skills (Arabacioglu & Unver, 2016; Feyzioğlu, Demirdağ, Akyıldız & Altun, 2012). Laboratory method is considered as one of the most effective teaching methods in growing lifelong learning skills of the students through various experiments. Wang (1993) revealed some ways of how students obtain information in laboratory activities, which are: (1) the manual, (2) oral instruction from the lecturer, (3) the concept that should be remembered, (4) the technique that should be remembered, (5) the new technique that should be learned, and (6) direct observation in the experiment. In laboratory activity, equipment and tools should not be provided directly, but students should be given the situation of a problem and asked to formulate hypotheses, design the experiment, verify, record the data, and assess and interpret the findings (Gultepe & Kilic, 2015). Wang (1993) also recommended that the lecturer needs to understand how students learn and help them learn more in a student-centered situation through more effective learning strategies. In this case, it is important for preservice teachers to learn how to apply knowledge and connect scientific problems into real-life contexts rather than directly learning concepts, theories, and laws (Konur & Yıldırım, 2016).

One of the crucial skills that support the students in understanding the scientific concept through laboratory work is critical thinking skills (CTS). In the literature, these skills have wide and various definitions. Critical thinking, as part of high-level thinking skills, is a reasonable reflective thinking that focuses on deciding what should be believed or done (Ennis, 2011). Critical thinking is considered as the process of analyzing, applying, conceptualizing, synthesizing, and evaluating information produced by observation, reasoning, experience, communication, or reflection, as the guidelines for belief and action (Allen, 2008; Canziani & Tullar, 2017; National Council for Excellence in Critical Thinking [NCECT], 2013; Scriven & Paul, 1987). It is based on universal intellectual values that cover clarity, precision, accuracy, consistency, good reasons, relevance, breadth, sound evidence, depth, and fairness (Gupta, Burke, Mehta & Greenbowe, 2015; NCECT, 2013; Oliver-Hoyo, 2003; Scriven & Paul, 1987). Furthermore, Facione (2011) also confirmed that critical thinking covers self-regulation, interpretation, analysis, explanation, inference, and ultimate evaluation. From that definition, critical thinking skills are needed by the graduates to practice what they had learned, make decision based on right consideration from available information, apply information in the new situation, and evaluate information that has been collected.

Critical thinking is one important element of scientific thinking skills that can be elicited through scientific process skills (Azar, 2010). Scientific process skills (SPS) are one part of lifelong learning skills that involve critical thinking that is used by students in solving problems, making them more actively involved, and consciously widen their abilities (Darus & Saat, 2014; Karsli & Şahin, 2009). In short, these two beneficial skills have a relationship with one to another. For this reason, we argue that scientific process skills are one relevant tool to manage information about the world around them, obtain new information, and process it critically (Žoldošová & Matejovičová,

2010). These skills can only be achieved if the lecturer considers SPS in each laboratory activity that they design for the students (Molefe, Stears & Hobden, 2016). Generally, Padilla (1990) classified SPS into 2 types; basic process skills (i.e., observing, inferring, predicting, measuring, classifying) and integrated process skills (i.e., formulating hypotheses, designing investigations, identifying and defining variables, experimenting, constructing tables and graphs, interpreting data, and drawing conclusions). Both of them are interrelated and their usage need to be adjusted with the educational level and the stage of students' cognitive development.

Nowadays, improving critical thinking and scientific process skills is one of the most important goals in many higher educational institutions. Nevertheless, surprisingly previous studies reported that acquisition of students' SPS and their CTS tend to be less satisfying (Aktaş & Ünlü, 2013; Aydogdu, 2017; Hardianti & Kuswanto, 2017; Irwanto, Rohaeti & Prodjosantoso, 2018a, in press; Irwanto, Rohaeti, Widjajanti & Suyanta, 2017a; Ongowo, 2017). Furthermore, in Indonesia, research which focused on developing both of these skills is also limited (Muhlisin, Susilo, Amin & Rohman, 2016; Osman & Vebrianto, 2013; Pradina & Suyatna, 2018). If these skills are not properly improved, we predict that students cannot construe knowledge and the concept that they obtained will not help them in understanding their surrounding world (Žoldošová & Matejovičová, 2010). It is necessary to consider some factors that influence students' performance in understanding science concepts. In their research, Boa, Wattanatorn and Tagong (2018), Brahler, Quitadamo and Johnson (2002) agreed that one of the external factors that influence the students' low critical thinking skills lie on the learning method applied by the lecturer.

According to the results of observation done by the researchers in science laboratory course at tertiary level, the lecturer tends to use conventional method (cookbook procedures). That finding is in line with what had been mentioned by Chairam, Somsook and Coll (2009), Koç, Doymuş, Karaçöp and Şimşek (2010) and Sari, Jasmidi, Kembaren and Sudrajat (2018). In this context, students only run the laboratory procedure prepared by the lecturer without instruction to inquire. Whereas in traditional learning, critical thinking skills are not improved optimally (Mahanal, Zubaidah, Bahri & Dinnurriya, 2016). Based on that reason, we contend that paradigm shift from a lecturer-centered to student-oriented is necessary in which the students investigate and conduct team work actively. Similarly, Brahler et al. (2002) also stated that student-centered non-traditional teaching method can improve critical thinking skills and higher learning outcomes. One of the proactive teaching methods that can facilitate students' achievement is Inquiry-based Laboratory Instruction (IBLI).

Inquiry-based laboratory is believed as the most relevant teaching method for promoting scientific concepts and scientific processes and developing research skills, covering asking research questions, formulating hypotheses, and arranging the test from the hypotheses (Casem, 2006; Tatar, 2012). In this approach, students investigate and evaluate critically the things around them and participate in learning to build the concept and long-term understanding like scientists (Löfgren, Schoultz, Hultman & Björklund, 2013; Lord & Orkwiszewski, 2006; Sağlam & Şahin, 2017). Inquiry-based learning environments emphasize active usage of critical thinking and scientific

process skills compared to memorizing the concept (Gehring & Eastman, 2008; Özgür & Yılmaz, 2017). In inquiry laboratory, students are not guided to conduct step-by-step instructions; however, they are provided with opportunities to understand science concepts, improve scientific process skills, and enhance their problem-solving skills (Irwanto, Rohaeti, Widjajanti & Suyanta, 2017b; Mutlu & Acar-Şeşen, 2018). Thereby, this method is considered to provide a positive impact on improving high-level thinking skills.

At tertiary level, inquiry becomes a very important approach in order to prepare the graduates to face the real world. Lord and Orkwiszewski (2006) reported that students that are taught by using inquiry teaching have higher achievement, scientific attitudes, and reasoning skills than control group students. Özgür and Yılmaz (2017) proposed that inquiry-based learning improved students' motivation and their conceptual understanding. Artayasa, Susilo, Lestari and Indriwati (2016) claimed the significance of different scientific process skills compared to conventional group. In USA, Maxwell, Lambeth and Cox (2015) explored the impact of inquiry learning towards fifth-grade students and reported that experimental group students showed improvement in their academic achievement, attitudes, and higher engagement compared to students who accepted conventional instruction. In Thailand, Chairam, Klahan and Coll (2015) found an increase in students' understanding of the chemical concepts between pre- and post-diagnostic tests. At the end of the intervention, students showed significant progress in phrasing scientific questions, designing experiments, identifying variables, drawing the concept list, presenting data, and analyzing results. These beneficial findings are also supported by Başer and Durmuş (2010), Duran and Dökme (2016), Ketpichainarong, Panijpan and Ruenwongsa (2010), Mutlu and Acar-Şeşen (2018), and Yakar and Baykara (2014). Based on various previous research findings, we conclude that IBLI method can improve students' cognitive, psychomotor, and affective learning outcomes even though their critical thinking and scientific process skills are not investigated yet.

As discussed previously, laboratory instruction has a vital role in science education because, actually, science concept is the investigation of the product in the laboratory using scientific method. For this reason, the lecturer should facilitate students to obtain factual information from various sources during the investigation and then encourage them to find new concepts and construct knowledge (Chiappetta, 1997). This research aimed to foster critical thinking and scientific process skills among preservice elementary teachers by using inquiry-based laboratory instruction. The research question underlying this study were:

1. Is there any significant difference on critical thinking skills score between control and experimental group students?
2. Is there any significant difference on scientific process skills score between control and experimental group students?
3. How is the improvement of students' performance between control and experimental groups after the laboratory course?

Method

Research Design

A quasi-experimental, pretest-posttest, non-equivalent control group design was employed in this research. It was conducted to investigate causal hypotheses about the causes that can be manipulated by comparing one or more experimental groups which were given treatment with one comparison group that was not given treatment (Creswell, 2009; Shadish, Cook & Campbell, 2002). Control group was taught by using traditional laboratory method, while the experimental group was taught by using Inquiry-Based Laboratory Instruction (IBLI). It was used to compare the impact of both teaching methods on critical thinking and scientific process skills. In brief, pretest-posttest control group design was presented in Table 1.

Table 1

Non-equivalent Pretest and Posttest Control Group Design

Groups	Pretest	Treatment	Posttest
Experimental	O_1	Inquiry-based laboratory	O_2
Control	O_3	Traditional laboratory	O_4

Research Sample

The sample consisted of 43 preservice elementary teachers (23 females, 20 males) who attended Teaching Science in Elementary School-II Laboratory Course in the Department of Elementary School Education, Muhammadiyah University of Ponorogo, in the fifth- semester academic year of 2017/2018. All participants were classified into two groups. A total of 21 students in the experimental group (13 females, 8 males) were randomly selected as the treatment group, and 22 students were in the control group (10 females, 12 males). They completed 5 experiment topics in one semester. The pariticipants were approximately aged 20-22 years old, chosen by using cluster random sampling. Given that this research selects groups rather than individuals, cluster random sampling is more appropriate (Fraenkel, Wallen & Hyun, 2012). All students in both groups followed the pretest, treatment, and posttest. 53.49% of the participants were females and the rest were male students.

Data Collection Instruments

The data were collected from students' practical work activities and their written reports during the laboratory course. Their performance in both groups was assessed by using the Oliver-Hoyo Rubric for Critical Thinking (OHRCT) and the Observation Checklist for Scientific Process Skills (OCSPS) as pretest and posttest. The OHRCT was developed by Oliver-Hoyo (2003), adapted and translated into Indonesian by Irwanto Rohaeti and Prodjosantoso (2018a), and obtained the coefficient of Cronbach's alpha reliability a=.84. The rubric was used to assess students' critical thinking skills based on their written laboratory reports. The instrument consisted of 6 traits; abstract, organization, the source of information, content, relevance, and presentation. Each trait targeted certain cognitive skills, including conceptualizing, analyzing, applying, synthesizing, and evaluating information. It is based on the universal intellectual

values (i.e. clarity, precision, accuracy, consistency, good reasons, relevance, breadth, sound evidence, depth, and fairness) that are embedded in the rubric (Oliver-Hoyo, 2003). In addition, each trait had a 5-point Likert scale with a range from 5 (all criteria fulfilled) to 1 (no criteria fulfilled).

The OCSPS was developed by Irwanto, Rohaeti and Prodjosantoso (2018b) to assess students' scientific process skills. The instrument consisted of 18 items, covering basic process skills (8 items) and integrated process skills (10 items). Basic process skills were observing, measuring, inferring, and communicating. Integrated process skills were investigating, identifying and controlling variables, formulating hypotheses, experimenting, and interpreting data, 2 items for each. All sub-skills were adapted from Arabacioglu and Unver (2016), Aydogdu (2017), and Padilla (1990). Each item had a 4-point Likert scale with a range from 4 (highly observed) to 1 (unobserved). Before conducting the main research, the instrument was tested with 176 students chosen randomly in Yogyakarta, and obtained the coefficient of Cronbach's alpha reliability a=.88. All students who participated in the pilot research were not involved in the main research.

Procedures

In the present research, research procedure that was implemented covered: (1) choosing control and experimental groups randomly, (2) giving the pretest, (3) conducting the course in both groups, (4) giving the posttest, and (5) analyzing and interpreting the results of the test. Students were instructed to work together in small groups and write laboratory report individually at the end of the experiment. The laboratory report was done outside the course as independent assignment. The course was conducted for 100 minutes per week by the same lecturer to avoid bias of the instructor. In the control group, students conducted the experiment by using cook-book procedure that had been provided by the lecturer. In this context, they proposed the question, and the lecturer explained and summarized the results of the experiment. While in the experimental group, all students were designed to go through each phase in guided inquiry which was adapted from Ješková et al. (2016), as illustrated in Table 2. Both groups completed all experiment topics at the same time.

Table 2

The Syntax of Guided Inquiry-Based Laboratory in Experimental Group

Steps	Activities
Presenting a contextual problem	The students observe and discuss a case given in small groups (4-5 students).
Planning and designing	The students formulate the question, define the problem, formulate the hypotheses, design the experiment, and predict the results of the experiment obviously and accurately.
Implementing	The students investigate, record the results, and make the decision about the experimental techniques.

Table 2 Continue

Steps	Activities
Analyzing and interpreting	The students present the results in the form of pictures, graphics, or table, determine the correlation between variables, compare the experimental data with the hypotheses, and conclude the experimental results.
Communicating	The students, in small groups, present their findings, discuss the results, and elaborate the written laboratory reports about the results obtained.
Conducting follow-up	The students predict the opportunity for future experiments, formulate the hypotheses to be followed-up, and apply the experimental techniques to new problems.

Experimental group students that were taught by using inquiry-based laboratory instruction and control group students that were taught by using confirmatory experiments method finished all experiment topics. The activities provided in this research were based on the topics illustrated in Table 3.

Table 3

Topics and Purposes of Experiments in Both Groups

Topics	Purposes
Units and measurement	Conducting measurement by using some measuring instruments and determining derived units based on the base units.
Static electricity	Observing the symptoms of static electricity and analyzing the phenomenon of electron transfer from one material to another material.
Dynamic Electricity	Investigating series, parallel, and mixed circuits; counting the resistance value in the electrical circuit; and determining the potential difference at the series, parallel, and mixed resistance by using a voltmeter.
Magnets and Electromagnets	Analyzing the relation between magnetic fields and electric currents; and analyzing the factors that influence the strength and the weakness of an induced magnetic field.
Conductors and Insulators	Investigating the materials that can and cannot deliver electricity; and explaining the characteristics of those things based on their ability in delivering electricity.

Data Analysis

All data obtained from observation and written laboratory reports were calculated. Descriptive statistics were administered to count the frequency and the percentage of the samples in control and experimental groups. In this research, quantitative data were analyzed by using non-parametric statistics because the sample size was small, which was less than thirty students (Bernard, 2000; Green & Salkind, 2008). Mann-Whitney *U* test was performed to examine the effect of instructional methods on students' critical thinking and scientific process skills. *N*-gain was employed to determine the increase in achievement scores between the pretest and posttest using the Hake's (1999) formula: *n*-gain = (posttest - pretest scores) / (maximum - pretest scores), with low (*n*-gain < .30), medium (.30 < *n*-gain < .70), and high criteria (*n*-gain > .70). This research used SPSS 17.0 at a significance level of .05.

Results

The findings of the research are explained in this section. The results of *U*-test, either pretest or posttest, to show the impact of IBLI on the preservice elementary teachers' critical thinking and scientific process skills were presented as follows (see Tables 4-8).

Table 4

Gap in Pretest Critical Thinking Skills Score between Experimental and Control Groups

Sub-Dimensions	Groups	*n*	Mean Rank	Sum of Ranks	Mann-Whitney *U* Test	
					U	*p*
Abstract	Experimental	21	22.71	477.00	216.000	.667
	Control	22	21.32	469.00		
Sources of Information	Experimental	21	20.38	428.00	197.000	.348
	Control	22	23.55	518.00		
Organization of the Report	Experimental	21	21.10	443.00	212.000	.576
	Control	22	22.86	503.00		
Relevance of the Ideas	Experimental	21	22.33	469.00	224.000	.837
	Control	22	21.68	477.00		
Content of the Report	Experimental	21	23.50	493.50	199.500	.313
	Control	22	20.57	452.50		
The Written Presentation	Experimental	21	22.60	474.50	218.500	.732
	Control	22	21.43	471.50		
Overall CTS	Experimental	21	21.90	460.00	229.000	.960
	Control	22	22.09	486.00		

Based on the mean rank shown in Table 4, pretest scores of experimental group students were slightly higher in terms of abstract, relevance of the ideas, the written presentation, and content of the report. Control group students were slightly higher in terms of sources of information and organization of the report. Nevertheless, overall, we did not find significant difference between both groups (U=229.000; p=.960). It indicated that prior to treatment, all students had similar prior knowledge in the six traits.

Table 5

Gap in Pretest Scientific Process Skills Score between Experimental and Control Groups

Sub-Skills	Groups	*n*	Mean Rank	Sum of Ranks	Mann-Whitney *U* Test	
					U	*p*
Observing	Experimental	21	21.86	459.00	228.000	.932
	Control	22	22.14	487.00		
Inferring	Experimental	21	21.83	458.50	227.500	.917
	Control	22	22.16	487.50		
Measuring	Experimental	21	24.02	504.50	188.500	.221
	Control	22	20.07	441.50		
Communicating	Experimental	21	23.29	489.00	204.000	.468
	Control	22	20.77	457.00		

Table 5 Continue

Sub-Skills	Groups	*n*	Mean Rank	Sum of Ranks	Mann-Whitney *U* Test	
					U	*p*
Identifying and Controlling Variables	Experimental	21	22.40	470.50	222.500	.770
	Control	22	21.61	475.50		
Investigating	Experimental	21	22.86	480.00	213.000	.553
	Control	22	21.18	466.00		
Formulating Hypotheses	Experimental	21	22.90	481.00	212.000	.532
	Control	22	21.14	465.00		
Experimenting	Experimental	21	23.14	486.00	207.000	.467
	Control	22	20.91	460.00		
Interpreting	Experimental	21	21.71	456.00	225.000	.865
	Control	22	22.27	490.00		
Overall SPS	Experimental	21	24.38	512.00	181.000	.217
	Control	22	19.73	434.00		

According to mean rank presented in Table 5, pretest scores of experimental group students were slightly higher in terms of identifying and controlling variables, measuring, formulating hypotheses, investigating, communicating, and experimenting skills. Control group students were slightly higher in terms of observing, inferring, and interpreting data skills. Similarly, overall, we also did not find significant difference between the scores of both groups (U=181.000; p=.217). It reflected that before the instruction, all students had equal prior scientific skills.

Table 6

Gap in Posttest Critical Thinking Skills Score between Experimental and Control Groups

Sub-Dimensions	Groups	*n*	Mean Rank	Sum of Ranks	Mann-Whitney *U* Test	
					U	*p*
Abstract	Experimental	21	27.33	574.00	119.000	.003
	Control	22	16.91	372.00		
Sources of Information	Experimental	21	26.90	565.00	128.000	.003
	Control	22	17.32	381.00		
Organization of the Report	Experimental	21	26.14	549.00	144.000	.019
	Control	22	18.05	397.00		
Relevance of the Ideas	Experimental	21	27.21	571.50	121.500	.002
	Control	22	17.02	374.50		
Content of the Report	Experimental	21	26.21	550.50	142.500	.002
	Control	22	17.98	395.50		
The Written Presentation	Experimental	21	27.00	567.00	126.000	.003
	Control	22	17.23	379.00		
Overall CTS	Experimental	21	32.43	681.00	12.000	.000
	Control	22	12.05	265.00		

At the end of the experiment, posttest was implemented. According to mean rank presented in Table 6, it showed that experimental group students were more superior in all sub-dimensions compared to control group students. Overall, we found out significant difference between the scores of both groups after the treatment (U=12.000;

p=.000). It indicated that implementation of IBLI had a significant effect on students' CTS. Furthermore, experimental group students obtained the highest mean rank in Abstract (*M*=27.33) and the lowest in Organization of the Report (*M*=26.14). Control group students obtained the highest mean rank in Organization of the Report (*M*=18.05) and the lowest in Abstract (*M*=16.91).

Table 7

Gap in Posttest Scientific Process Skills Score between Experimental and Control Groups

Sub-Skills	Groups	*n*	Mean Rank	Sum of Ranks	Mann-Whitney *U* Test	
					U	*p*
Observing	Experimental	21	26.00	546.00	147.000	.018
	Control	22	18.18	400.00		
Inferring	Experimental	21	27.29	573.00	120.000	.002
	Control	22	16.95	373.00		
Measuring	Experimental	21	26.36	553.50	139.500	.010
	Control	22	17.84	392.50		
Communicating	Experimental	21	26.40	554.50	138.500	.008
	Control	22	17.80	391.50		
Identifying and	Experimental	21	26.12	548.50	144.500	.003
Controlling Variables	Control	22	18.07	397.50		
Investigating	Experimental	21	27.40	575.50	117.500	.001
	Control	22	16.84	370.50		
Formulating	Experimental	21	27.88	585.50	107.500	.001
Hypotheses	Control	22	16.39	360.50		
Experimenting	Experimental	21	26.33	553.00	140.000	.011
	Control	22	17.86	393.00		
Interpreting	Experimental	21	27.10	569.00	124.000	.004
	Control	22	17.14	377.00		
Overall SPS	Experimental	21	32.57	684.00	9.000	.000
	Control	22	11.91	262.00		

According to mean rank presented in Table 7, it showed that experimental group students were more dominant in all sub-skills compared to control group students. Overall, we found out significant difference between the scores of both groups (*U*=9.000; *p*=.000). It confirms that implementation of IBLI had a significant impact on students' SPS. Moreover, experimental group students obtained the highest mean rank in Formulating Hypotheses (*M*=27.88) and the lowest in Observing (*M*=26.00). Meanwhile, control group students obtained the highest mean rank in Observing (*M*=18.18) and the lowest in Formulating Hypotheses (*M*=16.39).

Table 8

The Difference of N-Gain Scores between Experimental and Control Groups

	Critical Thinking Skills		Scientific Process Skills	
Groups	Control	Experimental	Control	Experimental
Pretest	16.32	16.24	15.91	16.62
Posttest	24.23	27.29	28.05	32.24
N-gain	.58	.80	.60	.81
Category	Medium	High	Medium	High

According to *n*-gain presented in Table 8, it showed that experimental group students had score improvement in high category, while control group students had score improvement in medium category. It can be concluded that there was enhancement of pretest to posttest in both groups, although experimental group students who were taught by using IBLI showed better performance compared to control group students that were taught by using traditional laboratory method in both dependent variables.

Discussion, Conclusion and Recommendations

In the context of 21st century learning, critical thinking and scientific process skills are considered as two superior skills that become the main purpose of teaching science at college or university level (Ahrari, Samah, Hassan, Wahat & Zaremohzzabieh, 2016; Boa et al., 2018; Irwanto, Rohaeti & Prodjosantoso, 2018c; Karsli & Şahin, 2009; Molefe et al., 2016). In order to develop both, we claim that science learning need to be designed to give opportunities to the students through inquiry activities. For this reason, this research aimed to improve students' cognitive and psychomotor domains by using IBLI method. Five practical works were done by control and experimental group students in one semester. Afterwards, we investigated students' performance at the beginning and at the end of the experiment by using two valid and reliable instruments.

According to the results, at the beginning of the instruction, all students showed less satisfying performance. Based on the analysis, we found out that there was no significant difference in pretest CTS scores of control and experimental groups. It indicated that prior to treatment, all students had similar prior knowledge. In addition, we also did not find out any significant differences among pretest SPS scores of both groups. It reflected that before the instruction, all students had equal prior scientific skills. Related to students' low skills, we found that they encountered difficulty in designing the experimental procedures, as reported by Yang and Park (2017). According to results of observation that we did before conducting the research, it was caused by the lecturer who just utilized traditional method during the laboratory instruction. Finally, we assumed that students taught by using conventional method tend to obtain less optimal achievement (Duran & Dökme, 2016; Quitadamo, Faiola, Johnson & Kurtz, 2008; Wartono, Hudha & Batlolona, 2018). The reason, as mentioned by Quitadamo et al. (2008), is that traditional method is not built from students' prior knowledge, it does not bridge how science was practiced in the real world, and it does not promote students' awareness to learn.

In order to compare the effectiveness of teaching methods, students in both groups were given different treatment. At the end of the instruction, all students showed better performance improvement compared to the pretest. Based on the findings, preservice teachers in the experimental group were more superior in all sub-dimensions of CTS compared to control group students. The fact is that students did not encounter any trouble in arranging written laboratory reports. It indicates that each step in inquiry can lead students to conceptualize, analyze, apply, synthesize, and

evaluate the information at once to decide what to trust and do. Overall, we concluded that there was significant difference in posttest scores of both groups. It can be claimed that IBLI has positive impact on students' CTS. Supportively, Maxwell et al. (2015) revealed that critical thinking can be developed through inquiry approach. In this regard, inquiry process is believed to facilitate students in explaining the concept, sharing the knowledge, explaining the opinion, listening to alternative opinion, discussing with others, and maintaining their ideas, as revealed by Tatar (2012). By participating in inquiry-based activities, students can recognize the nature of science, the phenomenon, and scientific concept; develop their ability in evaluating scientific data critically and participate in scientific community (Löfgren et al., 2013). This is the reason why IBLI can promote students' CTS in the current research.

We also underline that students' involvement in the inquiry process stimulate them to be active learners physically and mentally, help them acquire science and consolidate those processes with scientific knowledge, critical thinking, and scientific reasoning (Hsiao et al., 2017; Ozdem-Yilmaz & Cavas, 2016). Previous research also supports current findings. In Slovenia, Avsec and Kocijancic (2014) reported that inquiry-based learning had a large and positive effect on critical thinking skills. In Israel, Hugerat and Kortam (2014) investigated 28 undergraduate biology and chemistry students, and found out that inquiry had a significant effect on improving critical thinking skills. Furthermore, critical thinking skills of experimental group students that were taught by using inquiry-based learning were higher in all sub-skills than control group students who were taught by using traditional method (Duran & Dökme, 2016; Wartono et al., 2018). It is believed that the most effective way to promote critical thinking skills is through active participation in the laboratory experiments (Irwanto, Saputro, Rohaeti & Prodjosantoso, 2018; Lujan & DiCarlo, 2006). Thereby, we suggest the lecturers apply inquiry approach in laboratory instruction. In essence, the lecturer is not considered as the mentor but as the facilitator, and the students are not passive recipients but active learners in the instruction process (Škoda, Doulík, Bílek & Šimonová, 2015).

As another finding, we reported that experimental group students were more dominant in all sub-skills, either basic or integrated process skills compared to control group students. Overall, we found out significant difference between the scores of both groups. It can be claimed that IBLI had positive impact on students' SPS. Encouragingly, Gehring and Eastman (2008) and Sağlam and Şahin (2017) revealed that scientific process skills can be promoted through inquiry-based learning. For this reason, we emphasized that in inquiry, students also learn how to propose questions and find out the answer at once when they are involved in the intellectual activities covering observing, thinking, generalizing, and creating like a scientist (Hsiao et al., 2017; Maxwell et al., 2015). In this stage, all students are directed to collect and analyze the data and report their findings in scientific format, as expressed by Casem (2006). Various IBLI advantages stimulate the students to actively generate scientific ideas, conduct scientific investigation, and construct scientific concept. This is another reason why IBLI can promote students' SPS in the current research. The expected learning purpose in this research was promoting students' long-term knowledge.

As a result, experimental group students had CTS and SPS scores enhancement in high category, while control group students had score enhancement in medium category. In this case, the fact is that experimental group students were not focused only on hands-on experiences but also minds-on activities to improve their way of thinking, as informed by Žoldošová and Matejovičová (2010). According to that explanation, we consider IBLI as the most effective method to enhance students' achievement. In the laboratory, students can plan, design, and implement the experimental procedure. Furthermore, students analyze, interpret, and communicate the data much better than control group students. How can this happen? We assert that through inquiry-based laboratory instruction, all students learn how to optimize their problem-solving skills, foster their attitudes and skills, and also associate the knowledge in their daily experiences (Yakar & Baykara, 2014). All these activities possibly become the reasons why students' scientific skills in experimental group increased rapidly. Therefore, this is obvious that the students who were taught by using inquiry method had better performance enhancement. It can be asserted that there was an improvement in pretest and posttest scores of both groups although experimental group students showed better performance compared to control group students in both dependent variables.

Various evidence that support these current findings has been reported. In Indonesia, Hardianti and Kuswanto (2017) analyzed the achievement of 77 senior high school students, and reported that their SPS could be promoted by using inquiry-based learning. In Turkey, Şen and Vekli (2016) investigated activities of 24 preservice science teachers in biology laboratory and found out that scientific process skills can be developed through inquiry-based teaching. In another research that involved 30 preservice classroom teachers, Akben (2015) reported similar findings. Furthermore, students who are exposed to inquiry instruction also show more improvement in their critical thinking scores compared to conventional group students (Ekahitanond, 2013; Greenwald & Quitadamo, 2014). Moreover, students in inquiry laboratory group also showed better achievement and they enjoyed investigation compared to students in step-by-step directions group (Demircioglu & Ucar, 2015; Lord & Orkwiszewski, 2006; Rissing & Cogan, 2009). In brief, current research confirms that IBLI is a constructivist approach that provides a great impact on the development of the cognitive and psychomotor domains.

In this research, pretest-posttest critical thinking and scientific process skills scores of preservice elementary teachers were compared. Based on the *U*-test, there was a statistically significant difference in terms of CTS and SPS between experimental group students who were taught by using IBLI and control group students who were taught by using traditional laboratory instruction. At the end of the instruction, experimental group students showed domination on posttest scores in all sub-skills of critical thinking and scientific process skills. Subsequently, it was found out that gain CTS score of control and experimental group students were obtained .58 and .80 (difference .22), while gain SPS score of the students were obtained .60 and .81 (difference .21), respectively. It can be concluded that IBLI has a positive impact on the preservice elementary teachers' achievement compared to conventional group. The

findings indicated that laboratory activities significantly improve practical and generic skills. Through laboratory work, students are seen solving problems collaboratively and participating in the inquiry process actively in order to develop cognitive skills.

According to the findings, it can be further suggested that preservice elementary teachers need to be given opportunities to develop hands-on and minds-on experiences in the laboratory activities. The lecturer should lead the usage of various constructivist teaching methods that can propel the students to enhance various lifelong learning skills. As we know, teaching methods are one of the factors that influence students' achievement (Abdullah & Shariff, 2008; Akınoğlu & Tandoğan, 2007; Budsankom, Sawangboon, Damrongpanit & Chuensirimongkol, 2015). Furthermore, impact of IBLI on students' critical thinking and scientific process skills need to be investigated in other laboratory courses related to science learning at tertiary level.

This research has several limitations; First, we only involved preservice elementary teachers in the fifth semester as participants, thereby the findings could not be generalized. As such, we recommend further research to investigate the effectiveness of this method in branches, laboratory courses, gender, and other grade levels in order to strengthen claims. Second, current research involved limited samples, thereby that it can be improved by involving a wider sample in order to obtain detailed and comprehensive information. Moreover, we recommend that future researchers need to compare impact of different levels of inquiry (i.e., confirmation, structured, guided, and open) on preservice elementary teachers' performance. Equally important, it is also necessary to explore the effect of IBLI on other dependent variables.

References

Abdullah, S., & Shariff, A. (2008). The effects of inquiry-based computer simulation with cooperative learning on scientific thinking and conceptual understanding of gas laws. *Eurasia Journal of Mathematics, Science & Technology Education*, *4*(4), 387–398.

Ahrari, S., Samah, B. A., Hassan, M. S., Wahat, N. W., & Zaremohzzabieh, Z. (2016). Deepening critical thinking skills through civic engagement in Malaysian higher education. *Thinking Skills and Creativity*, *22*, 121–128.

Akben, N. (2015). Improving science process skills in science and technology course activities using the inquiry method. *Education and Science*, *40*(179), 111–132.

Akınoglu, O., & Tandoğan, R. Ö. (2007). The effects of problem-based active learning in science education on students' academic achievement, attitude and concept learning. *Eurasia Journal of Mathematics, Science & Technology Education*, *3*(1), 71–81.

Aktas, G. S., & Ünlü, M. (2013). Critical thinking skills of teacher candidates of elementary mathematics. *Procedia - Social and Behavioral Sciences*, *93*, 831–835.

Allen, M. (2008). Promoting critical thinking skills in online ınformation literacy

ınstruction using a constructivist approach. *College & Undergraduate Libraries, 15*(1-2), 21–38.

Arabacioglu, S., & Unver, A. O. (2016). Supporting inquiry based with mobile learning to enhance students' process skills in science education. *Journal of Baltic Science Education*, *15*(2), 216–231.

Artayasa, I. P., Susilo, H., Lestari, U., & Indriwati, S. E. (2016). The effectiveness of the three levels of inquiry in improving teacher training students' science process skills. *Journal of Baltic Science Education*, *16*(6), 908–919.

Avsec, S., & Kocijancic, S. (2014). The effect of the use of an inquiry-based approach in an open learning middle school hydraulic turbine optimisation course. *World Transactions on Engineering and Technology Education*, *12*(3), 329–337.

Aydogdu, B. (2017). A study on basic process skills of Turkish primary school students. *Eurasian Journal of Educational Research*, *67*, 51–69.

Azar, A. (2010). The effect of critical thinking dispositions on students achievement in selection and placement exam for university in Turkey. *Journal of Turkish Science Education*, *7*(1), 61–73.

Baser, M., & Durmuş, S. (2010). The effectiveness of computer supported versus real laboratory inquiry learning environments on the understanding of direct current electricity among pre-service elementary school teachers. *Eurasia Journal of Mathematics, Science and Technology Education*, *6*(1), 47–61.

Bernard, H. R. (2000). *Social research methods: Qualitative and quantitative approaches*. California: Sage Publications, Inc.

Boa, E. A., Wattanatorn, A., & Tagong, K. (2018). The development and validation of the blended socratic method of teaching (BSMT): An instructional model to enhance critical thinking skills of undergraduate business students. *Kasetsart Journal of Social Sciences*, *39*(1), 81–89.

Brahler, C. J., Quitadamo, I. J., & Johnson, E. C. (2002). Student critical thinking is enhanced by developing exercise prescriptions using online learning modules. *Advances in Physiology Education*, *26*(3), 210–221.

Budsankom, P., Sawangboon, T., Damrongpanit, S., & Chuensirimongkol, J. (2015). Factors affecting higher order thinking skills of students: A meta-analytic structural equation modeling study. *Educational Research and Review*, *10*(19), 2639–2652.

Canziani, B., & Tullar, W. L. (2017). Developing critical thinking through student consulting projects. *Journal of Education for Business, 92*(6), 271–279.

Casem, M. L. (2006). Student perspectives on curricular change: Lessons from an undergraduate lower-division biology core. *CBE - Life Science Education*, *5*, 65–75.

Chairam, S., Klahan, N., & Coll, R. K. (2015). Exploring secondary students' understanding of chemical kinetics through inquiry-based learning activities.

Eurasia Journal of Mathematics, Science and Technology Education, *11*(5), 937–956.

Chairam, S., Somsook, E., & Coll, R. K. (2009). Enhancing Thai students' learning of chemical kinetics. *Research in Science & Technological Education*, *27*(1), 95–115.

Chiappetta, E. L. (1997). Inquiry-based science: Strategies and techniques for encouraging inquiry in the classroom. *The Science Teacher*, *64*(7), 22–26.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. California: Sage Publications, Inc.

Darus, F. B., & Saat, R. M. (2014). How do primary school students acquire the skill of making hypothesis. *The Malaysian Online Journal of Educational Science*, *2*(2), 20–26.

Demircioglu, T., & Ucar, S. (2015). Investigating the effect of argument-driven inquiry in laboratory instruction. *Educational Sciences: Theory & Practice*, *15*(1), 267–283.

Duran, M., & Dökme, I. (2016). The effect of the inquiry-based learning approach on student's critical-thinking skills. *Eurasia Journal of Mathematics, Science and Technology Education*, *12*(12), 2887–2908.

Ekahitanond, V. (2013). Promoting university students' critical thinking skills through peer feedback activity in an online discussion forum. *Alberta Journal of Educational Research*, *59*(2), 247–265.

Ennis, R. H. (2011). *The nature of critical thinking: An outline of critical thinking dispositions*. Retrieved from http://faculty.education.illinois.edu/rhennis/documents/TheNatureof CriticalThinking_51711_000.pdf

Facione, P. A. (2011). *Critical thinking: What it is and why it counts*. Millbrae, CA: Measured Reasons and The California Academic Press.

Feyzioglu, B., Demirdağ, B., Akyıldız, M., & Altun, E. (2012). Developing a science process skills test for secondary students: Validity and reliability study. *Educational Science Theory and Practice*, *12*(3), 1899–1906.

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

Gehring, K. M., & Eastman, D. A. (2008). Information fluency for undergraduate biology majors: Applications of inquiry-based learning in a developmental biology course. *CBE Life Sciences Education*, *7*, 54–63.

Green, S. B., & Salkind, N. J. (2008). *Using SPSS for windows and macintosh (Analyzing and understanding data)*. New Jersey: Pearson Prentice Hall.

Greenwald, R. R., & Quitadamo, I. J. (2014). A mind of their own: Using inquiry-based teaching to build critical thinking skills and intellectual engagement in an undergraduate neuroanatomy course. *The Journal of Undergraduate Neuroscience Education*, *12*(2), A100–A106.

Gultepe, N., & Kilic, Z. (2015). Effect of scientific argumentation on the development of scientific process skills in the context of teaching chemistry. *International*

Journal of Environmental and Science Education, *10*(1), 111–132.

Gupta, T., Burke, K. A., Mehta, A., & Greenbowe, T. J. (2015). Impact of guided-inquiry-based instruction with a writing and reflection emphasis on chemistry students' critical thinking abilities. *Journal of Chemical Education*, *92*(1), 32–38.

Hake, R. R. (1999). *Analyzing change/gain scores*. Retrieved from http://www.physics.indiana.edu/~sdi/AnalyzingChange-Gain.pdf

Hardianti, T., & Kuswanto, H. (2017). Difference among levels of inquiry: Process skills improvement at senior high school in Indonesia. *International Journal of Instruction*, *10*(2), 119–130.

Hsiao, H.-S., Chen, J.-C., Hong, J.-C., Chen, P.-H., Lu, C.-C., & Chen, S. Y. (2017). A five-stage prediction-observation-explanation inquiry-based learning model to improve students' learning performance in science courses. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(7), 3393–3416.

Hugerat, M., & Kortam, N. (2014). Improving higher order thinking skills among freshmen by teaching science through inquiry. *Eurasia Journal of Mathematics, Science and Technology Education*, *10*(5), 447–454.

Irwanto, Rohaeti, E., & Prodjosantoso, A. K. (2018a). A survey analysis of pre-service chemistry teachers' critical thinking skills. *MIER Journal of Educational Studies, Trends & Practices*, *8*(1), 57–73.

Irwanto, Rohaeti, E., & Prodjosantoso, A. K. (2018b). Undergraduate students' science process skills in terms of some variables: A perspective from Indonesia. *Journal of Baltic Science Education, 17*(5), 751–764.

Irwanto, Rohaeti, E., & Prodjosantoso, A. K. (2018c). The investigation of university students' science process skills and chemistry attitudes at the laboratory course. *Asia-Pacific Forum on Science Learning and Teaching, 19*(2), [7].

Irwanto, Rohaeti, E., & Prodjosantoso, A. K. (in press). Analyzing the relationships between pre-service chemistry teachers' science process skills and critical thinking skills. *Journal of Turkish Science Education*.

Irwanto, Rohaeti, E., Widjajanti, E., & Suyanta. (2017a). Students' science process skill and analytical thinking ability in chemistry learning. In *AIP Conference Proceedings* (Vol. 1868, pp. 1–5). https://doi.org/10.1063/1.4995100.

Irwanto, Rohaeti, E., Widjajanti, E., & Suyanta. (2017b). The development of an integrated assessment instrument for measuring analytical thinking and science process skills. In *AIP Conference Proceedings* (Vol. 1847, pp. 1–6). https://doi.org/10.1063/1.4983907.

Irwanto, Saputro, A. D., Rohaeti, E., & Prodjosantoso, A. K. (2018). Promoting critical thinking and problem solving skills of pre-service elementary teachers through process-oriented guided-inquiry learning (POGIL). *International Journal of Instruction, 11*(4), 777–794.

Ješková, Z., Lukáč, S., Hančová, M., Šnajder, Ľ., Guniš, J., Balogová, B., & Kireš, M.

(2016). Efficacy of inquiry-based learning in mathematics, physics and informatics in relation to the development of students' inquiry skills. *Journal of Baltic Science Education*, *15*(5), 559–574.

Karsli, F., & Şahin, Ç. (2009). Developing worksheet based on science process skills: Factors affecting solubility. *Asia-Pacific Forum on Science Learning and Teaching*, *10*(1), [15].

Ketpichainarong, W., Panijpan, B., & Ruenwongsa, P. (2010). Enhanced learning of biotechnology students by an inquiry-based cellulase laboratory. *International Journal of Environmental & Science Education*, *5*(2), 169–187.

Koc, Y., Doymus, K., Karacop, A., & Simşek, Ü. (2010). The effects of two cooperative learning strategies on the teaching and learning of the topics of chemical kinetics. *Journal of Turkish Science Education*, *7*(2), 52–65.

Konur, K. B., & Yıldırım, N. (2016). Pre-service science and primary school teachers' identification of scientific process skills. *Universal Journal of Educational Research*, *4*(6), 1273–1281.

Löfgren, R., Schoultz, J., Hultman, G., & Björklund, L. (2013). Exploratory talk in science education: Inquiry-based learning and communicative approach in primary school. *Journal of Baltic Science Education*, *12*(4), 482–496.

Lord, T., & Orkwiszewski, T. (2006). Moving from didactic to inquiry-based instruction in a science laboratory. *American Biology Teacher*, *68*(6), 342–345.

Lujan, H. L., & DiCarlo, S. E. (2006). Too much teaching, not enough learning: What is the solution? *Advances in Physiology Education*, *30*(1), 17–22.

Mahanal, S., Zubaidah, S., Bahri, A., & Dinnurriya, M. S. (2016). Improving students' critical thinking skills through remap NHT in biology classroom. *Asia-Pacific Forum on Science Learning and Teaching*, *17*(1), [11].

Maxwell, D. O., Lambeth, D. T., & Cox, J. T. (2015). Effects of using inquiry-based learning on science achievement for fifth-grade students. *Asia-Pacific Forum on Science Learning and Teaching*, *16*(1), [2].

Molefe, L., Stears, M., & Hobden, S. (2016). Exploring student teachers' views of science process skills in their initial teacher education programmes. *South African Journal of Education*, *36*(3), 1–12.

Muhlisin, A., Susilo, H., Amin, M., & Rohman, F. (2016). Improving critical thinking skills of college students through RMS model for learning basic concepts in science. *Asia-Pacific Forum on Science Learning and Teaching*, *17*(1), [12].

Mutlu, A., & Acar-Şeşen, B. (2018). Pre-service science teachers' understanding of chemistry: A factorial design study. *Eurasia Journal of Mathematics, Science & Technology Education*, *14*(7), 2817–2837.

National Council for Excellence in Critical Thinking [NCECT]. (2013). *Critical thinking*. Retrieved from http://www.criticalthinking.org/pages/thenational-council-for-excellence-in-critical-thinking/406.

Oliver-Hoyo, M. T. (2003). Designing a written assignment to promote the use of critical thinking skills in an introductory chemistry course. *Journal of Chemical Education*, *80*(8), 899–903.

Ongowo, R. O. (2017). Secondary school students' mastery of integrated science process skills in Siaya county, Kenya. *Creative Education*, *8*, 1941–1956.

Osman, K., & Vebrianto, R. (2013). Fostering science process skills and improving achievement through the use of multiple media. *Journal of Baltic Science Education*, *12*(2), 191–204.

Ozdem-Yilmaz, Y., & Cavas, B. (2016). Pedagogically desirable science education: Views on inquiry-based science education in Turkey. *Journal of Baltic Science Education*, *15*(4), 506–522.

Ozgür, S. D., & Yilmaz, A. (2017). The effect of inquiry-based learning on gifted and talented students' understanding of acids-bases concepts and motivation. *Journal of Baltic Science Education*, *16*(6), 994–1008.

Padilla, M. J. (1990). Science process skills. National association of research in science teaching publication: Research matters - to the science teacher (9004). Retrieved from http://www.narst.org/publications/research/skill.cfm.

Pradina, L. P., & Suyatna, A. (2018). Atom core interactive electronic book to develop self efficacy and critical thinking skills. *The Turkish Online Journal of Educational Technology*, *17*(1), 17–23.

Quitadamo, I. J., Faiola, C. L., Johnson, J. E., & Kurtz, M. J. (2008). Community-based inquiry improves critical thinking in general education biology. *CBE Life Sciences Education*, *7*, 327–337.

Rissing, S. W., & Cogan, J. G. (2009). Can an inquiry approach improve college student learning in a teaching laboratory? *CBE Life Sciences Education*, *8*, 55–61.

Sağlam, M. K., & Şahin, M. (2017). Inquiry-based Professional Development Practices for Science Teachers. *Inquiry-Based Professional Development Practices for Science Teachers*, *14*(4), 66–76.

Sari, S. A., Jasmidi, Kembaren, A., & Sudrajat, A. (2018). The impacts of chemopoly-edutainment to learning activities and responses. *Journal of Education and Learning*, *12*(2), 311–318.

Scriven, M., & Paul, R. (1987). Defining critical thinking. In *The 8th Annual International Conference on Critical Thinking and Education Reform*. Retrieved from http://www.criticalthinking.org/pages/defining-critical-thinking/766

Sen, C., & Vekli, G. S. (2016). The impact of inquiry based instruction on science process skills and self-efficacy perceptions of pre-service science teachers at a university level biology laboratory. *Universal Journal of Educational Research*, *4*(3), 603–612.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. New York: Houghton Mifflin Company.

Škoda, J., Doulík, P., Bílek, M., & Šimonová, I. (2015). The effectiveness of inquiry based science education in relation to the learners' motivation types. *Journal of Baltic Science Education*, *14*(6), 791–803.

Tatar, N. (2012). Inquiry-based science laboratories: An analysis of preservice teachers' beliefs about learning science through inquiry and their performances. *Journal of Baltic Science Education*, *11*(3), 248–267.

Wang, C.-H. (1993). Improving chemistry experimental learning through process-mapping-hydrolysis of p-nitroacetanilide-. *Bull. Nat'l Taiwan Nor. Univ.*, *38*, 157–173.

Wartono, W., Hudha, M. N., & Batlolona, J. R. (2018). How are the physics critical thinking skills of the students taught by using inquiry-discovery through empirical and theorethical overview? *Eurasia Journal of Mathematics, Science and Technology Education*, *14*(2), 691–697.

Yakar, Z., & Baykara, H. (2014). Inquiry-based laboratory practices in a science teacher training program. *Eurasia Journal of Mathematics, Science and Technology Education*, *10*(2), 173–183.

Yang, H., & Park, J. (2017). Identifying and applying factors considered important in students' experimental design in scientific open inquiry. *Journal of Baltic Science Education*, *16*(6), 932–945.

Žoldošová, K., & Matejovičová, I. (2010). Finding out how the elementary school children manipulate with empirical material and how they process the obtained information. *International Electronic Journal of Elementary Education*, *2*(3), 327–347.

Eurasian Journal of Educational Research 80 (2019) 171-192

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

Implementation of Cognitive Diagnosis Modeling using the GDINA R Package

Jimmy de la TORRE[1], Lokman AKBAY[2]

ARTICLE INFO

Article History:
Received: 28 Dec. 2018
Received in revised form: 27 Feb. 2019
Accepted: 08 Mar. 2019
DOI: 10.14689/ejer.2019.80.9

Keywords
CDM, cognitive diagnosis modeling, GDINA package, R implementation

ABSTRACT

Purpose: Well-designed assessment methodologies and various cognitive diagnosis models (CDMs) to extract diagnostic information about examinees' individual strengths and weaknesses have been developed. Due to this novelty, as well as educational specialists' lack of familiarity with CDMs, their applications are not widespread. This article aims at presenting the fundamentals of CDM and demonstrating the various implementations using a freeware R package, namely, the GDINA. Present article explains the basics of CDM and provide sufficient details on the implementations so that it may guide novice researchers in CDM applications.

Research Methods: The manuscript starts with presenting the CDM terminology, including input and output of a CDM analysis. The introduction section is followed by generalized deterministic noisy and gate model framework. A brief description of the package GDINA is also provided. Then, numerical examples on various CDM analyses are provided using the R package with a graphical user interface. The paper is concluded by some additional functions and concluding remarks. **Results and Implications for Research and Practice:** Although other software programs are also available, using the GDINA package offers users some flexibilities such as allowing estimation of a wide range of CDMs and allowing nonprogrammers to benefit from this package through the GUI. In addition to ordinary CDM analyses, GDINA package further allows users to apply model selection at the test- and item-level to make sure that the most appropriate CDM (i.e., CDM that best explains the attribute interactions in the item) is fitted to the response data. Furthermore, to identify possible item-attribute specification mistakes in the Q-matrix, implementation of an empirical Q-matrix validation method is available in the GDINA package. Lastly, this package offers various handy graphs, which can be very useful in emphasizing important information and comparing various parameters and/or statistics.

[1] The University of Hong Kong, HONG KONG, e-mail: j.delatorre@hku.hk, ORCID: https://orcid.org/0000-0002-0893-3863
[2] Corresponding author, Burdur Mehmet Akif Ersoy University, TURKEY, e-mail: lokmanakbay@gmail.com, ORCID: https://orcid.org/0000-0003-4026-5241

Introduction

Assessments, when purposefully designed, can provide opportunities for collecting relevant diagnostic information. Interpretations based on such information then enable educational specialists to take precise remedial actions. Recently developed formative assessments are expected to provide students and teachers with detailed feedback on what students are able to do yielding information that can optimize instruction and learning. In other words, a formative assessment should identify individual strengths and weaknesses in a particular content, which results in enhanced teaching and learning environment (DiBello & Stout, 2007). For formative assessment to fulfil this task, in addition to well-designed assessment methodologies, various cognitive diagnosis models (CDMs) to extract diagnostic information from diagnostic assessments are needed. These models are regarded as latent class models, which can be used to detect mastery and nonmastery of multiple fine-grained skills or attributes in a particular content domain (de la Torre, 2009).

Although there are several software programs available to estimate CDMs, there are considerable amounts of benefits of using R for estimation purposes: (1) Conducting estimation of various CDMs are available in R; (2) Unlike the R packages, many software programs handle only one type of CDM (e.g., MDLTM for general diagnostic model [GDM; von Davier, 2006], Arpeggio Suite for noncompensatory-RUM [NC-RUM; Hartz, 2002], Mplus for log-linear CDM [LCDM; Henson, Templin & Willse, 2009]); (3) Many software programs are either commercial or only obtained by contacting to authors; and (4) Syntax preparation for some of these programs may require substantial effort. Although several R packages available for the CDM analyses; one of the most comprehensive packages is the GDINA package.

This article aims at explaining the fundamentals of CDMs as well as demonstrating the various implementations using GDINA package (Ma & de la Torre, 2018). The current article explains the basics of CDM and provides sufficient details on the implementations and may be used to guide novice researchers in CDMs related studies. We start by introducing CDM terminologies, and input and output of a CDM analysis. This will be followed by presentation of the G-DINA model framework including several specific CDMs, which can be derived from the G-DINA. A brief description of the package GDINA will then be provided. The fourth section will demonstrate how various analyses are conducted using the R package with a graphical user interface. In the fifth section, the paper summarizes additional features of the GDINA R package. The final section will provide some concluding remarks.

Input and Output in a CDM Analysis

Two input matrices are needed for a basic specification of a CDM. The first matrix consists of examinees' item responses, and may be called response matrix. This matrix is composed of examinees' binary (in the simplest form) responses to items on a test. This is typically an IxJ matrix $\boldsymbol{X}$, where the element x_{ij} indicates whether examinee i correctly responded item j ($x_{ij} = 1$) or not ($x_{ij} = 0$). The second matrix, which specifies relationship between each item on a test and content related attributes, is called Q-matrix (Tatsuoka, 1983). For instance, for a test consisting of $j = 1,2, \ldots, J$ items and

measuring $k = 1,2,\ldots,K$ attributes, the elements of a IxJ Q-matrix are composed of binary variables, where $q_{jk} = 1$ indicates that examinees must have the *kth* attribute to be able to achieve a correct response to *jth* item. Similarly, element of $q_{jk} = 0$ indicates that the *kth* attribute is not required for a successful response to *jth* item.

The aim of conducting a CDM analysis is to be able to make inferences about examinees' mastery status of each of the K attributes. Typically, but not always, attributes are dichotomous, and the K attributes form 2^K attribute patterns, which are called latent classes and denoted as $\boldsymbol{\alpha}_l = [\alpha_{l1},\ldots,\alpha_{lK}]$, where $l = 1,\ldots,2^K$. Each element $\boldsymbol{\alpha}_{lk}$ indicates whether members of latent class $\boldsymbol{\alpha}_l$ possess *kth* attribute. At the end of the analysis, each examinee is assigned with an attribute profile indicating which specific attributes the examinee has and has not mastered. In addition, the analysis provides information on (1) proportion of examinees mastered a specific attribute, and (2) proportion of examinees within each latent class.

Generalized DINA Model Framework

A wide range of saturated and reduced CDMs have been introduced in the literature. One way of distinguishing these various models pertains to attribute-effects considered in the response process. In a general (i.e., saturated) model, all main and interaction effects of measured attributes contribute to the item response function. The *generalized deterministic, noisy and gate* model (G-DINA: de la Torre, 2011) is a general CDM from which, more specific models can be derived. For example, *deterministic input, noisy "and" gate* model (DINA: Junker & Sijtsma, 2001), *deterministic input, noisy "or" gate* model (DINO: Templin & Henson, 2006), and Additive-CDM (ACDM: de la Torre, 2011) are derived from the G-DINA model.

Within the G-DINA model framework, a baseline probability (the probability of success when an examinee has not mastered any required attributes for item *j*), main effect terms (change in the success probability when a required attribute is mastered) and possible interaction terms (change in the success probability when more than one attribute is mastered) are specified. To derive the A-CDM, DINA model and DINO model from the G-DINA model, one needs to place specific constraints on the G-DINA item response function. Specifically, to obtain the A-CDM, all interaction terms in the G-DINA model item response function are set to zero. Likewise, to derive the DINA model, all item parameters but the baseline and highest order interaction are set to zero. Finally, to obtain the DINO model, the main and interaction effects are constrained to be equal with alternating signs. Above constraints result in $K_j^* + 1$ item parameters for the A-CDM and just two item parameters for the DINA and DINO models.

As explained in de la Torre (2011), in the specification of general CDMs, several link functions may be used. In saturated forms, the G-DINA model can be expressed using the *identity* link, *logit* link, and *log* link, all of which provide identical model-data fit. *Additive* models of log-linear CDM and log CDM models are the *linear logistic model* (LLM; Maris, 1999) and the *reduced reparametrized unified model* (R-RUM; Hartz, 2002),

respectively. Although ACDM, LLM, and R-RUM have the same number of item parameters, unlike their general models, they do not provide identical model-data fit as they assume different underlying processes. For details on saturated and reduced models and their parameter estimations, readers may refer to de la Torre (2011).

GDINA Package

The package 'GDINA' composed of a set of psychometric tools for cognitive diagnosis modeling. It has the capability of handling both dichotomous and polytomous response data. The G-DINA model, the sequential G-DINA (Ma & de la Torre, 2016) and many CDMs subsumed by these two can be estimated using this package. More specifically, in addition to the models subsumed by the G-DINA model, (i.e., the DINA, DINO, A-CDM, LLM, R-RUM), *multiple-strategy DINA model* [de la Torre & Douglas, 2008], many extensions of the G-DINA model such as the sequential G-DINA model for ordinal and nominal data [Ma & de la Torre, 2016], polytomous G-DINA model for polytomous attributes [Chen & de la Torre, 2013], multiple group G-DINA model for individuals from multiple groups [Ma, Terzi, Lee & de la Torre, 2017], and diagnostic tree model for polytomous response data with multiple strategies [Ma, 2018] can also be handled by the GDINA package. Marginal maximum likelihood estimation with expectation-maximization algorithm (MMLE/EM) is used for item parameter estimation. This package allows user to assign different CDMs to different items in a single test. It is also flexible in terms of handling independent, saturated, higher-order, and structured joint attribute distributions. Offering a *graphical user interface* is a notable feature of the package. Along with providing person and item parameter estimates and model-fit statistics, this package allows users to conduct various analyses including Q-matrix validation, item-fit evaluation, item and test-level model comparisons, and differential item functioning.

Demonstrations

To be able to follow the demonstrations below, R software (R 3.5.0 or higher version) along with R studio (a graphic user interface for R) need to be downloaded from https://cran.r-project.org/ and https://rstudio/com, respectively. After installing R and Rstudio, the GDINA package (version 2.2.0) with its all dependencies must be installed. Then, command > library(GDINA) is run to start working with the GDINA package. To help nonprogrammer users, we intend to apply the main functions of the package using *Shiny*, which is a web application framework that is used to build interactive web applications directly from R. To use Shiny with the GDINA function, one need to run the command > startGDINA(). Then the GDINA graphical user interface (GUI) opens in a new window.

Reading Input Data

Table 1
Generating Q-matrix

Item	A1	A2	A3	A4	A5	A6	Item	A1	A2	A3	A4	A5	A6
1	1	0	0	0	0	0	*10*	0	0	0	1	1	0
2	0	1	0	0	0	0	*11*	0	0	0	0	1	1
3	0	0	1	0	0	0	*12*	1	0	0	0	0	1
4	0	0	0	1	0	0	*13*	1	1	1	0	0	0
5	0	0	0	0	1	0	*14*	0	1	1	1	0	0
6	0	0	0	0	0	1	*15*	0	0	1	1	1	0
7	1	1	0	0	0	0	*16*	0	0	0	1	1	1
8	0	1	1	0	0	0	*17*	1	0	0	0	1	1
9	0	0	1	1	0	0	*18*	1	1	0	0	0	1

For simulation purposes, given a uniformly distributed attribute patterns, we generated a response data set based on the GDINA, ACDM, and DINA models. To generate the data, we used a hypothetical Q-matrix given in Table 1. Eighteen items requiring one, two, or three attributes constitute the Q-matrix measuring a total of six attributes. Thus, in the data generation, 1000 attribute patterns from 2^6 possible attribute patterns were drawn from a uniform distribution. Then, based on this sample, response data were generated following the DINA model, additive model, and G-DINA model. Responses for items with multiple attributes distributed among the three generating models. More specifically, items seven to 14 followed the DINA model, items 15 and 16 followed the ACDM, and last two items followed the G-DINA model. For the data generation purposes, lower and upper bound of probability corrects were drawn from a random uniform distribution of U(.05, .20). Then, final response data involved 1000 examinees' binary responses to 18 items. The R code used for data generation is given in Figure 1.

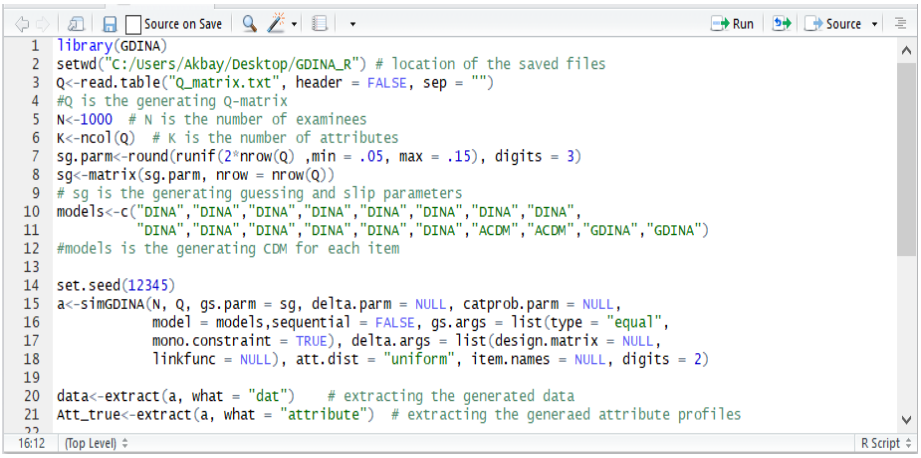


Figure 1. R Code for Data Generation

To read the input data (i.e., response data and Q-matrix) by the GUI, we need to go to the input tab and select our files from the folder where they are kept. There are three panels in this window. In the upper panel, shown in Figure 2, input files are read. Data in the input files need to be separated with one of the following: Tab, comma, semicolon, or space. If our input file has header, *Header* box must be checked. The remaining two panels (not given in the figure) are used to check whether the data are read properly. Specifically, middle and lower panels show the first six rows of the response data and the Q-matrix, respectively.

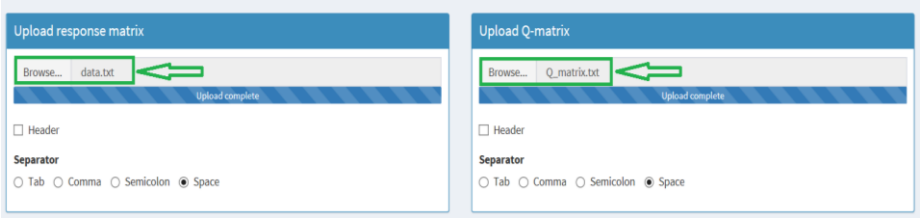


Figure 2. Upper Panel of the Input Tab.

Model Estimation

In many cases, item parameters along with examinees' attribute patterns are not known. In this situation, item parameters and examinee attribute profiles are estimated together. To fulfill this, GDINA package employs marginal maximum likelihood estimation (MMLE) through expectation-maximization (EM) algorithm. The *Estimation Settings* tab of the GUI provide us with several lists and options to select from. This tab is given in Figure 3. First, we need to select a CDM to fit among the eight available models: GDINA, logit GDINA, log GDINA, DINA, DINO, ACDM, LLM, and RRUM. There is a ninth option here, which allows user to specify a vector of models (comma delimited without quotation mark).

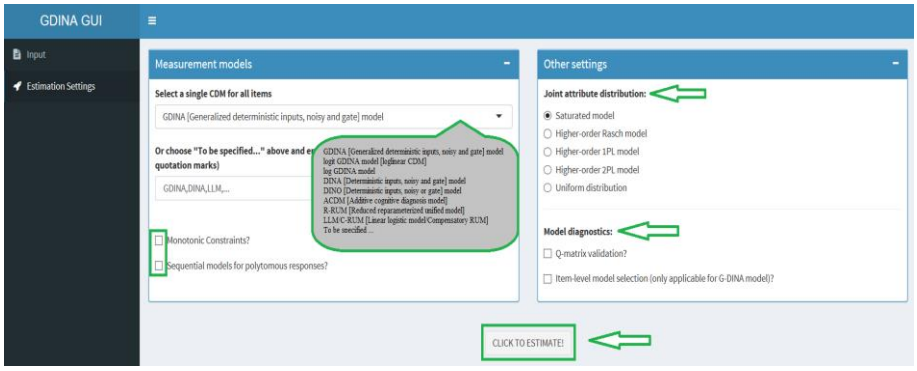


Figure 3. Estimation Specifications

Then, we need to make a decision on the attribute distribution. There are five options available in the GUI: Saturated, higher-order Rasch, higher-order 1PL, higher-order 2PL, and uniform distribution. The three higher-order attribute distribution types are defined based on the item response models for continuous ability. One- and

two-parameter logistic models, as well as the Rasch model are available options here. In addition, there are four more *optional* settings. We need to check the boxes corresponding to options such as applying *monotonic constraints* (i.e., mastering an additional attribute does not lead to a lower success probability) to the fitted model, *Q-matrix validation*, and *item-level model selection*. We then run model estimation by clicking the *Click to Estimate!* button.

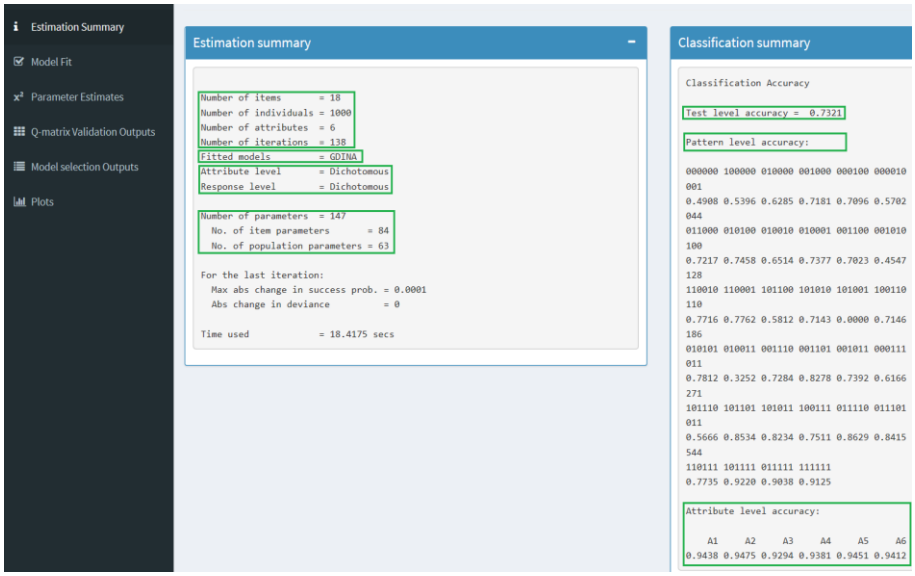


Figure 4. Estimation Summary

In this specific example, we fitted the GDINA model with saturated attribute distribution and applied monotonic constraints. We further requested Q-matrix validation and item-level model selection. Estimation results are found in the *Estimation Summary* tab when the estimation is completed. As shown in Figure 4, there are two panels on this tab. The left panel shows the estimation summary and right reports the classification summary. Estimation summary starts with some descriptive statistics such as the number of items, number of examinees, number of attributes, and number of iterations needed for the model to converge. For this specific example, 18 items, 1000 examinees, six measured attributes, and 138 iterations are reported. Then, information on the fitted model by item, attribute type, and response type are reported. For our example, we see that G-DINA model fitted to all items, and both attribute and response data are dichotomous.

In the estimation summary further information about the number of total, item, and person parameters are also displayed. Sixty-three person parameters and 84 item parameters were estimated. Notice that because we have a total number of six attributes to measure, in saturated attribute distribution, we must estimate 63 (2^6-1) latent classes, as well as all item parameters defined by the fitted model (84 delta

parameters defined by the GDINA model). Test and pattern level classification accuracy rates are reported on the right side under the classification summary panel.

Model-fit evaluation

Estimated CDM parameters are interpretable to the extent that model explains the data (Chen, de la Torre & Zhang, 2013). Thus, inferences based on any CDM are valid as long as the model fits to data. Data-model fit can be checked by examining *relative-* and *absolute- fit* statistics. The former statistics indicate the most appropriate model when competing models are available, whereas the latter shows whether model itself fits the data. Relative and absolute fit statistics are found under the *Model Fit* tab of the GUI. As shown in Figure 5, the relative test fit statistics, log-likelihood, Akaike information criterion (AIC; Akaike, 1974) and Bayesian information criterion (BIC; Schwarz, 1976) with the corresponding penalties are given on the relative test fit window.

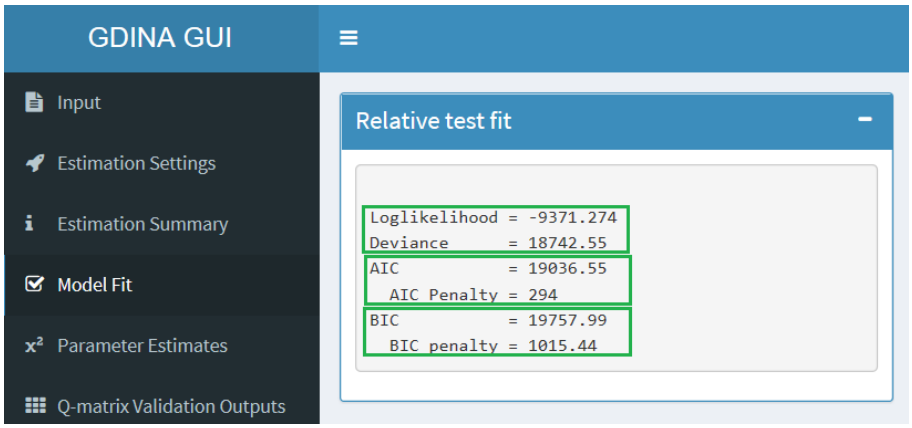


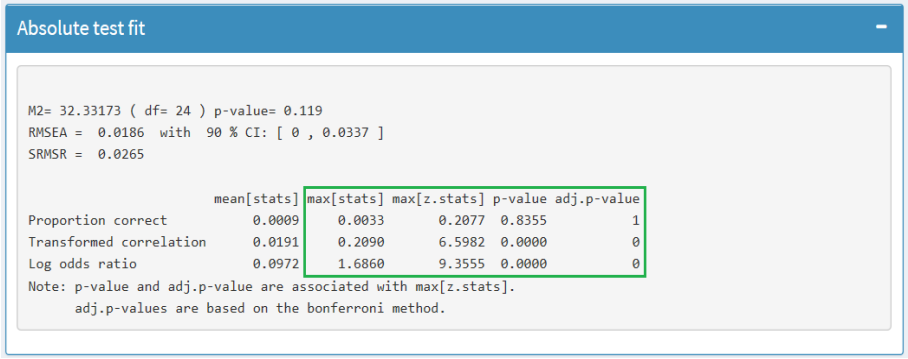
Figure 5. Relative Fit Statistics

When we have rival models that may fit the data, we need to check relative fit statistics to compare these models. To do so, we can look at the AIC and BIC values of nonnested models in which, the model with the least information criteria is preferred. When rival models are nested (e.g., GDINA vs. DINA), one can make a decision based on a likelihood ratio test: $LR = -2[LL_{reduced} - LL_{saturated}]$, where LR is the likelihood ratio of saturated and reduced models, which is compared against χ^2 distribution with a given significance level (e.g., .05) and degrees of freedom determined by the parameters difference in the compared models. Our null hypothesis for this test is the reduced model fits the data as good as the full model, and we reject the null when LR is larger than the critical χ^2 value.

Relative fit statistics of our simulation is given in Figure 5. For demonstration purposes, we also fitted the DINA model to the data that we have been analyzing. The log-likelihood, AIC, and BIC results when the DINA is fitted are: -9674.58, 19547.15, and 20033.02, respectively. AIC and BIC reported in Figure 5 are lower than these values obtained from the DINA model fit. Furthermore, computed $LR =$

$-2[-9674.58 + 9371.27] = 606.62$, given the model parameters 99 and 147 for the DINA and GDINA models, need to be compared against $\chi^2_{(.05;48)} = 65.17$. Because the computed LR is larger than the critical value, we reject the null hypothesis and conclude that the GDINA model fits the data better than the DINA model.

For absolute fit evaluation, GDINA package offers three statistics: proportion correct (*p*), log-odds ratio (*l*), and transformed correlation (*r*). Specifically, *p* stands for residuals between the observed and predicted proportion correct of individual items; *l* is computed based on log-odds ratio of item pairs; and *r* is computed based on the residuals between the observed and predicted correlations of item pairs (Chen, de la Torre & Zhang, 2013; de la Torre & Douglas, 2008; Sinharay & Almond, 2007). When these statistics are close to zero, it indicates that the model fits the data.


Absolute test fit

M2= 32.33173 (df= 24) p-value= 0.119
RMSEA = 0.0186 with 90 % CI: [0 , 0.0337]
SRMSR = 0.0265

	mean[stats]	max[stats]	max[z.stats]	p-value	adj.p-value
Proportion correct	0.0009	0.0033	0.2077	0.8355	1
Transformed correlation	0.0191	0.2090	6.5982	0.0000	0
Log odds ratio	0.0972	1.6860	9.3555	0.0000	0

Note: p-value and adj.p-value are associated with max[z.stats].
adj.p-values are based on the bonferroni method.

Figure 6. Absolute Fit Statistics

To use *p*, *l*, and *r*, their standard errors are also needed. Please refer to Chen, de la Torre and Zhang (2013) for derivation of the statistics and their respective standard errors. The *z*-scores of the statistics are used to test whether the residuals are significantly different from zero. Because *J*-1 proportion correct and *J*(*J*-1)/2 pairwise log-odds ratio and transformed correlations are available for a test with *J* items, only the maximum z-scores of each statistics are tested and the test results are provided by the GUI under the model fit tab. Here in our example, summary of absolute fit analysis is provided in Figure 6.

Means, maximums, maximum of z-scores, and corresponding p-values and adjusted p-values are reported in the figure. When we check maximum z-statistics for all three absolute fit statistics and corresponding p- and adj.p-values, we reject the null hypotheses that the model fit the data based on *l* and *r*. Only, *p* statistic retains the null hypothesis, which is claimed to be either unreliable or results in low rejection rates (Chen, de la Torre & Zhang, 2013). We should note there that adj.p-value is the bonferroni adjusted p-value as we implicitly test *J*-1 proportion correct and *J*(*J*-1)/2 pairwise log-odds ratio and transformed correlations by testing the maximum of the z-scores.

The above results indicate that there is at least one item that cannot be fitted to the model, which is expected because our data were generated based on the items

following the DINA, ACDM, and G-DINA models. However, if we fit DINA model to the data corresponding to the first 14 items using the first 14 rows of the Q-matrix, all three absolute fit statistics indicate that the model fit the data (p-values for the maximum z-statistics are .878, .133, and .124 for proportion correct, log-odds and transformed correlations, respectively. These results, therefore, indicate that if only the first 14 items confirming the DINA model constituted the test, all aforementioned absolute fit indices would yield model-data fit for all items. We should note that the GUI also provides a heatmap plot depicting the p-values log-odds and transformed correlation obtained from item pairs.

Parameter Estimation

Both the item and person parameters estimates are readily available under the *Parameter Estimates* tab of the GUI. This tab, given in Figure 7, offers a list of item parameter estimates with or without standard errors: *success probabilities of reduced latent classes*; *guessing and slip parameters*; *delta parameters;* and *success probabilities of all latent classes*. To obtain the standard errors along with the estimates, *Estimate S.E?* box must be checked. Figure 7 demonstrates that we requested guessing (*g*) and slip (*s*) parameters with corresponding standard errors. Under G-DINA model, *g* and 1-*s* parameters stand for lower and upper bound of success probabilities.

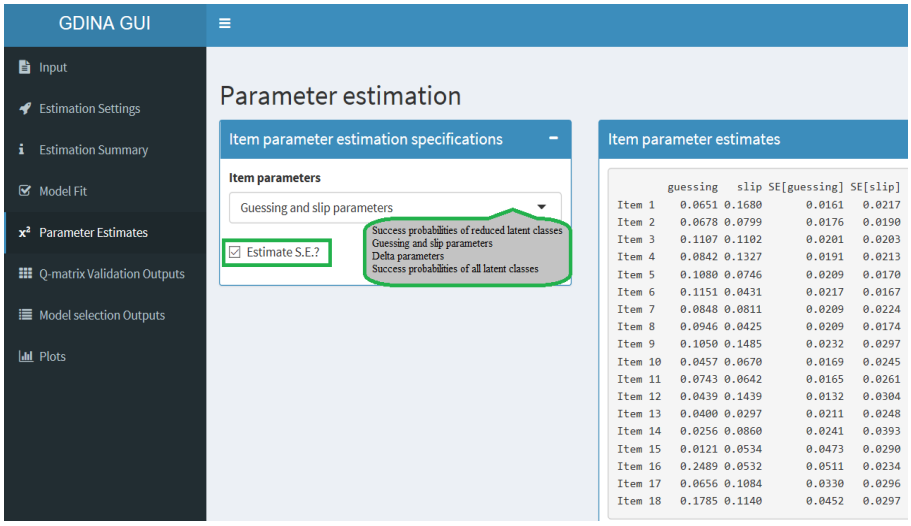


Figure 7. Item Parameter Estimation Specifications

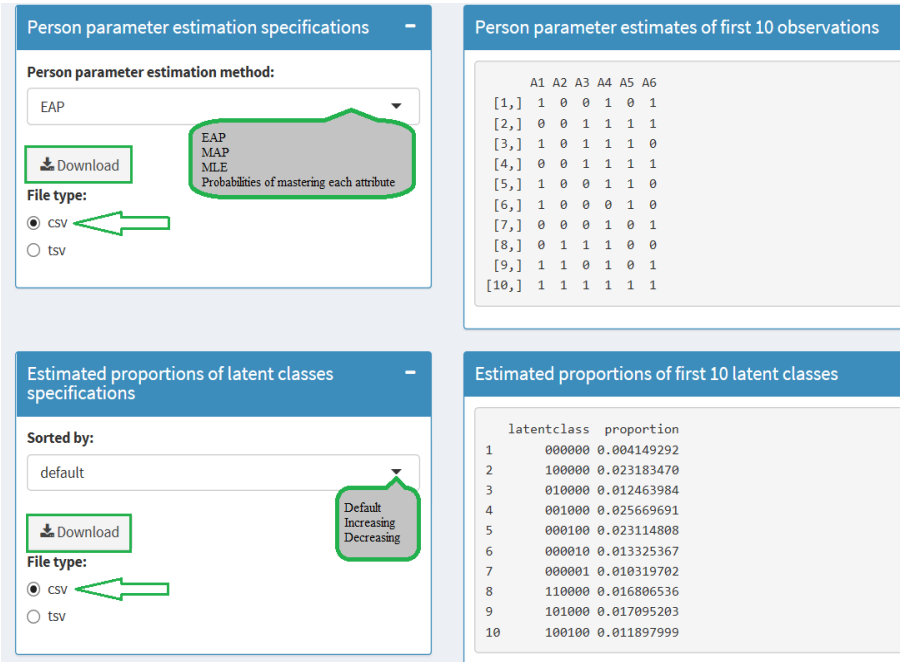


Figure 8. Person Parameter Estimation Specifications

In addition to item parameters, the *Parameter Estimates* tab of the GUI provides us with person parameters as well. Examinees' estimated attribute profiles for the first 10 examinees are shown in the middle panel of this tab. Attribute profile estimates based on maximum likelihood estimation (MLE), expected a priori (EAP), and maximum a priori (MAP) are computed and may be selected from the list given in the *Person parameter estimation method*. This list also includes the *Probabilities of mastering each attribute* option. In our example given in Figure 8, we see the first ten examinees' estimated attribute vectors. The GUI allows for all these types of person parameters to be downloaded as a .csv or .tsv file. To do so, we need to select one of the two file types, and hit the *Download* button.

Q-matrix Validation

In general, implementation of the CDMs requires a Q-matrix mapping attributes to items. This matrix embodies the cognitive specifications in test construction (Leighton, Gierl & Hunka, 2004). Because this input plays a crucial role in incorporating cognitive theories into psychometric practice, it needs to be correctly specified. Only then, CDMs can provide maximum diagnostic information (de la Torre, 2008). However, due to subjective judgments of content experts, Q-matrix construction process might not be fulfilled easily and successfully. This process may result in some misspecifications in the Q-matrix such that a required attribute may be missed, or an unnecessary attribute may be added in the matrix. These two types of

misspecifications are known as under- and over-specifications, respectively (de la Torre & Chiu, 2016).

Another possible error may emerge when both under- and over-specifications coexist for an item, which is referred to as over-and-under-specification (de la Torre & Chiu, 2016). Negative effects of misspecified Q-matrices in item calibration, examinee classification, and model misfit have been shown in the literature (e.g., Chiu, 2013; de la Torre, 2008; de la Torre & Chiu, 2016). To address misspecifications due to subjectivity in the process, a general Q-matrix validation procedure (i.e., de la Torre & Chiu, 2016) along with several others specifically developed for reduced models (e.g., Chiu, 2013; de la Torre, 2008) have been proposed.

An item specific discrimination index, ς_j^2, is proposed and used in the general Q-matrix validation method. ς_j^2 is a measure of weighted variance of success probability for a particular attribute distribution. De la Torre and Chiu (2016) showed that a correct q-vector yields homogenous latent groups in terms of probability of success. The idea behind use of this index for Q-matrix validation lies in the fact that correct q-vector should produce the highest variability in success probability. Their method suggests selecting a q-vector that approximates the maximum ς_j^2 with the fewest attribute specifications.

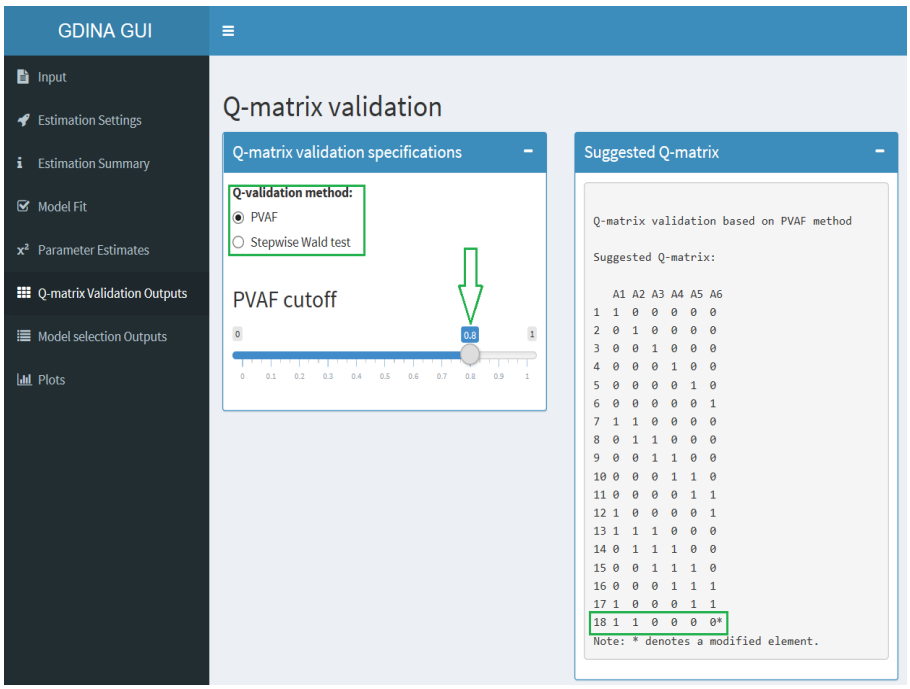


Figure 9. Setting PVAF Cutoff and Suggested Q-matrix

For the search algorithm to select a parsimonious q-vector producing approximately the highest ς_j^2 rather than the q-vector yielding the highest ς_j^2, a

stopping rule is implemented as a part of the algorithm. This stopping rule is defined by the proportion of variance accounted for (PVAF) by a particular q-vector relative to the highest ς_j^2. Because maximum ς_j^2 is produced by the q-vector requiring all measured attributes, a q-vector is selected by this method when PVAF is higher than a predetermined cutoff (i.e., ϵ).

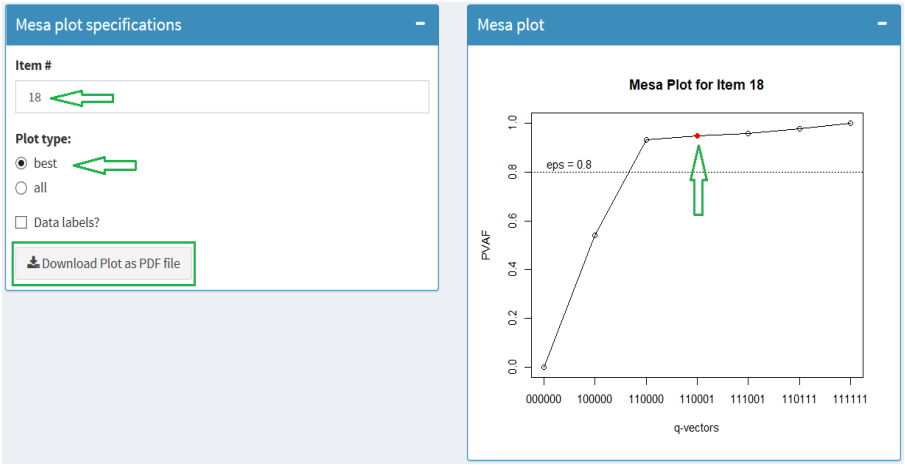


Figure 10. Mesa Plot with Specifications

The GDINA package implements the general Q-matrix validation method. Validation results (when requested in the estimation specifications tab) are found in the *Q-matrix Validation Outputs* tab in the GUI. On this output page, first PVAF cutoff needs to be specified so that suggested Q-matrix will be produced based on this cutoff. Here in our example given in Figure 9, we set the PVAF to .80 and the corresponding suggested Q-matrix yielded correction in one item (i.e., item 18). Stepwise implementation of the test is also available through the option *Stepwise Wald test*. Remember that although the validation method has suggested this q-vector, the final decision on correcting the Q-matrix should be based on the expert decision after careful consideration of related item, q-vector, and the attributes measured by the test.

Down on the same page, we are able to see Mesa Plot (de la Torre & Ma, 2018) for each item. Item 18 is shown in Figure 10 as an example, for which we plotted the *best q-vectors* option. We could also see *all q-vectors* option by selecting the plot type from *Mesa Plot Specifications*. PVAF values for this item indicates that a q-vector of 110000 is as good as the specified q-vector (110001). Notice that the original q-vector is marked with filled circle. Here we also have an option to print *data labels* by clicking the corresponding box in the specifications window. Furthermore, the GUI allows us to download Mesa plots as PDF file.

Item-level Model Comparison

GDINA package allows item-level model selection so that CDMs need not be specified a priori. To perform model selection between saturated and reduced models at the item-level, Wald test is used to compare G-DINA model against the fits of the

specific CDMs (i.e., DINA, DINO, ACDM, LLM, and RRUM). The Wald test proposed by de la Torre (2011) is an item level procedure, which can be performed for items requiring at least two attributes. Wald test requires a restriction matrix of **R** constraining the GDINA parameterization to derive the reduced models. This restriction matrix is used in computation of Wald statistic. For further information on the Wald statistics, readers may refer to de la Torre and Lee (2013) and Ma, Iaconangelo and de la Torre (2016).

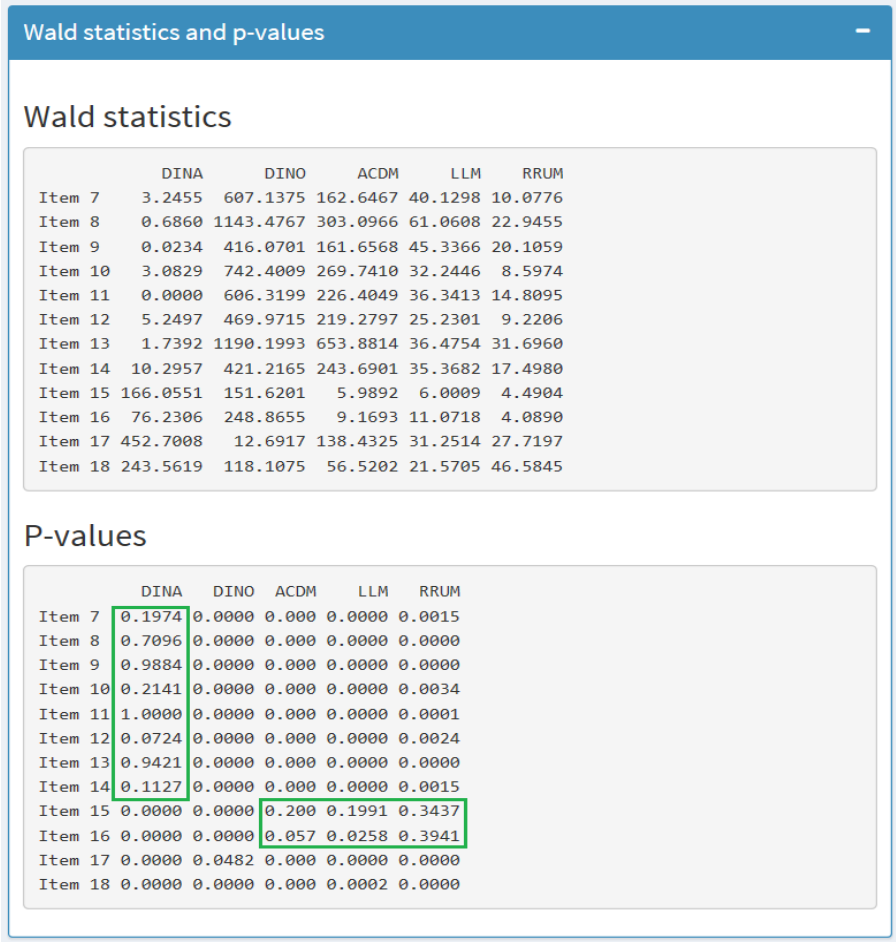


Figure 11. Item-level Model Comparison

For our case, the Wald statistics and corresponding p-values for the different reduced models, given in Figure 11, are reported under the *Model Selection Output* tab of the GUI. Notice that these statistics only printed for the items requiring at least two attributes for which more than one CDM can be fitted. Wald statistics and p-values indicates that items 7-14 are DINA and items 17-18 are the GDINA items. Remaining two items (i.e., 15-16) conform to at least two of the following CDMs: RRUM, ACDM

and LLM. Although the data for these two items are simulated using the ACDM, because these three models have the same number of item parameters under different *link functions* (see de la Torre, 2011 for different link functions), it may be hard to make a perfect selection.

Plots

The GUI allows users to generate various types of graphs via the *Plots* tab. These readily available graphs are useful tools to highlight important information and enable users to compare various parameters or statistics. For example, *Plot of probability of mastery for individuals* is given in Figure 12; which allows users to compare individuals' mastery probabilities. To do so, one need to enter a vector of individuals whose mastery probabilities are to be compared. This vector must be comma delimited, and no quatation marks is needed. Another type of graph is generated to display *individual posterior probabilities*. This bar graph allows user to display the mastery probabilities in increasing or decreasing order.

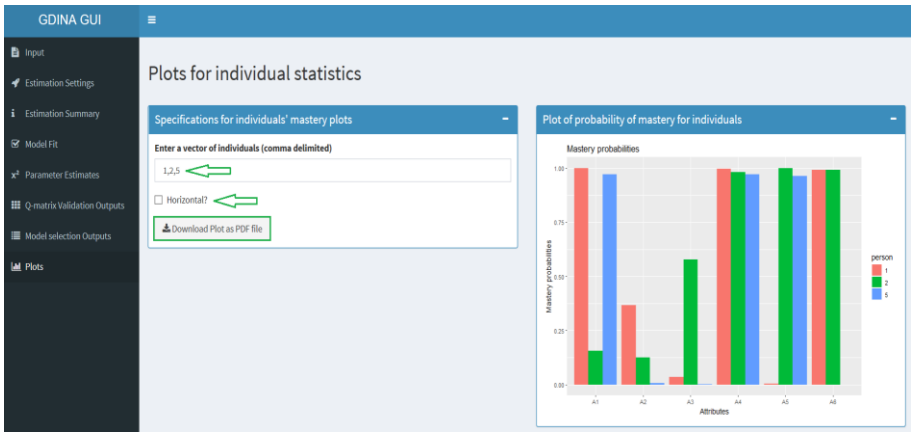


Figure 12. Individuals' Mastery Probability Plots

The next graph is the *Plot of proportions of latent classes*, which indicates the latent classes' mastery rates. The mastery proportions of the latent classes may also be displayed in decreasing or increasing order. All these graphs are displayed vertically as default; however, user can switch to the horizontal positioning by clicking the *Horizontal* box. The last graph is the *Item response function plot*, which displays the item statistics of the selected item. Error bars can also be added to this bar graph by clicking the *Error bars* box.

Additional Functions Offered by the Package

This section provides additional flexibilities offered by the package not covered by the GUI. R scripts are needed to use additional functions. By writing a script additional features of the package become available, and additional analyses that are not part of the GUI can be conducted. Note that, because our goal is to serve unexperienced

audience, we only focus on the fundamental functions used in dichotomous attribute cases. Information in this section is useful for readers who have some level of R programming skills.

We can start with the function *simGDINA( )* that is used to simulate examinee response data. Sample size (*N*), Q-matrix (*Q*), and lower and upper bound of probability correct (*gs.parm*) must be provided, where *gs.parm* is a matrix or data frame for lower and upper bound probability correct parameters (i.e., guessing and slip parameters). The contribution of the delta parameters of the additive models is specified by the *gs.args* argument for which two options, "equal" and "random", are available. The *model* argument allows users to assign a CDM to each item. *model* is a character vector for each item (or a scalar when all items are the same type), for which the available options are: "GDINA", "DINA", "DINO", "ACDM", "LLM", "RRUM", and "MSDINA". The *att.dist* argument is used to indicate the attribute distribution for simulation, where "uniform", "higher.order", "mvnorm", and "multinomial" are available options for the uniform, higher order, multivariate normal, and multinomial distributions, respectively. When attribute distribution happens to be multinomial, the probability of each attribute pattern needs to be specified through *att.prior*. If we would like attribute distribution to be higher-order, then, higher order distribution of attributes need to be specified through *higher.order.parm*. Lastly, a list of parameters for multivariate normal distribution is provided by *mvnorm.parm* when the attribute distribution to be generated is a multivariate normal.

The GDINA package allows users to generate hierarchical attribute structures and, to provide prior joint attribute distribution with specified hierarchical relations among the attributes. For this purpose, the *att.structure( )* function can be used. The *hierarchy.list* is a list used for specifying hierarchical structure among the attributes. Elements of this list specify direct prerequisite relations between the two attributes. For example, when three attributes (e.g., A1, A2, and A3) have a linear relationship where A1 is prerequisite for A2 (and A3), and A2 is also prerequisite for A3; this hierarchical structure is defined by the *list(c(1,2), c(2,3))*. The number of attributes *K* must also be specified to use this function. The *att.prob* argument is used in this function to set latent class probabilities as either "random" or "uniform."

The *GDINA( )* function calibrates the GDINA and reduced CDMs. Here we would like to focus on the features available using specific functions that are not available in the GUI. Along with the CDMs available in the GUI, multiple-strategy DINA model (MSDINA; de la Torre & Douglas, 2008) is also estimated by this function. This function allows hierarchically structured attributes in the estimation, for which *att.prior* must be specified such that prior weights 0 are assigned to impossible latent classes. This option is only available for calibration of the DINA, DINO, or G-DINA models.

To evaluate whether a specific CDM can replace the G-DINA model without significant loss in data-model fit, the *modelcomp( )* function can be used. Although model comparison can also be conducted in the item and test level using the GUI, use of this function allows one to specify the items for which the model comparison is

requested. Furthermore, we can also specify the reduced models to be considered in the model selection. Unlike in the GUI, Wald statistics is not the only option under this function. Two more statistics, namely, Lagrange multiplier and likelihood ratio, are also available for model comparison.

The *extract( )* function is used to extract information or data such as att.prior, discrim, prevalence, prosterior.prob, and delta.parm (i.e., attribute priors, discriminations, attribute prevalence, posterior probabilities, and delta parameters, respectively). To obtain such information, objects from class GDINA, itemfit, modelcomp, or simGDINA must be specified by the *object* argument. The *what* argument is then used to call the information we would like to extract.

The GDINA package permits one to evaluate differential item functioning (DIF), which may occur when the success probabilities on an item are different for the examinees from different groups with the same attribute mastery profile (Hou, de la Torre & Nandakumar, 2014). The dif() function is used to detect DIF based on the models in the GDINA function using the Wald test (Hou, de la Torre, & Nandakumar, 2014) and the likelihood ratio test (Ma, et. al, 2017). It should also be noted that the current version of the package only allows detection of DIF for two groups.

Concluding Remarks

In this manuscript, we first discussed relevant literature in cognitive diagnosis, which includes a saturated model and several reduced models. It was followed by presentation of the GDINA R package, which can be conveniently used for CDM analyses. We demonstrated step-by-step basic CDM analyses using the GUI with an aim of guiding novice researchers and practitioners in the field. Within the demonstration, we provided fundamental information on the statistics used and interpretations of the output. Although the analyses and options are limited when one uses the GUI, more advanced methods for CDMs are available in the package for users who are more familiar with R.

Several steps must be taken to implement cognitive diagnosis modeling successfully. The first step requires specification of the domain-specific attributes for which researchers may need to look into theories within the content domain, review the relevant literature, and conduct a protocol analysis with the think-aloud procedure (Akbay, Terzi, Kaplan, & Karaaslan, 2017; Leighton & Gierl, 2007; Tjoe & de la Torre, 2014). In general, finer grain sized attributes provide richer information (Alderson, 2005); however, an increase in the number of attributes measured by a single test requires a larger number of items which, in turn, requires much larger sample sizes for the model parameters to be estimated accurately.

Once the attributes are determined, items measuring single attributes and combinations of attributes are developed and the relationships between these items and the attributes must be carefully specified through the construction of a Q-matrix. It should be noted that the usefulness of the diagnostic information provided by the CDM analysis is highly dependent on the theoretical and empirical soundness of the Q-matrix (de la Torre, 2008; de la Torre, & Chiu 2016; Lee & Sawaki, 2009). Therefore,

item-attribute relationships must be specified carefully and correctly. In certain situations, researchers have expectations regarding how attributes interact in the item response process. In these situations, another fundamental step involves deciding a priori which CDMs should be fitted to item response data to test one's hypotheses.

Although other software programs are also available, using the GDINA package offers users some flexibilities such as allowing estimation of a wide range of CDMs and allowing nonprogrammers to benefit from this package through the GUI. Using the GDINA, one can obtain diagnostic information about individual examinees' mastery or nonmastery status of attributes; proportion of examinees who have mastered a specific attribute; and proportion of examinees that are in a specific latent class. In addition to ordinary CDM analyses, GDINA package further allows users to apply model selection at the test- and item-level to make sure that the most appropriate CDM (i.e., CDM that best explains the attribute interactions in the item) is fitted to the response data. Furthermore, to identify possible item-attribute specification mistakes in the Q-matrix, implementation of an empirical Q-matrix validation method is available in the GDINA package. Lastly, this package offers various handy graphs, which can be very useful in emphasizing important information and comparing various parameters and/or statistics.

References

Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. A&C Black.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Ed.), *Proceedings of the Second International Symposium on Information Theory* (pp. 267–281). Budapest, Hungary: Akad. Kiado.

Akbay, L., Terzi, R., Kaplan, M., & Karaaslan, K. G. (2017). Expert-based attribute identification and validation on fraction subtraction: A cognitively diagnostic assessment application. *Journal on Mathematics Education*, *8*, 103-120.

Chen, J. & de la Torre, J. (2013). A general cognitive diagnosis model for expert-defined polytomous attributes. *Applied Psychological Measurement*, *37*, 419-437.

Chen, J., de la Torre, J., & Zhang, Z. (2013). Relative and absolute fit evaluation in cognitive diagnosis modeling. *Journal of Educational Measurement*, *50*, 123-140.

Chiu, C. Y. (2013). Statistical refinement of the q-matrix in cognitive diagnosis. *Applied Psychological Measurement*, *37*, 598-618.

de la Torre, J. (2008). An empirically based method of Q-matrix validation for the DINA model: Development and applications. *Journal of Educational Measurement*, *45*, 343-362.

de la Torre, J. (2009). DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics*, *34*, 115-130.

de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika*, *76*, 179-199.

de la Torre, J., & Chiu, C. Y. (2016). General method of empirical Q-matrix validation. *Psychometrika*, *81*, 253-273.

de la Torre, J., & Douglas, J. A. (2008). Model evaluation and multiple strategies in cognitive diagnosis: An analysis of fraction subtraction data. *Psychometrika*, *73*, 595.

de la Torre, J., & Lee, Y. S. (2013). Evaluating the Wald test for item-level comparison of saturated and reduced models in cognitive diagnosis. *Journal of Educational Measurement*, *50*, 355-373.

de la Torre, J., & Ma, W. (2016, August). *Cognitive diagnosis modeling: A general framework approach and its implementation in R.* A Short Course at the Fourth Conference on Statistical Methods in Psychometrics, Columbia University, New York, NY.

DiBello, L. V., & Stout, W. (2007). Guest editors' introduction and overview: IRT-based cognitive diagnostic models and related methods. *Journal of Educational Measurement*, *44*, 285-291.

Hartz, S. M. (2002). *A Bayesian framework for the unified model for assessing cognitive abilities: Blending theory with practicality* (Unpublished doctoral dissertation).

Henson, R. A., Templin, J. L., & Willse, J. T. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika*, *74*, 191-210.

Hou, L., de la Torre, J. D., & Nandakumar, R. (2014). Differential item functioning assessment in cognitive diagnostic modeling: Application of the Wald test to investigate DIF in the DINA model. *Journal of Educational Measurement*, *51*, 98-125.

Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, *25*, 258-272.

Leighton, J. P., & Gierl, M. J. (2007). Defining and evaluating models of cognition used in educational measurement to make inferences about examinees' thinking processes. *Educational Measurement: Issues and Practice*, *26*(2), 3-16.

Leighton, J. P., Gierl, M. J., & Hunka, S. M. (2004). The attribute hierarchy method for cognitive assessment: A variation on Tatsuoka's rule-space approach. *Journal of Educational Measurement*, *41*, 205-237.

Ma, W. (2018). A diagnostic tree model for polytomous responses with multiple strategies. *British Journal of Mathematical and Statistical Psychology*. https://doi.org/10.111/bmsp.12137

Ma, W., & de la Torre, J. (2016). A sequential cognitive diagnosis model for polytomous responses. *British Journal of Mathematical and Statistical Psychology, 69,* 253-275.

Ma, W., & de la Torre, J. (2018). GDINA: The generalized DINA model framework, R package version 2.1.15. Retrived from https://CRAN.R-project.org/package=GDINA

Ma, W., Iaconangelo, C., & de la Torre, J. (2016). Model similarity, model selection, and attribute classification. *Applied Psychological Measurement*, *40*, 200-217.

Ma, W., Terzi, R., Lee, S. Y., & de la Torre, J. (2017, April). Multiple group cognitive diagnosis models and their applications in detecting differential item functioning. The annual meeting of American Educational Research Association, San Antonio, TX.

Maris, E. (1999). Estimating multiple classification latent class models. *Psychometrika*, *64*, 187-212.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461-464.

Sinharay, S., & Almond, R. G. (2007). Assessing fit of cognitive diagnostic models a case study. *Educational and Psychological Measurement*, *67*, 239-257.

Tatsuoka, K. K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement*, *20*, 345-354.

Team, R. C. (2015). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Templin, J. L., & Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods*, *11*, 287.

Tjoe, H., & de la Torre, J. (2014). The identification and validation process of proportional reasoning attributes: an application of a cognitive diagnosis modeling framework. *Mathematics Education Research Journal*, *26*, 237-255.

GDINA R Paketi Kullanarak Bilişsel Tanı Modelleri Uygulaması

Atıf:

de la Torre, J., & Akbay, L. (2019). Implementation of cognitive diagnosis modeling using the GDINA R package. *Eurasian Journal of Educational Research, 80,* 171-192, DOI: 10.14689/ejer.2019.80.9

Özet

Formatif değerlendirme için hazırlanmış olan ölçme araçları, amaçlı bir şekilde tasarlandığında, tanılayıcı bilgi toplamamıza imkân sağlarlar. Bu tür bilgilere dayanarak yapılan çıkarımlar eğitimcilerin telafi edici eylemler gerçekleştirmelerini sağlar. Son zamanlarda geliştirilen formatif değerlendirme yöntemleri öğrencilere ve öğretmenlere öğrencilerin hangi bilgi ve beceri parçacıklarına (bilişsel niteliklere) sahip oldukları ya da olmadıklarıyla ilgili ayrıntılı geri bildirim sağlayabilmektedir. Bu geribildirimlerin öğretimi ve öğrenmeyi optimize etmek amacıyla kullanılması beklenmektedir. Formatif ölçme ve değerlendirmelerin sonuçlarından tanısal bilgi edinebilmek için iyi tasarlanmış sınavların uygulanmasının yanı sıra çeşitli bilişsel tanı modellerinin kullanımına da ihtiyaç vardır. Ancak, bu modellerin alan yazına yeni kazandırılmış olması ve eğitimcilerin henüz yeterince aşina olmamaları nedeniyle, bilişsel tanı modellemelerinin (BTM) parametre kestirimlerinin ve diğer ilgili analizlerin uygulamaları yeterince yaygınlaşmamıştır.

Bu makalenin amacı BTM'ye aşina olmayan eğitimci ve araştırmacılara BTM'nin temel prensiplerini tanıtmak ve ücretsiz bir yazılım olan GDINA R paketi kullanılarak yapılabilecek çeşitli BTM uygulamalarını yeterince detaylı olarak göstermektir. BTM analizleri yapan bazı yazılım programlarının mevcut olmasına rağmen, alanla ilgili en kapsamlı paketlerden biri olan GDINA R paketinin kullanımının sağlayacağı avantajlar arasında şunlar sayılabilir: (1) Birçok bilişsel tanı modelinin kestiriminin R ile mümkün olması; (2) Diğer birçok yazılım programlarının tek tip modelin kestirimine imkan veriyor olması; (3) Diğer programların çoğunun ticari olması veya ancak yazarın kendisiyle irtibat kurularak temin edilebilir olması; ve (4) Sözdizimi (sintaks) hazırlamanın bazı programlarda oldukça zahmet verici olması.

Bu makalede öncelikle BTM analizinin girdi ve çıktıları da dahil olmak üzere BTM terminolojisi tartışılmakta, sonrasında GDINA model yapısı tanıtılmakta ve GDINA R paketi sunulmaktadır. Sonrasında ise nümerik bir örnek veri setinden yola çıkılarak, GDINA R paketi ve bu paketin sunduğu grafiksel kullanıcı ara yüzü (GUI: graphical user interface) kullanılarak yapılabilecek analiz türleri adım adım takip edilebilecek şekilde sunulmuştur. Ayrıca gerekli görülen noktalarda yapılan iş ve işlemlerin teorik bilgisine ve elde edilen sonuçların yorumlanmasına dair bilgiler verilmiştir. İlerleyen bölümlerde, GUI ile yapılamayan ancak GDINA paketinin sunduğu ve R kullanımına aşina olan kullanıcıların sözdizimi yazarak yapabilecekleri ilave analizlere ve paketin ek özelliklerine yer verilmiştir. Makalenin son bölümünde ise bazı hatırlatma ve tespitler yapılmıştır.

G-DINA modeli ve GDINA paketi

Genel ya da kısıtlanmış olarak birçok bilişsel tanı modelinin ayrımları öğrencinin maddeleri cevaplama sürecinde bilişsel niteliklerin (ölçülen bilgi ya da becerilerin) etkileşimlerine bağlı olarak yapılır. G-DINA gibi genel modellerde bütün ana etkiler ve etkileşim etkileri madde tepki fonksiyonuna katkıda bulunur. Bu tür genel modellerin parametrelerinde yapılacak kısıtlamalarla daha kısıtlı ya da daha sade modeller elde edilebilir. Bunun yanı sıra, genel modellerde birden farklı *link* fonksiyonlarının kullanımı da söz konusudur. Örneğin, G-DINA model identity, logit ve log linkleriyle farklı isimlerle sunulabilir, ancak farklı linklerle ortaya konulan genel modeller özdeş model-veri uyumuna sahiptirler.

GDINA paketi GUI aracılığıyla farklı linkler altında oluşturulabilecek olan GDINA, logit GDINA, log GDINA'nın yanı sıra, bu genel modellerde ortaya koyulacak kısıtlamalarla oluşturulabilen DINA, DINO, ACDM, LLM, ve RRUM modellerinin kestirimine imkan verir. Ayrıca, GUI aracılığıyla olmasa da, R kullanımına aşina olan ve sözdizimi yazabilen kullanıcılar GDINA R paketini kullanarak farklı stratejilerin kullanımı durumunu göz önünde bulunduran MS-DINA ve G-DINA'nın farklı uzantılarından oluşan modellerin (ör. sınıflama ve sıralama ölçekleriyle elde edilen veriler için *sequential* G-DINA ve kısmi puanlamalı nitelikleri için *polytomous* G-DINA) kestirimini yapabilirler.

GDINA'yı kullanarak, testi alanların nitelikleri kazanmış ya da kazanamamış olma durumlarıyla ilgili tanılayıcı bilgi; belirli bir niteliğin kazandırılma oranı; ve belirli bir profile sahip olan kişilerin oranı ile ilgili bilgiler elde edilebilir. Bunların dışında, temel CDM analizlerinin yanı sıra, madde ve nitelikler arasındaki etkileşimi en iyi açıklayan BTM'nin kullanılabilmesi adına GDINA paketi araştırmacıların test- ve madde-düzeyinde model seçimi yapabilmelerine imkan tanır. GDINA paketi aynı zamanda madde-nitelik eşleştirme matrisinde yapılabilecek muhtemel yanlışları ortaya çıkarmaya yardımcı olmak adına ampirik olarak Q-matrisin uygunluğunun doğrulamasını sağlayan metotların kullanılmasına imkan sağlar. Son olarak, bu paket, çeşitli kolay ve kullanışlı grafikler ortaya koyar. Bu grafikler elde edilen analiz sonuçlarında önemli noktaların vurgulanmasına ve elde edilen istatistik ve parametrelerin kolaylıkla karşılaştırılmasına olanak sağlar.

Anahtar sözcükler: Bilişsel tanı modelleri, GDINA, R paketi, BTM uygulamaları.

Eurasian Journal of Educational Research 80 (2019) 193-224

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

The Use of a Visual Image to Promote Narrative Writing Ability and Creativity

LISTYANI[1]

ARTICLE INFO

Article History:
Received: 26 Aug. 2018
Received in revised form: 26 Mar. 2019
Accepted: 23 Mar. 2019
DOI: 10.14689/ejer.2019.80.10

Keywords
narrative writing, visual images, pictures, comic strips, mystery

ABSTRACT

Purpose: Second language writing as an inherent part of ELT is no exception. One specific part of second language writing in which visual images can be used is narrative essay writing. Visual images, in this case, comic series and pictures, can be a very useful aid in stimulating students' ideas, creativity, as well as interest and ability in narrative essay writing. Students' writing activities in using pictures and comic strips are discussed in this study, including students' responses towards this particular topic, as well as the procedures of how these visual images are used. This study is therefore aimed at disclosing how pictures and comic strips, as forms of visual images, can be used to enhance students' narrative writing ability as well as creativity. In writing narrative essays with the help of visual images, students were helped in terms of generating ideas, developing logical and critical thinking, and improving reasoning skills.

Research Methods: The design of this study was qualitative in nature. The participants were comprised of 19 *Professional Narrative Writing* students. The data for this research was taken from documents, that is, students' essays and journals which were written after the writing activities were done, interviews with two students, and students' scores. There were also pre-tests and post-tests given at the beginning and the end of the semester, but the scores were descriptively presented. Interviews with two students were also conducted to validate the findings. These students, whose essays were used as analyzed documents, were enrolled in the *Professional Narrative Writing* class of the English Language Education Program, the Faculty of Language and Arts (FLA), Universitas Kristen Satya Wacana (UKSW), Salatiga, Indonesia. The class was conducted in Semester I of the 2017-2018 academic year. The students were fourth-semester students. The *Professional Narrative Writing* course taught students how to be professional in writing narratives.

Findings: The findings showed that pictures, as well as comic strips, were very useful in helping students to write narratives. They helped students generate ideas, delve into more creativity, as well as develop their imagination and motivation in writing, though some students experienced difficulties in some aspects like ideas, grammar, diction, and plot.

Implications for Research and Practice: The results of the study hopefully can inspire other narrative writing lecturers all over the globe to maximize the use of visual images, including pictures and comic strips. These visual aids can enhance students' writing abilities as well as their creativity. Students taking writing courses can hopefully be motivated to write better narratives.

[1] English Language Education Program (ELEP), Faculty of Language and Arts (FLA), UKSW Indonesia, email: lydia.listyani@gmail.com
ORCID id: 0000-0002-3284-9434, URL: https://orcid.org/0000-0002-3284-9434

Introduction

In this section, the background of the study, the rationale, the design, a literature review, previous studies, and research questions are discussed.

Background of the Study

Students' creativity and imaginations in writing should be triggered or encouraged in order to make them produce better writing. In a Narrative Writing class, the help of visual images – in this case, mystery pictures and picture series or sequential pictures in the form of comic strips, mean a lot for students. They are helped at least in one sense: developing their creativity and imagination. Thus, the aim of this study was to describe how visual images in the form of a picture series and mystery pictures helped students improve their narrative writing abilities, as well as expanded their imagination and creativity. From the findings of this research, it was found that students could be very imaginative and creative both in the language and content of the writing. OECD (2018, p.6) claims that

…to prepare for 2030, people should be able to think creatively, develop new products and services, new jobs, new processes and methods, new ways of thinking and living, new enterprises, new sectors, new business models and new social models… The constructs that underpin the competency include adaptability, creativity, curiosity, and open-mindedness.

Working collaboratively in small groups, students were successful in creating the plot of the story based on the provided pictures, as well as in other aspects of narrative writing, like coherence, unity, the development of ideas, problem formulation, and problem-solving. As mentioned earlier, there is a famous proverb which states that a picture is worth a thousand words. The meaning is, more or less, that a picture tells a story as well as, if not better than, a lot of written words. Everybody surely agrees with this; one picture can say a lot of things. This is what 19 students of a Narrative Writing class of the English Department of the Faculty of Language and Arts (FLA) of Universitas Kristen Satya Wacana (UKSW) Salatiga, Indonesia, were able to experience. They were given pictures and comic strips with no dialog at all, and they had to tell what happened in the stories in the form of a narrative essay. Most of them enjoyed this activity. They had to work in pairs or in a group of three, and they told their stories based on the comics or pictures given. These students were fifth-semester students who were left behind by their peer students. Their peer students had taken the *Professional Narrative Writing* class in the previous semester.

Rationale & Design

A significant shift has happened over the last century, from manufacturing to emphasizing information and knowledge services. "*Shared decision-making, information sharing, collaboration, innovation, and speed are essential in today's enterprises*" (Kamehameha Schools Research and Evaluation, 2010, p.1). Trilling and Fadel (2009), as cited in Kamehameha Schools Research and Evaluation, 2010, p 6, claim that today's students should be able to communicate clearly, collaborate with others, think critically, solve problems, be creative, and be innovative. Education has focused on the essences of communication, like speaking, writing, and reading, but the demands of

social relations and a global economy in this century call for a more diverse set of communication and collaboration skills. This becomes the underlying reason of the rationale of this study; that students of this century, including the ones learning English in English Language Education Program (ELEP), who later on are expected to be teachers, should be able to master these twenty-first century skills.

Comic strips, as a form of a picture series and mystery pictures, are a good media to use in narrative writing. Using these teaching aids in classroom activities can improve students' creativity and imagination in writing. From the student participants' journals, it can be clearly seen that they found this activity challenging. It successfully activated their imaginations and creativity. In our everyday lives, whether we realize it or not, we deal with narratives. They may not be in the form of a written essay, but they take the form of spoken narratives. When we talk about something that happened the other day, the previous week, month, or year, we use narratives. It is not something new in our lives. In this research, it is argued that to write narratives creatively and imaginatively, visual images need to be used. The design of this study is basically qualitative in nature, which is then backed up with some descriptive statistics. This will be discussed in length in the Methodology section.

Review of Literature

The use of pictures as an aid in teaching is not something new. Allford (2000, as cited by Rayo, 2015), claims that since the Renaissance era, Latin has been learned by using pictures. By the 1950s, pictures were widely used in ESL classrooms. The purpose was to stimulate learners' interest and reinforce materials in a textual form. Randle (n.d., as cited from Arnheim, 1964) stated that every picture and visual image that we perceive is a statement of thought. How skillful we are in cultivating our perception skills will determine how well we can understand the meaning from a visual image. Arnheim first coined the term 'visual thinking strategies' (VTS). VTS is a teaching strategy which is used to encourage critical and creative thinking by using visual images. There are three questions which are carefully worded to ask about an image during a group discussion to foster critical thinking. They are: *What's going on in the picture? What do you see which makes you say that? What more can we find?* The purposes of these questions are to construct meaning and understanding from the image. Silva, Santos and Bispo (2017) mentioned that comics have existed for more than 40,000 years, which began with cave paintings. They further revealed that there are several purposes of using comics in the educational field. First, they can be used as a teaching aid, and second, they can be applied to improve students' speaking abilities.

Rayo (2015) further stated that pictures require no or just a little explanation. As cited by Spanley and Peprnik (1967), Rayo (2015) explained that there are three guidelines for utilizing pictures in the classroom. First, they must be clear in showing a cause and effect. Next, pictures should contain examples of the actions that the authors want to emphasize. This is to avoid misinterpretations. The relationships between the elements in the pictures should also be clear so that the sequence of actions makes sense. Third, pictures should contain a break in the action to introduce an

element of surprise. By fulfilling these picture requirements, it can make pictures become successful aids for language learners.

Rayo (2015) claims that there are many benefits of using picture stories in the area of language learning. Among them is that it makes students engaged in learning when pictures are used. Besides that, pictures can also be an aid in the transfer of fictional stories to personalized life experiences. Pictures can be used with various degrees of difficulties. Pictures can also become a medium for students to express their thoughts and feelings. In line with Rayo, Aschawir (2014) postulates that the more interesting and varied the encouragement through visual aids, the faster and more effective the learning will be. Pictures, still according to Aschawir (2014), contribute to learners' interest and motivation. They also have positive effects on the sense of the context of the language and specific stimulus. Aschawir also introduced different kinds of pictures. They were wall pictures, wall charts, sequential pictures, flashcards, and board drawings.

Asrifan (2015) reinforced this idea. He mentioned that there are two types of materials used for language learning, visual and non-visual. He stated further that, "*A visual material offers an attractive and stimulating framework for writing practice and has great potential as an aid to develop writing skills since it provides both contexts and stimulation of a variety of activities. One of the visual materials is pictures*" (p. 245). Picture stories can help students in generating ideas and organizing those ideas in writing. Picture stories can also effectively stimulate students' imaginative ability. A similar insight also comes from Aschawir (2014). He stated that one of the main characteristics of a writing program is that learners should be given opportunities to communicate through writing, and also to simply enjoy writing. In building up learners' writing skills, a picture series can be used as a good medium or visual aid. Imastuti and Suparno (2012), as cited from Wright (1989), also claim that there are five roles in writing. First, pictures can motivate learners. Second, they contribute to the context in which learning takes place. Next, pictures can describe something in an objective way. The fourth one is that pictures can cue responses. Finally, pictures can stimulate and provide information in storytelling.

The use of pictures to help students write better is commonly known, especially in the area of narrative writing. Picture usage is a common practice in a performance-based assessment in writing. The next question may be: What is a performance-based assessment? From the name, one can guess that it is a kind of assessment that is based on performance. A performance-based assessment measures students' ability to apply skills and knowledge which are learned from a unit or units of study. The task challenges students to use higher-order thinking skills or create a product or complete a process (Chun, 2010, in Hilliard, 2015). Schweizer (1999) has a similar idea, mentioning that in writing tests, the pictures that are used should be the ones which foster the writing process. The underlying reason is that the quality of writing is affected by the pictorial stimulus which will be used to produce the writing (Schweizer, 1999). Hyvärinen (2007) also mentioned that a narrative is, as cited from Propp (1968 & 1984) about Aristotelian idea, a good tragedy which has a beginning, a

middle, and an end; while open, conversational, or artistic narratives emphasize a clear sequence of events.

There is a theory which states that both visual and verbal information are used to represent information (Thomas, 2014). According to this theory, human beings process and represent verbal and non-verbal information in separate yet related systems. Both visual and verbal information can be used to represent information. This theory was developed by Paivio (1991, p. 158) in the 1960s, and is known as the *Dual Coding Theory (DCT)*. DCT predicts that:

> *Word concreteness and imagery value should be the central variables in cognitive and educational tasks, which are related to meaning. Concreteness and imagery value have been measured empirically by ratings of the ease with which words, sentences, or larger units of text evoke a visual, auditory, or other mental picture (imagery value), or the degree to which they refer to tangible objects with concrete referents (concreteness) (Clark & Paivio, 1991).*

This theory also states that "mental representations" are related with both verbal and non-verbal symbolic modes. These mental representations retain the features of the concrete sensorimotor events, which they are based on. The verbal system contains "visual, auditory, articulatory, and other modality-specific verbal codes". In contrast, non-verbal representations include *modality-specific images for shapes*, for example, a chemical model; *environmental sounds* like a school bell; *actions*, for example, drawing lines or pressing keys; *skeletal or visceral sensations related to emotion* (e.g., a clenched jaw or a racing heart), and other non-linguistic objects and events (p. 151-2).

Paivio and some other researchers support the importance of imagery in cognitive operations. The human brain processes visual as well as verbal information in completely different ways; one is for the visual images, the other is for verbal language. Clark and Paivio (1991, p.162, p. 166) mentioned that DCT is concerned with the interaction between individual differences and the use of imagery in text comprehension. This theory of DCT assumes that people vary in both their ease as well as skill in which they use non-verbal and imaginative processes. They also stated that both imagery and concreteness play an important role in memory for a text. From their research, it was found that the subjects who received imagery remembered better than those who did not.

Beckley (2014) studied the relationship between visual and written narratives in student engagement. In her study, she collected data from direct observations of students, teacher notes, rubrics, student surveys, and student projects. She found that providing students with a variety of methods – especially image making - will make students produce stories which are richer, more elaborated, and engaging.

Related with visual images, Grainger (2004, p. 199) indicated that there are some key features of a visual image. They are the line, color, action and movement, size and location, and symbolism. Line and color are two basic elements of a picture, Grainger claims. On these two elements, more complicated features are built. Color is very important in one sense; that is, in creating the world as we expect readers to see it. Illustrated pictures also have action, movement, and the passing of time. However, artists describe the characters in a frozen form. Arhnheim (in Grainger, 2004, p. 204) indicated that sometimes the frozen form of the characters in the illustrations created by artists do not exist in the real world. It is the ability of the artists to describe the rapid movements of the characters' limbs in the motionless pictures that matter.

Readers or writers' perceptions and interpretations of the characters are influenced by the settings and locations where the characters are placed. Even the same picture may trigger different interpretations for different people. High, low, left, and right may show significant meanings. Certain symbols may be present in pictures which bring certain meanings. Dark clouds, rain, and night, for example, may symbolize sadness and a gloomy situation.

As summed up from the theories above, visual images can be used in the educational field. Two of them are pictures and comic strips. These two non-verbal representations can function as a media to help learners of English write narratives better. As stated earlier, we deal with narratives in our daily lives. We cannot avoid it. This is not exempt from the area of academic life, especially in the English language major. Students, as well as teachers, should deal with narrative writing. There are many strategies that teachers can use in order to motivate students to write better. Using pictures and a comic series is one of the strategies worth trying in narrative essay writing. Besides giving students visual images, pictures and comics also can develop students' motivation and eagerness to write.

Gutiérrez, Pueelo, and Galvis (2015) support this idea, saying that "*Through narrative writing, students organize ideas and experiences creatively and imaginatively by combining linguistic, pragmatic, and sociolinguistic competencies*." Incorporating pictures into language teaching and learning surely has positive impacts on learners' communicative skills. These teaching aids can help teachers to make language contextualized by "going out of the classroom walls".

Previous Research Reviewed

As mentioned above, the use of visual images in education is not something new. Recently, studies in language learning involving the use of visual images have been done by some researchers. In 2012, Imastuti and Suparno from UNS Solo Indonesia conducted a study on the use of a picture series in teaching writing in order to help students solve their problems in writing. They argued that there are five roles of pictures in writing. First, pictures can motivate students to take an active part in the writing process. Secondly, pictures help bring the 'world' into the classroom. Students will get visualization from pictures. Next, pictures can describe something objectively. The fourth benefit is that pictures can lead students' responses to a question. The last advantage is that pictures can provide information.

From their study, Ismatuti and Suparno (2012) concluded that their respondents, students of Public Junior High School (SMPN) 2 in Ampel, Indonesia, improved in their motivation after being given a picture series for their writing activity. The second conclusion was that there were improvements in students' writing skills as they wrote more detailed information, included longer paragraphs, had better organization, used sufficient supporting sentences, and could also change present tense into past tense. In short, these student respondents could write narrative texts better.

In the same year, 2012, Ayuningtyas and Wulyani also conducted action research on the eleventh graders of Public High School (SMAN) 1 in Srengat Blitar, Indonesia. They wanted to know if the ability of the eleventh graders in writing narrative texts could be improved by using picture sequences as instructional media. They obtained data through two observation checklists, questionnaires, and the students' writing products. The findings showed that there were good improvements in terms of students' attitudes and writing products. There is a need to conduct a similar study on tertiary-level students with participants of different age groups and English mastery level, which other pieces of research have not covered yet.

Aschawir (2014) also carried out a similar kind of research on second-semester students at the English Department of the Faculty of Letters of UMI, Indonesia. The experimental group was taught using a picture series, while the control group was taught without a picture series. Five writing components were emphasized, including content, organization, vocabulary, language use, and mechanics. From his research, it was found that the results of the use of a picture series increased experimental group's post-test writing scores.

Another research was conducted by Gutiérrez, Puello and Galvis, in 2015. The research was done with ten students in the experimental group, and ten students in the control group. The experimental group was taught with a pictures series technique and a process-based approach. In contrast, the other group was taught using the process-based approach only. From their research, it was found that the picture series technique which was applied to the process-based approach was suitable for EFL students of the eighth grade. They could improve their narrative writing skill in English. Besides that, students' motivation to practice their narrative writing was improved.

Rayo (2015) also contributed to the findings in this area. He examined the effectiveness of the use of picture stories to improve verb tenses in narrative writing of adult English learners. There were 36 participants from China. The results showed that there were no significant improvements in the experimental group. He found that these learners preferred the simple past tense form over other verb forms in their narrative writing.

Silva et al. (2017) also reported some research done in the past, dealing with the use of pictures. Results showed that the use of comics could enhance competence development, innovation, and flexibility in writing. Besides, it could reduce the gap between theory and practice. It also helped students to be more critical in writing. Jensen et al. (2007) conducted a study using drawings in management learning which

involved the use of comics as a way of representing knowledge. This study found that comics can help students to express and represent professional situations. In 2009, some high school physics teachers introduced Einstein's idea of using pictures to the students. The teachers thought that this teaching strategy could become a source of motivation to present words and images (as cited from Caruso & Freitas, 2009). Two years later, in 2011, Rosetto and Chiera-Macchia reported on the use of comics in teaching Italian. This activity helped students in developing memory, analysis, and reasoning skills. This also decreased their anxiety and encouraged their creativity. In 2012, Kiliçkaya and Krajka conducted research to identify whether students liked to create comics to facilitate their learning process. The findings showed that 24 out of 25 students claimed that they enjoyed participating in this activity.

With the aims of enriching the previous theories in the use of visual images in ELT and completing the existing literature in narrative writing using pictures, this study was thus conducted. This present study will confirm and complete the previous findings. The special addition lies in the additional visual materials used besides sequential pictures; that is, mystery pictures. Pedagogically, the results of this study are expected to provide new horizons and insights for writing teachers at any level about the use of visual images to enhance students' writing skills as well critical thinking and creativity.

Research Questions

The main focus of this paper is to describe and reveal how pictures and comics as a form of visual images can be used as a technique to teach students to write narrative essays. Basically, there are two questions to answer in this paper:

1. *How can pictures and comic strips be used to enhance students' creativity and abilities in narrative writing?*
2. *What are students' perspectives on the use of pictures and comic series in narrative writing?*

Research Setting

The research was conducted in the first semester of the 2017/2018 academic year. Nineteen students of the English Language Education Program, the Faculty of Language and Arts, UKSW, Salatiga, Indonesia, became the participants of this study. They were fourth-semester students.

Method

Research Design

The design of this study was basically qualitative, which was supported by quantitative data. A qualitative design was chosen as the main model of this research, mainly because in this kind of research, participants' perspectives are given priority (Cresswell, 2009). The underlying reason why participants' perspectives were given an important place in this research was that this study was *progressive* in nature, not

an *experimental* one. The bottom line was that the participants' perspectives constructed the social world, and their voices for changes were highly valued and very much appreciated; therefore, there were no right or wrong answers to a particular problem. What the researcher describes and what the participants say are both of equal importance (Halliday, 2002). The participants could freely express their opinions and feelings, including satisfaction or dissatisfaction with a certain treatment given to them. The perspectives could vary according to some factors, like cultures.

This research portrayed individuals – in this case, Narrative Writing class students - as they were building their social worlds; that is, the imaginative worlds they were building in their minds that later on were reflected in their narrative essays. As a researcher, I tried to build my own social world through my interpretation. Participants' perspectives, which represented human behavior, were then analyzed and interpreted in elaborate descriptions. This research could also be called a *narrative* study because all the data was interpreted in a "thick description", and later on presented in the form of narratives (Halliday, 2002).

As seen from the involvement of the research participants, this research can be called *participatory* and *interpretive*. It interprets human behavior from participants' perspectives, and the emphasis is on rich data collection and thick descriptions (Burns, 2003). Cresswell (2009, p. 9) reinforces this idea, mentioning that qualitative researchers tend to use open-ended questions so that participants can express their views. The researcher tries to understand the context of the participants through information gathering done personally. Quantitative data from the students' scores from the pre-tests and the post-tests were also used.

The data collection was done throughout Semester I/2017-2018 Academic Year, that is, from September till the end of November 2017. Students were given three days to finish one piece of writing. Data in the forms of essays, journals, and interviews were analyzed qualitatively. Meanwhile, data related to scores were quantitatively analyzed. The visual tools used in this study were five mystery pictures and two comic strips which represented sequential pictures. The selection of the pictures was done in collaboration with another lecturer who was teaching a parallel class in the same semester. Themes of the pictures were found based on a thought that those pictures had to give students ample opportunities to develop their imagination and critical thinking. At the beginning and the end of the semester, students were given pre- and post-tests. This was done to see whether the aspects of their narratives improved or not. The aspects graded were ideas, voice, diction, and convention. During the semester, they were asked to write two essays based on the pictures given, two email complaints, and one personal statement. Involving two kinds of data, qualitative and quantitative, the study can be considered as having a mixed-method design. About this mixed-method research, Gall, Gall and Borg (2007, p. 32) defined it as: "*A review of quantitative studies about a particular phenomenon combined with a review of qualitative studies about the same phenomenon that can provide richer insights and raise more interesting questions for future research.*" Due to the limitation of time and space, this paper only focuses on qualitative data.

According to Gall et al. (2007), in a one-group pre-test and post-test design, three steps are taken. The first step is the implementation of the pre-test. Second, treatment is given to all the participants. Last, the post-test is implemented. In this study, a pre-test was given at the beginning of the semester, while a post-test was given at the end of the semester. In a one-group pre-test and post-test design, the absence of a control group is not a problem. It does not constitute a threat to internal validity since the researchers can make a good estimate of the results of the pre-test as well as the post-test because of extraneous factors. A one-group pre-test and post-test design is the most justified design when extraneous factors can be estimated with a high degree of certainty or can be assumed to be minimal or non-existing.

During the narrative writing using pictures, there were two cycles. First, students were given two sequential pictures, and they were allowed to choose one sequence, each of which consisted of four comic pictures. They then worked in small groups, creating a narrative story based on the pictures. For the second step, students were given five mystery pictures. Just like the previous stage, they were given freedom to choose one among five pictures and asked to write a narrative based on the selected picture. Students' essays were read and graded based on certain grading rubrics. Journals were also qualitatively analyzed. After doing those two assignments, students were asked to write journals about their opinions and feelings in writing using pictures.

About qualitative research, Leung (2015) supported this idea, saying that "*The essence of qualitative research is to make sense of and recognize patterns among words in order to build up a meaningful picture without compromising its richness and dimensionality*." From Leung's statement, it can be clearly seen that in qualitative research, words are important as through words, a meaningful world is built. Leung further stated that similar to quantitative research, qualitative research seeks answers to questions of how, where, when, why, and who. In qualitative research, however, there is no need to assess any work numerically and statistically. Everything is described, explained, and elaborated through words.

Research Participants

All the data for this research was derived from documents, in the form of students' essays and journals. Altogether, there were 19 students who attended this Narrative Writing Class, Group A, in Semester I of the 2017-2018 academic year. They were all selected for this research. Before taking this class, they had previously taken *Creative Writing* and *Procedural Writing* courses. These students were late in taking this *Narrative Writing* class due to grammar classes that they had to repeat.

Research Instruments

The instruments used in this study were pre-tests, post-tests, interview protocol, as well as pictures and comic strips, which students had to base their essays on. Besides those two things, journals were also used to get students' perspectives on the use of pictures and comic strips.

Research Procedures

In detail, here are the steps of the procedures. At the beginning of the semester, on August 29, 2017, the students did their pre-tests. The post-tests were given at the end of the semester, on November 28, 2017. The scores were presented descriptively. For the narrative writing task of using pictures, there were two cycles which were explained afterward. First, I searched for the right pictures and comic strips without dialogs for the students to complete. After getting the right pictures from the right artists, I assigned students to work in groups of two or three, so they could start brainstorming their ideas about the pictures and the comic strips (they had to choose one comic strip and one picture). After imagining what had happened in both the pictures as well as comics, they began writing their stories in collaboration with others. The same stages happened when students had to write a narrative essay based on a mystery picture which they selected out of five pictures. Journals on students' opinions on the use of pictures and comics were also used as data in this research, besides interview results.

Validity and Reliability

Regarding the validity of qualitative research, Leung (2015) claims that the validity of qualitative research deals with the appropriateness of tools, processes, and data. There must also be several methods which are adapted to add validity, like first and second triangulation, and well-documented materials. This present study used appropriate tools of data collection like direct observations and a close look at students' journals and essays.

Reliability of qualitative research refers to the replicability of the process of research and the results as well (Leung, 2015). While getting original data from the participants, researchers have to verify with constant comparisons, the accuracy of the data related to the form and context. The scope and analysis of data should be comprehensive and inclusive. Reference to quantitative aspects should be made possible. This study also has a comprehensive and inclusive data analysis.

Another aspect of qualitative research is generalizability. Leung (2015) stated that the generalizability of qualitative research is usually not expected. This kind of research is meant to study specific phenomena in a certain population or ethnic group. However, Leung (2015) mentioned that there is a pragmatic approach to assess the generalizability of qualitative research. This includes systematic sampling, triangulation, constant comparison, documentation, and the multi-dimensional theory.

All data, including students' pre-tests, post-tests, essays, journals, and interviews were analyzed qualitatively. Pre-tests were given at the beginning of the semester, while post-tests were given at the end of the semester to see in what aspects students improved. Two pieces of students' essays written collaboratively, one based on comic series and the other on mystery pictures, were also graded based on certain aspects. Students' journals and interview results were then conducted to support the findings from tests and essays.

Data Analysis Procedures

All the qualitative data derived for this research were then interpreted in the form of *narratives*. The focus of the interpretation was the participants' narrative essays and perspectives. The improvements of several aspects were then seen from students' narrative essays, in terms of ideas, creativity, and fluency of ideas. The aspects graded from the essays included their ideas, the writers' voice, (reflected in dialogs in the essays), word choice or diction, and grammar and mechanics. The grading sheet is attached as an appendix.

The participants' behavior and motivation in this series of activities were also observed from what they said in the journals as well as interviews with two students who were randomly selected. The data were then interpreted in a "rich description". After the findings were presented, the data were then qualitatively interpreted. Finally, conclusions were drawn.

In the discussion section, the process of the essay writing based on visual images from the beginning till the end, the strengths, the main challenges, as well as the solutions to the problems are discussed. Some samples of students' completed dialogs in comic strips and students' completed essays are also presented. Students worked in collaboration with one or two other classmates in a small group. After working hard to make the outlines of their essays, they began writing with their group members, and they submitted the drafts to the lecturer.

Results

Presented below are students' essays which were written collaboratively with their peers. Based on comic strips with empty bubbles and mystery pictures, they worked in small groups of two or three students. Here their creative and imaginative minds could be seen. They managed to develop their writing based on the visual images given.

The visual images used in this study were comic strips with sequences, or usually called as sequential pictures, and five mystery pictures, which were single pictures. For the comics, two comic strips without dialogs were given to the students. The first one showed a picture of a car speeding up and a policeman on the right side of the street, sitting on his motorbike. Next, the policeman was chasing the car. In the third picture, the car stopped, and the policeman got off his motorbike and approached the car. The last picture showed a group of clowns sitting closely next to each other in the small car.

The second comic strip was about a boy and his father. In the first picture, it was shown that a boy was squatting near his father who was making something from wood. Next, the boy was conversing with his father about a wooden board. In the third picture, the boy was talking to his friend, and the boy was sitting on a small wooden wagon. In the fourth picture, both the boy and his friend were enjoying the ride. However, in the fifth and sixth pictures, the boy and his friend began to lose control of

the wagon, as the ground started to slope downwards and the wagon was speeding up uncontrollably. The seventh picture showed the two kids being thrown hard and jumped into a valley. In the last picture, they landed hard on the ground headfirst.

For the mystery pictures, there were five choices for the students. The first one was a picture of an older sister with her younger brother climbing a cliff and they were at the top. The second picture was of a girl running into a mysterious room with stairs going down into the dark room. The next one, picture 3, was of two men in a kind of treehouse, with water under it. The fourth one was a picture of a man writing in the dark and it was raining. He was at the top of a cliff. The last picture, picture 5, showed a woman/ man (unclear because of the storm) walking alone in the snow. Presented below are samples of students' collaborative essays based on the visual images they got.

Text 1: Group 1's Completed Collaborative Essay Based on Comic Strip 1 (Unedited)

WAGON'S ACCIDENT

An accident could be happened to anyone who was careless. One day, there was a naughty little boy called Bob. He came to his father, Edward. He felt curious about what his father did and then he asked him, "What's that? What are you doing?" His father answered, "This is a wagon. I'm just fixing it."

Bob felt more curious about it and wanted to try it. An hour later, his father took a nap. Besides, Bob took the wagon secretly and brought it to jungle. He met a tiger, Toddy. He asked Toddy to play together. Toddy agreed and said that he would control the speed through the speedometer in his hand. Bob sit on the wagon and Toddy started push the wagon while running, so that it could go fast.

They enjoyed that game, but they worried about the speed of the wagon went faster and faster. Toddy looked to the speedometer and he shocked because the speed was out of limit. They became panic and didn't know what they had to do. They afraid and screamed. Finally, they fell at the brink and laid down on the ground. Bob said, "Huft! That was awful." Then Toddy replied, "Yes, absolutely awful and my speedometer is broken."

The accident made their body sick and the speedometer and the wagon was broken. We can conclude that we have to be careful while doing something.

Text 2: Group 2's Competed Collaborative Essay Based on Comic Strip 2 (Unedited)

THE CLOWNS

A policeman chased a small car full of unexpected people. Long time ago, there was a funny accident in Noting Hill. It was a little town in London that popular with the high discipline traffic system. There was a firm cop at that time, Mr. Danish. He has been working as a cop for about fifteen years. He was really obedient to the rules.

One day, he looked a small car drove very fast. He said, "Hi, the ugly small car, stop please!" However, the car didn't stop and went faster than before. He chased the car by his motorcycle. He screamed out the car and said, "Stop! I will kill you by my hand if you don't stop right now!"

After that the car stopped immediately. He walked to the car and said, "Stop little kid! Or I will slap your face! Get out of the car!" He knocked the drive's window and the driver opened it. He shocked after looked who the passenger is.

They were not naughty little kids, but they are group of clowns. The driver said to him, "What's up? Seriously? You want to slap twenty of us? Our face one by one?" Mr. Danish astonished and just said, "Oh, no!". From that story we can learn that we have to think first before we take an action.

From Text 1 and Text 2, it was clear that the students really worked hard to make the best stories that they could, based on the comic strips given. They worked collaboratively with other friends in small groups. After such conflicting moments with their peer students in the groups to decide the plot of the story, they came up with interesting ideas for their essays. This is one of the benefits of this activity: students practiced and sharpened their soft skills while they were negotiating meaning and finding the best solutions for any disagreement.

Text 3: Students' Collaborative Narrative Essay 1 Based on Mystery Picture 5 (Unedited)

A Choice

Once upon a time, there was a man named Edward. He was a kind person, patient, loyal, and always help each other. Edward had a beautiful wife named Diane. He loved her so much with all of his heart, but Diane betrayed him. When Edward knew that his wife betrayed him, Edward tried to stay and kept struggle for his relationship. Edward always be patient in facing his wife. Overtime, as a husband that had been betrayed by the woman that he loved, he realized that every patient had limit. Because of that, he should chose a bitter choice for a better life.

Edward and Diane were the best couple in their city. Edward was a kind person and Diane was one of the most beautiful women in the city. Everyone who looked them will be felt jealous, because they were really harmonious and romantic couple. Everything had changed when Diane met another man named Alfonso. She fell in love with Alfonso, "Oh God! I falling in love with Alfonso and wanna live with him, but how about my husband? What I should says to him ?" Diane thinks. Edward tried to be patient and defend his relationship, but Diane still chose to left him. Edward's heart was so broke and he did not have other choises except to let Diane go. He realized that this is the most bitter choice that he ever made, "I hope this is a better choice for myself," he said. Edward felt that he should

did something to forgot his problems, "I need some vacation to refresh my mind," he thinks. So he decided to roam places that he never knew to found his happiness again.

Edward strated his new journey to forgot about his problems. He just started his journey alone, without someone accompenied him. Sometimes he felt so lonely. His eyes looked so empty and right now he was so thin. He lost his spirit of life and did not eat anything for around a week. He just drank the drinks that he brought from his city. However, his drinks was not enough for his stock. In this condition, he was in a desert and felt thirsty, "I need some water. Who can help me?" he whimpering. At the same time, he also looked at the sand at the front of him. He imagine and wonder to himself, "Is there any hope for me?" he tried to not cried. It was a hard journey for edward, but he still tried to walking even though it was so hard to kept his two feet on the ground. Luckily, he finally found a city. Edward felt saved and the citizen in that city were so kind. Edward think that it was a perfect city to started his new life. Then, he built a new house in the city and made good relations to the other citizens. In that city, Edward became rich as the greatest sword maker in the city. It was made him so grateful.

Edward realized that his choice to leave his wife was the best for him, "maybe it is bitter, but I know it is better," he said. Edward so grateful for the result that he had got now. He tried to not hate Diane, he wish that she could be happy on her new life with her new man. Edward wanted to forgot about his internal problem with Diane, and he forgave her, "everyone has a choice and Diane too," he said in his heart. Right now, Edward was so happy with his new life. He also enjoyed his new activity as a sword maker. Edward realized that God had the best for him.

We should chose a painful choice for a better life. No matter if it so hard, no mater if it made you hurt. We should made a choice, because a choice was an important aspect in life. We should brave in made a choice, and remember that life is a choice.

Text 4: Students' Collaborative Narrative Essay 2 Based on Mystery Picture 3 (Unedited)

One day, Jenny and Jerry would go to Raja Ampat Island with a boat. They have been prepared all of the things that they need for a long time. They were so excited about it. Unfortunately, there was an accident. Suddenly, the storm has come and hit the boat into the secret island. They were so sad, but Jenny had a good idea to build a tree house. They were working together to make a tree house. Usually, people like to work independently. However, working with partners is better.

There are several things that they needed to build a tree house but they have not got it in the secret island. They needed some woods, nails and tools to build the tree house. They have tried to search anything that they need around the secret island but they could not find anything. Jenny said, "I have not got anything that we need to build the tree house." Jerry answered, "I think, we can use the tree and the grass." So, they used the tree and the grass to build a tree house."

The process to build a tree house was difficult. There were some things that happened to them. The wind was very heavy at that time, so they struggled to build the tree. Also, they

found the difficulties to build the tree house, but they could not solve the problem. Jenny said "build the tree house is very difficult, so I cannot do that." She did not continue on making the tree house and stopped helping Jerry on making the tree house. However, Jerry did not complain about that. The thing that he wanted was the tree house must be finished so both of them could stay longer in the secret island.

Jerry was building the tree house with a spirit. However, Jerry built the tree house alone by himself and Jenny just looked at him. At first, Jenny did care about it anymore. She gave up because she thought that the tree house would not be finished. On the other hand, Jenny felt pity to Jerry because Jerry made the tree house alone. Finally, Jenny helped him to build the tree house. They were building the tree house together. They did it. The tree house looks beautiful and amazing.

In conclusion, they are willing to help each other and they will sacrifice for one another whenever necessary. They promised to each other that they would always be together in any cases. At last, friends will always help you and they will not let you down no matter what the conditions are. A good friendship starts from how you can accept the positives and the negatives sides of your friends.

Students had to work hard in a collaborative team, thinking creatively and logically. Text 3 and Text 4 were two sample essays that they wrote based on the pictures given to them. They could make up such interesting stories with very good moral lessons. These students to my surprise were able to create such an interesting plot for each story that they chose. It is another benefit of this activity, developing creativity and critical thinking, which are two among other twenty-first century skills, being able to communicate clearly, collaborate with others, think critically, solve problems, be creative, and be innovative (Binkley et al, 2012).

In this section, the detailed stages of the picture-based narrative writing will also be discussed. The stages of writing narratives using pictures and comic strips are as follows. First was the selection of pictures and comics. After selecting the desired pictures and comics, students began interpreting them in groups. For the third step, students had to imagine what was happening and complete the comic strips which had no dialogs. After doing this fourth step, they had to proceed to the fifth step, which was making an outline of their essay. If they were confused, they could consult with the lecturer. Finally, they were ready to write their first draft. After they received input from the lecturer, they did revisions and submitted their final draft. These steps are described in Figure 2.

From the sample texts above, which were written by the students in collaboration, it can be seen that students' creativity and imagination developed throughout the learning process. The proofs can be seen from students' pre-tests, post-tests, journals and interviews results, all of which were presented in this study. In doing the picture-based activity, students were supposed to think creatively, logically, and imaginatively to produce a sensible piece of writing. These were the strengths of this writing process. Despite the grammatical errors that they still made, these students

showed that they were able to write narrative essays well. Not only did they have to deal with the plot of their essay, but also how interesting the story was. When done collaboratively, this work became lighter. Besides learning the skills of writing narrative essays, students also learned about soft skills, working in collaboration with others, delving into their imagination, using their creativity, applying logical thinking, and synthesizing those ideas together with other group members. Figure 1 shows the flow of stages in this research. Most of the time, students worked collaboratively with their peers in writing the narratives based on the pictures given.

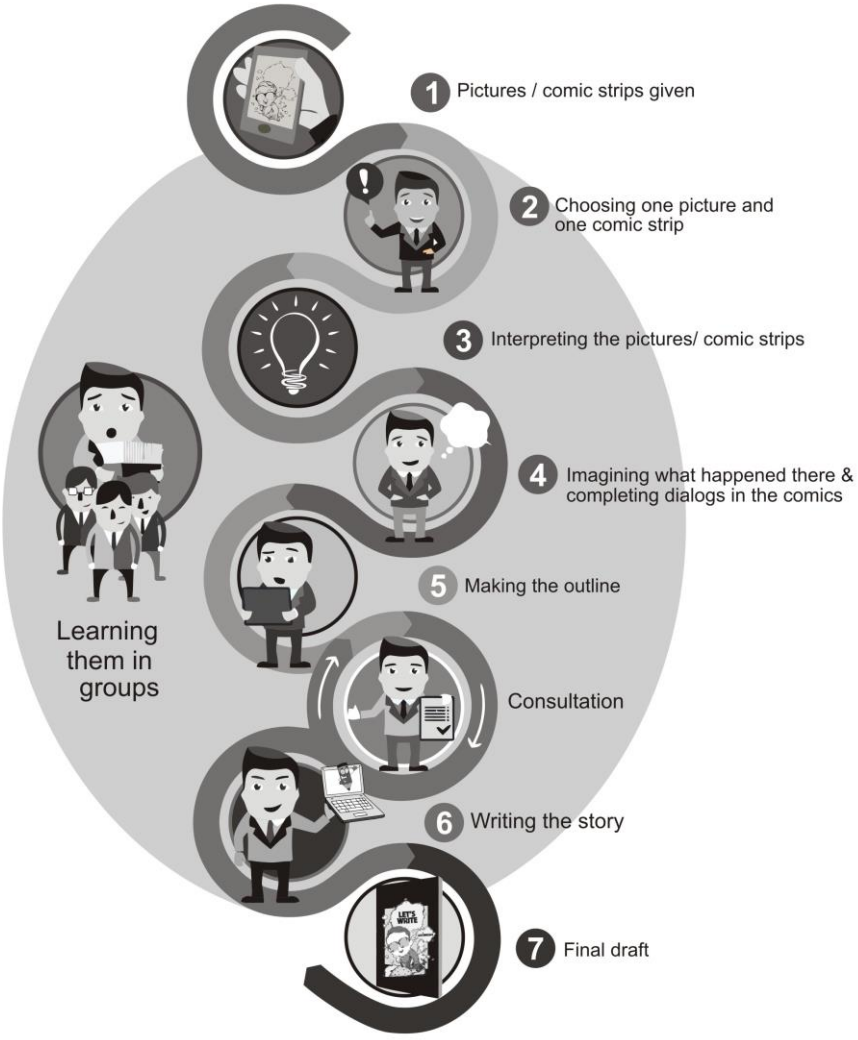


Figure 1. Infographic Picture of the Procedures of Narrative Essay Writing Based on Pictures/ Comic Strips

No matter how good a process is, there must be some drawbacks. The same case happened with this narrative writing using visual images. The main problem encountered by the students was when they did not have anything in mind. They were blank and just looked at the pictures as well as the comic strips with no ideas at all. Thanks to collaborative work, because when some students did not have any ideas in their minds, other group members could help them by sharing what they had in mind.

From students' journals, which were submitted after these series of activities were finished, various responses were given. Mostly, out of 19 students taking the class, 15 students said positive things about it. They mentioned that writing a narrative essay based on pictures or comic strips was enjoyable, fun, interesting, and challenging. It developed their imagination as well as creativity. This activity also improved their grammar and vocabulary. They were in a way 'forced' to use various adjectives to describe a situation or a person. However, due to space limitations, opinions from only five students who were randomly selected are discussed. Described in excerpts 1, 2, 3, 4, 5, and 6 are some statements from these students.

Excerpt 1: Student A's Journal on Picture-Based Stories

> *It made me play with my imagination more than usual, but we worked in a group of two. It was easier than working alone. My friend helped me to make the dialogue into a story. When we played with our imagination, it increased our new vocabulary and our creativity. It also helped me to improve my writing skills with the correct grammar.*

Student A seemed to enjoy this activity, even though this student was not a talented writer. She got help from her friends, which was very meaningful for her. Different from Student A, Student C felt that she enjoyed this activity mainly because she had to play with her imagination and try hard to make her story interesting for people.

Excerpt 2: Student C's Journal on Picture-Based Stories

> *It's a great idea and I like this activity. It forces you to use your imagination before you write a story from the picture that you've chosen. You must see the pictures first and choose your favorite one. There is no explanation what the picture is about. You have to imagine the right story for your picture, and after you decide it, write down your story idea. This activity is really fun and challenging. The pictures are interesting also contain a mystery. All you need here is just an imagination and creativity to make it good and people will be interested with your story.*

Different from his peer students, Student F emphasized that creativity was the key to doing this exercise. For him, this activity was not a difficult one. This can be seen from Excerpt 3.

Excerpt 3: Student F's Journal on Picture-Based Stories

> *Learning narrative writing through games is very fun and makes learning become easy to understand. For example, during the picture-based story activities we had to make a topic sentence based on the picture. By looking at the picture, we could arrange the story. I think that was easy to do for everyone. I understood a lot by using this method because the games could make me think creatively. I liked this method and I think it was enjoyable for everyone in this class. I think that is my opinion towards my impression about learning narrative writing through games.*

What this student meant by 'games' was completing the comic strips with dialogs. As mentioned earlier, two comic strips were given to the students. One was a series of pictures of a policeman chasing a speeding car, and the other was a picture of a man fixing a stroller for his kid. Besides comic strips, mystery pictures were also given to them. The next opinions presented here come from Student I and Student O.

Excerpt 4: Student I's Journal on Picture-Based Stories

> *I was very cheerful to work together with my team members. I felt very happy taking this class. Gladly, I kept praying to calm me down, practicing and studying hard on the difficult parts or when I lacked in doing those parts. Someday, I will become a better writer in the future.*

Student I felt excited and happy taking this class, mainly because her close friends were in the same class she was in. That was why she had no problems in adjusting herself with her peer students in the group as well as in interrelating ideas with them. Another student, Student O, mentioned that this activity made her defeat her boredom in the classroom. Being a student who was good at fashion design, Student O was basically an imaginative and creative student. This activity reinforced her being an imaginative and creative person.

Excerpt 5: Student O's Journal on Picture-Based Stories

> *In my opinion, using pictures in the Narrative Writing class was very interesting and also very helpful to me. This was interesting because these pictures made me not feel bored. These pictures also made the*

> *atmosphere more fun. The use of pictures was also very helpful to me because from the pictures, I came to know about the picture of the story that I would make. These pictures also helped to provoke my imagination when I was writing. So in my opinion, these pictures were very important in the writing class. In addition to being fun, it helped reduce my boredom and also increased my creativity in composing a story.*

Apart from the students who enjoyed the learning process using pictures and comic strips, there were also students who did not enjoy this activity fully. Only three students mentioned the drawbacks of these activities. Student A, Student H, and Student P mentioned several things which made them uncomfortable about this, even though they also admitted that they enjoyed the class. Here is Student A's statement about it: "*Sometimes I felt confused with the pictures entitled the kid. I thought it made me confused because in the picture there was a tiger, who was the boy's friend*" (Student A's statement in Journal 1).

Student H emphasized the difficulties that she faced. It dealt with interrelating her ideas with her friends': "*Sadly, when I made the story, I often found some difficulties. The difficulty that I often found in writing the story was to combine my ideas with my friend's and make the story interesting to read.*" (Student H's statement in Journal 1). The last student who did not really like writing based on pictures or comic strips was Student P. She admitted that actually it was fun. However, she felt tired because she felt that she was not a creative person:

Excerpt 6: Student P's Journal on Picture-Based Stories

> *For me it's fun sometimes but sometimes it's tiring because I'm not the typical person who is creative enough to make a story. It needs a long time for me to think really hard about the idea, plot, and characters for the story. Besides that, it's a bit complicated for me to put the grammatical structures in writing. I feel that every writing that I make still has some wrong grammatical structure.*

From these three students who did not really like the idea of narrative writing based on pictures or comic strips, there is at least one thing in common. They all admitted that the activity was fun and challenging. However, their limitations like their fear of lack of creativity, difficulty in interrelating ideas, and confusion in making the plot became obstacles for them to enjoy all the class sessions.

From interviews with two Narrative Writing students, in January 2018, Student B and Student M who were randomly selected, it was revealed that they faced some challenges writing narratives based on the pictures given. Student M mentioned that the main difficulties lied in language problems and coherence in writing the story. She and her friends tried to seek help from the lecturer, as described in the excerpt below:

Excerpt 7: Student M's Journal on Picture-Based Stories

> *Sometimes, when the students were trying to produce an essay based on the picture, they found some difficulties. For example, they had already written an essay well, but the grammar was wrong. People that read that essay maybe could not understand, because the sentences were very confusing. The second one, when the students were writing an essay, they needed conjunctions to connect one sentence with the other sentence. Sometimes the students could not choose the right conjunction. If the students found a difficulty like that, they could meet with the lecturer to ask advice about an essay in how to use good grammar and conjunctions.*

Student B also mentioned that he encountered difficulties in grammar, especially when there were flashbacks and progressive plots in the same essay. He also got confused with the tenses to be used, and the diction he should use. Being stuck in finding ideas was something which made him frustrated. To solve these problems, Student B tried to use a thesaurus and consulted the lecturer to confirm his sentences. Besides that, he tried to analyze the pictures carefully, "*I tried to analyze and imagine the pictures in real life. I also tried to connect the pictures with films or animations which I saw before. Finally, I tried to share my ideas with my partners in my group.*"

In summary, it can be concluded that pictures can be used as an effective teaching aid in narrative writing. This is in line with what Carry (2016) mentioned, in that students need visual images to help them read and understand texts. This is also supported by Asrifan (2015), asserting that to boost students' writing abilities, students need to be creative. Teachers, therefore, need to create an environment which allows for students' creativeness, so that students can get involved deeply into the writing process, and have both passion and compassion to write. They will have no fear or doubt and will not think that they cannot write. From the quantitative data, these findings say a lot about students' improvements in their narrative writing scores, which will be discussed in detail in the next part.

Discussion, Conclusion, and Recommendations

In this part, the process of the essay writing based on visual images from the beginning till the end, the strengths, the main challenges, as well as the solutions to the problems are discussed. First, the answers to the two research questions are discussed. The first question was: *How can pictures and comic strips be used to enhance students' creativity and ability in narrative writing?* It was answered in the previous section. Starting from the first stage, students worked in pairs or groups of three. They brainstormed ideas, made outlines, wrote the first drafts, and revised the drafts together. Each group was given the task to write one narrative essay based on one comic strip, and one narrative essay based on a mystery picture. The results show that

these students could be very creative and imaginative in their ideas, as reflected in their essays.

The second question was: *What are students' perspectives on the use of pictures and comic series in narrative writing?* Out of 19 students in the Professional Narrative Writing class, only three students had mixed feelings of both excitement and confusion or tiredness. Their main problems were a lack of ideas, confusion in creating dialogs, and tiredness of thinking too hard about the characters, plot, conflicts, and dialogs in the stories. The first and second research questions have thus been answered.

As previously mentioned, similar studies have been done in the area of narrative writing using pictures. Gutiérrez et al. (2015) conducted experimental research on 20 EFL ninth graders in an urban area of Colombia. After those 16 meetings, the experimental group improved in their narrative writing. Using pictures helped these students increase their vocabulary usage. They could utilize words that explained and evaluated the pictures. The students in the experimental group outperformed the ones in the control group.

Two other researchers, Ismatuti and Suparno (2012) also conducted research on junior high school students in Solo, Indonesia. They concluded that their respondents, SMPN 2 Ampel students, improved in their motivation after being given a picture series in their writing activity. In the meantime, Aschawir (2014) also did a similar kind of research on second-semester students of the English Department of the Faculty of Letters of UMI, Indonesia. From his research, it was found that using picture series increased experimental group members' post-test writing scores.

Another piece of research was also conducted by Rayo (2015). He examined the effectiveness of the use of picture stories to improve verb tenses in the narrative writing of adult English learners. There were 36 participants from China. From his study, the results showed that there were no significant improvements in the experimental group. He found that these learners preferred using the simple past form over other verb forms in their narrative writing. Another researcher, Asrifan (2015), conducted an experimental study on 78 senior high school students to find out whether the use of picture stories could improve students' writing abilities or not. The findings showed that using pictures enhanced students' abilities to write narrative compositions at Public Senior High School (SMAN) 3 in Pare Pare, South Sulawesi, Indonesia.

Ayuningtyas and Wulyani (2012) conducted action research on the eleventh graders of Public Senior High School (SMAN) 1 in Srengat Blitar, Indonesia. They wanted to know how the abilities of eleventh graders in writing narrative texts could be developed using picture sequences as instructional media. Data were obtained through two observations checklists, questionnaires, and the students' writing products. The findings showed that there were good improvements in terms of students' attitudes and writing products. Silva et al. (2017) also reported several studies involving the use of pictures in different fields of study like physics and language. Pictures were found to be beneficial in facilitating learning. Students became motivated and more creative. This present study confirms the previous

findings that pictures can enhance students' creativity and critical thinking in tertiary-level writing.

There were some similarities between this present study and research conducted by Gutiérrez et al. (2015), Ismatuti and Suparno (2012), Aschawir (2014), Ayuningtyas and Wulyani (2012), and Rayo (2015). All these studies dealt with the same thing, that is, the use of picture series in narrative essay writing. This was the red line of both: all the pieces of research made use of pictures as an aid to boost students' creativity and abilities in writing. This present study was different from the previously conducted ones in terms of the design. This present study was qualitative in nature, while the others were quantitatively implemented. The educational levels of the participants in all these different studies were also different. In the above-mentioned studies, all the participants were from the secondary level, high school students; while in this study, the participants were tertiary level students.

As seen from the pre-tests and post-tests which were given to the students, there were improvements in students' writing. From the pre-tests, the lowest score was 40, while the highest was 87.5, and the average was 54.21. In the post-tests, the lowest was 52.50, the highest was 92.5, with an average of 70.66. It meant that the difference between the average of the pre-test and post-test results was very big, which was 16.45. However, there were three students whose pre-tests and post-tests did not have big differences; they were Student I, Student P, and Student R. Further research needs to be done to determine the reasons behind this fact.

Table 1
Students' Pre-test & Post-test Scores

NAMES	Pre-test	Post-test
Student A	52.50	70.00
Student B	60.00	82.50
Student C	57.50	77.50
Student D	70.00	82.50
Student E	50.00	80.00
Student F	40.00	80.00
Student G	45.00	55.00
Student H	65.00	92.50
Student I	87.50	90.00

Table 1 Continue

Student J	**40.00**	**60.00**
Student K	50.00	57.50
Student L	50.00	70.00
Student M	55.00	70.00
Student N	45.00	55.00
Student O	55.00	65.00
Student P	57.50	60.00
Student Q	50.00	80.00
Student R	50.00	52.50
Student S	50.00	62.50

The student participants in this study also faced some challenges as they revealed in their journals, and they tried hard to solve their problems. In general, participants admitted the benefits and enjoyed this picture-based writing activity. They mentioned that it improved their imagination and creativity, as well as their vocabulary and writing skills. This activity was also fun, interesting, and engaging.

Besides the benefits of writing narratives using pictures, there were some drawbacks too. The main difficulties were in interrelating individual ideas with those of other students. Besides that, some students, especially those who felt that they were not creative enough, admitted that they faced difficulties in finding ideas, creating the plot, and developing the characters. They also got confused in writing the story. Thus, this activity was tiring for them and increased their anxiety. In overcoming those challenges, students received help from their peer students in the group.

This study acquired data from various methods: students' texts, journals, scores, and interviews. From all the data derived in this study, several conclusions can be drawn. First, visual images are needed to build up students' imaginative minds. Visual images here are comic strips and pictures. They are there to facilitate students in writing narratives. These visual images can be used to stimulate students to be creative and think logically.

Second, comic strips with empty bubbles and mystery pictures are good media to enhance their creativity in narrative writing and to sharpen students' soft skills as well. Students need to train their soft skills while cooperating with their peers. Expressing ideas, refusing one's idea politely, interrelating one's own idea with others' ideas, showing disagreements, suggesting new things, and other skills involving special social abilities need to be sharpened.

The last conclusion is that the lecturer should help these learners boost their creativity, imagination, and logic. Two types of media that can be used are pictures and comic strips. Hopefully, this piece of research can be useful for other lecturers of narrative writing, giving a new horizon for lecturers in giving performance-based assessments as well as sharpening students' soft skills. Future researchers can conduct similar research on other writing courses, like Procedural Writing or Creative Writing. They can also delve into other types of pictures more deeply, or ask students' opinions on what types of pictures they like best as an aid for them to write narratives.

What was impressive from these fifth semester students was how they could be creative and imaginative in their writing. I suggest that future researchers conduct research on narrative writing using more kinds of images to help motivate students to be creative and think critically, including pictures and videos. Other designs of research can also be applied like experimental or action research, as conducted by Gutiérrez et al. (2015), as well as Rayo (2015), and Ayuningtyas and Wulyani (2015).

Further research on collaborative work can also be conducted involving more participants. Future researchers can also carry out similar studies on different types of writing, namely *Creative Writing* or *Procedural Writing,* to see how effective the teaching can be when aided with pictures or other forms of visual images. There are some possible implications in the narrative essay writing teaching and learning process. Students' imagination and creativity will be built up and enhanced when they are tasked to write with the help of pictures. Lecturers can also maximize these students' abilities to transform what is in their minds into a written form.

Acknowledgments

Special thanks are due to UKSW for funding this research, also to the students of Professional Narrative Writing Class of UKSW Salatiga, Indonesia, Semester I, 2017-2018. I am also indebted to Mr. Andrew Thren, my language advisor, as well as Mr. Lourentius Febby, whose comic strips and pictures were used in this study.

References

Aschawir, A. (2014). Using series pictures to develop the students' ideas in English narrative writing. *Scholarly Journal of Education*, 3(7), 88-95. Retrieved from http://scholarly-journals.com/sje/archive/2014/October/pdf/Aschawir.pdf

Asrifan, A. (2015). The use of picture story in improving students' ability to write narrative composition. *International Journal of Language and Linguistics*, 3(4), 244-251. Retrieved from http://article.sciencepublishinggroup.com/html/10.11648.j.ijll.20150304.18.html

Ayuningtyas, B. D. & Wulyani, A. N. (2012). *Using picture sequences to improve the ability of eleventh graders at SMAN 1 Srengat Blitar in Writing Narrative Texts*. Retrieved from http://jurnal-

online.um.ac.id/data/artikel/artikel637D0BAC482D981925210404FD9CCF72.pdf

Beckley, N. M. (2014). *Relationship between visual and written narratives in student engagement*. (MA thesis). University of Iowa. Retrieved from https://ir.uiowa.edu/etd/4574/

Binkley et al. (2012). Griffin et al. (eds.). *Assessment and Teaching of 21st Century Skills*. Retrieved from https://www.uxgreece.com/uploads/4/3/9/9/43997425/21st_century_skills.pdf

Burns, A. (2003). *Collaborative action research for English language teachers*. Cambridge: CUP.

Carry, D. D. (2016). *Visual literacy in English language teaching*. Cambridge: CUP. Retrieved from: http://languageresearch.cambridge.org/images/Language_Research/CambridgePapers/CambridgePapersInELT_VisualLiteracy_2016_ONLINE.pdf

Clark, J. & Paivio, A. (1991). Dual Coding Theory and education. *Educational Psychology Review*, 3(3), 143-211. Retrieved from http://www.csuchico.edu/~nschwartz/Clark%20%26%20Paivio.pdf

Cresswell, J. (2009). *Research designs*. London: Sage Publications.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction*. 8th Ed. Boston: Pearson Education.

Grainger, T. (2004). *Language and literacy*. NY: Routledge and Palmer.

Gutiérrez, K. G. C., Puello, M. N., & Galvis, L. A. P. (2015). Using picture series techniques to enhance narrative writing among ninth grade students at Institución Educativa Simón Araujo. *English Language Teaching*, 8(5), 45-71. Retrieved from https://files.eric.ed.gov/fulltext/EJ1075242.pdf

Halliday, A. (2002). *Doing and writing qualitative research*. London: Sage Publications.

Hilliard, P. (2015). *Performance-based assessment: Reviewing the basics.* Retrieved from: https://www.edutopia.org/blog/performance-based-assessment-revewing-basics-patricia-hilliard

Imastuti, M. W., Sujoko, & Suparno. (2012). *Improving students' writing ability in narrative text using picture series*. (An unpublished paper). English Education Study Program of UNS Surakarta Indonesia. Retrieved from https://media.neliti.com/media/publications/60129-EN-improving-students-writing-ability-in-na.pdf

Kamehameha Schools Research & Evaluation. (2010). *21st century skills for students and teachers*. Retrieved from http://www.ksbe.edu/_assets/spi/pdfs/21_century_skills_full.pdf.

Leung, L. (2015). Validity, reliability, and generalizability in qualitative research. *Journal of Family Medicine and Family Care.* Vol. 4(3), 324-327. July-September 2015. Retrieved July 29, 2018, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4535087/

OECD. (2018). *The future we want*. Retrieved from https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf

Torres, P. (n.d.). *Personal narrative – Six trait rubric.* Retrieved from https://www.pinterest.com/pin/429249408208038168/

Randle, K. (n.d.). *Visual Arts: Effective means to enhance creative writing quality*. https://www2.cortland.edu/dotAsset/122273.pdf

Rayo, D. S. (2015). *Writing narratives with the aid of pictures.* (Thesis for Master of Arts Degree). University of Western Ontario. Retrieved from http://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=4365&context=etd

Silva, A. B., Santos, G. T., & Bispo, A. C. K. A. (2017). The comics as teaching strategy in learning of students in an undergraduate management program. *Mackenzie Management Review*, 18(1), 40-65. Retrieved from www.scielo.br>pdf.>ram>1678-6971-ram-18-01-0040.pdf.

Schweizer, M. L. (1999). *The effect of content, style, and color of picture prompts on narrative writing: An analysis of fifth and eighth grade students' writing*. (A paper). The Faculty of the Virginia Polytechnic Institute and State University. Retrieved from https://theses.lib.vt.edu/theses/available/etd-021899-102353/unrestricted/etd.pdf

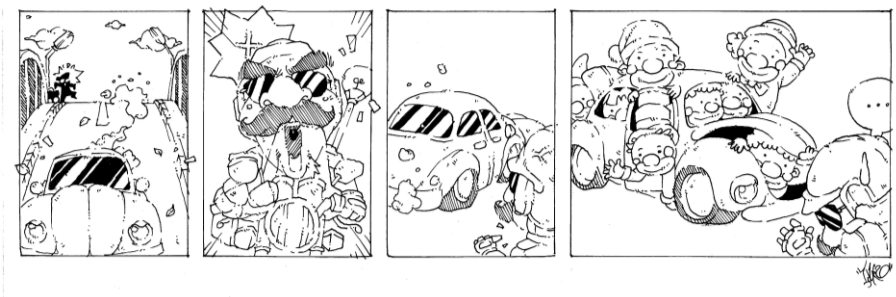
Thomas, N. J. T. (2014). Dual coding and common coding theories of memory. *Stanford Encyclopedia of Philosophy*. Retrieved from https://plato.stanford.edu/entries/mental-imagery/theories-memory.html on November 7, 2017.

Appendix A: The Six-Trait Rubrics (Torres, P., n.d.)

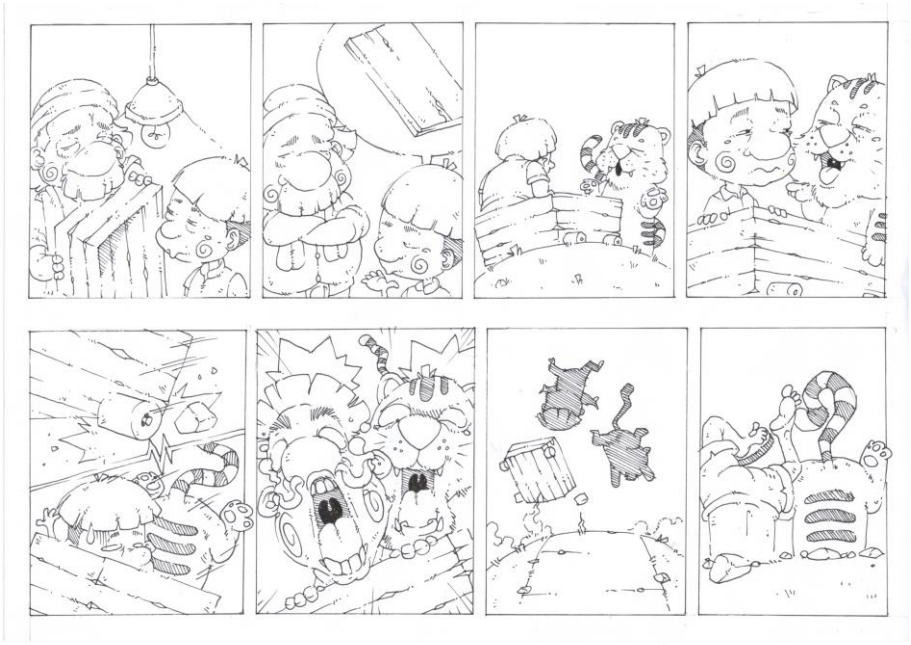
Ideas	Voice	Word Choice	Conventions
5 The narrative tells about an unforgettable experience. The details make the story truly memorable.	**5** The writer's voice creates an unforgettable experience for the reader.	**5** The writer's exceptional word choice captures the experience.	**5** Punctuation and grammar are correct. The writing is free from spelling errors.
4 The writer tells about an interesting experience. Details help create the interest.	**4** The writer's personal voice creates interest in the story. Dialogue is used.	**4** Specific nouns, strong verbs, and well-chosen modifiers create vivid pictures and express clear feelings.	**4** The narrative has a few minor errors in punctuation, spelling, or grammar.
3 The writer tells about an interesting experience. More details are needed.	**3** The writer's voice creates interest in the story. More dialogue is needed.	**3** Specific nouns, strong verbs are used. Modifiers are needed to create a clearer picture.	**3** The writing has several errors in punctuation, spelling, or grammar.
2 The writer needs to focus on one experience. Some details do not relate to the story.	**2** A voice can usually be heard. More dialogue is needed.	**2** Strong nouns, verbs and modifiers are needed to create a clear picture.	**2** Some errors confuse the reader.
1 The writer needs to focus on one experience. Details are needed.	**1** The voice is weak. Dialogue is needed.	**1** General and overused words do not create a clear picture.	**1** Many errors make the writing confusing and hard to read.

Appendix B: The Comic Series

Comic strip 1: The Clowns



Comic strip 2: The Stroller



Appendix C: The Mystery Pictures

Picture 1: On a Cliff

Picture 2: The Mysterious Room

Picture 3: The Tree House

Picture 4: The Cliff

Picture 5: Walking in the Storm

Eurasian Journal of Educational Research 80 (2019) 225-250

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

Assessment Scale of Academic Enablers: A Validity and Reliability Study*

Melike AKBIYIK[1], Murat SENTURK[2]

ARTICLE INFO

Article History:
Received: 13 Sep. 2018
Received in revised form: 21 Jan. 2019
Accepted: 15 Mar. 2019
DOI: 10.14689/ejer.2019.80.11

Keywords
academic competencies, academic qualification, academic enablers, university education

ABSTRACT

Purpose: This study aims to determine how university students evaluate the academic enablers they have acquired. To this end, the Assessment Scale of the Academic Enablers (ASAE) was developed and applied to 5,208 university students to test its validity and reliability.

Research Methods: The study employed a quantitative research design during the data collection and the analysis phases.

Findings: The ASAE consists of 20 items and three sub-factors: (i) learning competencies, (ii) communication and social competencies, and (iii) homework and responsibility competencies. Item-total correlations were found to vary from 0.44 to 0.70 and factor-loading values from 0.44 to 0.82. The three factors explain 51% of the total variance and the scale's reliability coefficient is 0.90. Results show that gender makes a significant difference in ASAE scores in favor of women. Also, the mean scores differed statistically depending on faculty. Particularly, students in the Faculty of Education obtained the highest scores in all factors; on the other hand students in the Faculty of Law earned mostly the lowest scores.

Implications for Research and Practice: The results show that the ASAE is a valid and reliable measurement tool that universities in Turkey can use to evaluate their success in using academic enablers for increasing students' academic success.

*This study was a part of the Istanbul University Students Profile Project that was funded by Scientific Research Projects Coordination Unit of Istanbul University with the project number of SBG-2016-21161.

[1] Corresponding Author, İstanbul University, TURKEY, E-mail: melike.akbiyik@istabul.edu.tr, ORCID: https://orcid.org/0000-0002-1838-380X

[2] İstanbul University, TURKEY, E-mail: murat.senturk@istanbul.edu.tr, ORCID: https://orcid.org/0000-0002-9628-6300

Introduction

Factors impacting students' academic success and performance constitute indicators of the quality of countries' education systems and determiners of educational policies in need of change (Alnabhan, Al-Zegoul & Harwell, 2001). It is, then, only pertinent that educators and policy makers examine these factors. There are several studies (Laidra, Pullman & Allik, 2006; Rohde & Thompson, 2007; Stanovich, Cunningham & Freeman, 1984) that accept cognitive capacity as the main factor influencing students' academic achievement while others (Jenkins & Demaray, 2015) focus on attitudes and behaviors, such as motivation (Linnenbrink & Pintrich, 2002; Schunk & Zimmerman, 1994; Zimmerman, 1998), self-efficacy (Elias & Loomis, 2002; Vrugt, Langereis & Hoogstraten, 1997; Wood & Locke, 1987), study behaviors (Devine, 1987; Gettinger & Seibert, 2002; Hoover & Patton, 1995), class participation (Cobb, 1972; Greenwood, Horton, & Utley, 2002; Willingham, Pollack & Lewis, 2002), and positive social behaviors (Malecki & Elliott, 2002; Wentzel & Watkins, 2002).

The literature indicates that higher education should be compatible with environmental needs and economic development (Chryssolouris, Mavrikios & Mourtzis, 2013; Davies, 2017; Marchello, 1987; Sohal, 2013) and should develop learning enablers, such as personal productivity, flexibility, and lifelong learning (Avargil, Herscovitz & Dori, 2012; Deaconu, Osoian, Zaharie & Achim, 2014; Mulder, Gulikers, Wesselink & Biemans, 2009). Competency-based systems first emerged in the USA in the 1970s (Winterton, 2009, as cited in Deaconu et al., 2014) and Mulder et al. (2009) state that currently, the US educational system takes the following three areas as its base: (i) students' acquisition of behavioral learning, (ii) their acquisition of the basic skills needed for all jobs, and (iii) performance improvement. Competency understanding in vocational, technical, and higher education also gained importance in Europe during the 1980s. In the 1990s, significant steps were taken in Europe through processes implemented first in Lisbon and then Bologna. In 2008, the European Qualifications Framework for Lifelong Learning (EQF) was formed. Thirty-nine countries, including Turkey, determined their own national qualifications regulations based on the EQF depending on education level.[1]

The National Qualifications Framework for Higher Education in Turkey (NQF-HETR)[2] determined qualifications to be the knowledge, skills, and competency (i.e., responsibility and autonomy) that universities should provide students of any level, area, and program. The EQF defined knowledge as theoretical and factual, skills as either cognitive (including logical, intuitive, and creative thinking) or practical (including manual skills and methods, materials, tools, and the use of tools), and competency as the student's ability to apply knowledge and skills independently and responsibly.

[1] For detailed information on the EQF, see: http://www.cedefop.europa.eu/en/events-and-projects/projects/european-qualifications-framework-eqf
[2] For detailed information about Turkey's National Qualifications Framework, see: http://tyyc.yok.gov.tr/?pid=10

Academic competencies indicate a student's performance as well as the standards that are used in assessing academic performance (Cole, 1991). DiPerna and Elliott (1999) describe academic competencies as a multi-dimensional structure necessary for academic success and include students' skills, attitudes, and behaviors as being among said competencies. Academic competencies are divided into two main components, namely, academic knowledge and skills (i.e., academic skills) and academic enablers.[3] Academic knowledge and skills and academic enablers are complementary components that work as long as they exist together. In other words, academic competence is not simply achieved by acquiring academic success through academic knowledge and skills; instead, it is achieved by also acquiring the tools that will support learning and the formation and application of academic knowledge and skills. Those attitudes and behaviors considered to be academic enablers have been identified as study skills, academic motivation, social relations, and participation (DiPerna & Elliott, 2002). Similar distinctions are also found in the form of hard and soft skills, where mostly soft skills are related to business life (Laker & Powell, 2011; Andrews & Higson, 2008; Schulz, 2008; Kumar & Hsiao, 2007).

During the development of the scale used in this study, several prior studies related to academic competency and enablers (e.g., Avargil et al. 2012; Deaconu et al., 2014; DiPerna & Elliot, 1999; Gettinger & Seibert, 2002; Linnenbrink & Pintrich, 2002; Mulder et al., 2009; Wentzel & Watkins, 2002) were examined, including the 8 key competencies of the European Commission for lifelong learning[4] (i.e., communication in one's native language, communication in a foreign language, basic competencies in mathematic and science/technologies, digital competencies, learning to learn, social and civic-related competencies, taking initiative and entrepreneurship, and cultural awareness and expression) and competencies at the undergraduate level in the NQF-HETR[5] (i.e., being able to work independently and take responsibility, learning competency, communication and social competency, and field-specific competencies).

Higher education must never cease to improve on and strengthen countries' human resources infrastructure by imparting on students' not only academic skills but also skills that will serve them in a holistic manner. To this end, Turkey's higher education system engages in many activities and supports studies aimed at improving universities and service quality. A secondary objective is to gauge how students perceive the many changes that have been made and the many services that have appeared as a result.

[3] In Turkish, there is no word that corresponds to the English word enablers. It is expressed in Turkish in a way that means something similar to *making it possible* or *facilitating an opportunity*.
[4] For detailed information, see: https://ec.europa.eu/education/sites/education/files/document-library-docs/factsheet-key-competences-lifelong-learning_en.pdf
[5] For more information on Level 6 (undergraduate education) competencies of the NQF-HETR, see: http://tyyc.yok.gov.tr/?pid=33

This study seeks to develop a scale measuring students' perceptions regarding how deeply they have acquired or improved in the academic enablers that higher education institutions endeavor to impart on them. Concordantly, this study is important in three aspects. Firstly, it will serve as a tool that universities may use to evaluate students' perceptions on the academic enablers intended to be imparted on them during their undergraduate education. Secondly, measuring students' perceptions will afford practitioners and policymakers valuable data to use in their respective fields. Thirdly, performing the study with 5,208 students in Istanbul University renders its reliability and validity strong.

Method

Research Design

This research was planned following a survey model seeking determining the certain characteristics of a group.

Research Sample

The universe of the research consists of undergraduate students who attended Istanbul University during the 2015-2016 academic-year. The minimum sample number required for accurate assessment was calculated as 3,914 with a 99% confidence level and a 2% margin of error. The sample size consisted of 5,208 students, with an average age of 22, the oldest being born in 1960 and the youngest being born in 2000. Of the students, 2,742 were female (52.65%) and 2,405 were male (46.18%).

Research Instruments and Procedures

Development of ASAE

Creation of item pool: A comprehensive literature review was completed on the qualifications, competencies, and enablers required for and pertaining to higher education. We furthermore investigated a measurement tool developed for freshmen students (CIRP; Astin, 1966) in America by the Higher Education Research Institute (HERI) (Eagen et al., 2015; HERI, 2016a, b, c, & d; Pryor, Hurtado, Saenz, Santos, & Korn, 2007) and measures used in profile studies performed at Cornell, Iowa State, Oregon State, Michigan, and Indiana universities and Carleton, Bowdoin, and Amherst colleges (Cornell University, 2015; Kuh, 2009; Massachusetts Institute of Technology, 2014). We also examined studies conducted in various European countries (Bargel, Ramm, & Multrus, 2001; Busse, 2015; University College of London, n.d.; The Higher Education Economy, 2013), Canada, and Australia (Baik, Naylor, & Arkoudis, 2015; Canadian University Survey Consortium & Prairie Research Associates, 2013). Finally, we analyzed studies examining higher education conducted at universities in China, Russia, and Kyrgyzstan (Centre of Development and Resources for Students, 2012; Ivanov Devlet Universitesi, n.d.; Moskova Devlet Universitesi, n.d.; Pomor Devlet Universitesi, n.d.).

Despite the vast number of internationally authored studies on this subject, similar studies are quite scarce in Turkey (Atasever, 2007; Cicek, Baykul & Keles, 2014; Gizir et al., 2010; Hatipoglu, Acar, Vural Akar & Binay, 2012; Kustepeli & Gulcan, 2002; Sencar, 2013; Sevuktekin, Nargelecekenler & Cetin, 2012; Yaylali et al., 2006), conducted mostly in different faculties (Akyurt, 2009; Cevik & Yigit, 2009; Ozel, 2006; Sahin, 2005; Senol & Tufekci, 2007; Tekin, 2014) and departments (Ekiz, 2006; Yigit, Esenay & Derebent, 2007, Ilgaz & Akdol, 2009; Issi, 2008; Kaya & Buyukkasap, 2005; Kizilcaoglu, 2003; Senses, 1999).[6] The theoretical and practical reviews have been provided to identify a large number of themes and items to pool from. The draft form has been created from the item pool.

Receiving and implementing expert opinion: In the second stage, five experts in the fields of scale evaluation and the educational sciences were consulted to determine not only the scale's linguistic and expressive appropriateness but also the suitability level of each item measuring the selected academic enablers.

Rewriting the items of the scale: In the third stage, the scale items were rewritten according to the field experts' recommendations.

Application of the pilot study: In the fourth stage, we performed a pilot study with 314 students after receiving approval from the ethics committee.

Item analysis, exploratory factor analysis: In the fifth stage, the raw state of the scale was applied to 5,208 students (see Table 1 for sampling design) in Istanbul University. Consequently, we conducted an exploratory factor analysis in order to learn its validity and reliability.

Table 1

Sampling Design[7]

No	Faculty	Universe	Calculated Sample	Target Sample	Respon-dents	Frequency Distribution (%)
1.	Physical Edu. & Sports	601	42.43	42	43	0.8
2.	Cerrahpasa Medicine	2,867	200.93	201	229	4.4
3.	State Conservatory	283	18.78	19	21	0.4
4.	Dentistry	1,028	72.44	72	72	1.4

[6] This is not an exhaustive list of studies on the topic in question.
[7] Numbers are based on the 2015-2016 academic year.

Table 1 Continue

No	Faculty	Universe	Calculated Sample	Target Sample	Respon-dents	Frequency Distribution (%)
5.	Pharmacy	1,144	80.06	80	88	1.7
6.	Literature	13,008	887.94	888	998	19.2
7.	Science	4,056	249.57	250	295	5.7
8.	F. Nightingale Nursing	1,038	73.28	74	79	1.5
9.	Hasan Ali Yucel Education	3,257	227.05	227	239	4.6
10.	Law	6,916	487.63	488	515	9.9
11.	Economics	10,992	757.40	757	611	11.7
12.	Theology	4,173	257.62	258	313	6.0
13.	Communication	3,775	263.76	264	173	3.3
14.	Istanbul Medicine	3,111	219.50	220	225	4.3
15.	Management	2,726	182.08	182	203	3.9
16.	Engineering	7,504	500.06	500	558	10.7
17.	Forestry	1,863	131.53	132	139	2.7
18.	Health Sciences	1,372	96.51	97	106	2.0
19.	Political Sciences	2,114	132.09	132	157	3.0
20.	Fisheries	370	18.92	19	27	.5
21.	Transport & Logistics	401	28.24	28	29	.6
22.	Veterinary Medicine	1,005	70.95	71	88	1.7
	Total	70,987	4,998.76	5,001	5,208	100

Validity and Reliability:

Exploratory Factor Analysis (EFA): The data obtained from respondents were subject to both a Kaiser-Meyer-Olkin (KMO) Test of Sampling Adequacy and Barlett's Test of Sphericity to assess suitability. The KMO index ranges from 0 to 1. So that respondent data may be considered suitable for factor analysis, not only should the KMO index be .50 or greater, Barlett's Test of Sphericity should be significant ($p < .05$) (Hair, Anderson, Tatham & Black, 1995). The KMO value was .93 and Barlett's test results ($\chi^2 = 7{,}521.998$; $SD = 190$; $p < .001$) were significant, meaning that the correlation matrix is suitable for an exploratory factor analysis to be conducted on it. While a factor load of .45 or greater is considered, .30 is often accepted (Otrar & Argin, 2015). In this study, .30 was accepted as the lower cut-off point for factor loading. A three-factor ASAE explaining 51% of the total variance emerged as a result of the factor analysis. The scale was found to consist of 20 items and item-total correlations ranged from .44 to .77. The factor loads related to sub-factors are given in Table 2.

Table 2

Factor Loads Related to the Sub-factors (n=5,208)

Factor 1: Learning Competencies	Rotated Factor Load	Item-Total Correlation
I have established a relationship between the events I encounter in daily life and what I have learned.	.56	.44
I have had the opportunity to learn how to work independently.	.48	.52
I have developed a positive attitude toward life-long learning.	.67	.59
I check what I have written in order to develop my writing skills.	.61	.66
I question the reliability and quality of the information I receive.	.74	.51
I try to find alternative solutions to problems.	.77	.63
I have reviewed scientific research and articles.	.60	.62
I have had the opportunity to take responsibility.	.64	.54
Eigenvalue = 7.17 Variance explained = 35.85%		

Table 2 Continue

Factor 2: Communication & Social Competencies	Rotated Factor Load	Item-Total Correlation
I communicate with one or more faculty members through email.	.53	.59
I work with my friends on class projects.	.55	.65
I have received advice from a faculty member after class.	.61	.50
I participate in in-class discussions.	.67	.62
I have worked with students on a project outside of class.	.63	.60
I have had the opportunity to develop my computer skills.	.44	.60
I have given an oral presentation on an issue facing society.	.65	.58
I have asked a faculty member questions in class.	.50	.64
Eigenvalue = 1.75 Variance explained = 8.76%		
Factor 3: Homework & Responsibility Competencies		
I have done homework after gathering information and ideas from different sources.	.67	.62
I have done homework in an electronic environment.	.66	.70
I come to class with my homework complete.	.74	.64
I turn in my homework on time.	.82	.58
Eigenvalue = 1.31 Variance explained = 6.53%		
KMO = 0.93 Bartlett's Test of Sphericity (χ^2 = 7,521.998; *SD* = 190)		

Using factors' content as a basis, the first sub-factor was named *Learning Competencies* (LC) and consists of 8 items. Items' factor loading ranged from .48 to .77. The factor's eigenvalue was 7.17, which corresponded to 35.85% of the total variance. The second sub-factor was named *Communication and Social Competencies* (C&SC) and consisted of 8 items. Items' factor loadings ranged from .44 and .67. The eigenvalue of the factor was 1.75, which corresponds to 8.76% of the total variance. The third sub-factor was named *Homework and Responsibility Competencies* (H&RC) and consisted of 4 items. Items' factor loads ranged between .58 and .70. The eigenvalue

of the factor was 1.31, which corresponded to 6.53% of the total variance. These three factors together explained 51% of the total variance.

Cronbach Alpha values related to the sub-factors are given in Table 3. In Table 3, it is seen that the Cronbach Alpha values for the sub-factor LC was .842, the sub-factor C&SC was .802, and the sub-factor H&RC was .813. Additionally, Cronbach's alpha value for ASAE was .904.

Table 3
Cronbach's Alpha Values for the Scale's Sub-factors (n=5208)

Factor	Cronbach's Alpha Value
1. LC	.842
2. C&SC	.802
3. H&RC	.813
Total	.904

After the reliability tests were conducted, an independent sample *t*-test was performed with the scores of those students who had scored in both the upper and lower 25 percentile. The *t*-test sought to determine both items' discriminating power and whether participants' answers to the items differed by group (Ergin, 1995). The results showed the differences for all groups to be statistically significant ($p < .001$) and that this difference favored the upper 25-percentile group (see Table 4).

Table 4
ASAE Scores by the Upper and Lower 25 Percentiles

Score	Groups	*n*	χ	S	S hx	*t*- test		
						t	*SD*	*p*
1. LC	Lower	1,152	17.72	3.82	.11	-143.13	2,436	.000
	Upper	1,286	35.91	2.35	.07			
2.C&SC	Lower	1,247	11.89	2.27	.06	-156.20	2,687	.000
	Upper	1,442	30.03	3.51	.09			
3. H&RC	Lower	1,197	6.30	1.93	.06	-177.12	2,774	.000
	Upper	1,579	17.62	1.43	.04			
Total Scale	Lower	1,310	40.12	7.59	.21	-146.42	2,580	.000
	Upper	1,272	81.21	6.63	.19			

Lastly, the Pearson product-moment correlation analysis showed a positive and significant relationship ($p < 0.001$) between factors. In other words, all factors contain the same structure (see Table 5).

Table 5

Pearson Product-Moment Correlations

Factors	C&SC	H&RC	Total
LC	.603*	.534*	.858*
C&SC		.616*	.888*
H&RC			.790*

* $p < .001$

Data Analysis

The data obtained were analyzed using IBM SPSS 21. We performed a descriptive analysis to devise evaluation criteria for academic enablers, an independent *t*-test to determine whether the mean scores differed by gender, and an ANOVA to ascertain whether the scores differed by faculty. Moreover, we conducted a post hoc Bonferroni[8] test to determine the origin of the differences observed.

Results

The mean and standard deviation scores related to the ASAE are given in Table 6. ASAE scores were, by sub-dimension, $\bar{X}$=27.52, s=6.93) in LC, $\bar{X}$=21.13, s=7.21) in C&SC, and $\bar{X}$=12.75, s=4.45) in H&RC. ASAE scores in general were $\bar{X}$=61.17, s=15.92).

[8] Post hoc results are not mentioned due to the limit on words allowed. However, the researchers are able to share them upon request.

Table 6

ASAE Results (n=5,208)

Sub-dimensions	$\bar{X}$	**SD**
1. LC	27.5	6.93
2.C&SC	21.1	7.21
3. H&RC	12.7	4.45
ASAE (Total Scale)	61.1	15.92

Table 7 illustrates that there is a significant difference in ASAE scores (t = 7.49; p < .001) by gender in favor of women ($\bar{X}_{women}$ = 62.74; $\bar{X}_{men}$ = 59.43). There are also significant differences in LC (t = 6.58; p < .001) in favor of women ($\bar{X}_{women}$ = 28.13; $\bar{X}_{men}$ = 26.86) and in H&RC (t = 13.44; p < .001) also in favor women ($\bar{X}_{women}$ = 13.53; $\bar{X}_{men}$ = 11.87). That being said, however, no significant difference between men or women was found in C&SC (t = 1.62; p > .05).

Table 7
ASAE Scores by Gender (n=5,208)

						***t*-Test**		
Measure	**Groups**	***n***	$\bar{X}$	***SS***	$Sh_{\bar{x}}$	***t***	***SD***	***P***
ASAE	Females	2,742	62.74	15.80	.30	7.49	5,145	.000
	Males	2,405	59.43	15.83	.32			
LC	Females	2,698	28.13	6.89	.13	6.58	5,060	.000
	Males	2,364	26.86	6.89	.14			
C&SC	Females	2,686	21.29	7.20	.14	1.62	5,024	.105
	Males	2,340	20.96	7.22	.15			
H&RC	Females	2,699	13.53	4.33	.08	13.44	5,048	.000
	Males	2,351	11.87	4.42	.09			

Table 8 depicts the ANOVA results showing that mean scores differed statistically by faculty in the ASAE (F = 42.82; p < .001). Specifically, students in the Faculty of Education obtained the highest scores whereas students attending the Faculty of Law earned the lowest scores.

Table 8
ANOVA Results of the ASAE by Faculty (n=5,208)

f, $\bar{X}$ & *SS* Values				ANOVA Results					
Group (Faculty)	***n***	$\bar{X}$	***SS***	**Var. *K.***	***KT***	***SD***	***KO***	***F***	***p***
Cerrahpasa Medicine	229	53.29	16.95	**Between groups**	194,960.61	21	9,283.84	42.82	.000
State Conservatory	21	65.38	17.78	**Within group**	1,124,355.39	5,186	216.81		
Dentistry	72	56.56	13.76	**Total**	1,319,316	5,207			
Pharmacy	88	54.15	15.08						
Literature	998	65.62	15.40						
Science	295	59.80	16.61						
Nursing	79	67.66	15.10						
Education	**239**	**71.79**	**13.19**						
Law	**515**	**48.89**	**14.06**						
Economy	611	55.60	15.27						
Theology	313	67.43	13.01						
Communication	173	62.66	15.05						
Istanbul Medicine	225	59.05	15.01						
Management	203	65.29	13.32						
Engineering	558	63.28	13.71						
Forestry	139	66.92	13.31						
Health Sciences	106	62.44	13.78						
Political Sciences	157	63.26	14.92						
Sports Sciences	43	67.44	12.09						
Fisheries	27	65.37	13.90						
Shipping & Logistics	29	59.76	13.40						
Veterinary	88	60.44	15.47						
Total	5,208	61.17	15.92						

Sub-dimensions' mean scores were also evaluated by faculty and were found to differ statistically by faculty in LC (F = 16.68; p < .001) (see Table 9) in C&SC (F = 43.58; p < .001) (see Table 10), and in H&RC (F = 66.61; p < .001) (see Table 11). Students of the Faculty of Education scored the highest in all three sub-dimensions. On the other hand, pharmacy students scored the lowest in LC and law students the lowest in C&SC and H&RC.

Table 9
One-Way ANOVA Results for LC by Faculty

f, $\bar{X}$ & SS Values				ANOVA Results					
Group (Faculty)	***n***	$\bar{X}$	***SS***	**Var. K.**	***KT***	***SD***	***KO***	***F***	***P***
Cerrahpasa Medicine	226	24.72	6.98	**Between groups**	15,813.32	21	753.02	16.68	.000
State Conservatory	21	27.14	8.28	**Within Groups**	230,206.51	5,100	45.14		
Dentistry	72	25.51	6.88	**Total**	246,019.83	5,121			
Pharmacy	**86**	**24.67**	**7.37**						
Literature	982	29.38	6.86						
Science	288	26.72	7.19						
Nursing	78	29.12	6.74						
Education	**238**	**30.65**	**5.73**						
Law	510	25.85	7.23						
Economy	598	25.89	6.88						
Theology	306	30.37	5.99						
Communication	169	27.74	6.66						
Istanbul Medicine	225	26.60	6.80						
Business	201	27.50	6.34						
Engineering	542	26.56	6.51						
Forestry	134	28.25	5.59						
Health Sciences	105	26.87	6.61						
Political Sciences	156	28.89	6.45						
Sports Sciences	42	28.79	5.94						
Fisheries	26	28.77	5.79						
Transport & Logistics	29	24.86	7.44						
Veterinary	88	27.31	6.74						
Total	5,122	27.52	6.93						

Table 10

One-Way ANOVA Results for C&SC by Faculty

f, $\bar{X}$, & SS Values				ANOVA Results					
Group (Faculty)	***n***	$\bar{X}$	***SS***	**Var. K.**	***KT***	***SD***	***KO***	***F***	***P***
Cerrahpasa Medicine	229	17.74	7.12	**Between groups**	40,486.86	21	1,927.95	43.58	.000
State Conservatory	20	24.10	7.89	**Within groups**	224,018.75	5064	44.24		
Dentistry	72	17.90	6.40	**Total**	264,505.61	5085			
Pharmacy	87	18.09	6.56						
Literature	973	22.28	7.09						
Science	284	21.56	7.29						
Nursing	78	24.24	6.48						
Education	**231**	**25.99**	**6.52**						
Law	**509**	**15.38**	**6.02**						
Economics	593	18.77	6.89						
Theology	305	23.19	6.09						
Communication	168	21.91	7.40						
Istanbul Medicine	223	19.70	6.51						
Business	201	23.53	6.24						
Engineering	539	23.45	6.02						
Forestry	135	24.53	6.33						
Health Sciences	102	22.22	5.68						
Political Sciences	153	21.14	6.86						
Sports Sciences	43	25.21	5.24						
Fisheries	26	23.15	6.89						
Transportation & Logistics	28	21.46	6.87						
Veterinary	87	20.59	7.37						
Total	5,086	21.13	7.21						

Table 11

One-Way ANOVA Results for H&RC by Faculty

***f*, $\bar{X}$ & *SS* Values**				**ANOVA Results**					
Group (Faculty)	***n***	$\bar{X}$	***SS***	**Var. K.**	***KT***	***SD***	***KO***	***F***	***p***
Cerrahpasa Medicine	226	10.82	5.16	**Between Groups**	21,796.59	21	1,037.93	66.61	.000
State Conservatory	21	14.24	4.43	**Within groups**	79,269.22	5,087	15.58		
Dentistry	71	13.20	3.67	**Total**	101,065.81	5,108			
Pharmacy	85	11.75	3.41						
Literature	972	14.31	4.03						
Science	289	11.98	4.32						
Nursing	77	14.82	3.67						
Education	**236**	**15.21**	**3.28**						
Law	**510**	**7.73**	**4.02**						
Economics	598	11.28	4.45						
Theology	306	14.28	3.29						
Communication	171	13.21	4.04						
Istanbul Medicine	223	12.78	3.87						
Business	202	14.26	3.32						
Engineering	544	13.61	3.50						
Forestry	135	14.70	3.49						
Health Sciences	106	13.55	3.35						
Political Sciences	152	13.53	4.08						
Sports Sciences	43	13.49	3.10						
Fisheries	27	13.67	3.46						
Transport & Logistics	28	13.86	4.16						
Veterinary	87	12.59	4.23						
Total	5,109	12.75	4.45						

Discussion, Conclusion and Recommendations

This study has aimed to develop a valid and reliable scale that can be used to measure competencies and, more specifically, the academic enablers identified by the NQF-HETR that students are expected to acquire during their undergraduate education. To determine the structure validity of the 20-item ASAE an exploratory factor analysis was conducted using varimax rotation. As a result, we ascertained there to be three factors that account for 51% of the total variance. Factor load values for each item in the scale vary between 0.44 and 0.82. These dimensions were defined as *leaning competencies*, *communication and social competencies*, and *homework and responsibility competencies*. Cronbach's alpha reliability value related to the ASAE was 0.90 and the three factors have a Cronbach's alpha value greater than 0.80, indicating both the scale as a whole and its dimensions to be internally consistent. The differences for all groups were found to be statistically significant ($p < .001$) as a result of the factor-based discriminant analysis we conducted. Item-total correlation coefficients ranged between 0.44 and 0.70, and item-remainder correlation coefficients between 0.43 and 0.65. In other words, they are all above the general acceptance of 0.20. An examination of the correlation among the three factors led to the meaningful and positive relations to be stated statistically. The existence of high and positive relations not only indicates that the scale consists of independent factors but also proves that they have the same structure.

As a result, the 20-item ASAE was prepared in the form of a five-point Likert scale from 1 = "never" to 5 = "always." In other words, no items were reverse coded. A total score is obtained from the scale, and this score shows at what level students are considered to have acquired the academic enablers in question from their university.

According to findings, the Assessment Scale of Academic Enablers is valid, reliable, and suitable for understanding to what degree students have acquired or developed the academic enablers that the NQF-HETR requires universities to impart on them during their undergraduate education. The scale will provide universities with detailed information on how to identify the strengths and weaknesses of their institutions in imparting academic enablers and how to establish or reestablish the link between academia and employment.

References

Akyurt, N. (2009). Meslek Yuksekokullari ve Marmara Universitesi Saglik Hizmetleri Meslek Yuksekokulu ogrencilerinin genel profili [Vocational schools and Marmara University's Health Services vocational school general student profile]. *Firat Saglik Hizmetleri Dergisi*, *4*(11), 175–189.

Alnabhan, M., Al-Zegoul, E., & Harwell, M. (2001). Factors related to achievement levels of education students at Mu'tah Universtiy. *Assessment and Evaluation in Higher Education*, *26*(6), 593–604.

Andrews, J., & Higson, H. (2008). Graduate employability: 'Soft skills' versus 'hard' business knowledge: A European study. *Higher Education in Europe*, *33*(4), 411–422.

Atasever, S. (2007). *Gazi Universitesi ogrenci profili* [Gazi University student profile]. Master's thesis, Gazi Universitesi, Egitim Bilimleri Enstitusu, Ankara. Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/

Avargil, S., Herscovitz, O., & Dori, Y. J. (2012). Teaching thinking skills in context-based learning: Teachers' challenges and assessment knowledge. *Journal of Science Education and Technology*, *21*, 207–225.

Baik, C., Naylor, R., & Arkoudis, S. (2015). The first year experience in Australian Universities: Findings from two decades, 1994-2014. Melbourne: Melbourne Centre for the Study of Higher Education, The University of Melbourne. Retrieved from

http://melbourne-cshe.unimelb.edu.au/__data/assets/pdf_file/0016/1513123/FYE-2014-FULL-report-FINAL-web.pdf

Bargel, T., Ramm, M., & Multrus, F. (2001). Studiensituation und studentische Orientierungen [Study situation and student orientations]. 7. Studierendensurvey an Universitaten und Fachhochschulen, Bonn: Bundesministerium für Bildung und Forschung.

Busse, B. (2015). Studierendensurveys in Landerndes Europaischen Hochschulraumes [Student surveys in the lands of the European Higher Education Area]. Heftezur Bildungs – und Hochschulforschung 80, Universitat Konstanz, Arbeitsgruppe Hochschulforschung.

Canadian University Survey Consortium & Prairie Research Associates. (2013, June). *CUSC 2013 First-Year University Student Survey-Master Report.* Retrieved from http://www.cusc-ccreu.ca/publications/2013_CUSC_FirstYear_master%20report.pdf

Centre of Development and Resources for Students. (2012). *Graduate employment survey, 2012 – Bachelor's degrees.* Pokfulam: CEDARS & The University of Hong Kong. Retrieved from http://www.cedars.hku.hk/sections/careersplacement/GraduateEmploymentSurvey/report/FD/FD_report_2012.pdf

Chryssolouris, G., Mavrikios, D., & Mourtzis, D. (2013). Manufacturing systems: Skills & competencies for the future. *Procedia CIRP*, *7*, 17–24.

Cobb, J. (1972). Relationships of discrete classroom behaviors to fourth grade academic achievement. *Journal of Educational Psychology*, *63*, 74–80.

Cole, D. (1991). Change in self-perceived competence as a function of peer and teacher evaluation. *Developmental Psychology*, *27*, 682–688.

Cornell University. (2015). *PULSE survey instrument 2015* [Survey Form]. New York: Cornell University. Retrieved from

http://irp.dpb.cornell.edu/wp-content/uploads/2012/03/2015-Cornell-PULSE-survey-instrument-for-web.pdf

Cevik, O., & Yigit, S. (2009). Egitim Fakultesi ogrencilerinin profillerinin belirlenmesi - Amasya Universitesi ornegi [Determining student profiles for the Faculty of Education: The case of Amasya University]. *Cukurova Universitesi Sosyal Bilimler Dergisi, 33*(1), 89–106.

Cicek, H. S., Baykul, A., & Keles, M. K. (2014). Ogrenci profilinin degerlendirilmesine yonelik bir arastirmasi: Suleyman Demirel Universitesi ornegi [A research on student profile evaluations: The case of Suleyman Demirel University]. *SDU Fen Edebiyat Fakultesi Sosyal Bilimler Dergisi, 33*, 171–186.

Davies, H. (2017). Competence-based curricula in the context of Bologna and EU higher education policy. *Pharmacy, 5*, 17.

Deaconu, A., Osoian, C., Zaharie, M., & Achim, S. A. (2014). Competencies in higher education system: An empirical analysis of employers' perceptions. *Contemporary Priorities in Business Education, 16*(37), 857–873.

Devine, T. (1987). *Teaching study skills: A guide for teachers*. Boston: Allyn and Bacon.

DiPerna, J., & Elliott, S. (1999). Development and validation of the academic competence evaluation scales. *Journal of Psychoeducational Assessment, 17*, 207–225.

DiPerna, J., & Elliott, S. (2002). Promoting academic anablers to improve student achievement: An introduction to the mini-series. *School Psychology Review, 31*(3), 293–297.

Ekiz, D. (2006). Sinif ogretmenligi meslegine yonelen adaylarin profilleri ve gelecege yonelik beklentilerinin incelenmesi [Analysis of teacher candidate profiles on the teaching profession and their expectations of the future]. *Firat Universitesi Sosyal Bilimler Dergisi, 16*(1), 131–147.

Elias, S., & Loomis, R. (2002). Utilizing need for cognition and perceived self efficacy to predict academic performance. *Journal of Applied Social Psychology, 32*(8), 1687–1702.

Ergin, D. E. (1995). Olceklerde gecerlik ve guvenirlilik [Validity and reliability in scales]. *M.U. A.E.F. Egitim Bilimleri Dergisi, 7*.

Gettinger, M., & Seibert, J. (2002). Contributions of study skills to academic competence. *School Psychology Review, 31*(3), 350–365.

Gizir, C. A., Gizir, S., Aktaş, M., Gocer, S., Omur, S., Yuce, G., & Kirik, C. (2010). *Mersin Universitesi ogrenci profili* [Mersin University student profile]. Mersin: MEU PDR Merkezi, Yayin No: 1.

Greenwood, C., Horton, B., & Utley, C. (2002). Academic engagement: Current perspectives on research and practice. *School Psychology Review, 31*, 328–349.

Hair, J., Anderson, R.E., Tatham, R. L., & Black, W.C. (1995). Multivariate data analysis (4th ed.). New Jersey: Prentice Hall Inc.

Hatipoglu, A., Acar, E., Vural Akar, R., & Binay, B. (2012). *Ogrenci profili arastirmasi-2011* [Student profile research-2011]. Aydin: Adnan Menderes Universitesi Genclik Sorunlari Arastirma ve Uygulama Merkezi.

Higher Education Economy (2013). Retrieved from https://www.heacademy.ac.uk/sites/default/files/resources/engagement_for_enhancement_case_studies.pdf

HERI (2016a). Retrieved from https://heri.ucla.edu/

HERI (2016b). *CIRP Freshman survey 2016* [Survey Form]. Los Angeles: HERI & UCLA. Retrieved from https://www.heri.ucla.edu/PDFs/surveyAdmin/tfs/2016-CIRP-Freshman-Survey.pdf

HERI (2016c). *CIRP your first college year 2016* [Survey Form]. Los Angeles: HERI & UCLA. Retrieved from http://www.heri.ucla.edu/PDFs/surveyAdmin/YFCY/2016/YFCY2016.pdf

HERI (2016d). *CIRP diverse learning environments - Coresurvey 2016* [Survey Form]. Los Angeles: HERI & UCLA. Retrieved from http://www.heri.ucla.edu/PDFs/surveyAdmin/dle/2016/DLE-2016-Core-Instrument.pdf

Hoover, J., & Patton, P. (1995). *Teaching students with learning problems to use study skills: A teacher's guide*. Austin : TX: Pro-Ed.

Ilgaz, Y., & Akdol, B. (2009). İ. U. İktisat Fakultesi Isletme Bolumu ogrencilerine yonelik bir profil calismasi [A profile study on students in Istanbul University's Faculty of Economics]. *Istanbul Universitesi Iktisat Fakultesi Mecmuasi, 59*, 155–183.

Issi, A. D. (2008). *Gazi Universitesi Egitim Fakultesi Guzel Sanatlar Egitimi Bolumu Muzik Ogretmenligi Ana Bilim Dali ogrenci profili* [Gazi University student profiles for the Faculty of Education Fine Arts Department and Department of Music Education]. (Yuksek lisans tezi, Gazi Universitesi, Egitim Bilimleri Enstitusu, Ankara). Retrieved from

https://tez.yok.gov.tr/UlusalTezMerkezi/

Ivanov Devlet Universitesi. (n.d.). *Anket formu*. Retrieved from https://docs.google.com/forms/d/e/1FAIpQLSdpXQa7gA5lwVC9p6V5m-yQkQOFWqtNGIlNuP-MYtGZ4vQQyQ/viewform?formkey=dDdwTjcxUlQVVWUVFMbExncUlhX3c6MQ

Jenkins, L., & Demaray, M. (2015). An investigation of relations among academic enablers and reading outcomes. *Psychology in the Schools, 52*(4), 379–389. DOI: 10.1002/pits.21830.

Kaya, A., & Buyukkasap, E. (2005). Fizik Ogretmenligi Programi ogrencilerinin profilleri, ogretmenlik meslegine yonelik tutum ve endiseleri: Erzurum ornegi [Physics teaching program student profiles and attitudes and concerns toward the teaching profession: The case of Erzurum]. *Kastamonu Egitim Dergisi, 12*(2), 367–380.

Kizilcaoglu, A. (2003). Necatibey Egitim Fakultesi sosyal bilgiler ogretmeni adaylarinin profili [Necatibey Education Faculty's social sciences teacher candidate profile]. *Balikesir Universitesi Sosyal Bilimler Enstitusu Dergisi, 6*(10), 87–105.

Kuh, G. D. (2009). The national survey of student engagement: Conceptual and empirical foundations. *New Directions for Institutional Research* [Special Issue: Using NSSE in InstitutionalResearch], *141,* 5–20. doi: 10.1002/ir.283

Kumar, S., & Hsiao, J. (2007). Engineers learn "soft skills the hard way": Planting a seed of leadership in engineering class. *Leadership and Management in Engineering, 7*(1), 18–23.

Kustepeli, Y., & Gulcan, Y. (2002). Dokuz Eylul Universitesi Isletme Fakultesi ogrenci profili: Yansimalar ve oneriler [Dokuz Eylul University Faculty of Business student profile: Reflections and suggestions]. *Dokuz Eylul University Journal of Faculty of Business, 3*(2), 98–114.

Laidra, K., Pullman, H., & Allik, J. (2006). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elemantary to secondary school. *Personality and Individual Differences, 42,* 441–451. doi: 10.1016/j.paid.2006.08.001

Laker, D., & Powell, J. (2011). The differences between hard and soft skills and their relative impact on training transfer. *Human Resource Development Quarterly, 22*(1), 111–122.

Linnenbrink, E., & Pintrich, P. (2002). Motivation as enablers for academic success. *School Psychology Review, 31*(3), 313–327.

Malecki, C., & Elliott, S. (2002). Children's social behaviors as predictors of academic achievement: A longitudinal analysis. *School Psychology Quarterly, 17*(1), 1–23.

Marchello, J. M. (1987). Universtiy and industrial development. *Industry and Higher Education,* 10–14.

Massachusetts Institute of Technology (2014). *MIT survey of new student 2014* [Survey Form]. Boston, MA: MIT. Retrieved from http://web.mit.edu/ir/surveys/pdf/New_MIT_2014-instrument.pdf

Moskova Devlet Universitesi. (n.d.). *Anket formu* [Survey form] (L. Allahverdiyeva,

Trans.). Retrieved from http://webanketa.com/forms/6cs3adhh5wt6ashs60r32d8/ru/

Mulder, M., Gulikers, J., Wesselink, R., & Biemans, H. (2009). The new competence concept in higher education: Error or enrichment? *Journal of European Industrial Training, 33*(8/9), 755–770.

Otrar, M., & Argin, F. S. (2015). Ogrencilerin sosyal medyaya iliskin tutumlarini belirlemeye yonelik bir olcek gelistirme calismasi [A scale development study on determining students' attitudes on the relationship to social media]. *Egitim ve Ogretim Arastirmalari Dergisi, 4*(1), 391–403.

Ozel, A. (2006). Dumlupinar Universitesi Egitim Fakultesinde ogrenim goren ogrencilerin sosyo-ekonomik ve kulturel yapisi [Socio-economic and cultural structure of students in Dumlupinar University's Faculty of Education]. *Dumlupinar Universitesi Sosyal Bilimler Dergisi, 16*, 317–336.

Pomor Devlet Universitesi. (n.d.). *Anket formu* [Survey form] (L. Allahverdiyeva, Trans.). Retrieved from http://narfu.ru/pomorsu.ru/www.pomorsu.ru/_doc/quality/education/questionnaire_09.pdf

Pryor, I. H., Hurtado, S., Saenz, V. B., Santos, J. L., & Korn, W. S. (2007). *The American freshman: Forty year trends*. Los Angeles: Higher Education Research Enstitute & UCLA. Retrieved from http://www.heri.ucla.edu/PDFs/pubs/TFS/Trends/Monographs/TheAmericanFreshman40YearTrends.pdf

Rohde, T., & Thompson, L. (2007). Predicting academic achievement with cognitive ability. *Intelligence, 35*(1), 83–92.

Schulz, B. (2008). The importance of soft skills: Education beyond academic knowledge. *Journal of Language and Communicatio*n, 146–154.

Schunk, D., & Zimmerman, B. (1994). *Self-regulation of learning and performance: Issues and educational applications*. Hillside, N.J.: Erlbaum.

Sencar, O. (2013). *Universite gencligi arastirmasi: Turkiye'deki universite gencliginin profili* [University youth research: Youth profile at Turkey's universities]. Ankara: MetroPOLL Stratejik ve Sosyal Arastirmalar Merkezi.

Sevuktekin, M., Nargelecekenler, M., & Cetin, L. (2012). Uludag Universitesi ogrencilerinin sosyo-ekonomik profil arastirmasi [Socio-economic profile research of Uludag University students]. *Uludag Universitesi Iktisadi ve Idari Bilimler Fakultesi Dergisi, 31*(2), 99–128.

Sohal, A. S. (2013). Developing competencies of supply chain professionals in Australia: Collaboration between businesses, universities and industry associations. *Supply Chain Management: An International Journal, 18*(4), 429–439.

Stanovich, K., Cunningham, A., & Freeman, D. (1984). Intelligence, cognitive skills,

and early reading progress. *Reading Research Quarterly, 19*(3), 278–303.

Sahin, O. (2005). Yeni bir universite ve yeni ogrenciler: Mugla Universitesi ogrenci profili [A new university and new students: Mugla University student profile]. *Egitim Toplum Bilim, 3*(10), 18–39.

Senol, H., & Tufekci, O. K. (2007). Muhasebe programinda egitim goren ogrencilerin profili ve beklentileri: SDU ornegi [Accounting program students' profile and expectations]. *Suleyman Demirel Universitesi Sosyal Bilimler Enstitusu Dergisi, 3*(5), 151–163.

Senses, F. (1999). Yuksekogretimde ogrenciler: ODTU Iktisat Bolumu ogrenci profili [Students in higher education: ODTU Economics Department student profile]. *ODTU Gelisme Dergisi, 26*(1-2), 179–201.

Tekin, M. E. (2014). Selcuk Universitesi Veteriner Fakultesi ogrenci profili [Selcuk University Veterinary Faculty student profile]. *Eurasian Journal of Veterinary Science, 30*(4), 181–187. DOI:10.15312/EurasianJVetSci.201447374.

University College London. (n.d.) National student survey (NSS). Retrieved from https://www.ucl.ac.uk/srs/governance-and-committees/qme/qmea-z/nss

Vrugt, A., Langereis, M., & Hoogstraten, J. (1997). Academic self efficacy and malleability of relevant capabilities as predictors of exam performance. *The Journal of Experimental Education, 66*, 61–72.

Wentzel, K., & Watkins, D. (2002). Peer relationships and collaborative learning as contexts for academic enablers. *School Psychology Review, 31*, 366–377.

Willingham, W., Pollack, J., & Lewis, C. (2002). Grade and test scores: Accounting for observed differences. *Journal of Educational Measurement, 39*, 1–37.

Wood, R., & Locke, E. (1987). The relationship of self efficacy and grade goalsto academic performance. *Educational and Psychological Measurements, 47*, 1013–1024.

Yaylali, M., Oktay, E., Ozen, U., Akan, Y., Ozer, H., Kiziltan, A., … & Akturk, E. (2006). *Universite gencliginin sosyo-ekonomik profili arastirmasi* [Research on university youths' socio-economic profile]. Erzurum: Ataturk Universitesi Iktisadi ve Idari Bilimler Fakultesi Yayinlari.

Yigit, R., Esenay, F. I., & Derebent, E. (2007). Turkiye'de hemsirelik son sinif ogrencilerinin profili [Final class nursing student profiles in Turkey]. *Cumhuriyet Universitesi Hemsirelik Yuksek Okulu Dergisi, 11* (3), 1–12.

Zimmerman, B. (1998). Academic studying and the development of personal skill: A self-regulatory perspective. *Educational Psychologist, 33*, 73–86.

Akademik Yetkinlik Araçları Değerlendirme Ölçeği: Geçerlik ve Güvenirlik Çalışması

Atıf:

Akbiyik, M., & Senturk, M. (2019). Assessment scale of academic enablers: A validity and reliability study. *Eurasian Journal of Educational Research, 80,* 225-250, DOI: 10.14689/ejer.2019.80.11

Özet

Problem Durumu: Ülkelerin eğitim sistemlerinin kalitelerinin bir göstergesi ve eğitim politikalarında değişim ihtiyacının belirleyicileri olmaları sebebiyle öğrencilerin akademik başarılarını ve akademik performanslarını etkileyen faktörler incelenmelidir. Akademik yetkinlik (*academic competencies*) hem bir öğrencinin performansını hem de bu performansı değerlendirmek üzere kullanılan standartları işaret ederken; akademik yetkinlik akademik başarı için gerekli olan ve öğrencinin beceri, tutum ve davranışlarını içeren çok boyutlu bir yapıdır. Akademik yetkinlik, "akademik bilgi ve beceri" (yani temel akademik ve(ya) uygulamalı bilişsel bilgi ve beceriler) ile "akademik yetkinlik araçları" (yani bu bilgi ve beceriyi edinmesine katkı sağlayacak tutum ve davranışlar) şeklinde iki temel bileşenden oluşmaktadır. Yetkinlik temelli sistem ilk defa 1970'lerde ABD'de iş performansını ölçmeye yönelik uygulamaların mevcut ekonomik çerçevede başarısız olması sebebiyle ortaya çıkarken; 1990'larda, önce Lisbon ardından Bologna süreçleri ile Avrupa'da yükseköğretimde yeterlilik konusunda önemli adımlar atılmıştır. 2008'de ise Avrupa Yaşam Boyu Öğrenme Yeterlilikler Çerçevesi (*European Qualifications Framework for Lifelong Learning_EQF*) oluşturulmuştur. Türkiye'de dahil olmak üzere 39 ülke, kendi ulusal yeterlilik çerçevelerini EQF'i temel alarak, farklı eğitim kademelerine göre belirlemiştir. Türkiye Yükseköğretim Yeterlilikler Çerçevesi (TYYÇ) ile öğrencilerin herhangi bir programdan mezun olana kadar kazanmaları gereken bilgi (*knowledge*), beceri (*skills*) ve yetkinlikler (*responsibility & autonomy*) oluşturulur. EQF'te sorumluluk ve özerklik olarak adlandırılan bu başlık Türkiye'de yetkinlik şeklinde kullanılmaktadır. Ölçek geliştirme sürecinde, TYYÇ'de yer alan bilgi ve beceri alt başlıkları yerine sadece "yetkinlikler" alt başlığı ve altında yer alan ifadeler/maddeler dahil edilmiştir. Bunun sebebi hem bilgi ve becerilerin alan ve program temelli olması ve genellenebilirliğinin düşük olması hem de yetkinlikler altındaki tutum ve davranışların üniversite ve istihdam ilişkisini güçlendirmesi, bu yetkinliklerin özellikle 21. yüzyıl becerileri ile temelden ilişkili olmasıdır. Ölçeği isimlendirirken akademik yetkinlikler yerine "akademik yetkinlik araçları" ifadesini kullanmayı tercih edilmesindeki sebep ise ölçekte yer alan maddelerin uluslararası literatürdeki yetkinlik (competencies) kavramının akademik bilgi ve beceri ile birlikte tamamlayıcısı olan akademik yetkinlik araçları (enablers) kavramına karşılık gelmesidir.

Araştırmanın Amacı: Bu araştırmanın amacı, üniversite öğrencilerinin yükseköğretimde kazandırılması ya da geliştirilmesi hedeflenen akademik yetkinlik araçlarını ne düzeyde kazandıkları ya da geliştirdiklerine yönelik algılarını ölçen bir ölçek geliştirmektir.

Araştırmanın Yöntemi: Araştırmanın evrenini 2015-2016 yılında İstanbul Üniversitesi'nde dört yıllık fakültelerde (tıp fakülteleri de dahil edilmiştir) öğrenim gören Türkiye Cumhuriyeti vatandaşı lisans öğrencileri oluşturmaktadır. %99 güven seviyesi, %2 hata payı dikkate alınarak yapılan hesaplamada asgari örneklem sayısı 3914 olarak hesaplanmıştır. Uygulama sonucunda, en yaşlı öğrencinin 1960 ve en genç öğrencinin 2000 doğumlu olduğu ve ortalama yaşın 22 bulunduğu 5208 öğrenciden oluşmaktadır. Öğrencilerin 2742'sinin (%52,65) kadın, 2405'inin (%46,18) erkek olduğu görülmektedir. Türkiye Yükseköğretim Kurumunun EQF temelinde geliştirdiği ulusal yeterlilikler çerçevesinde belirlediği ve öğrencilerin lisans eğitimleri sürecinde geliştirmesini beklediği yetkinlikleri ölçmede kullanılabilecek geçerli ve güvenilir bir veri toplama aracının geliştirilmesi amacıyla hazırlanmış olan Akademik Yetkinlik Araçları Değerlendirme Ölçeği (AYADÖ) 20 maddeden oluşmaktadır. "1=Hiç" ve "5= Her Fırsatta" şeklinde puanlanmaktadır. Ters puanlanan bir madde yer almamaktadır. Ölçekten toplam bir skor elde edilmekte, bu skor öğrencinin üniversitesinden akademik yetkinlik araçlarını ne düzeyde kazandığını/edindiğini düşündüğünü göstermektedir.

Araştırmanın Bulguları: Yapılan açımlayıcı faktör analizi sonucunda ölçekte yer alan maddeler; Öğrenme Yetkinliği, İletişim ve Sosyal Yetkinlik ve Ödev ve Sorumluluk Yetkinliği başlıkları altında üç faktöre yüklenmiştir. Bu üç faktör toplam varyansın %51'ni açıklamaktadır. Varimax rotasyon sonucunda maddelerin faktör yükleri 0.44 - 0.82 arasında değişmektedir. Ölçeğin bütünü için Cronbach Alpha değeri 0.90'dır. Alt başlıkların Cronbach Alpha değerinin 0.80 üzerinde olması hem ölçeğin bütün olarak hem de alt boyutların kendi içinde tutarlı olduğunu göstermektedir. Faktör bazında ayırt edicilik analizinde tüm gruplar için farklılıkların istatistiksel olduğu ($p<.001$) görülmüştür. Madde toplam korelasyon katsayıları 0.44-0.70 arasında; madde kalan korelasyon katsayıları 0.43-0.65 arasındadır. Faktörler arası ilişkileri belirlemek üzere yapılan korelasyon analizi sonucunda tüm faktörler kendi arasında ve tüm faktörlerle toplam puan arasında pozitif yönde $p<.001$ düzeyinde anlamlı bir ilişki olduğu bulunmuştur. Bu sonuç da ölçekteki tüm faktörlerin aynı yapı içinde olduklarını kanıtlamaktadır. Ölçek ve alt ölçek puanlarının cinsiyet ve fakülte bazında karşılaştırmalı analizleri de yapılmıştır. AYADÖ puanları cinsiyet değişkenine göre anlamlı bir fark göstermiştir (t=7,49; $p<$,001). Söz konusu farklılık kadınların lehinedir ($\bar{X}_{kadın}$=62,74; $\bar{X}_{erkek}$=59,43). Öğrenme yetkinliği alt ölçeği (t=6,58; $p<$,001) ile ödev ve sorumluluk yetkinliği alt ölçeği (t=13,44; $p<$,001) puanlarının da cinsiyete göre anlamlı şekilde farklılaştığı görülmüştür. Söz konusu farklılık öğrenme yetkinliği alt ölçeği ($\bar{X}_{kadın}$=28,13; $\bar{X}_{erkek}$=26,86) için de, ödev ve sorumluluk yetkinliği alt ölçeği ($\bar{X}_{kadın}$=13,53; ($\bar{X}_{erkek}$=11,87) için de kadınların lehinedir. Ancak iletişim ve sosyal yetkinlik alt ölçeği puanlarının cinsiyete göre istatiksel olarak anlamlı bir farklılık göstermemiştir (t=1,62; $p>$,05). Fakültelere göre Akademik Yetkinlik Araçları

Değerlendirme Ölçeği puanları dikkate alındığında fakülte farkı gözetmeksizin örneklemdeki öğrencilerin akademik yeterlilikler ölçeğine ilişkin aritmetik ortalamaları ($\bar{X}$) 61,17 (*ss*=15,92) olarak elde edilmiştir. AYADÖ puanları fakülte değişkenine göre anlamlı fark göstermiştir (F=42,82; *p*<,001); en yüksek ortalama Hasan Ali Yücel Eğitim Fakültesi, en düşük ortalamaya Hukuk Fakültesi öğrencilerinindir. Fakültelere göre de alt ölçekler yine ayrı ayrı değerlendirilmiştir. Öğrenme yetkinliği (F=16,68; *p*<,001), iletişim ve sosyal yetkinlikler (F=43,58; *p*<,001) ve ödev ve sorumluluklar yetkinliği (F=66,61; *p*<,001) alt ölçekleri puanları için fakültelerin aritmetik ortalamaları arasındaki fark istatistiksel olarak anlamlıdır.

Araştırmanın Sonuç ve Önerileri: Ölçeğe ilişkin verilen istatistiksel skorlar ile cinsiyet ve özellikle fakültelere yönelik karşılaştırmalı analizler, ölçeğin TYYÇ'ne bağlı olarak lisans düzeyinde eğitim veren yükseköğretim kurumları tarafından, bu kurumların öğrencilere kazandırmakla (ya da öğrencilerde hali hazırda var olan bu beceri, tutum ve davranışlarını geliştirmekle) sorumlu oldukları temel akademik yetkinlik araçlarını öğrenciye ne düzeyde kazandırdıklarını anlamak, yine bu kurumların akademik başarıyı destekleyen araçları kazandırma sürecindeki zayıf ve güçlü yanlarını belirlemek ve akademi-istihdam ilişkisindeki halkaları işlevsel olarak oluşturmak/yeniden yapılandırmak amacıyla kullanılabilecek geçerli ve güvenilir bir araç olduğunu göstermektedir.

Anahtar Kelimeler: Akademik yetkinlik, akademik yeterlilik, yükseköğretimde yeterlilikler, üniversite, kalite.

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 3**00 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 İngilizcedir.
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 ..... : 120 TL*

Institution */ Kurum Aboneliği ..... : 280 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