JESEH

VOLUME
8
ISSUE
1

YEAR
2022

Journal of Education in Science, Environment and Health

e-ISSN 2149-214X

Journal of Education in Science, Environment and Health
Volume 8, Issue 1, 2022 e-ISSN:2149-214X

EDITORIAL BOARD

Journal of Education in Science, Environment and Health (JESEH)

The Journal of Education in Science, Environment and Health (JESEH) is a peer-reviewed and online free journal. The JESEH is published quarterly in January, April, July and October.The language of the journal is English only. As an open access journal, Journal of Education in Science, Environment and Health (JESEH) does not charge article submission or processing fees. JESEH is a non-profit journal and publication is completely free of charge.

The JESEH welcomes any research papers on education in science, environment and health using techniques from and applications in any technical knowledge domain: original theoretical works, literature reviews, research reports, social issues, psychological issues, curricula, learning environments, book reviews, and review articles. The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission to the JESEH.

Abstracting/ Indexing

Journal of Education in Science, Environment and Health (JESEH) is indexed by following abstracting and indexing services: SOBIAD, Scientific Indexing Service (SIS), Education Resources Information Center (ERIC).

Submissions

All submissions should be in electronic (.Doc or .Docx) format. Submissions in PDF and other non-editable formats are not acceptable. Manuscripts can be submitted through the journal website. All manuscripts should use the latest APA style. The manuscript template for formatting is available on the journal website.

Contact Info

Journal of Education in Science, Environment and Health (JESEH)
Email: jesehoffice@gmail.com
Web : www.jeseh.net

Journal of Education in Science, Environment and Health
Volume 8, Issue 1, 2022 e-ISSN:2149-214X

CONTENTS

Journal of Education in Science, Environment and Health
Volume 8, Issue 1, 2022 https://doi.org/10.21891/jeseh.1055590

Extending Technology Acceptance Model with Scientific Epistemological and Science Teaching Efficacy Beliefs: A Study with Preservice Teachers

Tezcan Kartal, Ibrahim Serdar Kiziltepe, Busra Kartal

Article Info

Article History

Published:
01 January 2022

Received:
07 March 2021

Accepted:
18 October 2021

Keywords

Technology acceptance model
Science teaching efficacy beliefs
Epistemological beliefs
Preservice teachers

Abstract

The technology acceptance model (TAM) is a widely used framework to investigate factors influencing technology use in education. TAM refers to a person's technology-related attitudes and beliefs influencing intention to use and actual use of technology and seeks predictors of behaviors whether to accept or reject using technology. There are various external variables extended to TAM to increase the predictivity of the model and the generalizability of findings. However, what is not yet clear is the impact of teacher-related variables such as teaching efficacy and epistemological beliefs on teachers' technology acceptance and behavioral intention. This study examined 710 preservice teachers' technology acceptance using an extended-TAM with scientific epistemological and science teaching efficacy beliefs. Data were collected through a self-reported measurement tool. Structural equation modeling was used to analyze data. Results revealed that the research model explained 59% of the variance in behavioral intention, and perceived usefulness is the most prominent determinant of behavioral intention. The subdimension of scientific epistemological beliefs, justification, is the strongest determinant in influencing TAM constructs among the external variables (epistemological and science teaching efficacy beliefs). Science teaching efficacy beliefs had small effects on technology acceptance constructs. Recommendations were made based on the findings.

Introduction

Technology has fast become a key instrument in teaching and learning as it has the potential of improving knowledge acquisition and transfer (Eksail & Afari, 2020). Incorporating new technologies in teaching and learning is a continuing concern within educational research (Granić & Marangunić, 2019; Teo et al., 2015). Teachers are the agents of effective technology integration (Siyam, 2019; Teo, 2009; Wong et al., 2012). Therefore, many researchers are interested in factors influencing teachers' technology use (Akar, 2019; Scherer, Siddiq, & Tondeur, 2019). Perceptions of technology integration, beliefs regarding teaching and learning, and efficacy beliefs are examples of the factors influencing teachers' technology use in education (Siyam, 2019). Teachers' decision to use technology in their teaching is closely related to their technology acceptance, which refers to the teachers' willingness to use technologies to accomplish their teaching-related tasks (Akar, 2019; Avcı Yucel & Gulbahar, 2013).

Among the frameworks investigating users' technology acceptance, the most commonly used is Technology Acceptance Model (TAM). The factors determining the success or failure of technology integration have been studied by many researchers using TAM (Avcı Yucel & Gulbahar, 2013; Scherer et al., 2019). TAM refers to a person's technology-related attitudes and beliefs influencing intention to use and actual use of technology (Davis, 1985) and seeks predictors of behaviors whether to accept or reject using technology (Granić & Marangunić, 2019). Among the studies regarding TAM, education studies exist extensively (Avcı Yucel & Gulbahar, 2013; Granić & Marangunić, 2019).

A large body of TAM studies has investigated preservice teachers' technology acceptance (Bardakcı & Alkan, 2019; Teo et al., 2015; Wong et al., 2012). These studies focused on different technologies such as mobile applications (Al-Azawei, & Alowayr, 2020; Bano et al., 2018), interactive whiteboards (Bardakcı & Alkan, 2019). However, various samples and contexts might lead to diverse findings regarding the relations among the constructs (Scherer & Teo, 2019). It is essential to consider several external variables to understand better the factors influencing technology acceptance (Avcı Yucel & Gulbahar, 2013). There are various external variables extended to TAM to increase the predictivity of the model and generalizability of findings, such as TPACK (Bardakcı & Alkan, 2019), individual innovativeness (Akar, 2019), and teacher efficacy (Joo et al., 2018). It is suggested to replicate TAM studies with different modeling approaches and larger samples (Scherer et al., 2019). What is not yet clear is the impact of teacher-related variables such as teaching efficacy and epistemological beliefs on teachers’ technology acceptance and behavioral intention.

Turkey is among the countries that conducted large national projects to incorporate digital technologies into teaching and learning processes (Bardakcı & Alkan, 2019), and the project of Movement of Enhancing Opportunities and Improving Technology, called FATIH was developed in 2010. FATIH project aims to ensure equality in terms of technological resources among students all around the country. To achieve the project goals, most schools and classrooms have been re-designed to increase the availability of technological resources through interactive whiteboards, internet access, tablets, and specific portals. Keeping up with the developing technologies is considered a competence for the teaching profession in Turkey (Akar, 2019). Therefore, teachers’ technology acceptance is crucial within the context of the Turkish teacher education programs to avoid wasting these investments. Considering the importance of adding various external variables to TAM, we investigated the impact of science teaching efficacy beliefs and scientific epistemological beliefs on 710 preservice teachers’ technology acceptance. Therefore, this study is supposed to make a major contribution to research on TAM by demonstrating the effect of external variables that were not examined previously in TAM studies.

Theoretical Framework

Technology Acceptance Model

Since the inclusion of technology in business and education, the reasons for accepting or rejecting technology have sparked the attraction of researchers growingly (Granić & Marangunić, 2019). TAM is the most commonly used model to explain teachers’ intention to use technology in education by examining users’ beliefs and attitudes because of its simplicity and understandability (Eksail & Afari, 2020; Scherer et al., 2019; Siyam, 2019). TAM adopts the idea that individuals tend to use new technology if they believe it would improve their performance and be free of effort (Akar, 2019). Recent review studies concluded that TAM is a relevant model in examining factors influencing technology use (Granić & Marangunić, 2019; Scherer et al., 2019).

Davis (1989) adapted TAM from the Theory of Reasoned Action developed by Fishbein and Ajzen (1975) to investigate the determinants influencing behavioral intention that leads to actual usage. TAM deals with the relationship between attitude, intention, and behavior. The main factors determining the level of acceptance of technology are perceived ease of use (PEU) and perceived usefulness (PU) (Granić & Marangunić, 2019; Wong et al., 2012; Scherer et al., 2019). TAM posits that PEU and PU significantly influence attitude toward using (ATU) and, in turn, behavioral intention (BI). The relations between PEU, PU, and ATU and the predictive role of PU and ATU on BI are the particular concerns of the original TAM (Siyam, 2019). Figure 1 presents the constructs and relationships among these constructs in TAM.

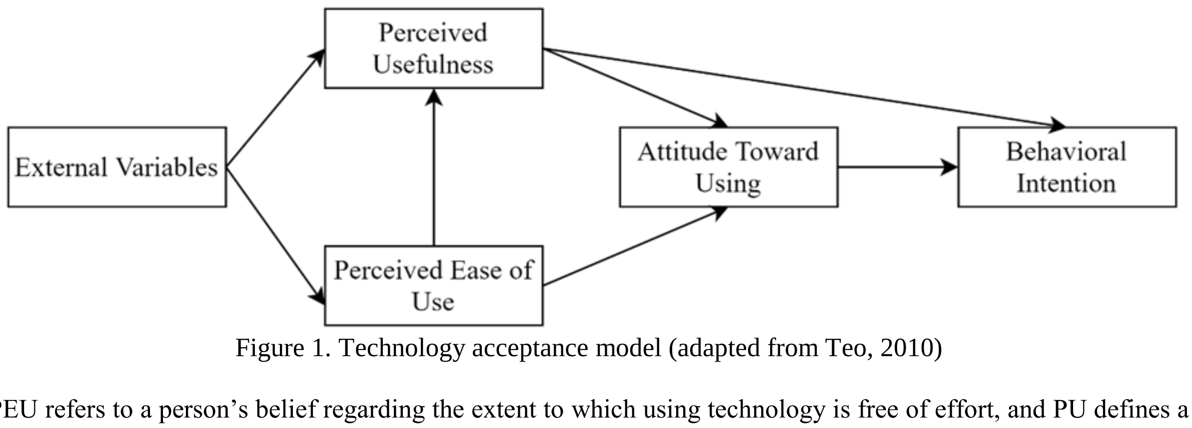

Figure 1. Technology acceptance model (adapted from Teo, 2010)

PEU refers to a person’s belief regarding the extent to which using technology is free of effort, and PU defines a person’s belief regarding the degree to which using technology would improve the person’s performance (Avcı Yucel & Gulbahar, 2013). PEU also proposes that regarding technology as easy to use and believing in their ability to manage technology make people use technology (Teo et al., 2015). PEU significantly impacts PU, and both have influences on ATU. In other words, a person with positive attitudes toward using technology also perceives that using technology is effortless and improves her/his performance. PU was found to be the most significant determinant in a review study by Avcı Yucel and Gulbahar (2013). Similarly, most research concluded that PU had the strongest effect on BI (Akar, 2019; Granić & Marangunić, 2019). Besides, Bardakcı and Alkan (2019) proposed that believing the impact of interactive whiteboards on teaching performance promoted preservice teachers' intentions to use, consistent with given findings.

Attitude toward using technology consists of feelings about technology use (Eksail & Afari, 2020; Kartal, 2019) and determines how teachers respond to technologies and the extent to which technology integration would be successful (Teo et al., 2015). TAM adapts the idea that attitude is the major determinant in accepting or rejecting technology (Davis, 1989) and is influenced by PEU and PU. BI influenced by PU and ATU is closely related to teachers' intrinsic motivations to use technology (Anderson, Groulx, & Maninger, 2011; Kartal, 2019) and the actual use of technology (Teo, 2010; Teo et al., 2015). Researchers attempted to achieve higher percentages of explained variance in BI as the behavioral intention is a key instrument to predict actual technology usage behavior in preservice teachers' future classrooms.

Science Teaching Efficacy Beliefs

Bandura (1977) regarded self-efficacy as a belief in a person's capability for performing a specific task. According to Bandura (1977), if a person feels confident in performing a specific task and believes in the favorable result, he/she feels efficacious in performing the given task. Teaching efficacy is a teacher's belief regarding his/her capacity to promote student learning (Gagnier, Holochwost, & Fisher, 2021). Gibson and Dembo (1984) mentioned two distinct dimensions of teaching efficacy based on Bandura's social cognitive theory: self-efficacy beliefs and outcome expectations. The former is teacher confidence in own teaching abilities. The latter assumes that effective teaching influences student learning.

Teachers with a high level of science teaching efficacy are supposed to be open and willing to innovations, new teaching methods, and new ideas such as using instructional technologies (Blonder et al., 2013; Gagnier et al., 2021; Kartal & Dilek, 2021; Woolfolk-Hoy & Spero, 2005). Teachers' efficacy levels determine the effort they put on and the time they spent to achieve their teaching-related goals (Tschannen-Moran & Woolfolk-Hoy, 2001). Teaching efficacy beliefs impact teachers' in-class behaviors and teaching methods. Teachers with high teaching efficacy might be resistant to deal with challenges in the classroom and insist on promoting all students' learning (Ekici, 2016; Kartal, 2020). To achieve these teaching objectives, teachers' use of technology has undoubtedly great benefits. Therefore, it is supposed that teachers with high science teaching efficacy beliefs are supposed to tend to use technology in their teaching.

The implementation and success of reforms promoting technology use are influenced by teacher beliefs (Gagnier et al., 2021; Kartal & Çınar, 2018). Blonder and colleagues (2013) reported that opportunities to develop teaching efficacy promoted teachers' tendency to use new technologies. Teachers improving their science education with technologies also promote students' learning motivation (Al-Azawei & Alowayr, 2020; Huang et al., 2020), collaborative learning (Kartal & Dilek, 2021), and cognitive gains (Becker et al., 2020). Some teachers might consider technology an essential component in effective science teaching. Kartal and Dilek (2021) found that a technology-supported teaching method course and microteaching promoted preservice teachers' science teaching efficacy beliefs. Similarly, preservice teachers reported that using technology promoted students' learning, and technology was a crucial constituent of effective science teaching (Min et al., 2020).

Scientific Epistemological Beliefs

Epistemological beliefs are beliefs about knowledge and knowing (Conley et al., 2004; Hofer & Pintrich, 1997). Beliefs about knowledge consist of the source of knowledge and justification for knowing. On the other hand, beliefs about knowing include certainty and simplicity of knowledge (Lee et al., 2021). Strong epistemological beliefs might be considered sophisticated as well as weak beliefs might be considered naïve. Someone who perceives himself/herself as being able to think and act like a scientist has sophisticated beliefs, and someone who believes scientists constructed nearly all almost of actual knowledge has naïve beliefs (Demirbag & Bahcivan, 2021; Kızıltepe & Kartal, 2021).

Epistemological beliefs are related to teaching and learning beliefs (Bahcivan, 2014; Cheng et al., 2009; Deng et al., 2014; Kızıltepe & Kartal, 2021). Sophisticated beliefs that assume knowledge evolves in nature with constructions by self or anyone are more likely to lead teachers to teach in a constructivist way (Deng et al., 2014). Teachers' epistemological beliefs have a crucial role in teaching effectiveness (Bondy et al., 2007) and teaching-related behaviors (Schommer-Aikins, 2004). Besides, research also showed the interrelatedness of epistemological beliefs with digital literacy (Demirbag & Bahcivan, 2021; Güneş & Bahçivan, 2018), attitude toward computer use (Teo, 2008), and type of technology use such as traditional or constructivist (Deng et al., 2014). Nevertheless, it is still underresearched the effect of epistemological beliefs on technology acceptance.

Literature Review

This section highlights the preservice teachers' technology acceptance studies, mainly focusing on their external variables. Avcı Yucel and Gulbahar (2013) reviewed TAM studies based on variables used in the study, working areas, measurement items, and results. They found that PU was the most effective variable, followed by PEU. A vast majority of research investigated the structural relationships between TAM constructs and technological complexity, social norms, computer self-efficacy, and facilitation conditions (Baydas & Goktas, 2017; Huang & Teo, 2019; Aypay et al., 2012; Kabakçı-Yurdakul et al., 2014; Lee et al., 2010; Teo, 2009). Granić and Marangunić (2019) addressed the gap of incorporating new external variables into the TAM and studying with larger samples in their review study.

Different from the mentioned studies, Wong and colleagues (2012) explored the effect of computer teaching efficacy and gender on student teachers' technology acceptance and concluded that computer teaching efficacy was the strongest determinant of ATU. Joo and colleagues (2018) added teacher efficacy as an external variable to TAM. Siyam (2019) extended TAM by adding job relevance, time, self-efficacy, and access to technology as external variables to investigate special education teachers' technology acceptance. Individual innovativeness and the social norm were the external variables in the study of Akar (2019), regarding primary and secondary teachers' technology acceptance. Bardakcı and Alkan (2019) investigated student teachers' intentions to use interactive whiteboards, investigating the effect of traditional and constructive teaching beliefs, individual innovativeness, pedagogical, technological, and technological pedagogical knowledge, interactive whiteboard self-efficacy, and effort and performance expectancy. The result revealed that performance expectancy is the one variable that significantly influences respondents' intention to use interactive whiteboards.

Teacher beliefs and attitudes play a major role in determining the extent to which technology would be used in education, with a more significant influence on technology use than first-order barriers such as access and availability (Kartal, 2019; 2020; Siyam, 2019). Teachers' pedagogical beliefs are amongst the most frequently investigated teacher beliefs in TAM (Gurer & Akkaya, 2021; Gyamfi, 2016; Huang & Teo, 2021; Li et al., 2019; Teo & Zhou, 2017; Teo et al., 2008). To our knowledge, the external variables in this study, scientific epistemological beliefs and science teaching efficacy beliefs, are not investigated in a TAM study.

Research Model and Hypotheses

Understanding student teachers' technology acceptance is crucial as their level of acceptance would provide insight into the effective and efficient use of technology in future classrooms (Wong et al., 2012). It is essential to test various research models to increase the predictive validity of TAM in educational settings (Parkman, Litz, & Gromik, 2018; Siyam, 2019). Examining the effect of external variables is crucial since the patterns in the impacts of teacher-related factors might promote the design of teacher preparation programs (Siyam, 2019). Most of the research highlighted the structural relationships among the external variables and PEU and PU (Gurer & Akkaya, 2021; Gyamfi, 2016; Siyam, 2019; Wong et al., 2012), but the direct effects of external variables on BI are still underresearched.

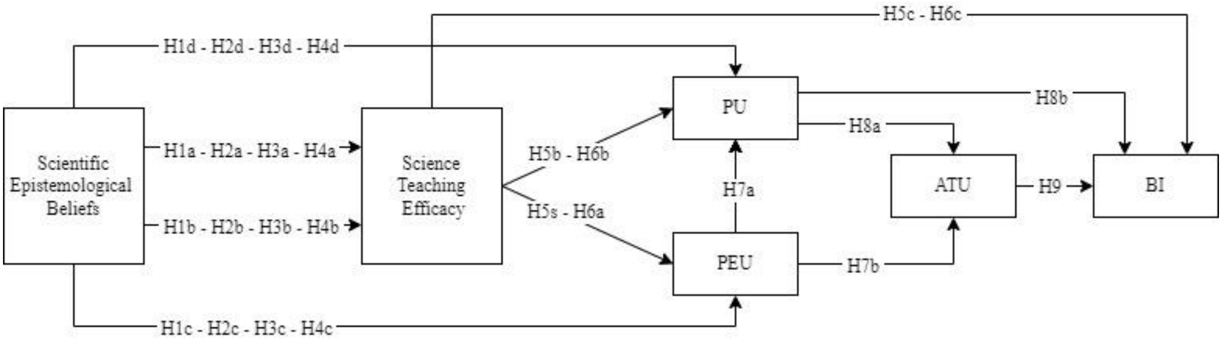

Figure 2. The research model

To reveal the chain of the influence of external variables on BI, researchers need to incorporate various external variables into TAM (Joo et al., 2018; Huang & Teo, 2021; Sang, Valcke, Van Braak, & Tondeur, 2010). We incorporated two interrelated constructs to reveal preservice teachers' technology acceptance: scientific epistemological beliefs and science teaching efficacy beliefs. Therefore, this study makes a major contribution to research on TAM by demonstrating the effect of scientific epistemological and science teaching efficacy beliefs on preservice teachers' technology acceptance. Figure 2 represents the hypothesized research model.

The hypotheses generated based on the literature review are as follows:

H1a: Source will significantly influence personal science teaching efficacy.
H1b: Source will significantly influence science teaching outcome expectancy.
H1c: Source will significantly influence PEU.
H1d: Source will significantly influence PU.
H2a: Certainty will significantly influence personal science teaching efficacy.
H2b: Certainty will significantly influence science teaching outcome expectancy.
H2c: Certainty will significantly influence PEU.
H2d: Certainty will significantly influence PU.
H3a: Justification will significantly influence personal science teaching efficacy.
H3b: Justification will significantly influence science teaching outcome expectancy.
H3c: Justification will significantly influence PEU.
H3d: Justification will significantly influence PU.
H4a: Development will significantly influence personal science teaching efficacy.
H4b: Development will significantly influence science teaching outcome expectancy.
H4c: Development will significantly influence PEU.
H4d: Development will significantly influence PU.
H5a: Personal science teaching efficacy beliefs will significantly influence PEU.
H5b: Personal science teaching efficacy beliefs will significantly influence PU.
H5c: Personal science teaching efficacy beliefs will significantly influence BI.
H6a: Science teaching outcome expectancy beliefs will significantly influence PEU.
H6b: Science teaching outcome expectancy beliefs will significantly influence PU.
H6c: Science teaching outcome expectancy beliefs will significantly influence BI.
H7a: PEU will significantly influence PU.
H7b: PEU will significantly influence ATU.
H8a: PU will significantly influence ATU.
H8b: PU will significantly influence BI.
H9: ATU will significantly influence BI.

Method

Research Design

This study uses a structural equation modeling (SEM) to reveal the structural relationships between preservice teachers' scientific epistemological beliefs, science teaching efficacy beliefs, and technology acceptance. SEM allows researchers to see the extent to which data is consistent with the hypothesized model, employing a simultaneous analysis of the entire system of the variables (Byrne, 2016). This study tested an extended-TAM, including preservice teachers' science teaching efficacy beliefs and epistemological beliefs as external variables, through model fit indices, path analysis, and hypothesis testing.

Participants

The participants were 710 preservice teachers whose major programs were science, elementary, and early childhood teaching. One of the external variables was science teaching efficacy beliefs; therefore, we invited preservice teachers trained for science teaching in their teacher preparation programs and supposed to teach science when they begin teaching. The respondents were informed about the purpose of the study and their rights to withdraw from the administration whenever they wished. The instrument was administered to 738 preservice teachers, and 710 of them were entirely completed. The return rate was 96%, demonstrating a higher rate than the suggested rate range of 70%-80% by Cresswell (2015) to make valid interpretations. Table 1 demostrates that just over half of the sample (56.2%) was fourth-grade, and over three-quarters of the sample

(83.8%) was female. In terms of technical proficiency and opportunity, more than half of the respondents reported that they used technology daily, and almost 90% of them perceived their competency either at the intermediate level (56.3%) or at the advanced level (31.4%).

Table 1. Demographic information of participants

Variable	Number	%
Grade level		
3.Grade	311	43.8
4.Grade	399	56.2
Gender		
Female	595	83.8
Male	115	16.2
Major programs		
Science Education	333	46.9
Elementary Education	182	25.6
Early Childhood Education	195	27.5
Computer ownership		
Yes	495	69.7
No	215	30.3
Hours of computer usage		
Less than an hour a day	210	29.6
1-3 hours a day	120	16.9
More than three hours a day	75	10.6
Less than an hour a week	118	16.6
1-3 hours per week	134	18.9
More than three hours a week	53	7.5
Computer competency		
Basic level	54	7.6
Intermediate level	400	56.3
Advanced level	223	31.4
Proficient	33	4.6
Age	M=21.997 (SD=1.734)	

Instruments

A survey questionnaire was distributed to respondents in a paper-pencil environment, and it took approximately 30 minutes for respondents to complete the instrument. Participants were asked to provide demographic information and rate their level of agreement to statements on the ten constructs in the research model; personal science teaching efficacy (PSTE), science teaching outcome expectancy (STOE), source, certainty, development, justification, perceived ease of use (PEU), perceived usefulness (PU), attitude toward using (ATU), and behavioral intention (BI). The instrument has three sections in addition to the demographic information, which are (i) Science Teaching Efficacy Beliefs Instrument (STEBI-B), (ii) Scientific Epistemological Beliefs Scale (SEBs), and (iii) Technology Acceptance Scale.

STEB-I was developed by Enochs and Riggs in 1990 and adapted into Turkish by Tekkaya, Cakiroglu, and Ozkan in 2004. The instrument has two factors, namely PSTE and STOE, with 13 and 10 items, respectively. On the other hand, Conley and colleagues (2004) developed SEBs in 2004, and Bahcivan (2014) adapted the scale into Turkish. The scale has four factors; source (5 items), certainty (6 items), development (6 items), and justification (9 items). The high scores obtained from these scales demonstrate that respondents have a high level of science teaching efficacy beliefs and sophisticated epistemological beliefs.

Lastly, the technology acceptance scale was designed to measure the following fundamental TAM constructs; PEU, PU, ATU, and BI. Most of the items were derived from the study of Ursavaş and colleagues (2014), which has adapted items into Turkish. It may be challenging for studies to result in similar findings related to intention to use technology as the target behavior might have overly broad definitions. Ajzen (2006) proposed that it would be better to identify the behavior at an appropriate level of specificity because individuals focus on these specific definitions, and more valuable results would be yielded. Besides, Sang and colleagues (2010) argued that measuring constructs such as BI by two items might hinder understanding the constructs clearly. Therefore, we added items to the PU, ATU, and BI from the existing research (Sang et al., 2010; Teo, 2009). The added items were prepared by using appropriate translation-back-translation procedures. All statements were measured on a 5-point Likert scale, ranging from "strongly disagree" (1) to "strongly agree" (5). The reliability coefficients were calculated for this study with the obtained data. Table 2 includes the sample items, number of items, sources of the items, and the reliability coefficients for each factor.

Table 2. Number of items, sample items, and Cronbach's alpha for each construct

Variables	Items	Sample item
PSTE (α=.868)	13	I will continually find better ways to teach science.
STOE(α=.779)	10	The teacher is generally responsible for the achievement of students in science.
Source(α=.841)	5	Only scientists know for sure what is true in science.
Certainty(α=.834)	6	Scientists always agree about what is true in science.
Development(α=.794)	6	New discoveries can change what scientists think is true.
Justification(α=.849)	9	A good way to know if something is true is to do an experiment.
Perceived usefulness(α=.831)	7	Using computers will increase my productivity.
Perceived ease of use(α=.768)	3	I find computers easy to use.
Attitudes towards using(α=.795)	8	I like using computers.
Behavioral intention(α=.875)	10	I plan to use computers in the future.

Data Analysis

Data were examined in terms of missing data and outliers before data analysis, and the negatively-worded items were reverse-coded. Descriptive data were generated for all variables, including mean, standard deviation, and minimum and maximum scores. Normality was ensured by checking to what extent the skewness and kurtosis values are in the recommended range (|3| and |10|, respectively). The next step was to assess convergent and discriminant validity and the goodness of the model fit. Lastly, SEM was carried out using AMOS, version 21, to test the hypotheses in the research model. Path coefficients, direct, indirect, and total effects were investigated.

Results and Discussion

This section comprised the descriptive statistics, convergent and discriminant validity of data, path analysis, and hypothesis testing.

Descriptive Statistics

The means, minimum and maximum values, standard deviations, and skewness and kurtosis values were generated for all variables to confirm the normal distribution (Table 3). The mean scores were above the midpoint of 3.00, ranging from 3.46 to 4.03 in all factors. The standard deviations ranged between .478-.708, indicating a narrow spread around the mean.

Table 3. Descriptive statistics, skewness, and kurtosis values for all constructs

	N	Min	Max	Mean	Std. Deviation	Skewness	Kurtosis
Source	710	1.40	5.00	3.74	.686	-.422	.076
Certainty	710	1.00	5.00	3.58	.708	-.475	.162
Development	710	1.17	5.00	3.90	.573	-.721	1.184
Justification	710	1.22	5.00	4.03	.573	-1.071	1.243
PSTE	710	2.00	5.00	3.65	.569	.061	-.486
STOE	710	1.80	5.00	3.46	.478	-,125	.673
PEU	710	1.67	5.00	3.67	.667	-.165	.073
PU	710	2.00	5.00	3.84	.556	-.378	.511
ATU	710	2.00	5.00	3.71	.587	-.013	-.159
BI	710	1.10	5.00	3.86	.553	-.529	1.089

The descriptive statistics imply that respondents had sophisticated epistemological beliefs and a high level of science teaching efficacy beliefs. Besides, they reported positive perceptions regarding the ease of use and usefulness of technology, positive attitudes towards and responses to using technology in education. The skewness (ranging from |.013| to |1.071|) and kurtosis (ranging from |.073| to |1.243|) values were within the recommended value range (|3| and |10|, respectively) by Kline (2011). The results demonstrated that the normality of data was confirmed.

Convergent validity

There is three procedures to assess convergent validity: (1) item reliability, (2) composite reliability index (CRI), and (3) the average variance extracted (AVE) (Fornell & Larcker, 1981). Convergent validity deals with the extent to which different items measure the same construct. Item reliability is associated with the factor loadings, and it is recommended that the correlations between observed and latent variables (factor loadings) be .50 and above (Hair et al., 2019; Kline, 2011). The factor loadings of items ranged between .50-.82, indicating the reliability of items (Table 4). The composite reliability index should be .70 and above to be adequate, and Table 4 demonstrates that all values are above the threshold. The benchmark for AVE to be acceptable is .50 and above. It is worth noting that the adequate levels of CRI might be adequate for convergent validity if AVE values are not within the recommended range (Fornell & Larcker, 1981). It was found that a few of the AVE values were higher than the benchmark of .50. As anticipated by Fornell and Larcker (1981), the factor loadings and CRI values confirm the convergent validity of the research model.

Table 4. Results of convergent validity for the measurement model

Latent Variable	No. of items	Range of the factor loadings	CRI	AVE	Cronbach's α
Source	5	.56-.74	.78	.44	.84
Certainty	6	.53-.66	.77	.39	.83
Development	6	.52-.74	.78	.37	.79
Justification	9	.51-.71	.84	.37	.85
PSTE	13	.58-.82	.93	.51	.87
STOE	10	.53-.64	.81	.30	.78
PEU	3	.56-.69	.67	.40	.77
PU	7	.62-.76	.88	.51	.83
ATU	8	.52-.76	.85	.54	.80
BI	10	.51-.66	.85	.56	.88

Note. CR=$(\sum\lambda)^2/(((\sum\lambda)^2+(\sum(1-\lambda^2)))$
Note. AVE= $(\sum\lambda^2)/(\sum\lambda^2+\sum(1-\lambda^2))$

Discriminant validity

Kline (2011) stated that "a set of variables presumed to measure different constructs show discriminant validity if their intercorrelations are not too high" (p.72.). The intercorrelations ranged between .022-.728, smaller than the benchmark of .90 (Kline, 2011). To assess discriminant validity, the square root of the average variance extracted of each construct is compared to inter-construct correlations of the given construct. The evidence of the discriminant validity is that the square roots of the average variance extracted of the constructs are higher than the intercorrelation coefficients between the given construct and other constructs (Fornell & Larcker, 1981; Kline, 2011). The bold diagonal elements are the square roots of the AVEs, and it is seen that the diagonal elements (the square roots of the AVEs) are higher than the off-diagonal elements (inter-construct correlations), confirming the discriminant validity of the research model for further analysis.

Table 5. Intercorrelation matrix

		1	2	3	4	5	6	7	8	9	10
1. Source	r	**.660**									
2. Certainty	r	.607**	**.624**								
3.Development	r	.226**	.267**	**.608**							
4.Justification	r	.252**	.308**	.593**	**.608**						
5.PSTE	r	.296**	.311**	.320**	.451**	**.714**					
6. STOE	r	.022	.085*	.325**	.359**	.330**	**.547**				
7. PEU	r	.060	.030	.232**	.270**	.214**	.144**	**.632**			
8. PU	r	.196**	.216**	.421**	.523**	.396**	.336**	.387**	**.714**		
9. ATU	r	.204**	.212**	.387**	.436**	.395**	.297**	.372**	.703*	**.734**	
10. BI	r	.120**	.146**	.412**	.467**	.329**	.258**	.425**	.710*	.728**	**.748**

*p<.05, **p<.01

Model fit

The fit between the research model and obtained data were assessed before examining the structural model, employing a maximum likelihood estimation procedure in AMOS, version 21. The most commonly used indices to evaluate the goodness-of-fit were the ratio of the minimum fit function to its degree of freedom (χ^2/df), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Standardized Root Mean Residual (SRMR), and Root Mean Square Error of Approximation (RMSEA). The desirable values of acceptable fit for these indices are lower than 5.0 for χ^2/df, greater than .90 for CFI and TLI, and less than .08 for SRMR and RMSEA, respectively. The results of the research model (χ^2/df= 3.187, CFI=.935, TLI=.907, RMSEA=.056, and SRMR=.734) indicate that the measurement model satisfied that recommended thresholds and has an acceptable level of fit (Hair et al., 2019; Kline, 2011).

Hypothesis testing and path analysis

Twenty-seven hypotheses were generated, and 15 of them were supported. Table 6 reveals the path coefficients between constructs and the results of the hypothesis testing. Preservice teachers' beliefs regarding that the source of knowledge is not always authority significantly influenced personal science teaching efficacy beliefs (β=.163, $p<.01$). The sophisticated epistemological beliefs assuming that there is more than one answer (certainty) had a positive influence on PSTE (β=.180, $p<.001$). Additionally, the justification dimension has positive influences on PSTE (β=.377, $p<.001$), STOE (β=.417, $p<.001$), PEU (β=.4117, $p<.05$) and PU (β=.252, $p<.001$). Lastly, there is a significant positive relationship between the development dimension and STOE (β=.232, $p<.001$).

Table 6. Hypothesis testing results

Hypotheses	Path	Path coefficient	Results
H1a	Source→PSTE	.163**	Supported
H1b	Source→STOE	-.077	Not supported
H1c	Source→PEU	.067	Not supported
H1d	Source→PU	.022	Not supported
H2a	Certainty→PSTE	.180***	Supported
H2b	Certainty→STOE	-.004	Not supported
H2c	Certainty→PEU	-.085	Not supported
H2d	Certainty→PU	.027	Not supported
H3a	Justification→PSTE	.377***	Supported
H3b	Justification→STOE	.417***	Supported
H3c	Justification→PEU	.117*	Supported
H3d	Justification→PU	.252***	Supported
H4a	Development→PSTE	-.038	Not supported
H4b	Development→STOE	.232***	Supported
H4c	Development→PEU	.083	Not supported
H4d	Development→PU	.022	Not supported
H5a	PSTE→PEU	.083	Not supported
H5b	PSTE→PU	.102*	Supported
H5c	PSTE→BI	-.136**	Supported
H6a	STOE→PEU	.104	Not supported
H6b	STOE→PU	.235***	Supported
H6c	STOE→BI	.033	Not supported
H7a	PEU→PU	.234***	Supported
H7b	PEU→ATU	.426***	Supported
H8a	PU→ATU	.651***	Supported
H8b	PU→BI	.670***	Supported
H9	ATU→BI	.361***	Supported

*p<.05, **p<.01, ***p<.001

There were significant relationships between PSTE and PU (β=.102, $p<.05$). Interestingly, the significant relationship between PSTE and BI was negative (β=-.136, $p<.05$). Hypotheses testing revealed that STOE significantly influenced PU (β=.235, $p<.001$). TAM hypotheses were also supported with medium to large effects.

Figure 3 provides the relationships among the latent variables and the explained total variance in each endogenous variable. As Figure 3 shows, STOE and ATU were significant positive predictors of BI. The research model accounts for approximately 59% of the variance in BI, indicating a high level of explained variance. The predictors significantly determined PSTE, STOE, PEU, PU, and ATU by the percentages of 29%, 35%, 9%, 40%, and 46%, respectively.

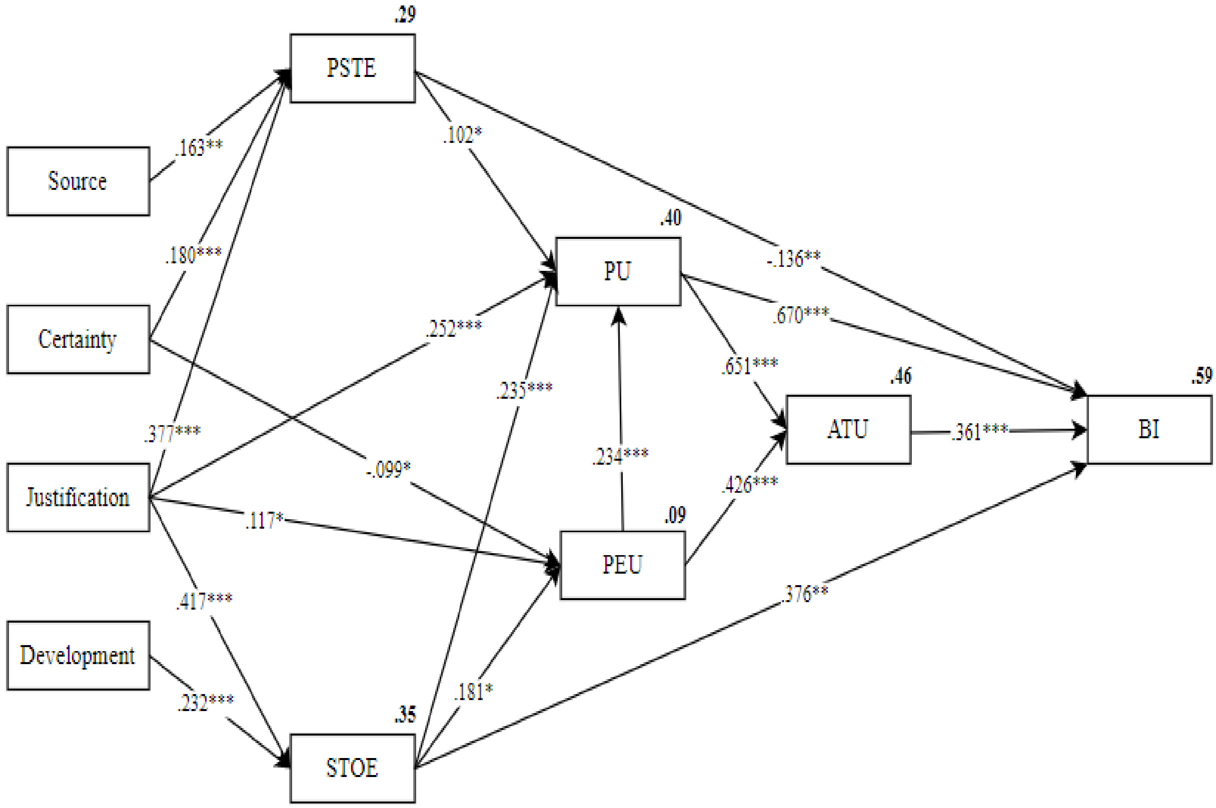

Figure 3. The results of the structural model

Table 7 presents the results of the direct, indirect, and total effects on each endogenous variable. The benchmarks proposed by Cohen (1988) (<.1 as small, <.3 as medium, and <.5 as large) were used in evaluating the size of the effect of a determinant on an outcome. The strongest determinant of BI is PU, with a total effect size of .898, and PEU follows PU with a total effect size of .363, which is entirely an indirect effect. Justification is also the most prominent determinant of BI among the external variables to TAM. The nine determinants in the research model accounted for approximately 59% of the variance in BI.

Similarly, the determinants explained 46% of the variance of ATU. The most vital determinant of ATU was PU, with a total effect size of .632, which is a large effect. PEU and justification were the other strong determinants of ATU, with total effect sizes of .574 and .355, respectively. Lastly, justification was also the strongest determinant of PU and PEU, with total effect sizes of .433 and .192, respectively. The justification was the most dominant determinant of TAM constructs among the variables external to TAM.

Conclusion and Discussion

The present study was designed to test an extended TAM including scientific epistemological beliefs and science teaching efficacy beliefs as external variables. After ensuring the convergent and discriminant validity of the measurement tools, the goodness of the model fit was assessed, and the model-fit analysis has shown that the research model has an acceptable level of fit to the obtained data. Respondents had the highest mean score in the justification dimension and the lowest mean score in science teaching outcome expectancy beliefs. The mean scores in each dimension were above the midpoint of 3, indicating that respondents had sophisticated epistemological beliefs, high levels of science teaching efficacy, positive perceptions regarding the use of technology in education.

Table 7. Direct, indirect, and total effects of the model

Outcome	Determinant	Standard estimates: Direct effect	Indirect effect	Total effect
PSTE (R^2=.29)	Source	.163	-	.163
	Certainty	.180	-	.180
	Justification	.377	-	.377
	Development	-.038	-	-.038
STOE (R^2=.35)	Source	-.077	-	-.077
	Certainty	-.004	-	-.004
	Justification	.417	-	.417
	Development	.232	-	.232
PEU (R^2=.09)	Source	.067	.005	.072
	Certainty	-.085	.015	-.070
	Justification	.117	.075	.192
	Development	.083	.021	.104
	PSTE	.083	-	.083
	STOE	.104	-	.104
PU (R^2=.40)	Source	.022	.015	.037
	Certainty	.027	.001	.028
	Justification	.252	.181	.433
	Development	.022	.075	.097
	PSTE	.102	.019	.121
	STOE	.235	.024	.259
	PEU	.234	-	.234
ATU (R^2=.46)	Source	-	.054	.054
	Certainty	-	-.012	-.012
	Justification	-	.355	.355
	Development	-	.105	.105
	PSTE	-	.112	.112
	STOE	-	.208	.208
	PEU	.426	.148	.574
	PU	.632	-	.632
BI (R^2=.59)	Source	-	.020	.020
	Certainty	-	-.010	-.010
	Justification	-	.381	.381
	Development	-	.116	.116
	PSTE	-.136	.121	-.015
	STOE	.033	.248	.281
	PEU	-	.363	.363
	PU	.670	.228	.898
	ATU	.361	-	.361

The structural model consisted of 27 hypotheses, and the results of the path analysis showed that data supported 15 of them. Path analysis demonstrated that beliefs about the source and certainty of the knowledge were significantly related to PSTE. The justification dimension significantly influenced PSTE, STOE, PEU, PU. In other words, respondents who believed knowledge should be justified by experiments and multiple sources, also feel efficacious in science teaching and have positive perceptions of the technology's ease of use and usefulness. Beliefs about the nature of knowing (source and justification) are found to be positively related to students' efficacy beliefs in learning in science (Kapucu & Bahçivan, 2015). The results consistently showed that preservice teachers' sophisticated beliefs regarding the nature of knowing are positively related to their personal science teaching efficacy beliefs. Surprisingly, respondents' epistemological beliefs, except for justification, did not have significant relationships with PEU and PU. Contrary to this finding, Demirbag and Bahcivan (2021) found that certainty and development dimensions were positively related to digital literacy.

The second belief system added to TAM was science teaching efficacy beliefs, which have two distinct dimensions regarding teaching efficacy and the impact of effective science teaching on students' learning. These distinct dimensions (PSTE and STOE) significantly influenced the PU. Preservice teachers who felt efficacious in science teaching and believed effective teaching would promote students' learning perceived technology as an effective and productive tool in their teaching practices. Additionally, all TAM hypotheses were supported with medium to large effect sizes. This finding supports the existing research addressing the positive relationship between PEU, PU, ATU, and BI (Baydas & Göktas, 2017; Gurer & Akkaya, 2021; Joo et al., 2018; Siyam, 2019; Wong et al., 2012).

This study has also shown the direct, indirect, and total effects of determinants on each endogenous variable. The determinants in the research model explained 59 % of the variance in BI, indicating a greater explained variance than in other studies (Eksail & Afari, 2020; Wong et al., 2012; Sang et al., 2010; Siyam, 2019; Teo, Ursavaş, & Bahçekapılı, 2012). The strongest determinant of BI was PU, with a total effect size of .898. PU has a large direct effect on BI, and ATU also moderated its effect. This finding implies that it is more likely for preservice teachers to use technology in education when they perceive using it would improve their teaching and feel positive emotions regarding its use. The justification dimension had the largest effect on BI among the external variables, and STOE followed it. The effect of justification was entirely indirect. The moderating effect of PSTE, STOE, PEU, and PU promoted the effect of justification. Only believing in knowledge should be justified does not ensure to intend using technology; instead, these beliefs should be supported with science teaching efficacy beliefs and positive perceptions regarding technology use. Interestingly, the direct effect of the PSTE on BI was negative and smaller than the indirect effect. It is possible to imply that the combined effect of PSTE, PEU, PU, and ATU on PSTE leads to positive responses to BI.

The determinants of ATU accounted for approximately 46% of its variance. The most prominent determinants were PU, PEU, and justification, respectively. This finding supports the idea of Venkatesh (2000), indicating that PEU and PU are fundamental constructs in TAM. When it comes to determinants of PU and PEU, justification was the strongest determinant of both. It is also worth noting that science teaching efficacy beliefs and scientific epistemological beliefs explained 9% of PEU variance, leaving 91% unexplained. PEU might be considered a sort of competence to use technology (Wong et al., 2012). Therefore, respondents might not have perceived teaching efficacy and epistemological beliefs as related to PEU. Justification had the largest total effects on PEU, PU, ATU, and BI among the external variables. An implication of this is the possibility that technology is a productive and valuable tool in justifying knowledge, providing multiple opportunities.

To sum up, this study extends our knowledge of TAM and shows that incorporating teaching efficacy and scientific epistemological beliefs into TAM explained more than half of the variance in BI. Preservice teachers' beliefs about justifying knowledge were significant determinants of their technology acceptance and behavioral intention. Teaching efficacy beliefs had small effects on technology acceptance. This finding has important implications for developing preservice teachers' understanding of the relatedness of teaching efficacy and technology and needs to be further examined to reveal the underlying reasons. Besides, an unexpected finding was that PSTE has a negative direct and positive indirect effect on BI. It is difficult to explain this result, but it might be implied that beliefs regarding effective science teaching should be supported with positive perceptions of and attitudes toward technology use to increase preservice teachers' willingness to teach with technology. Further work is required the establish the reasons for the negative effect.

Finally, a number of important limitations need to be considered. Data was collected through self-reported measures, which might lead respondents to overestimate their beliefs, perceptions, and attitudes and give responses to meet the desired outcome. Therefore, it is crucial to collect qualitative data further to give a detailed insight into the findings. Secondly, the current study has only examined the technology acceptance of preservice teachers. To reveal how science teaching and scientific epistemological beliefs impact technology acceptance and behavioral intention, studies with larger samples and in-service teachers might allow to compare and establish the relationships among the observed constructs. Lastly, the research model explained 59% of the variance in BI, and 41% of the variance remained unexplained. Further research should include various external variables to increase the predictivity of the model.

Scientific Ethics Declaration

The authors declare that the scientific ethical and legal responsibility of this article published in JESEH journal belongs to the authors.

References

Ajzen, I. (2006). Constructing a theory of planned behavior questionnaire. *University of Massachusetts Amherst,* 1-7. Retrieved from http://people.umass.edu/aizen/pdf/tpb.measurement.pdf on 1 October 2021.

Akar, S. G. M. (2019). Does it matter being innovative: Teachers' technology acceptance. *Education and Information Technologies*, *24*(6), 3415-3432. https://doi.org/10.1007/s10639-019-09933-z

Al-Azawei, A., & Alowayr, A. (2020). Predicting the intention to use and hedonic motivation for mobile learning: A comparative study in two Middle Eastern countries. *Technology in Society*, *62*, 101325. https://doi.org/10.1016/j.techsoc.2020.101325

Anderson, S. E., Groulx, J. G., & Maninger, R. M. (2011). Relationships among preservice teachers' technology-related abilities, beliefs, and intentions to use technology in their future classrooms. *Journal of Educational Computing Research*, *45*(3), 321-338. https://doi.org/10.2190%2FEC.45.3.d

Avcı Yucel, U., & Gulbahar, Y. (2013). Technology acceptance model: A review of the prior predictors. *Ankara University Journal of Faculty of Educational Sciences (JFES)*, *46*(1), 89-109. https://doi.org/10.1501/Egifak_0000001275

Aypay, A., Celik, H. C., Aypay, A., & Sever, M. (2012). Technology acceptance in education: A study of pre-service teachers in Turkey. *Turkish Online Journal of Educational Technology-TOJET*, *11*(4), 264-272.

Bahcivan, E. (2014). Examining relationships among Turkish pre-service science teachers' conceptions of teaching and learning, scientific epistemological beliefs and science teaching efficacy beliefs. *Journal of Baltic Science Education*, *13*(6), 870-882.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change. *Psychological Review*, *84*(2), 191-215.

Bano, M., Zowghi, D., Kearney, M., Schuck, S., & Aubusson, P. (2018). Mobile learning for science and mathematics school education: A systematic review of empirical evidence. *Computers & Education*, *121*, 30-58. https://doi.org/10.1016/j.compedu.2018.02.006

Bardakcı, S., & Alkan, M. F. (2019). Investigation of Turkish preservice teachers' intentions to use IWB in terms of technological and pedagogical aspects. *Education and Information Technologies*, *24*(5), 2887-2907. https://doi.org/10.1007/s10639-019-09904-4

Baydas, O., & Goktas, Y. (2017). A model for preservice teachers' intentions to use ICT in future lessons. *Interactive Learning Environments*, *25*(7), 930-945. https://doi.org/10.1080/10494820.2016.1232277

Becker, S., Klein, P., Gößling, A., & Kuhn, J. (2020). Using mobile devices to enhance inquiry-based learning processes. *Learning and Instruction*, *69*, 101350. https://doi.org/10.1016/j.learninstruc.2020.101350

Blonder, R., Jonatan, M., Bar-Dov, Z., Benny, N., Rap, S., & Sakhnini, S. (2013). Can you tube it? Providing chemistry teachers with technological tools and enhancing their self-efficacy beliefs. *Chemistry Education Research and Practice*, *14*(3), 269-285. https://doi.org/10.1039/C3RP00001J

Bondy, E., Ross, D., Adams, A., Nowak, R., Brownell, M., Hoppey, D., ... & Stafford, L. (2007). Personal epistemologies and learning to teach. *Teacher Education and Special Education*, *30*(2), 67-82. https://doi.org/10.1177%2F088840640703000202

Byrne, B. M. (2016). *Structural equation modeling with Amos: Basic concepts, applications, and programming* (3rd Ed.). Taylor & Francis Group: New York. http://www.routledge.com/9781138797031

Cheng, M. M., Chan, K. W., Tang, S. Y., & Cheng, A. Y. (2009). Pre-service teacher education students' epistemological beliefs and their conceptions of teaching. *Teaching and Teacher Education*, *25*(2), 319-327. https://doi.org/10.1016/j.tate.2008.09.018

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

Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology*, *29*(2), 186-204. https://doi.org/10.1016/j.cedpsych.2004.01.004

Creswell, J. W. (2015). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (5th Ed.). Boston: Pearson Education.

Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems: Theory and results* (PhD). Cambridge, MA: Massachusetts Institute of Technology. Retrieved from https://dspace.mit.edu/bitstream/handle/1721.1/15192/14927137-MIT.pdf?sequence=2.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319-340. https://doi.org/10.2307/249008

Demirbag, M., & Bahcivan, E. (2021). Comprehensive exploration of digital literacy: Embedded with self-regulation and epistemological beliefs. *Journal of Science Education and Technology*, *30*(3), 448-459. https://doi.org/10.1007/s10956-020-09887-9

Deng, F., Chai, C. S., Tsai, C. C., & Lee, M. H. (2014). The relationships among Chinese practicing teachers' epistemic beliefs, pedagogical beliefs and their beliefs about the use of ICT. *Journal of Educational Technology & Society*, *17*(2), 245-256. https://www.jstor.org/stable/10.2307/jeductechsoci.17.2.245

Ekici, D. I. (2018). Development of pre-service teachers' teaching self-efficacy beliefs through an online community of practice. *Asia Pacific Education Review*, *19*(1), 27-40. https://doi.org/10.1007/s12564-017-9511-8

Eksail, F. A. A., & Afari, E. (2020). Factors affecting trainee teachers' intention to use technology: A structural equation modeling approach. *Education and Information Technologies*, *25*(4), 2681-2697. https://doi.org/10.1007/s10639-019-10086-2

Enochs, L. G., & Riggs, I. M. (1990). Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale. *School Science and Mathematics*, *90*(8), 694-706. https://doi.org/10.1111/j.1949-8594.1990.tb12048.x

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

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39-50. https://doi.org/10.1177%2F002224378101800104

Gagnier, K. M., Holochwost, S. J., & Fisher, K. R. (2021). Spatial thinking in science, technology, engineering, and mathematics: Elementary teachers' beliefs, perceptions, and self-efficacy. *Journal of Research in Science Teaching*. https://doi.org/10.1002/tea.21722

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology*, *76*, 569-582. https://psycnet.apa.org/doi/10.1037/0022-0663.76.4.569

Granić, A., & Marangunić, N. (2019). Technology acceptance model in educational context: A systematic literature review. *British Journal of Educational Technology*, *50*(5), 2572-2593. https://doi.org/10.1111/bjet.12864

Güneş, E., & Bahçivan, E. (2018). A mixed research-based model for pre-service science teachers' digital literacy: Responses to "which beliefs" and "how and why they interact" questions. *Computers & Education*, *118*, 96-106. https://doi.org/10.1016/j.compedu.2017.11.012

Gurer, M. D., & Akkaya, R. (2021). The influence of pedagogical beliefs on technology acceptance: a structural equation modeling study of pre-service mathematics teachers. *Journal of Mathematics Teacher Education*, 1-17. https://doi.org/10.1007/s10857-021-09504-5

Gyamfi, S. A. (2016). Identifying Ghanaian pre-service teachersâ readiness for computer use: A Technology Acceptance Model approach. *International Journal of Education and Development Using ICT*, *12*(2), 105-122.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. (2019). *Multivariate data analysis* (8th Ed.). United Kingdom: Pearson Education, Inc.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, *67*, 88-140. https://doi.org/10.3102%2F00346543067001088

Hoy, A. W., & Spero, R. B. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and teacher education*, *21*(4), 343-356. https://doi.org/10.1016/j.tate.2005.01.007

Huang, F., & Teo, T. (2019). Influence of teacher-perceived organisational culture and school policy on Chinese teachers' intention to use technology: An extension of technology acceptance model. *Educational Technology Research and Development*, *68*, 1547-1567. https://doi.org/10.1007/s11423-019-09722-y

Huang, F., & Teo, T. (2021). Examining the role of technology-related policy and constructivist teaching belief on English teachers' technology acceptance: A study in Chinese universities. *British Journal of Educational Technology*, *52*(1), 441-460. https://doi.org/10.1111/bjet.13027

Huang, S. Y., Kuo, Y. H., & Chen, H. C. (2020). Applying digital escape rooms infused with science teaching in elementary school: Learning performance, learning motivation, and problem-solving ability. *Thinking Skills and Creativity*, *37*, 100681. https://doi.org/10.1016/j.tsc.2020.100681

Joo, Y. J., Park, S., & Lim, E. (2018). Factors influencing preservice teachers' intention to use technology: TPACK, teacher self-efficacy, and technology acceptance model. *Journal of Educational Technology & Society*, *21*(3), 48-59. https://www.jstor.org/stable/26458506

Kapucu, S., & Bahçivan, E. (2015). High school students' scientific epistemological beliefs, self-efficacy in learning physics and attitudes toward physics: A structural equation model. *Research in Science & Technological Education*, *33*(2), 252-267. https://doi.org/10.1080/02635143.2015.1039976

Kartal, B., & Çınar, C. (2018). Examining pre-service mathematics teachers' beliefs of TPACK during a method course and field experience. *Malaysian Online Journal of Educational Technology (MOJET)* *6*(3), 11-37. http://dx.doi.org/10.17220/mojet.2018.03.002

Kartal, B. (2019). Examining pre-service teachers' attitudes beliefs, and intentions related to using educational technologies. *YYÜ Journal of Education Faculty*, *16*(1), 436-463. http://dx.doi.org/10.23891/efdyyu.2019.132

Kartal, B. (2020). Pre-service science and mathematics teachers' teaching efficacy beliefs and attitudes toward teaching: A partial correlation research. *Australian Journal of Teacher Education (Online)*, *45*(9), 42-61. http://dx.doi.org/10.14221/ajte.2020v45n9.3

Kartal, T., & Dilek, I. (2021). Developing pre-service elementary science teachers' science teaching efficacy beliefs through microteaching by promoting efficacy sources. *International Journal on Social and Education Sciences*, *3*(4), 710-731. https://doi.org/10.46328/ijonses.124

Kızıltepe, İ. S., & Kartal, T. (2021). Examination of pre-service teachers' epistemological beliefs in terms of multiple variables. *Journal of Kırşehir Education Faculty, 22*(3), 2036-2069. https://doi.org/10.29299/kefad.990289

Kline, R. B. (2011). Convergence of structural equation modeling and multilevel modeling. In M. Williams (Ed.), *The SAGE Handbook of Methodological Innovation*. Thousand Oaks, CA: Sage.

Lee, J., Cerreto, F. A., & Lee, J. (2010). Theory of planned behavior and teachers' decisions regarding use of educational technology. *Journal of Educational Technology & Society*, *13*(1), 152-164. https://www.jstor.org/stable/10.2307/jeductechsoci.13.1.152

Lee, S. W. Y., Luan, H., Lee, M. H., Chang, H. Y., Liang, J. C., Lee, Y. H., ... & Tsai, C. C. (2021). Measuring epistemologies in science learning and teaching: A systematic review of the literature. *Science Education*, *105*, 880-907. https://doi.org/10.1002/sce.21663

Li, Y., Garza, V., Keicher, A., & Popov, V. (2019). Predicting high school teacher use of technology: Pedagogical beliefs, technological beliefs and attitudes, and teacher training. *Technology, Knowledge and Learning*, *24*(3), 501-518. https://doi.org/10.1007/s10758-018-9355-2

Min, M., Akerson, V., & Aydeniz, F. (2020). Exploring preservice teachers' beliefs about effective science teaching through their collaborative oral reflections. *Journal of Science Teacher Education*, *31*(3), 245-263. https://doi.org/10.1080/1046560X.2019.1690818

Parkman, S., Litz, D., & Gromik, N. (2018). Examining pre-service teachers' acceptance of technology-rich learning environments: A UAE case study. *Education and Information Technologies*, *23*(3), 1253-1275. https://doi.org/10.1007/s10639-017-9665-3

Sang, G., Valcke, M., Van Braak, J., & Tondeur, J. (2010). Student teachers' thinking processes and ICT integration: Predictors of prospective teaching behaviors with educational technology. *Computers & Education*, *54*(1), 103-112. https://doi.org/10.1016/j.compedu.2009.07.010

Scherer, R., & Teo, T. (2019). Unpacking teachers' intentions to integrate technology: A meta-analysis. *Educational Research Review*, *27*, 90-109. https://doi.org/10.1016/j.edurev.2019.03.001

Scherer, R., Siddiq, F., & Tondeur, J. (2020). All the same or different? Revisiting measures of teachers' technology acceptance. *Computers & Education*, *143*, 103656. https://doi.org/10.1016/j.compedu.2019.103656

Schommer-Aikins, M. (2004). Explaining the epistemological belief system: Introducing the embedded systemic model and coordinated research approach. *Educational Psychologist*, *39*(1), 19-29. https://doi.org/10.1207/s15326985ep3901_3

Siyam, N. (2019). Factors impacting special education teachers' acceptance and actual use of technology. *Education and Information Technologies*, *24*(3), 2035-2057. https://doi.org/10.1007/s10639-018-09859-y

Tekkaya, C., Cakiroglu, J., & Ozkan, O. (2004). Turkish pre-service science teachers' understanding of science and their confidence in teaching it. *Journal of Education for Teaching*, *30*(1), 57-68. https://doi.org/10.1080/0260747032000162316

Teo, T. (2008). Pre-service teachers' attitudes towards computer use: A Singapore survey. *Australasian Journal of Educational Technology*, *24*(4). https://doi.org/10.14742/ajet.1201

Teo, T. (2009). Modelling technology acceptance in education: A study of pre-service teachers. *Computers & Education*, *52*(2), 302-312. https://doi.org/10.1016/j.compedu.2008.08.006

Teo, T. (2010). A path analysis of pre-service teachers' attitudes to computer use: applying and extending the technology acceptance model in an educational context. *Interactive Learning Environments*, *18*(1), 65-79. https://doi.org/10.1080/10494820802231327

Teo, T., & Zhou, M. (2017). The influence of teachers' conceptions of teaching and learning on their technology acceptance. *Interactive Learning Environments*, *25*(4), 513-527. https://doi.org/10.1080/10494820.2016.1143844

Teo, T., Chai, C. S., Hung, D., & Lee, C. B. (2008). Beliefs about teaching and uses of technology among pre-service teachers. *Asia-Pacific Journal of Teacher Education*, *36*(2), 163-174. https://doi.org/10.1080/13598660801971641

Teo, T., Fan, X., & Du, J. (2015). Technology acceptance among pre-service teachers: Does gender matter?. *Australasian Journal of Educational Technology*, *31*(3). https://doi.org/10.14742/ajet.1672

Teo, T., Ursavaş, Ö. F., & Bahçekapili, E. (2012). An assessment of pre-service teachers' technology acceptance in Turkey: A structural equation modeling approach. *Asia-Pacific Education Researcher*, *21*(1), 191-202. http://researchrepository.murdoch.edu.au/id/eprint/48415

Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783-805. https://doi.org/10.1016/S0742-051X(01)00036-1

Ursavaş, Ö., Şahin, S., & McIlroy, D. (2014). Technology acceptance measure for teachers: T-TAM. *Journal of Theory and Practice in Education*, *10*(4), 885-917. Retrieved from https://dergipark.org.tr/tr/pub/eku/issue/5462/74152

Venkatesh, V. (2000). Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Information Systems Research*, *11*(4), 342-365. https://doi.org/10.1287/isre.11.4.342.11872

Wong, K. T., Teo, T., & Russo, S. (2012). Influence of gender and computer teaching efficacy on computer acceptance among Malaysian student teachers: An extended technology acceptance model. *Australasian Journal of Educational Technology*, *28*(7). https://doi.org/10.14742/ajet.796

Author(s) Information

Tezcan Kartal
Kırşehir Ahi Evran University
Department of Mathematics and Science Education, Kırsehir, Turkey
ORCID ID: https://orcid.org/0000-0001-7609-3555

Ibrahim Serdar Kiziltepe
Kırşehir Ahi Evran University
Department of Mathematics and Science Education, Kırsehir, Turkey
ORCID ID: https://orcid.org/0000-0002-6210-5372

Busra Kartal
Kırşehir Ahi Evran University
Department of Mathematics and Science Education, Kırsehir, Turkey
Contact e-mail: busra.kartal@ahievran.edu.tr
ORCID ID: https://orcid.org/0000-0003-2107-057X

JESEH
Journal of Education in Science, Environment and Health
Volume 8, Issue 1, 2022 https://doi.org/10.21891/jeseh.1059124

The Effect of STEM Applications on the Scientific Creativity of 9th-Grade Students

Seyide Eroglu, Oktay Bektas

Article Info

Article History

Published:
01 January 2022

Received:
23 August 2021

Accepted:
01 January 2022

Keywords

Experiential learning theory
STEM education
Scientific creativity
Atomic and periodic system

Abstract

The study aims to investigate the effect of STEM applications designed for the atomic system and periodic system unit on the scientific creativity of 9th-grade students. The exploratory sequential design, mixed research, is used in the study. The pretest-posttest quasi-experimental design has been preferred for the quantitative part of the study and the phenomenology design for the qualitative part. The study lasted 12 weeks during the first semester of the 2016-2017 academic year. The sample of the study consists of 133 ninth-grade high school students. The study uses an experimental and a control group. The Scientific Creativity Test was used as a data collection tool. Quantitative data were analyzed using the independent samples t-test through the package program SPSS 22. On the other hand, the qualitative data were analyzed using content analysis. As a result of the study, a statistically significant difference was observed to exist between the groups in terms of scientific creativity in favor of the experimental group. In addition, the participants were determined to have put forward many different and extraordinary thoughts with different perspectives. Based on these results, including activities that will improve students' scientific creativity in chemistry lessons is suggested.

Introduction

Creativity is a very old concept and cannot be easily defined because it cannot be observed directly (Lewis, 2005). Although there is no universal definition of creativity, there are some common points where different definitions come together. Considering these common points, the statements "unusual, original, cannot be evaluated with certain standards" draw attention to the concept of creativity (Boden, 1994). Creativity appears in different fields and includes different process steps (Glück et al., 2002). For example, creativity in the field of science is expressed as scientific creativity.

Torrance (1984) defined creativity as the ability individuals have to be sensitive to disruptions, changes, or incompatibilities in their environment; to identify the problems in their environment, and to produce solutions to these problems. Torrance (1984) highlighted issues related to scientific creativity such as presenting the problem situation clearly and listing different solution suggestions in the solution process. Scientific creativity can be expressed as creating new theories, organizing new experiments, and putting forward new ideas based on existing scientific knowledge (Hu & Adey, 2002; Moravcsik, 1981). Problem-solving hypothesizing, experimental design, and technical innovation are expressed as components specific to scientific creativity (Alexander, 1992; Lin et al., 2003). The steps in the process of scientific creativity are listed as recognizing the problem, ordering the solution suggestions, testing and determining the most appropriate solution, applying it, and finally accepting, rejecting, or revising it (Hu et al., 2013; Moravcsik, 1981).

To develop scientific creativity, teachers should support the use of open-ended questions. Therefore, students should be able to ask enough open-ended questions to solve a problem. Teachers should also encourage their students to use techniques such as brainstorming (Siew, et al. 2015). In addition, teachers should develop their students' critical thinking and product creation skills. In other words, to encourage creative thinking, teachers should enable their students to form hypotheses for the solution of a problem, design experiments, and enable them to follow technological developments (Lin, et al., 2003). For all these reasons, activities based on STEM education were used in this study.

Different studies are found in the literature on determining creativity (Guilford 1988; Torrance, 1990). The most widely known of these is the Torrance Tests of Creative Thinking Test (TTCT). It consists of two verbal and two modal parts, four in total. In addition, TTCT consists of the sub-dimensions of fluency, flexibility, and originality (Chen et al., 2015). Because determining creativity in a specific field such as a scientific field is not possible with TTCT, the Scientific Creativity Test for Secondary School Students (SCT) was developed by Hu and Adey (2002) to determine students' creative thinking. SCT consists of seven open-ended questions, each measuring an aspect of scientific creativity (i.e., unusual uses, problem finding, product development, scientific imagination, problem-solving, science experimentation, and product design). Each question is evaluated in terms of the sub-dimensions of fluency, flexibility, and originality (Lin, et al., 2003).

Hu and Adey (2002, p.391) created the three-dimensional Scientific Structure Creativity Model (SSCM) to explain scientific creativity (see Figure-1). The SSCM consists of three basic dimensions: process, personal feature, and product. The dimension of the process is based on imagination and thinking. Imagination is one of the most basic features found in creative individuals. In the process, individuals think about the solution to the problem using their imagination. Thinking here is not in the form of convergent thinking (suggesting the most appropriate solution according to the available information) but divergent thinking (proposing the untested, different, or outside the existing solution). The dimension of the personal feature includes the sub-dimensions of fluency, flexibility, and originality. Fluency is defined as the ability to produce many solutions to the problem situation, while flexibility is defined as the ability to produce different solutions by looking at the problem situation from different angles. Finally, originality is expressed as being able to produce unique individual solutions for the problem situation. The dimension of the product consists of four sub-dimensions: technical product, scientific knowledge, scientific fact, and scientific problem. The creative product presented as a solution to the problem situation should have a technical infrastructure, scientific knowledge, and a structure that includes scientific facts (Hu & Adey, 2002).

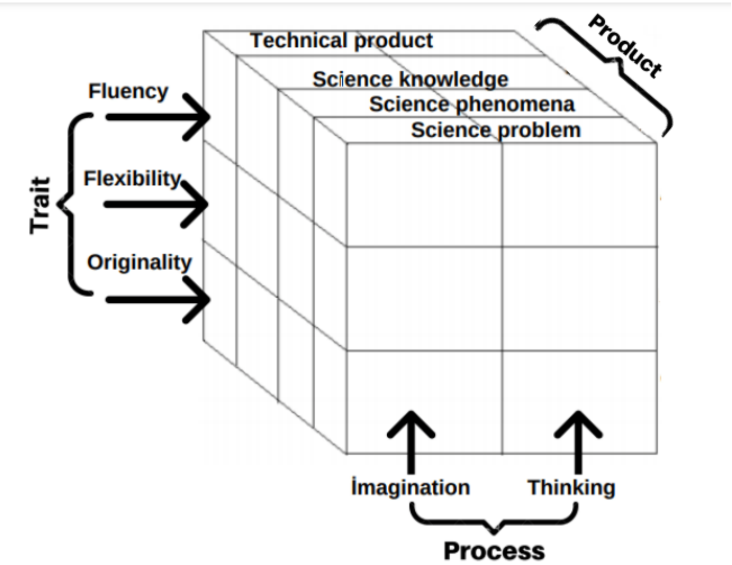

Figure 1. Scientific structure creativity model (Hu & Adey, 2002)

Scientific creativity can be expressed as a driving force in the emergence of scientific developments. Therefore, young people with creative skills are needed for countries to develop economically and progress in the fields of science and technology by increasing their competitiveness (Martins & Terblanche, 2003; Selby et al., 2005). Educational environments need to be rearranged to raise young people with scientific creativity. In other words, environmental conditions must be made suitable for developing individuals' creativity skills. Improving individuals' environmental conditions is necessary for the classroom environment. When examining the related literature, classroom environments enriched with various tools are stated to be important for developing students' creativity skills (Kang et al., 2015; Urban, 2005).

Science classes have an important role in developing scientific creativity skills. In science lessons, students learn that science begins with curiosity, that imagination has an important place in scientific studies, and that a problem can have more than one solution. In this respect, science classes support the development of scientific creativity (Hadzigeorgiou et al., 2012; Moravcsik, 1981).

As a science class, chemistry is also an important lesson that can contribute to developing scientific creativity. One of the chemistry course aims is to raise creative individuals who learn by doing, question, are solution-oriented and curious and have scientific literacy (Daud et al., 2012). Therefore, this study has adopted the idea that students' scientific creativity can be developed in chemistry.

When examining the related literature, different learning strategies are seen to have been used to develop scientific creativity in science lessons (e.g., Semmler & Pietzner, 2017). As an interdisciplinary approach, STEM education is among the strategies used to develop scientific creativity (Barry & Kanematsu, 2006; Barry et al., 2017; Henriksen, 2014; Kanematsu & Barry, 2016; Walsh et al., 2013).

STEM is the abbreviation for the disciplines of science, technology, engineering, and mathematics (Dugger, 2010, p. 2). STEM is an innovative approach in the field of education based on the integration of these four main disciplines (English, 2017; White, 2014). The philosophical foundations of STEM education are based on the constructivist approach (Kelley & Knowles, 2016). Constructivism puts the student at the center of the learning process and is an approach in which the teacher guides the process. Constructivism wants students to take an active role in the learning process and advocates the development of creativity skills (Fosnot, 2013). This philosophical approach is suitable for the nature of STEM, which is based on learning by experience (Nuangchalerm, 2018), and the nature of scientific creativity based on generating many different solutions to real-life problems (Barry & Kanematsu, 2006). In addition, because the 5E (engage, explore, explain, elaborate, evaluate) learning cycle model is a useful and frequently preferred learning model especially in science and chemistry lessons, STEM applications have been decided to be integrated into the 5E learning cycle model (Bybee & Landes, 1990).

Many studies are found in the literature on the positive effects of STEM in the field of education (Cole & Mathilde, 2008; Ostler, 2012; Van Soom & Donche, 2014). These studies stated the integration of different disciplines into courses will increase students' academic achievement, contribute to their meaningful learning, and increase their motivation toward lessons (Van Soom & Donche 2014). STEM has also been underlined to play an important role in the development of 21^{st}-century skills such as critical thinking, problem-solving, and creativity (Wang, 2012).

In addition to those mentioned above, the literature has emphasized learning environments supported by STEM education to increase scientific creativity in students (e.g., Walsh et al., 2013). Kanematsu and Barry (2016) argued that all countries wanting to get stronger should develop scientific creativity with STEM education. Barry and Kanematsu (2006) stated the field of chemistry to be important for society and many things in daily life to be related to chemistry. This is why chemistry teachers need to turn their lessons into scientific experiences and encourage scientific creativity in the laboratory. Based on all these, STEM education is used in this study to improve students' scientific creativity. As the subject, the "Atomic and the Periodic System" course has been chosen. When examining the contents of the atomic and periodic system unit, most topics are seen to be abstract and to not allow exceeding the ideas put forth (e.g., coming up with many different and original ideas about atomic models and the periodic table is not possible in a scientific framework) (Azizoğlu et al., 2015). The aim is to examine whether students' scientific creativity can be improved in a subject involving abstract concepts and limitations using STEM-based lesson activities. Conducted to reveal the effect of STEM-based lesson activities on students' scientific creativity, the study seeks answers to the questions "What is the effect of STEM applications on ninth-grade students' scientific creativity scores?" and "What are the students' thoughts about scientific creativity?"

Method

Research Model

This study uses the explanatory sequential design, a mixed research method, taking into account the classifications of Creswell and Plano Clark (2007). The study uses the pretest-posttest control group quasi-experimental design for collecting the quantitative data. The phenomenology design has been used to collect the qualitative data of the study. Phenomenology offers the researcher the opportunity to examine people's existing event experiences in depth (Van Manen, 2007). Researchers can interpret participants' views of the participants using the hermeneutic model of phenomenology (Van Manen, 2007). This study prefers hermeneutic phenomenology to interpret the effect of STEM-based applications on students' scientific creativity based on the participants' experiences.

Participants

The target population of this study is all 9th-grade students in the Kocasinan district of Kayseri Province in Turkey. The accessible population is the 9th-grade students in the Kocasinan district. To generalize the accessible population, the number of 9th-grade students in the accessible population was determined first, and at least 10% of this number was reached (Tabachnick & Fidell, 2013). Cluster sampling was preferred in the study. Cluster sampling is based on the inclusion of groups that are similar in some respects in the accessible population (Thompson, 1990). In this study, the existing classes in the accessible population were accepted as clusters. Therefore, four classes from the clusters in a school in the accessible population were included in the study.

The sample of the study consists of 133 9th-grade students studying at an Anatolian High School that the first researcher could easily access; the experimental group has 68 students and the control group has 65. Qualitative data were collected using easily accessible case sampling, a type of purposeful sampling. With this sampling, researchers prefer participants who are appropriate in terms of time and location and who also want to participate in the study voluntarily (Merriam, 2013). In the findings section, participants in the experimental group are represented by the letter "E". In this study, the researchers used the statements of the participants in the experimental group for qualitative data, as they tried to prove the effectiveness of the practices in the experimental group on scientific creativity.

Data Collection Tools

This study uses the 7-item Scientific Creativity Test for Secondary School Students (SCT) developed by Hu and Adey (2002). The researchers examined the Turkish version of the test (Atasoy et al., 2007) and adapted the test questions to the "Atomic and Periodic System" unit. The pilot application of the draft test was carried out by the first author to determine the understandability of the questions. The pilot study was applied to 69 tenth-grade high school students (48 girls and 21 boys).

The results obtained from the pilot study were presented to the opinions of three science educators, and the necessary corrections and changes were made to the test questions. For example, regarding Question 7, the students were determined to be unable it present creative ideas and to have been affected by the examples given under the questions. Accordingly, the question "Please design an atomic model that can best describe the internal structure of an atom by considering the model that appears in your mind when the atom is mentioned. Draw the picture of the model you have designed and indicate the name and function of each part" was changed completely to "Design an experimental setup for proving the conservation of mass. Describe each step of the experiment using pictures or figures. Write down the names of each implement or base substance in the setup".

In addition, the statement "How did people from ancient times treat their illnesses?" was removed from the second question and the statement "How will people treat their illnesses in the future?" was added. In addition, in the study, the Cronbach alpha reliability coefficient was calculated as .823 to determine the reliability of the scores obtained from the BYT. Therefore, researchers have determined that the scores of the participants reflect reality at the level of .823 (Pallant, 2016).

Data Collection Process

The necessary official permissions have been obtained for each stage of the study. The data were collected in the first semester of the 2016-2017 academic year. SCT was applied as a pretest (pre-SCT) at the beginning of the study and as a posttest at the end (post-SCT). Course applications were carried out by a chemistry teacher, not the researchers. The first researcher conducted teacher training on STEM activities before the pretest was applied. The practicing teacher did not have any previous knowledge about STEM education. The first researcher gave him one-week training.

The teacher was also supported with written documents related to STEM education. Throughout the application, the first researcher helped the teacher of each group. A research assistant in science education took classes as an observer to observe the teacher. Lessons were conducted in both the experimental and control groups as prescribed by the chemistry course curriculum. The treatment took eight weeks. Teacher training, pretests, treatment, project presentations, and posttests were conducted over a total of 12 weeks.

Treatment

The application process consisted of 16-course hours, two hours per week. In addition, the presentation of the project studies was made in the ninth week. In the experimental group, lessons on the atomic and periodic system were conducted using the disciplines of STEM education. While integrating STEM into the lessons, chemistry was focused on, and at least one of the other disciplines (technology, mathematics, and engineering) was included with the course integration being carried out according to the steps in the 5E learning cycle model. In the STEM activities created to develop students' creativity, the attempt was made to present learning environments with visuals and the support of digital resources. In addition, the attempt was made to support students' creativity through activities that encourage students to research and view a situation from different angles. The weekly distribution of STEM activities associated with the 5E learning cycle model, dimension of STEM, and scientific creativity is shown in Table 1.

Table 1. Distribution of STEM activities by week

Week	STEM activity	5E	Dimension of STEM	Association with Scientific Creativity
1	Black Box	Explore	Science, technology, engineering, and mathematics	Identifying numerous solution suggestions for predicting objects in the box (fluency) Model creation (originality)
2	Einstein's Big Idea	Elaborate	Science and technology	Emphasis on how scientists use their creativity in their work (fluency, flexibility, and originality)
3	Democtritus' ideas about the atom	Engage	Science and mathematics	Making inferences about how Democtritus' ideas about the atom were formed (fluency and originality)
4	Project_1 (subatomic particles)	Evaluate	Science, technology, engineering, and mathematics	Creating products using different materials (originality)
5	Atomic models (Dalton, Thomson, and Rutherford)	Explore - Explain	Science and technology	Supporting students in making original drawings that contain many ideas about the structure of the atom (fluency and originality)
6	Classification of the elements	Explore - Explain	Science, technology, and mathematics	Students are expected to make different classifications based on elements' different physical/chemical properties (fluency, flexibility, and originality)
7	4D element	Elaborate	Science and technology	Enabling students to create different compounds by combining different elements (fluency and originality)
8	Project 2 (My favorite element)	Evaluate	Science, technology, engineering, and mathematics	Asking students to choose one of the first 20 elements in the periodic table and design a model for their chosen element (fluency, flexibility, and originality)
9	Project 3 (Periodic table design)	Evaluate	Science, technology, engineering, and mathematics	Asking students to design periodic tables using different materials (fluency, flexibility, and originality)

This study is based on the Scientific Structure Creativity Model (SSCM). Activities were designed by taking into account the dimensions expressed in this model. In the control group, courses were carried out by traditional methods in line with the objectives in the chemistry curriculum. The methods, techniques, and learning tools used by the teacher in her class were used. Although the science program adopted a student-centered approach, the teacher took a more teacher-centered approach. For instance, only textbooks and smartboards were employed in the lessons. In addition, no experiments were performed in the courses. The teacher explained the lessons with lecturing and solved the sample questions in the textbook. The teacher wrote some parts on the board, then the students wrote what was on the board in their notebooks.

Power and Effect Size

The minimum number of samples required in the study was calculated as 102 using GPower 3.1.9.4 program. For this aim, the level of significance (α = .05), the calculated power (p = 0.80), and the effect size (d = 0,5) was determined at the beginning of the study (Cohen et al., 2003; Hinkle et al., 2003). In this study, the number of participants was 133 and therefore, more than the minimum number of samples. The calculated power was compared with the observed power of the study obtained by the GPower 3.1.9.4 program. Also, the effect size at the beginning of the study was compared with the effect size at the end of the study. Therefore, the generalisability of the study to the accessible population was discussed in terms of external validity. The effect size at the end of the study was calculated using the effect size calculator excel file on the cem.org site.

Data Analysis

Quantitative data analysis. Quantitative data were analyzed using the SPSS-22 package program. To decide whether parametric tests can be used or not, the normal distribution was first checked (Pallant, 2016). For this purpose, the Kolmogorov Smirnov test was used as the sample size is over 50 (Pallant, 2016). Secondly, whether the arithmetic means, mode, and median values (measures of central tendency) are close to or equal to each other was checked to test the normal distribution (George & Mallery, 2001). Thirdly, whether or not the skewness and kurtosis values were between -1 and +1 was checked (Tabachnick & Fidel, 2013). In addition, the assumptions of adequate sample and homogeneity of the variances between groups (Levene test) were met (Pallant, 2016). After the assumptions were seen to be met, the research question was answered using the independent samples t-test (Pallant, 2016). A significance level of $p > 0.05$ was taken as the basis in the statistical analysis.

Qualitative data analysis. Qualitative data was analyzed obtained from the participant expressions and drawings in the SCT using content analysis (Marshall & Rossman, 2006). Codes and categories were created for the students’ responses. Taking into account the studies in the literature (e.g., Deniş Çeliker & Balım, 2012), the codes for each question were collected together. Therefore, they coded openly. At the next stage, they collected the codes under categories, taking into account the causal connections between the codes. Thus, they did axial coding (Corbin & Strauss, 1990). The categories were created taking into account the scientific creativity dimensions of Hu and Adey (2002). Therefore, themes are fluency, flexibility, and originality. Authors have reached a consensus on categories and codes. Since the scientific creativity test covers the trait dimension, the other two dimensions of creativity were ignored.

Three different sub-scores were calculated for Questions 1, 2, 3, and 4 about fluency, flexibility, and originality. While calculating the fluency sub-score, the answers given by the participants unrelated to their qualifications regarding the questions were counted. Participant expressions were categorized to calculate the flexibility sub-score. Participants received a score of 1 for each category in which their answers (codes) were included. To calculate the originality sub-score, the participant codes were scored based on the percentage of repetition. A sub-score of 2 was assigned for answers with a repetition frequency of less than 5%, of 1 for answers with a repetition frequency of 5-10%, and of 0 for answers with a repetition frequency of more than 10% (Deniş Çeliker & Balım, 2012). In Tables 2-5 below, the categories and codes for the fluency and flexibility for Questions 1-4 are listed in each respective table.

Table 2. Question 1: Categories and codes created for fluency and flexibility

Category	Codes
Modeling	Chemistry modeling
	Modeling states of matter
	Atomic structure
	Creating an atom model
Experimental	Establishing an experimental setup
Features of the ping-pong ball	Physical-chemical properties of a ping-pong ball
Used in Physics, Biology, and	Use in physics
Astronomy	Use in biology
	Use of Astronomy
Material making	Chemistry material making
Lecture	Atomic calculations
	Historical development of the atom

Table 3. Question 2: Categories and codes created for fluency and flexibility

Category	Codes
Past life	Dinosaurs
	Daily life in ancient times
Future life	Life in the future
	Technology in the future
Discovery	Past discoveries
	Future discoveries
The beginning of life	The beginning of humanity and the world
Health	Health in the past
	Health in the future
Scientists and scientific studies	Structure of the atom
	Scientists
Mysterious events	Egyptian pyramids
Past events	History of the Turks
	Ottoman period
	The beginning of Islam

Table 4. Question 3: Categories and codes created for fluency and flexibility

Category	Codes
Elemental Properties	Changing elements' location
	Examples of elements
	Changing the names of the elements
Visually impaired	Braille for the visually impaired.
	With sound and light
Functional / electronic	Electronic
	Tactile feature
	Coded
	Smart device application
	Interactive
	Usage in different places
Different criteria	Different size and dimension
	Different shapes
	Secret alphabet
	Fragrant
Instructive	Game
	Instructive
	Humorous
Visual richness	Different coloring
	Enriched with visuals

Table 5. Question 4: Categories and codes created for fluency and flexibility

Category	Codes
Earth / Planet	Changes regarding Earth and planets.
...it breaks down	The balance would be disrupted.
	The structure of matter would deteriorate.
Developments (inventions/technology)	Technology would not improve.
	Different studies and inventions could not be made.
Creatures and life	Vital activities get harder.
	Item variety decreases.
Structure of the atom	There would be no atom and subatomic particles.
	Electron-related change.
Gravity	There would be no gravitational force
	There would be no magnetic field.

... does not occur	Elements, compounds, and mixtures would not have been formed. Integrity/ objects would not be formed There would be no chemical reactions. There would be no electricity/electrification. Magnetism would not have formed. Machines would not be used / would not work Bonds would not be formed. There is no change of state. Nothing would happen.

Question 5 was scored as the sum of the fluency and originality sub-scores. Participants were asked to divide a square into four equal parts, and the participants' drawings were scored. According to this, a sub-score of 3 is given for each drawing with less than a 5% repetition rate, 2 points for a 5-10% repetition rate of drawings, and 1 for each drawing with a repetition rate greater than 10%. Below are the drawings encountered in the Pre-SCT in Figure 2.

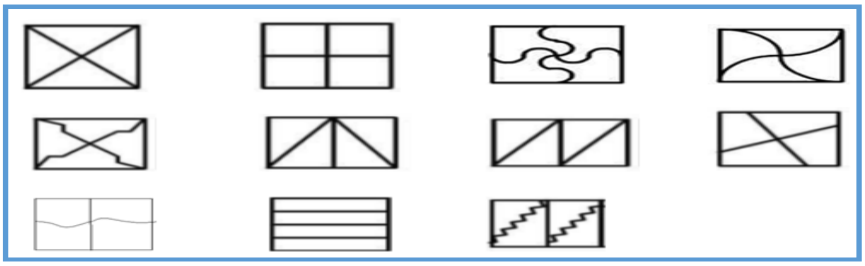

Figure 2. Pre-SCT drawings.

Question 6 asks the participants to suggest methods for comparing the reactivity of two metals. Scoring for this question was carried out in over two different sub-scores: fluency + flexibility and originality. Participants received two points for each method they suggested, and an additional two points when they made explanations about the method. The total score was expressed as the method score. Also, each method was sub-scored separately while giving points for originality. While giving scores for originality, a sub-score of 4 was given for each answer with less than a 5% repetition rate for the drawings and of 2 for a 5-10% rate of repetition. Some of the methods suggested by the participants are listed below.

- Reaction state.
- Melting /Heating.
- Ability to attract other metals.
- Resistance.
- Magnetic attractiveness.
- Electrical conductivity.
- Electron donation tendency.
- Position on the periodic table.
- Mass.
- Combustion reactions.
- Reactivity with nonmetals.
- Mixing with different substances.
- Density.
- Using parentheses.
- A neutral bar.
- Reactivity with water.
- Oxidation (Reactivity with Oxygen [O_2]).
- Luminosity.
- Measuring precise angles.
- Machinability.
- Research.
- The damage they cause when thrown to the ground.

Question 7 asks the participants to create an experimental setup showing the conservation of mass. Regarding this question, students were expected to make drawings and explanations about their drawings. For the experimental drawing setup, a fluency + flexibility sub-score was given for a maximum of 3 points for each criterion (a maximum of 15 points total). In addition, a sub-score of originality between 1 and 5 points was given based on the overall impression. Unrealistic expressions were also scored by consensus as 1 sub-score for the originality. Table 6 lists the criteria that formed for Question 7.

Table 6. Criteria for Question 7

Criteria number	Criteria name
1	Particle-sized display
2	Expressing through numbers
3	Establishing an experimental setup
4	Giving examples of elements and using compounds
5	Explanation of the process
6	Other (Unrealistic Statements)

Validity and Reliability

Regarding external validity; the sample size was chosen to reflect the accessible population and to be generalizable to the accessible population. To generalize to the accessible population, the effect size and power of the study were calculated. In addition, regarding the number of participants in the sample, it was acted by the rule of working with at least 10% of the accessible population.

Regarding internal validity; a content validity study was conducted for the data collection tools used in the study. For this purpose, the literature was reviewed and the available tests were listed and evaluated together. The current test was partially arranged according to the "Atom and Periodic System" subject contents. At this stage, the expert opinion of three science educators who are experts in their fields was sought. A pilot study was conducted to test the suitability of the items. Changes made according to the results of the pilot study are mentioned in the data collection tools section.

To prevent elements that threaten internal validity; bias was prevented in assigning the participants to the groups, the sample size was kept large considering the possibility of data loss, sufficient time was left between the pretest and the posttest to prevent the pretest effect, and the implementation was made by a teacher to prevent the researcher from affecting the practice. Reliability analysis was performed for BYT and Cronbach's alpha reliability coefficient was calculated.

Results

Quantitative Findings from the Scientific Creativity Test

Table 7. Pre-SCT and post-SCT normality test analysis results

Tests	Group	Kolmogorov-Smirnov Statistic	*df*	*p*
Pre-SCT	Experiment	.060	68	.200*
	Control	.145	65	.002
Post-SCT	Experiment	.057	68	.200*
	Control	.069	65	.200*

Table 7 shows the analysis results from the pre-SCT and post-SCT normality tests. The scores can be said to show normal distribution except for the pretest scores from the control group. An evaluation was made for normal distribution by taking into account the criteria mentioned below.

Table 8. SCT pretest-posttest descriptive statistics values

Group	n	Mean	Mode	Median	Skewness	Kurtosis	Min	Max
Pre-Experiment	68	24.220	24	24	-.061	-.689	8	40
Pre-Control	65	22.076	20	21	.705	.892	10	39
Post-Experiment	68	30.147	23	30	.351	-.296	13	55
Post-Control	65	22.646	17	23	.393	-.334	12	39

Table 8 includes the SCT pretest and posttest arithmetic means, modes, medians, skewness, and kurtosis values. When examining the pretest and posttest results separately in terms of the experimental and control groups, the arithmetic means, modes, and medians appear approximately equal to each other. Also, the skewness and kurtosis values were determined to be between +1 and -1. As a result of these values, the SCT pretest and posttest scores are considered to be normally distributed (Field, 2013).

To meet the assumption of sufficient sampling, a study was conducted with 133 participants from among the accessible population of 1,394 (at least 10% of the accessible population) (Pallant, 2016). Because the same sample is valid for the dimensions of fluency, flexibility, and originality, the assumption of sufficient sampling is met in both cases. Levene's test has been used for the other assumption of homogeneity of variances. According to the Levene test results ($t_{(131)} = 5.490$; $p = .004$), the alternative *p-value* ($p = .000$) is used due to the significant difference between the groups' variances (Pallant, 2016).

Table 9. Independent samples t-test results regarding the pre-SCT and post-SCT scores

Tests	Class	*n*	Mean	*SD*	t	*df*	*p*
Pre-SCT	Experiment	68	24.220	7.915	1.827	131	.070
	Control	65	22.076	5.440			
Post-SCT	Experiment	68	30.147	9.234	5.490	131	.000
	Control	65	22.646	6.308			

$p < 0.05$

Table 9 is analyzed separately in terms of the pretest and posttest. Accordingly, no statistically significant difference is found between the Pre-SCT scores for the experimental and control groups ($t_{(131)} = 1.827$; $p = .070$). However, a statistically significant difference is found between the Post-SCT scores of the experimental and control groups ($t_{(131)} = 5.490$; $p = .000$). The average of the scientific creativity scores for the experimental group ($M = 30.147$) on the posttest is significantly higher than the average of the control group ($M = 22.646$).

The observed power is 0.89 and the effect size is .94. The observed power value is greater than the calculated power (0.80). Moreover, the effect size indicates that the learning strategy represents 94% of the variance in the post-test scores. In addition, the effect size was accepted as 0.5 (medium) at the beginning of the study. This value is less than the effect size value (.94) at the end of the study. Therefore, since the observed power is greater than the calculated power and the effect size at the end of the study is greater than the initial effect size, the difference in favor of the experimental group can be generalized to the accessible population and this generalisability is practically significant to provide the external validity (Cohen et al., 2003).

To apply parametric tests in terms of the dimensions of fluency, flexibility, and originality dimensions related to the Post-SCT, the assumptions were checked to see if they were met. Kurtosis and skewness values were examined for normal distribution, and these values were determined to be between +1 and -1. Therefore, the normality assumption for the scores for the sub-dimensions from the Post-SCT has been met.

Table 10. Independent samples t-test results regarding the groups in terms of the sub-dimensions

Tests	Class	*n*	Mean	*SD*	t	*df*	*p*
Post-Fluency	Experiment	68	13.441	4.433	5.439	131	.000
	Control	65	9.908	2.941			
Post-Flexibility	Experiment	68	11.132	3.515	5.104	131	.000
	Control	65	8.508	2.319			
Post-Originality	Experiment	68	9.941	4.121	3.746	131	.000
	Control	65	7.523	3.294			

$p < 0.05$

Levene's test was used for the other assumption of homogeneity of variances. However, the variances between the groups are not homogeneous in all three dimensions, significance (*p*) values in Table 10 are used to indicate the homogeneity of the variances (Pallant, 2016). When examining Table 10, a statistically significant difference was found between the groups in terms of fluency, flexibility, and originality regarding the scientific creativity scores (Fluency: $t_{(131)} = 5.439$, $p = .000$; Flexibility: $t_{(131)} = 5.104$, $p = .000$; Originality: $t_{(131)} = 3.746$, $p = .000$). Therefore, the mean scores for the fluency (13.441), flexibility (11.132), and originality (9.941) sub-scores of the experimental group in the posttest were found to be higher than the mean scores of the control group.

The observed power is 0.89 and the effect size values were calculated as .94 for fluency, .88 for flexibility, and .65 for originality. Therefore, the effect size indicates that the learning strategy represents 94% of the variance in the fluency post-test scores. The learning strategy also represents 88% of flexibility posttest scores and 65% of originality post-test scores. Therefore, since the observed power is greater than the calculated power and the effect size values at the end of the study is greater than the initial effect size, the difference in favor of the experimental group can be generalized to the accessible population and this generalisability is practically significant to provide the external validity (Cohen et al., 2003).

Qualitative Findings of the Scientific Creativity Test

This section involves tables containing the frequency values of participants' expressions and examples of expressions under the codes created for each question in SCT. A limited number of participants' statements are included in the tables, with detailed examinations being made in some cases by including participant expressions not given in the table.

Table 11. Pre-test/post-test codes and fluency scores of the experimental group participants for the 1st question

	Fluency (Frequency)	
Codes	Pre-test	Post-test
Chemistry modeling	18	33
Modeling states of matter	4	2
Atomic structure	15	45
Creating an atom model	-	30
Establishing an experimental setup	6	11
Physical-chemical properties of a ping-pong ball	9	3
Use in physics	16	7
Use in biology	-	1
Use in astronomy	31	20
Chemistry material making	1	2
Atomic calculations	-	4
Historical development of the atom	-	1
Total Fluency Score	100	159

Table 12. Pre-test/post-test flexibility and originality scores of experiment group participants for the 1st question

E5	Pre-test	Post-test
Direct quote	"Modeling of electrons, neutrons, and atoms can be done with the ping-pong ball."	"We can take advantage of the structural, that is, the physical feature of the ping pong ball. With the ping pong ball; We can model many things like atoms, protons, neutrons, electrons, elements, etc. In the laboratory, a ping pong ball is used as a tool to concretize things."
Number of categories	1	3
Categories name	Modeling	Modeling Material making Features of the ping pong ball
Flexibility score	1	3
Originality score	0	3

Question 1 is examined in terms of three sub-dimensions: fluency, flexibility, and originality. Tables 11 and Table12 shows that the post-test fluency sub-scores (159) of the participants in the experimental group were higher than the pre-test (100) because they expressed many ideas under more than one code. Looking at the expression of participant E5 from the experimental group, he presented many ideas embedded under many codes, such as "Modelling, Material making and Features of the ping pong ball". Therefore, in addition to fluency scores, post-test flexibility scores (3) are higher than pre-test (1). Finally, when the situation is examined in terms of originality sub-scores; E5 preferred the expressions that he used less frequently in the post-test instead of the expressions he used very frequently in the pre-test (*In the laboratory, a ping pong ball is used as a tool to concretize things*). This increased the originality score in the post-test.

Table 13. Pre-test/post-test codes and fluency scores of the experimental group participants for the 2nd question

	Fluency (Frequency)	
Codes	Pre-test	Post-test
Dinosaurs	1	4
Daily life in ancient times	33	23
Life in the future	15	8
Technology in the future	7	8
Past discoveries	23	23
Future discoveries	8	13
The beginning of humanity and the world	2	4
Health in the past	5	9
Health in the future	8	7
Structure of the atom	-	15
Scientists	-	31
Egyptian pyramids	5	4
History of the Turks	4	2
Ottoman period	4	2
The beginning of Islam	2	1
Total Fluency Score	117	154

Table 14. Pre-test/post-test flexibility and originality scores of experiment group participants for the 2nd question

E8	Pre-test	Post-test
Direct quote	"How did they communicate and talk in ancient times? Will there be teleportation in the future?"	"How was the atom found? What's the last thing we'll learn about Atom? In addition, I would like to go back to Mendeleev's time and ask how he imagined the location of elements that are not in the periodic table."
Number of categories	2	3
Categories name	Past life Scientists and scientific studies	Discovery Past life Scientists and scientific studies
Flexibility score	2	3
Originality score	1	3

Table 15. Pre-test/post-test codes and fluency scores of the experimental group participants for the 3rd question

	Frequency	
Codes	Pre-test	Post-test
Changing elements' location	8	13
Examples of elements	4	7
Changing the names of the elements	7	11
Braille for the visually impaired	5	13
With sound and light	11	18
Electronic	3	4
Tactile feature	-	3
Coded	-	3
Smart device application	-	1
Interactive	-	4
Usage in different places	3	1
Different size and dimension	4	12
Different shape	6	4
Secret alphabet	-	1
Fragrant	-	1
Game (puzzle, riddle)	-	5
Instructive	13	6
Humorous	-	3
Different coloring	11	9
Enriched with visuals	14	12
Total Fluency Score	89	131

Table 16. Pre-test/post-test flexibility and originality scores of experiment group participants for the 3rd question

E9	Pre-test	Post-test
Direct quote	"I used to make embossed pictures with special functions for the visually impaired."	"I used to emboss for the visually impaired. I'd do 3D. I'd write in the secret alphabet."
Number of categories	1	3
Categories name	For the visually impaired	For the visually impaired Different size and size Hidden alphabet
Flexibility score	1	3
Originality score	1	4

Table 17. Pre-test/post-test codes and fluency scores of the experimental group participants for the 4th question

	Fluency (Frequency)	
Codes	Pre-test	Post-test
Changes regarding Earth and planets	15	13
The balance would be disrupted	-	4
The structure of matter would deteriorate	2	6
Technology would not improve	-	9
Different studies and inventions could not be made	2	8
Vital activities get harder	-	13
Item variety decreases	-	1
There would be no atom and subatomic particles	2	15
Electron related change	-	11
There would be no gravitational force	6	10
There would be no magnetic field	5	7
Elements, compounds, and mixtures would not have been formed	5	5
Integrity/objects would not be formed	-	3
There would be no chemical reactions	5	5
There would be no electricity/electrification	-	5
Magnetism would not have formed	10	8
Machines would not be used / would not work	16	13
Bonds would not be formed	7	11
There is no change of state	-	1
Nothing would have happened	-	1
Total Fluency Score	75	149

Table 18. Pre-test/post-test flexibility and originality scores of experiment group participants for the 4th question

E31	Pre-test	Post-test
Direct quote	"... There would be no electricity."	"Since the earth would not have a magnetic field, weather events would not develop. There would be no deviation in the cathode rays. Electrons would be in the nucleus. Protons would revolve around electrons. There would be no Crocks experiment."
Number of categories	1	3
Categories name	... does not occur	... does not occur Earth/ Planet Structure of the atom
Flexibility score	1	3
Originality score	0	5

Table 13 and Table 14 show the post-test scores of fluency, flexibility, and originality for question 2. Participants asked more detailed questions in the post-test. For example, E8 asked a question about Mendeleev's imagination and increased the flexibility score in the post-test. Fluency scores increased in the post-test. Eight different codes were mentioned in the post-test (see Table 15). Similarly, flexibility scores of participants increased such as E9. In addition, the participants presented very detailed and unusual thoughts. Thus, their originality scores increased as well in the post-test (see Table 16). For instance, E9 made rare statements by mentioning 3D features and the secret alphabet in the post-test.

Fluency scores increased in the post-test. For example, the E31's fluency score was 1 in the pre-test and 5 in the post-test. Nine different codes regarding fluency were mentioned in the post-test (see Table 17). Similarly, flexibility scores of participants increased such as E31. The participants mostly associated the question with “atom and subatomic particles”, and they used the names of scientists or the experiments they conducted while explaining their thoughts. For example, E31 put forward the extraordinary view that *“the Crocks experiment could not be done”*. This increased the originality score in the post-test.

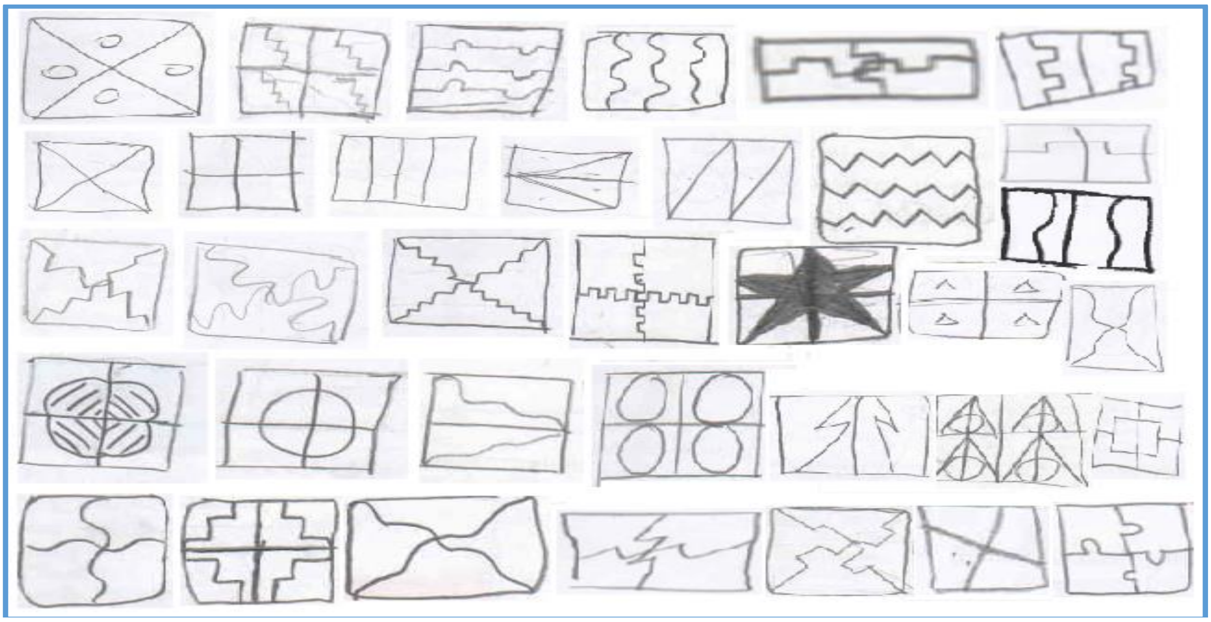

Figure 3. Posttest drawings from the experimental group related to SCT Question 5

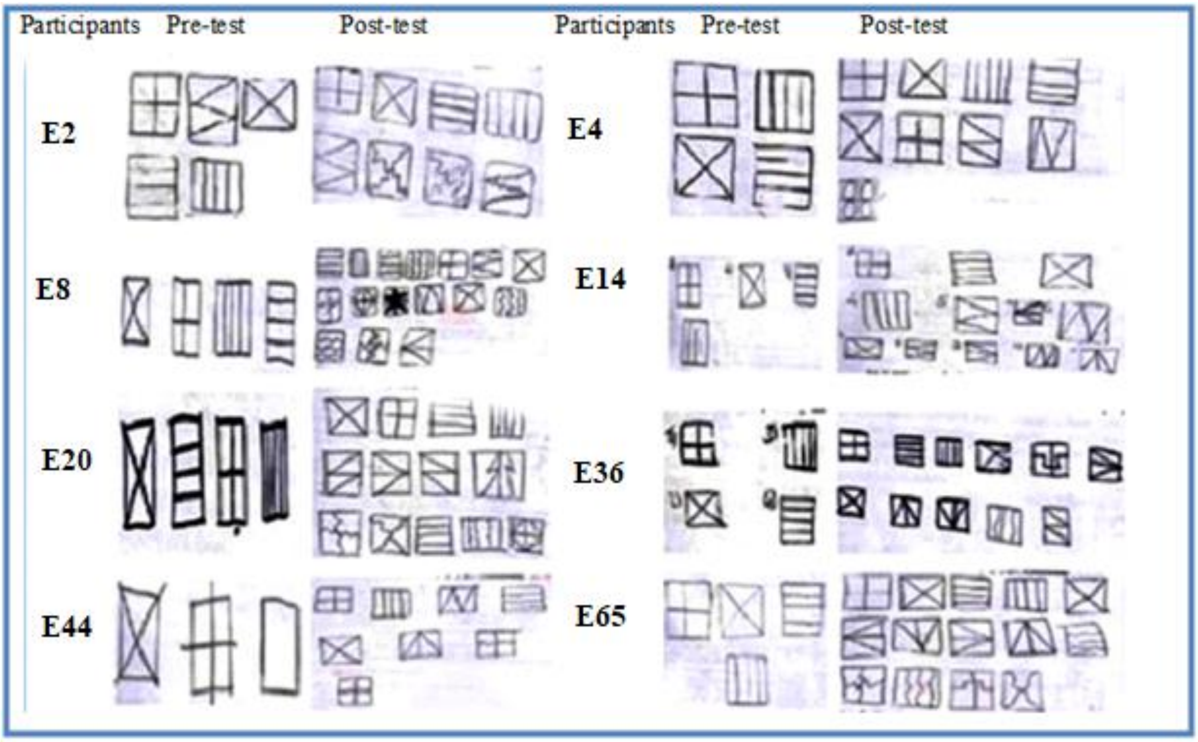

Figure 4. Comparison of SCT experimental group’s pretest and posttest drawings

Question 5 of the SCT asks participants to divide a square into four equal parts to test their problem-solving skills. The participants in the experimental group suggested 36 different methods for a division operation. They offer many different drawing suggestions (see Figure 3). From this point of view, participants suggested multiple (fluency) and uniquely different (originality) drawings. Experimental group participants increased the number of suggested methods (fluency) in the posttest drawings and produced more original drawings.

Table 19. Pre-test/post-test codes and frequency scores of the experimental group participants for the 6th question

	Frequency	
Methods suggest	Pretest	Posttest
Reactive states	8	Melting/heating
Melting/heating	7	2
Can be attracted by a magnet	16	12
Don't look at electrical conductivity	21	17
Electron donation tendency	-	7
Its place on the periodic table	2	25
Mass	-	2
Reactivity with acids	-	8
Reactivity with water	-	4
Luminosity	7	2
Machinability	6	4
Physical state	-	1
Reaction rate	-	14
According to the product at the end of the reaction	-	2
Compounding	-	3
Durability	4	-
Other	4	2
Total	75	115

Table 20. Pre-test/post-test method (fluency + flexibility) and originality scores of experiment group participants for the 6th question

E38	Pretest	Posttest
Direct quote	“I'll try to shoot it with a magnet. Whichever one holds it well is good. I'd electrocuted.”	“I would experiment with the magnet. The metallic character increases as you go down the periodic table. I would put them both in a liquid where the metal would react and look at their reaction times. It reacts more quickly and is more active.”
Number of methods suggested	1	3
Methods suggested	Being able to be pulled by a magnet.	Being able to be pulled by a magnet. The place of the element on the periodic table. The state of reacting or not.
Method score	4	8
Originality score	0	6

For Question 6, the qualitative findings of the experimental group are presented in Tables 19 and 20. These tables show the pretest/posttest comparisons of the experimental group in terms of the dimensions of scientific creativity. While evaluating 6 questions, the statements of the participants were evaluated in terms of method score and originality sub-score. Unlike the pretest, it is seen that the participants conduct experiments to test metals for reactivity. In addition, in the posttest, different from the pretest, the participants presented different statements under codes such as "according to the final product of the reaction, reaction rate, and compound formation". When the statement of E38 is examined, it is seen that the participant suggested many methods. This situation caused the participant's method score in the pretest (4) to increase in the posttest (8). In addition, since the expressions such as *"looking at the reaction rate"* among the suggested methods are rare, the originality score of the participant also increased in the posttest (6).

In the post-test for this question, it was observed that some participants drew detailed figures regarding the method suggestions. The drawings were accepted as explaining the proposed method and this caused their method scores to increase. Quotations of the experimental group participants' methods of comparing metals are shown in Figure 5 below.

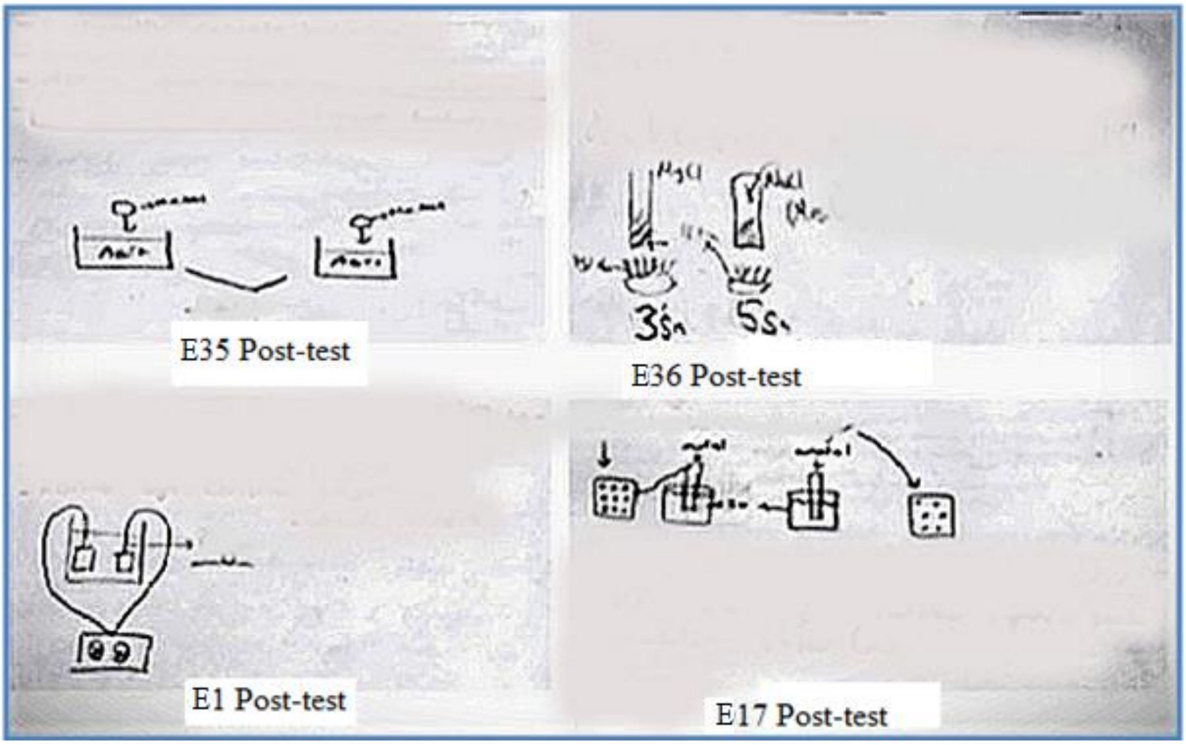

Figure 5. Drawings from some of the experimental group participants for Question 6

According to Figure 5 when looking at the experimental group participants' expressions, they were observed to make use of the properties of metals, explain the method they defend with detailed expressions and support their thoughts with drawings to compare the reactivity's of metals. This situation supports the increase in their method score compared to the control group.

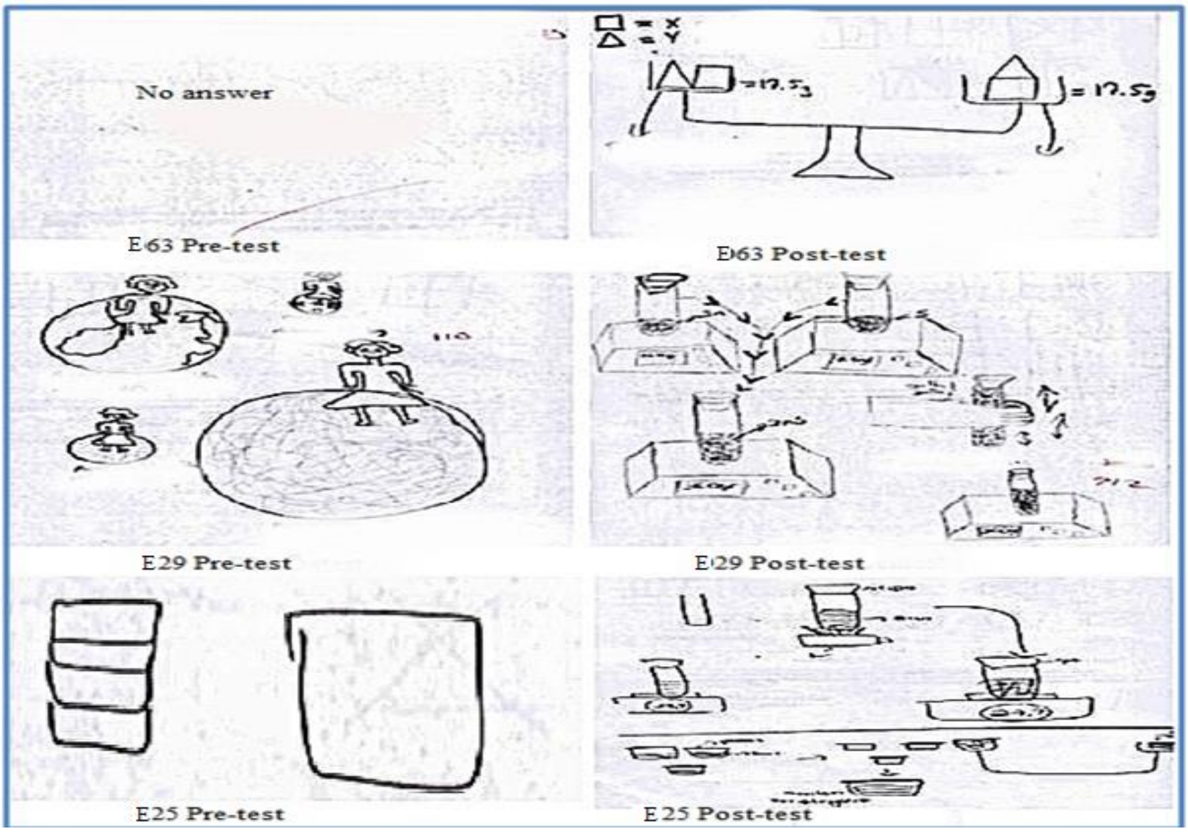

Figure 6. Question 7: pretest and posttest drawings from the experimental group participants.

When examining the sample drawings in Figure 6, all participants in the pretest are seen to have generally made drawings relating the conservation of mass to making measurements on the earth and the moon. However, when examining Figure 7 in terms of the experimental group participants' drawings in the posttest, they were found to have tried to show the conservation of mass using different experimental setups; they were observed to have symbolized elements and compounds at the micro-level and taken into account gas releases that may occur as a result of the reaction. When examining the experimental setups from the experimental group participants in Figure 7, the method / experimental setup scores are seen to be high due to meeting the previously determined criteria (criteria: *Particle-sized display, expressing through numbers, establishing an experimental setup, and Giving examples of elements and using compounds*).

Discussion and Conclusions

The research shows that STEM-based learning activities have a significant effect on the scientific creativity skills of the experimental group. A significant difference is found between the groups in terms of scientific creativity scores in favor of the experimental group. When examining the relevant studies in the literature, many studies are seen on attempts to develop participants' scientific creativity (e.g. Özkök, 2005; Siew, et al., 2015; Walsh, Anders, & Hancock, 2013). Among these studies, Özkök (2005) stated the use of technology and the interdisciplinary approach to increase students' scientific creativity. This situation supports the results from this study, which uses technology and attempts to present an interdisciplinary perspective. However, when examining the studies in the literature, studies are seen to have been conducted on the effect of STEM education on scientific creativity (Barry & Kanematsu, 2006, 2007; Henriksen, 2014; Kanematsu & Barry, 2016; Walsh, Anders, & Hancock, 2013). When considering the results from Walsh, Anders, and Hancock's (2013) study, they stated creativity to be able to be developed through education and the environment factor to be important in developing creativity. In addition, they emphasized that creative learning environments developed about STEM will increase students' creativity. This study has attempted to present students with STEM-based activities developed as well as rich and creative learning environments to improve students' creativity. In this way, the students thought differently and used their imaginations. Similarly, the results from Henriksen's (2014) study, which associated STEM and creativity, stated STEM disciplines to be related to creativity, and the creativity of successful people in these fields to also be high. The results from the current study support the results from Henriksen's (2014) study.

In conclusion, to sum up, in this study, students sought many different solutions to the problems they encountered in their daily lives to improve their scientific creativity. In addition, rich learning environments were created with STEM-based activities and students could easily express their thoughts and imaginations. In addition, students were allowed to demonstrate their productivity. In other words, STEM-based activities led students to generate many ideas (fluency), to look at the same problem situation from different perspectives (flexibility), and to produce original solutions (originality).

This study has shown students' scientific creativity to have improved in the sub-dimensions of fluency, flexibility, and originality. In the literature, Barry and Kanematsu's (2007) study on scientific creativity and STEM stated that students' creativity would improve through STEM activities. This study involved many STEM activities (e.g., STEM activities that enable students to design with different materials enabled them to develop their creativity skills). Through these activities, students were able to generate a large number of ideas (fluency) and design products by revealing different perspectives (flexibility and originality). The participants' posttest statements given in the findings section also support all three sub-dimensions being developed.

According to Hu and Adey (2002), fields such as science, mathematics, engineering, and technology require the use of imagination by their nature and provide rich thinking opportunities, enable the development of scientific creativity, and form creative products. According to Hu and Adey (2002), scientific creativity is defined as the thinking skill that enables original ideas to be produced with an interdisciplinary perspective. Based on the scientific creativity model developed by Hu and Adey (2002), this study developed students' scientific creativity using STEM-based activities, an approach that uses different disciplines such as science, mathematics, engineering, and technology together. In this way, the STEM approach has been suggested as a way for students to generate original ideas.

Recommendations

In this study, students' scientific creativity skills were developed for fluency, flexibility, and originality using STEM activities. For this reason, STEM activities that enable students to generate ideas, view events from different angles and make original designs should be planned, and such activities should be frequently included in chemistry lessons. In addition, STEM activities should be supported with open-ended questions.

Limited STEM activities are found to have been developed due to the subject being studied. Again, due to the nature of a subject's contents, activities for design creation (e.g., periodic table design) are not in the form of a design from scratch but are limited to different additions. Therefore, creating STEM activities for different subjects that allow students to design from scratch is recommended.

Scientific Ethics Declaration

We, the authors, declare that the scientific ethical and legal responsibility of this article published in the JESEH journal belongs to the authors.

Acknowledgments

This study is produced from the doctoral dissertation of the first author in the advisory of the second author.

References

Alexander, P. A. (1992). Domain knowledge: Evolving themes and emerging concerns. *Educational Psychology, 27*, 33-51.

Atasoy, B., Kadayıfçı, H., & Akkuş, H. (2007). Öğrencilerin çizimlerinden ve açıklamalarından yaratıcı düşüncelerinin ortaya konulması (Çizimler ve açıklamalar yoluyla yaratıcı düşünceler). *Journal of Turkish Educational Sciences, 5*(4), 679-700.

Azizoğlu, N., Aslan, S., & Pekcan, S. (2015). The periodic system and teaching with analogies model: the effects of teaching method, gender, and motivation on students' achievement. *Elementary Education Online*, *14*(2), 472-488. http://ilkogretim-online.org.tr http://dx.doi.org/10.17051/io.2015.39450

Barry, D. M., & Kanematsu, H. (2006). The international program to promote creative thinking in chemistry and science. *The Chemist*, *83*(2), 10-14.

Barry, D. M., Kanematsu, H., Lawson, M., Nakahira, K., & Ogawa, N. (2017). Virtual STEM activity for renewable energy. *Procedia Computer Science*, *112*, 946-955. http://dx.doi.org/10.1016/j.procs.2017.08.130.

Boden, M.A. (1994). What is creativity? In M.A. Boden (Ed.), *Dimensions of creativity* (pp. 75-117).

Bybee, R. W., & Landes, N. M. (1990). Science for life & living: An elementary school science program from biological sciences curriculum study. *The American Biology Teacher*, *52*(2), 92-98.

Chen, Q. L., Xu, T., Yang, W. J., Li, Y. D., Sun, J. Z., Wang, K. C., ... & Qiu, J. (2015). Individual differences in verbal creative thinking are reflected in the precuneus. *Neuropsychologia*, *75*, 441-449.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum Associates.

Cole, D., & Espinoza, A. (2008). Examining the academic success of Latino students in science technology engineering and mathematics (STEM) majors. *Journal of College Student Development*, *49*(4), 285-300.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, *13*(1), 3-21.

Creswell, J. W., & Clark, V. L. P. (2015). *Designing and conducting mixed methods research*. Sage publications.

Daud, A. M., Omar, J., Turiman, P., & Osman, K. (2012). Creativity in science education. *Procedia-Social and Behavioral Sciences*, *59*, 467-474. http://dx.doi.org/10.1016/j.sbspro.2012.09.302.

Deniş Çeliker, H., & Balım, A. (2012). Adaptation of scientific creativity test to Turkish and its assessment criteria *Usak University Journal of Social Sciences, 5*(2), 1-21.

Dugger, W. E. (2010). Evolution of STEM in the United States (Paper) Presented at the 6th *Biennial International Conference on Technology Education Research* on Dec 8-11, 2010 in Australia.

English, L. D. (2017). Advancing elementary and middle school STEM education. *International Journal of Science and Mathematics Education*, *15*(1), 5-24.

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics, 4th edition,* Washington, DC: Sage.

Fosnot, C. T. (2013). *Constructivism: Theory, perspectives, and practice*. Teachers College Press.
George, D., & Mallery, P. (2001). *SPSS for Windows Step by Step, 3rd edition,* Boston, MA, USA: Allyn and Bacon.
Glück, J., Ernst, R., & Unger, F. (2002). How creatives define creativity: Definitions reflect different types of creativity. *Communication Research Journal*, *14*(1), 55-67.
Guilford, J. P. (1988). Some changes in the structure-of-intellect model. *Educational and Psychological Measurement*, *48*(1), 1-4.
Hadzigeorgiou, Y., Fokialis, P., & Kabouropoulou, M. (2012). Thinking about creativity in science education. *Creative Education*, *3*(5), 603-611. http://dx.doi.org/10.4236/ce.2012.3508.
Henriksen, D. (2014). Full STEAM ahead: Creativity in excellent STEM teaching practices. *The STEAM Journal*, *1*(2), 1-7.
Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Correlation: A measure of a relationship. In Applied statistics for the behavioral sciences* (5th ed.). Houghton Mifflin Company.
Hu, W., & Adey, P. (2002). *A* scientific creativity test for secondary school students. *International Journal of Science Education*, *24*(4), 389-403.
Hu, W., Wu, B., Jia, X., Yi, X., Duan, C., Meyer, W., & Kaufman, J. C. (2013). Increasing students' scientific creativity: The “learn to think” intervention program. *The Journal of Creative Behavior*, *47*(1), 3-21.
Kanematsu, H., & Barry, D. M. (2016). STEM and creativity. *In STEM and ICT Education in Intelligent Environments* (pp. 15-23). Springer, Cham.
Kang, D., Park, J. &Hong, H. (2015). Changes in the number of ideas depend on the time when conducting scientific creativity activities. *Journal of Baltic Science Education, 14* (4), 448-459.
Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM education*, *3*(1), 1-11. http://dx.doi.org/10.1186/s40594-016-0046-z.
Lewis, T. (2005). Creativity is a framework for the design/problem-solving discourse in technology education. *Journal of Technology Education*, *17*(1), 35-52.
Lin, C., Hu, W., Adey, P., & Shen, J. (2003). The influence of CASE on scientific creativity. *Research in Science Education*, *33*(2), 143-162.
Marshall, C. & Rossman, G. B. (2006). *Designing qualitative research* (4th Edition). USA: Sage Publications.
Martins, E.C., & Terblanche, F. (2003) “Building organizational culture that stimulates creativity and innovation”, *European Journal of Innovation Management*, *6*(1), 64-74.
Merriam, S.B. (2013). *Qualitative research: A guide to design and implementation. (S. Turan, Trans). Nobel Academic Publishing (Original work published 2009).*
Moravcsik, M. J. (1981). Creativity in science education. *Science Education*, *65*(2), 221-27.
Nuangchalerm, P. (2018). STEM education and Kolb’s learning styles. *Journal of Community Development Research (Humanities and Social Sciences)*, *11*(2), 11-17.
Ostler, E. (2012). 21st century STEM education: A tactical model for long-range success. *International Journal of Applied Science and Technology*, *2*(1), 28-33.
Özkök, A. (2005). Effects of interdisciplinary creative problem-solving teaching program on creative problem-solving skills. *Hacettepe University Journal of Education*, *28*, 159-167.
Pallant, J. (2016). *SPSS kullanma klavuzu spss ile adım adım veri analizi*. (S. Balcı ve B. Ahi, Trans). Anı Yayıncılık (Original work published 2013).
Selby, E. C., Shaw, E. J., & Houtz, J. C. (2005). The creative personality. *Gifted Child Quarterly*, *49*(4), 300-314.
Semmler, L., & Pietzner, V. (2017). Creativity in chemistry class and in general–German student teachers’ views. *Chemistry Education Research and Practice*, *18*(2), 310-328. http://dx.doi.org/10.1039/c6rp00230g.
Siew, N. M., Chong, C. L., & Lee, B. N. (2015). Fostering fifth graders’ scientific creativity through problem-based learning. *Journal of Baltic Science Education*, *14*(5), 655-669. https://doi.org/10.33225/jbse/15.14.655.
Tabachnick, B. G., & Fidell, L. S. (2013). Using multivariate statistics (6theds). Pearson.
Thompson, S. K. (1990). Adaptive cluster sampling. *Journal of the American Statistical Association*, *85*(412), 1050-1059.
Torrance, E. P. (1984). The role of creativity in the identification of the gifted and talented. *Gifted Child Quarterly*, *28*(4), 153-156.
Torrance, E.P. (1990). *Torrance test of creative talent.* Beaconville, IL: Scholastic Testing Service.
Urban, K. K. (2005). Assessing creativity: The test for creative thinking-drawing production (TCT-DP). *International Education Journal*, *6*(2), 272-280.
Van Manen, M. (2007). Phenomenology of practice. *Phenomenology & Practice*, *1*(1), 11-30.

Van Soom, C., & Donche, V. (2014). Profiling first-year students in STEM programs based on autonomous motivation and academic self-concept and relationship with academic achievement. *PloS one*, *9*(11), 1-13. https://doi.org/10.1371/journal.pone.0112489.

Walsh, E., Anders, K., & Hancock, S. (2013). Understanding, attitude, and environment: The essentials for developing creativity in STEM researchers. *International Journal for Researcher Development*, *4*(1), 19-38.

White, D. W. (2014). What is STEM education and why is it important. *Florida Association of Teacher Educators Journal*, *1*(14), 1-9.

Author(s) Information

Seyide Eroglu
Nuh Mehmet Baldoktu Anatolian High School
Kayseri, Turkey
Contact e-mail: *seyideeroglu@gmail.com*
ORCID ID: *0000-0002-7363-6638*

Oktay Bektas
Erciyes University
Department of Science Education, Kayseri, Turkey
ORCID ID: *0000-0002-2562-2864*

Journal of Education in Science, Environment and Health
Volume 8, Issue 1, 2022 https://doi.org/10.21891/jeseh.1056757

Design Considerations of Online Infertility Prevention Training (OIPT): Development and Evaluation

Bahar Baran, Sirin Nur Yaci, Murat Ocal, Digdem Muge Siyez

Article Info

Article History

Published:
01 January 2022

Received:
20 May 2021

Accepted:
20 November 2021

Keywords

Usability
Design considerations
Interactive environments
Infertility

Abstract

Health-related internet sources should contain accurate, reliable and updated information, and their usability should be high to enable people to navigate successfully. This study aimed to develop interactive Online Infertility Prevention Training (OIPT) and to reveal design principles by evaluating the online training. The development, design, and redesign processes of this study included four iterations: 1) needs analysis, 2) implementation as a usability test, 3) implementation with the target group in a class, and 4) comparison of iteration II and III, and decision; and two outputs. The research process included the participation of the target audience in the design process and there were continuous corrections to the design. In the study, a design-based research approach was used. In the qualitative dimension, the data was collected with "User Information Form", "System Usability Observation Form", and the system usability interview form, and task cards. In the quantitative dimension, Infertility Knowledge Test, System Usability Scale, and also database records were used. The data obtained from tools were analyzed descriptively and predictively. As a result, this research claimed that design, development, and evaluation belong to health-related sources on the Internet should include field design principles and techniques and specific methods, although designers use prevalent techniques. The results indicated that gender's effect on media and material design, the importance of the use of expert authority in the presentation of content knowledge that system usability significantly increased from iteration two to three. Therefore, gender differences in media and material preferences except for content knowledge and also expert views may be considered in the design of health-related online training. In addition, while teaching concrete parts of the human body, interactive animations and visuals may be useful to motivate and increase engagement students to use the content again and again. Furthermore, the results of this study show that feedback design is crucial in the navigation of online learning environments. Overall, this study strengthens the idea that while designing online education environments, it is important to determine users' needs and expectations at first, afterwards conducting usability tests to specify the design challenges.

Introduction

A significant part of the information searched on the Internet is related to health information (Harvey et al., 2017). Prior research (TUIK, 2018) showed that the ratio of households who have connected to the Internet in the last three months and searched for some issues about health is 68.8%. Several studies showed that patients used the Internet to search for health-related information (Harvey et al., 2017; Trotter & Morgan, 2008). Thus, people can learn health-related information not only from doctors but also from the Internet, which is a relatively recent development. For example, people might consider the symptoms they experience and search the combination of symptoms on the Internet to obtain information, consulting a doctor afterwards if their search yielded a result that may be of concern. Indeed, it seems that accessing information on the Internet is very advantageous on the one hand due to the rapid accessibility of information, however, if the information is inaccurate, this could be harmful. In this respect, it is very important to create reliable information sources about health and to work on accuracy.

The Internet Sources about Infertility

Another advantage of easily accessing health-related information on the Internet is the opportunity to obtain information about issues that people might feel uncomfortable asking their friends, families, and even health professionals. Because of social labelling, infertility, is one of the health-related issues that individuals are often reluctant to ask questions about and express themselves to others. Infertility is defined as the failure to produce a clinical pregnancy after 12 months or more of regular unprotected sexual intercourse (Zegers-Hochschild et al., 2009). In a study conducted with 549 infertile individuals, 87.9% of the participants (93.7% of female, 80% of male) searched for information about infertility on the Internet (Zelkowitz et al., 2016). A different study showed that the Internet is an important source of information for patients receiving infertility treatment (Zillien et al., 2011). However, it is known that information about health on the Internet is insufficient. For example, a study evaluated 107 websites according to the criteria of credibility, accuracy and ease of navigation showed all websites received low scores on the criteria, especially the accuracy criterion (Marriott et al., 2008).

Today, the Internet has become a key part of the education system and online training is now being effectively used due to the advantages they provide. Studies showed that the Internet can also be used as an alternative information source for sexual health education (Borzekowski et al., 2006; von Rosen et al., 2017). For example, it can serve as a mediator to obtain sexual health education that provides the important benefit of increasing the knowledge of disadvantaged groups that are difficult to reach using traditional sources (von Rosen et al., 2017). Sexual health education programs are available through face-to-face training (Chi et al., 2015; Thato & Penrose, 2013), leaflets (Wojcieszek & Thompson, 2013), online training (Deng et al., 2017) or video-based training (Conceição et al., 2017). However, infertility-related subject matter is limited in the content of these programs. On the other hand, some of the causes of infertility (e.g., smoking, alcohol use, sexually transmitted diseases, advanced age, caffeine consumption, being overweight or weak, and stress) can be prevented by participating in a training which may be effective at reducing these risk factors. Epstein et al. (2002) stated that when used correctly, the Internet can be used as an effective source to prevent traumatizing emotions related to infertility. In view of all that has been mentioned so far, the Internet may be an effective method to deliver an infertility prevention program, which was the goal of the current investigation. In addition, designing of this online education environment is a significant aspect of this investigation. The following part will address the user-based design of the current study as the base.

User-based Design in Online Education

The success and completion of an online education is closely related to the characteristics and goals of learners (Williams et al., 2018). Recent studies indicated that although enrollment rates for online courses are high, the finishing rates are low; researchers have been investigating reasons for low completion rates (El Said, 2017; Rieber, 2017). Reasons for not finishing an online course include pedagogical distance and design problems between the student and the content. A student's perception, values, judgments, and personality structure (Travers et al., 1993) affect the motivation of using the learning environment. To reduce the pedagogical distance between students and the environment, it is important that e-learning environments are designed according to the characteristics of the students, including involving students in the design processes. Therefore, in the design of online environments, user experiences should be collected before, during and after the design, and formative and summative evaluations should be made. Designing these environments based on student characteristics can increase student satisfaction and enable them to actively participate in their learning processes, ultimately affecting their success (Çağıltay, 2011; Gülbahar et al., 2008). After having addressed the importance of user-based design in online education, the next section will be discussed design-based research in multimedia learning environments, importance of usability tests, and cognitive theory of learning.

Design and Development in Multimedia Based Online Education

In the current study, a design-based research method was used to consider student knowledge and multimedia preferences in the design of an “Online Infertility Prevention Training” website. To be innovative in the design of this learning environment it was necessary to use a research method that included the target user on the design team, and constantly updated the design. Research studies that design, develop, and redesign innovative learning environments are termed Design-Based Research (Barab & Squire, 2004; Zheng, 2015). Design-based research can be characterized in terms of the designed artifact and resultant theory (Barab & Squire, 2004). This research involves a series of development and correction processes on the basis of design-based research, such as the development of the environment, analyzing design problems with usability tests, redesign, implementation with target users and evaluation. The International Organization for Standardization (ISO 9241-11, 1998) defines usability as a specific group of users to perform certain tasks effectively, efficiently and with satisfaction (Çağıltay, 2011). Adding the usability tests to design cases helps to ensure seeing the interaction between the user and design directly.

Problems related to usability in multimedia learning environments can lead students to move away from their learning objectives (Virvou & Katsionis, 2008). This may be explained by the cognitive theory of multimedia learning (Mayer, 2009), based on dual coding theory (Paivio, 2001), cognitive-load theory (Sweller, 2005) and limited capacity theory (Baddeley, 1999). This theory states that during learning with multimedia, a student's cognitive processing capacity consists of 1) extraneous processing, 2) essential processing and 3) generative processing. Extraneous processing is the cognitive process that occurs during learning that does not serve instructional purposes and is caused by design problems. Essential processing is the cognitive process related to the content that needs to be learned and occurs when content becomes complex. Generative processing is cognitive processing that aims to make sense of the presented material and occurs as a result of focused learning (Mayer, 2009). Designers need to pay particular attention to the control of extraneous cognitive processing among them (Davids et al., 2015). By using the cognitive theory of learning, designers try to prevent users' from becoming lost during navigation in e-learning or multimedia learning environments (Altun, 2000). Therefore, the involvement of target users in the design process of online education environments is important in order to improve usability. Content is another important aspect of multimedia learning environments. When we examined the literature, there is no comprehensive online education program for university students aimed at informing about infertility and sexual health in Turkey. Generally, ads of hospitals or doctors are prevalent as online education resources. This finding has informed us in shaping the subject content in this research.

The Aim of the Study

The current study aims to investigate the design, development, and redesign processes of "Online Infertility Prevention Training" (OIPT) website. These processes are 1) Need analysis, 2) Design and Development of "OIPT", 3) Testing usability, 4) Redesign, 5) Implementation in a class, and 6) Comparison of Iteration II and III, and Decision (Figure 1). We steadily improved the design interventions as a part of the iterative design process to make them more practicable and enhance the knowledge.

Based on these iterations, the current study aimed to reveal answers to the following research questions:

- What prior knowledge do university students have about infertility and what are their preferences related to multimedia learning? (Iteration I)
- How usable is the online environment (OIPT) as a system in the usability test? (Iteration II)
- How usable is the online environment as a system for a real implementation in a class? (Iteration III)
- Is there any significant increase in system usability between the usability test and real implementation in a class? (Iteration IV)

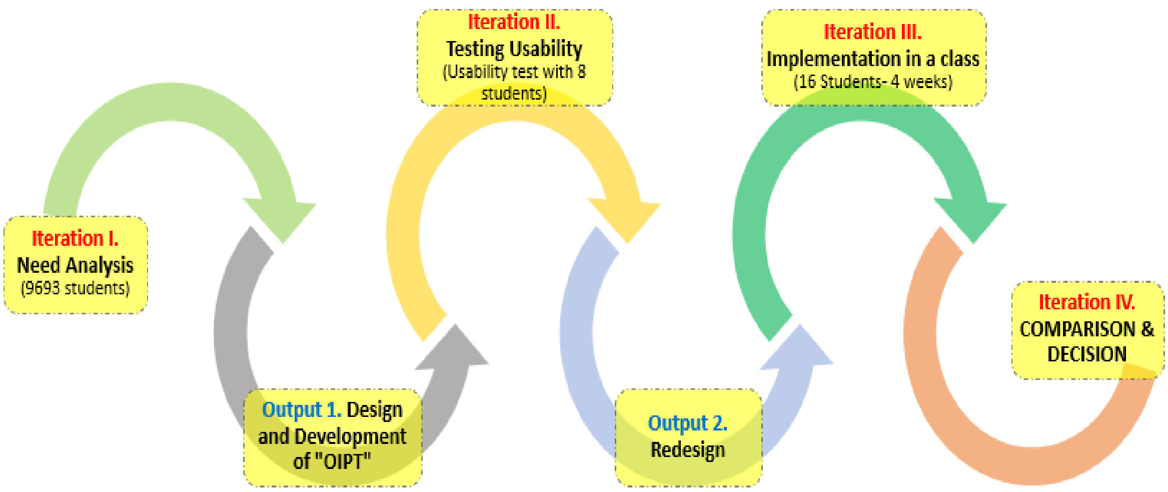

Figure 1. Four research iterations and their outputs

In the following sections, the method of study (research design, sample, data collection tools etc.), iterations and findings of each of them are indicated.

Method

Research design

This study was designed using the design-based research approach, which consists of an iterative design and evaluation process. This approach can help designers by generating, testing, and refining design and design concepts iteratively to move the research further. Multimedia design and usability are complex phenomena (Alwashmi et al., 2019). So, researchers advocate for continuous and iterative testing in order to respond to user needs, preferences, technical challenges, and faults (Adam et al., 2019; Hattink et al., 2016). More specifically, it is also important to take into consideration to target users of online environments while designing these environments to enhance learning and engagement. In this study, as a part of the iterative design process, there were continuous corrections. In the next sections, each iteration of the study, methods, challenges, and results (findings) of it, and outputs is explained in detail.

Sample

This study has consisted of four research iterations and their outputs as mentioned before. Therefore, the sample consists of different participants. Iteration I phase has consisted of 9693 undergraduate students (5002 females, 4691 male) from 12 regions in Turkey in the 2016-2017 academic year. To determine regions Statistical Region Units Classification-1 criteria were used. Iteration II phase has consisted of 8 undergraduate students from the Department of Computer Education and Instructional Technologies (n=4) and the Department of Guidance and Psychological Counseling (n=4) who were selected by convenience sampling in Dokuz Eylül University in the 2017-2018 academic year. Lastly, Iteration III has consisted of 16 undergraduate students who were studying in the Department of Guidance and Psychological Counseling at Dokuz Eylül University in the 2017-2018 academic year. These students selected an elective course titled “Sexual Health Knowledge” voluntarily and implementation was carried out within the course.

Data Collection Tools

In the qualitative dimension, the data was collected with “User Information Form”, “System Usability Observation Form”, and the system usability interview form, and task cards developed by the researcher. In the quantitative dimension, System Usability Scale (SUS), was developed by Bangor et al. (2008) and adapted to Turkish by Çağıltay (2011), Infertility Knowledge Test (IKT), developed by Seymenler (2017), and also database records were used. Detailed information on the tools is given in the related iteration steps in the following parts.

Data Analyses

The chi-square test was used to determine university students’ multimedia learning preferences based on their gender (Iteration I). In Iteration II, the usability of the online education environment was investigated by using different data collection methods and tools. The think-aloud method and system usability observation form were applied during the experiment. The system usability interview and system usability scale were used after it to collect quantitative and qualitative data (Iteration III). Quantitative data obtained from the usability test was analyzed descriptively. One of the nonparametric statistical methods, the Mann-Whitney U test, was used to compare the system usability in Iteration IV. In addition, the collected qualitative data were recorded to the MS Word program and transcribed; the data were then analyzed using descriptive and content analysis techniques. Detailed information on how to data was analyzed is given in the related iteration steps in the following parts.

Steps of the Study

Iteration 1. Need Analysis

Research Question 1

What prior knowledge do university students have about infertility and what are their preferences related to multimedia learning?

Method of Iteration 1

The development of an online infertility prevention training (OIPT) was a part of a larger project, Examining University Students' Infertility Knowledge and Attitudes towards Infertility and Developing and Evaluating Infertility Prevention Psycho-Education Program and Online Education Program, which was supported by The Scientific and Technological Research Council of Turkey (TUBITAK). In the context of the project, researchers developed Infertility Knowledge Test (IKT) and Attitudes toward Infertility Scale (ATIS) as data collection tools, conducted a detailed needs analysis and then developed a Psycho-Education curriculum and an online education environment in this study. The prior knowledge levels of university students about infertility were available in another published article (Siyez et al., 2018). The multimedia learning preferences part of the need analysis is the issue of this article. The data related to multimedia design preferences of groups was analyzed with chi-square test.

Iteration I. Challenges and Results

In this study, the needs analysis is one of the components of design cases. User needs and expectations are important for the interface design. In this context, to determine popular subjects and prior knowledge about infertility, the needs analysis about infertility knowledge was conducted on 9693 university students. Besides, several challenges had to be addressed during the needs analysis. Before starting the project, we examined the literature to identify how studies conducted about online environments related to infertility subject. But we couldn't find a satisfactory study for university students about it. Studies that are conducted were mostly to advertise hospitals or doctors, and also not planned as a training. Therefore, the first design challenge was the lack of a sample study. The second design challenge is the topic. As we stated in the 'Introduction' section, infertility is a sensitive issue, particularly in societies social rules are dominant (Gungor et al., 2013; Şahin, 2012). For instance, asking questions about it is seen as prejudicial or disturbing.

The results from the published study showed that students had some misconceptions about infertility and sexual behaviors (Siyez et al., 2018). This study was conducted with 9693 undergraduate students (5002 females, 4691 males) from 12 regions in Turkey. In addition, results indicated that more than 90% of the participants would like to learn about subjects related to "Factors affecting infertility", "Sexual health and reproductive health" and "Sexually transmitted diseases" (Baran et al., 2017; Siyez et al., 2018).

The students' multimedia learning preferences were also investigated based on their gender. The chi-square test results (Table 1) indicated that they expected rich content from their online education including "questions and answers about infertility, expert video, animation, different content by sex and different designs according to sex". Interestingly, the percentage of participants that preferred "different contents according to sex" did not differ by gender, χ^2 (1, n = 9693) = 3.69, p < .056).

Table 1. Students' content preferences in the online environment

Multimedia design preferences		Sex				χ^2	p
		Female		Male			
		f	%	f	%		
Questions and answers about infertility	Yes	4598	95.9	4273	94.2	14.61	.000
	No	195	4.1	262	5.8		
Expert (doctor) video	Yes	4491	93.4	4157	91.6	11.70	.001
	No	316	6.6.	383	8.4		
Animation	Yes	4067	84.9	3941	86.9	8.09	.004
	No	724	15.1	592	13.1		
Different contents by sex	Yes	3757	78.3	3623	79.9	3.69	.056
	No	1043	21.7	912	20.1		
Different designs according to sex	Yes	3168	66.1	3292	72.7	48.17	.000
	No	1628	33.9	1237	27.3		

Output 1. Design and Development of Online Infertility Prevention Training (OIPT)

The Online Infertility Prevention Training (OIPT) based on the results from Iteration I was divided into four units. Unit 1 included male and female reproductive systems with organs and menstruation. Unit 2 included infertility, treatment methods and biological and preventable factors (age, body mass index, sex time, nutrition habits). Unit 3 continued with other preventable factors such as sexually transmitted diseases (STD), protection methods from STD. Unit 4 was related with the other preventable factors such as cancer, substance use, stress, chemicals, and radiation (Figure 2).

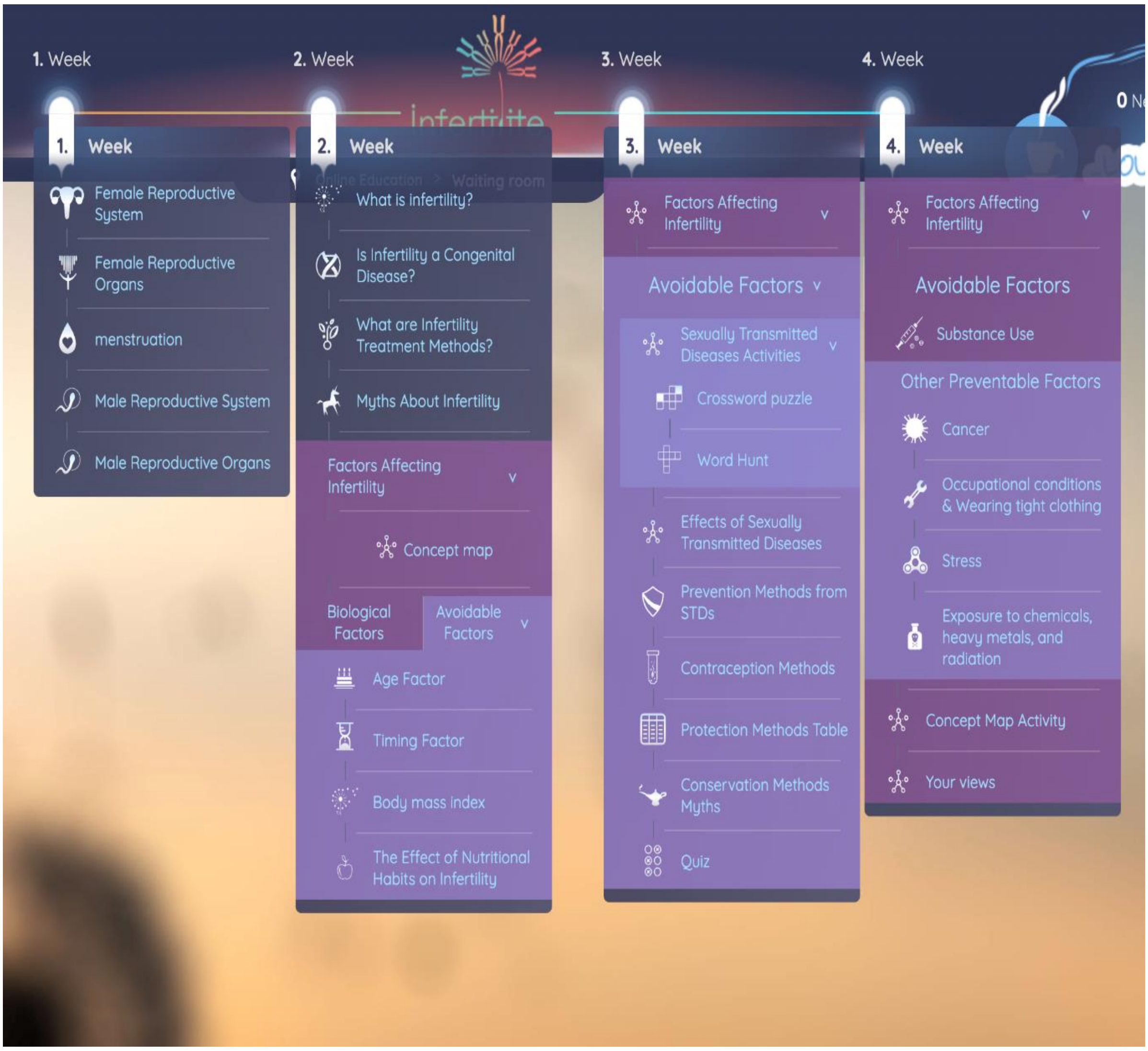

Figure 2. The content of the online infertility prevention training (OIPT)

Male and Female Reproductive System

Two interactive animations of the female and male reproductive systems were used to teach the 1) functions, 2) spelling and 3) places of the reproductive organs (Figure 3). The side and front view of the reproductive systems was also designed. More specifically, if the aim is to teach a specific organ (e.g., Fallopian tubes) in the female reproductive system, this organ blinks on the page to attract the attention of users. The aim was to teach the function of an organ, and the verbal and vocal information (with a speaker symbol) were presented to describe its function to users. To teach the spelling of an organ, a drag and drop game was used. Users had to drag letters to the correct place in the game. Then, they received positive verbal and visual feedback if they succeeded. To teach the locations of an organ in the reproductive systems, another drag and drop game was designed. The users identified the location of an organ in the female reproductive system and received positive verbal and visual feedback.

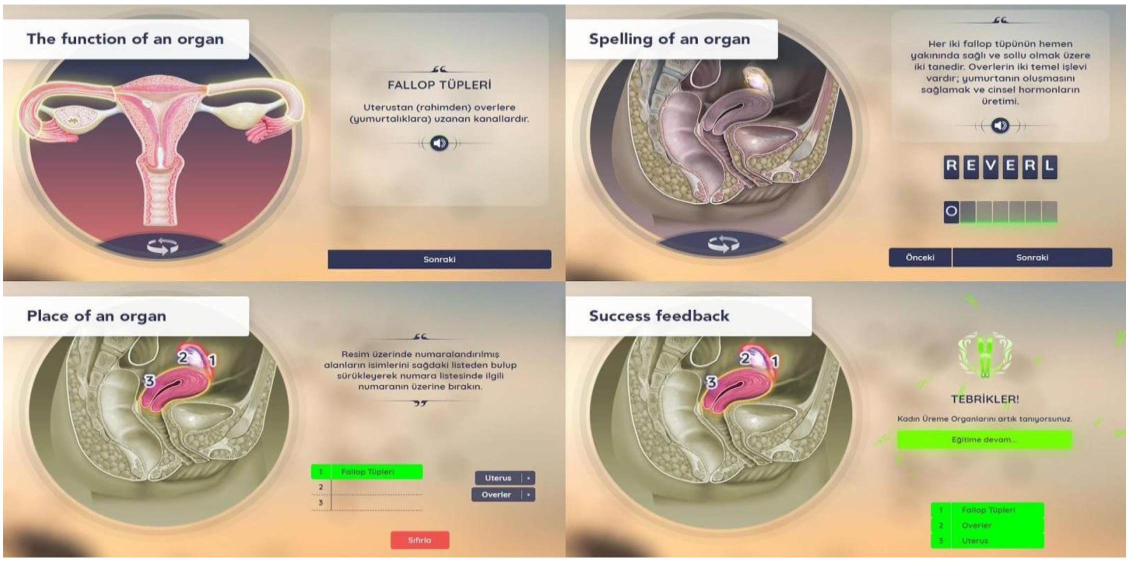

Figure 3: Teaching the female reproductive system: 1) the function of an organ, 2) spelling of an organ, 3) location of an organ, 4) feedback design when a user succeeded

Video

Doctor videos and 3D animation videos showing the menstrual cycle were used for subject narration (Figure 4). The videos included answers from students' misconception about infertility, which we have learned from iteration I (need analysis). Multiple choice questions were posed before and after the video. When a user wanted to skip the video, the skipped video was marked on the video page, and users were not allowed to complete the activity without watching the video.

Body Mass Index tool

This interactive tool was designed to teach the importance of infertility prevention factors such as sex, age, weight, and height (Figure 4). Users received different feedback while the values of risk factors (age, gender, height, weight) changed. For example, the infertility risk rate is higher when a woman's age is 40 than when her age is 34.

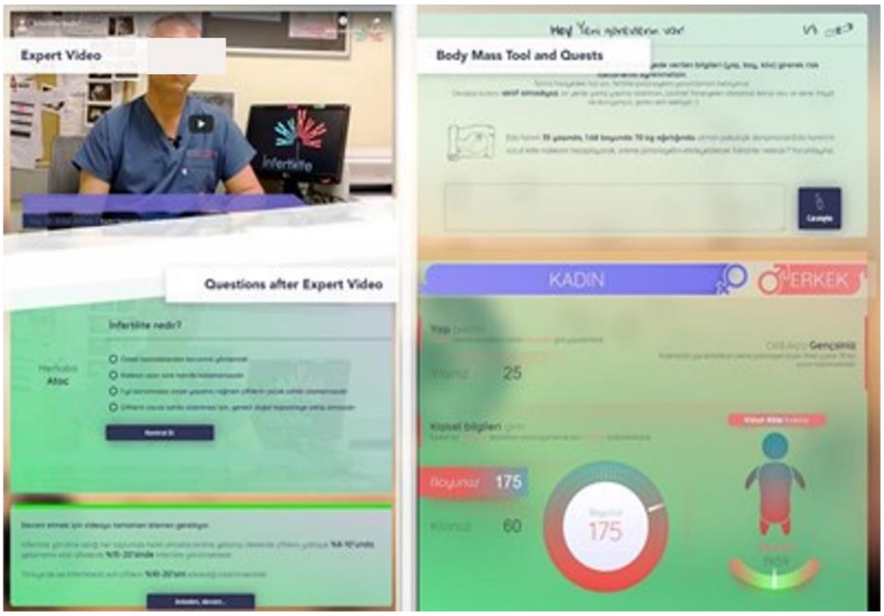

Figure 4: Example of the expert video and a multiple-choice question asked at the end of the video and the body mass tool with questions

Iteration II. Testing Usability

Research Question 2

How usable is the online environment as a system in the usability test?

Method of Iteration II

Eight undergraduate students (six female and two male) from a public university in Turkey volunteered to participate in the experimental study evaluating the usability of the online environment. This study used Nielsen's (2012) recommendation to determine the number of participants. He suggested including at least five users to reveal basic problems related to a system in usability studies. The characteristics of the participants should be defined before a usability study because experienced and inexperienced users show different behaviors (Çağıltay, 2011). This current study defined the participants' personal computer ownership, computer and internet usage experiences, internet usage level, online education experience which may affect the results of system usability with the User Information Form. Seven of the participants have a personal computer; computer usage experience range included 10 years (n =2), 7-9 years (n = 4), 4-6 years (n=1) and 1-3 years (n = 1); internet usage experience range included 10 years (n=1), 7-9 years (n=4), 4-6 years (n=2), 1-3 years (n=1). Three of the participants defined the level of internet use as "sufficient", one of them as "very sufficient", and three of them as "somewhat sufficient". Also, two participants stated that had an online learning experience before.

The usability of the online education environment was investigated by using different data collection methods and tools. Before the experiment, user information form was applied to the participants. The think aloud method, task cards and system usability observation form were applied during the experiment. The system usability interview and system usability scale were used after it.

- *Think Aloud Method:* The aim of this method is to explore the interaction between participants and the system (Boren & Ramey, 2000). The method requires that users vocally express their thoughts while performing their assigned tasks.

- *User Information Form:* This form was used to collect detailed data about participants. With this form, participant number, gender, department, computer ownership, computer experiences of participants, level of using internet etc. data were collected.

- *System Usability Observation Form*: Start/end time, duration of the task, number of errors, completion status of the task, and the need for assistance was recorded.

- *System Usability Interview Form*: Opinions of the participants about the tasks, their experiences, and suggestions about the online education environment were obtained.

- *System Usability Scale (SUS):* The scale was developed by Bangor et al. (2008) and adapted to Turkish by Çağıltay (2011). It uses a 5-point Likert type scale consisting of 10 items. Five items are inversely scored. The points for each user were summed and the total score was multiplied by 2.5 to obtain the scale score for each user. The possible range of scores was 0(lowest)-100(highest). Bailey (2006) stated that if the SUS score was between 65-70, the environment had moderate satisfaction.

- *Database records:* Participants' data related to tasks during the experiment was recorded in the database.

The participants performed seven tasks during the experiment. The tasks included both a simple task such as finding a piece of organ on a visual and a complex task such as realizing a task by using multiple interactive tools. The tasks were completed on a desktop computer in a research lab. A task started with the researchers giving the task card and ended with either the successful or unsuccessful completion of the task.

- *Task 1. Signing in the online education:* The participants were expected to log in to the website with the e-mail address and password given to them. http://infertiliteyionleme.deu.edu.tr/course/.

- *Task 2. Listening to the definition of "Cervix" in the female reproductive system:* The participants were expected to click on the speaker icon next to the "Cervix" (Figure 4).

- *Task 3. Answering the question at the end of the video titled "What are the methods of infertility treatment?":* The participants were expected to watch the doctor video (Figure 4) and answer the multiple-choice question at the end of the video.

- *Task 4. Completing the first mission on the "Body Mass Index".* The participants should click on the text "Hey, you have a new task!" to start the mission. Then, the mission task window opened (Figure 4). Participants were expected to carry out the mission by using the "Body Mass Index" tool.

- *Task 5. Examining "Tubal Ligation" method, one of the birth control methods.*

- *Task 6. Finding and crossing out the word "Syphilis" on the Word Hunt puzzle on the sexually transmitted diseases page* (Figure 5).

- *Task 7. Matching the function of "Penis" with the figure of it:* This drag and drop activity (Figure 3) required correctly matching male reproductive organs with the related numbers on the right side of the image.

In addition, while implementing the usability test, task duration was examined. Although a usability test conducts qualitatively mostly, sometimes it can be really helpful to keep track of design progress and users' behaviors. With this perspective, we investigated the task duration and applied the System Usability Scale to the participants.

Tasks

When database records and observation form results are examined, it was found that tasks 1, 3, 5, and 7 were completed successfully (Table 2). The participants' uncompleted tasks are shown as X1, X2, X3, X4 and X5 in Table 2. Effort duration for the uncompleted tasks was 128 seconds (X1), 168 seconds (X2), 85 seconds (X3), 75 seconds (X4) and 234 seconds (X5).

Table 2: Time distribution of the participants while performing the tasks

	Time of task (second)							
Participant Number	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7	Total
1	40	**X1**	228	**X2**	32	85	28	413
2	37	23	219	**X3**	17	32	29	357
3	41	62	238	254	43	32	12	682
4	26	24	169	100	33	**X4**	9	361
5	70	44	134	280	35	36	16	615
6	50	100	221	**X5**	30	80	30	511
7	49	85	295	135	24	26	41	655
8	21	46	256	105	25	24	25	502
Total time (second)	334	384	1760	874	239	315	190	4096
Mean time (second)	41.75	54.86	220	174.8	29.88	45	23.75	

System Usability Scale

The SUS scores of the first working group are shown in Table 3. The highest score was 87.5 and the lowest score was 35. The mean SUS score was 69.38 (SD = 18), indicating that the online infertility prevention training had moderate satisfaction.

Table 3: SUS scores in Iteration II

K1	K2	K3	K4	K5	K6	K7	K8	SD	X
70	87.5	57.5	85	87.5	62.5	35	70	18	69.38

Output 2. Challenges and Redesign of the online environment

The roles of the design team in this implementation are as follows; 8 participants, 1 web designer, 1 content designer, and 1 expert in multimedia learning. The following challenges emerged while performing tasks and then the following updates as a solution were made after the implementation as the usability test of the system.

- In the old design, email and password text boxes were visible for logging on but these text boxes were hidden for registering. Because, while performing task 1, a participant clicked the wrong button ("Sign up"). Also, three of the participants performed task 1 above the average task completion time. This was the first challenge in the design of the task. We determined the reasons of this challenge as navigating of the "Sign up" button. As we stated before, the email and password fields were open on the page, so no direct option was offered to the user. The other reason of it can be about button color because the button color is striking. According to the data obtained from the observations and the interview, we think that these design problems misled the user to the wrong button, therefore we changed the design of the login page. In the new design, textboxes were removed, and "Login" and "Register" buttons were added to the page (Figure 5). When a user clicks, the text boxes appear just above the button.

- In the old design of reproductive systems, users could listen to the information about the organs in the reproductive systems when they clicked on the speaker icon. However, if they clicked the icon again, the voice started again (Figure 3). While participants performed task 2, we observed challenges about the design. The first challenge is that participants often tried to click the speaker icon again to stop the sound. Also, some users were hesitant about whether to click the icon as well. The second challenge is that one participant couldn't complete the task, and the average completion time of task 2 is 54.86, three of the participants performed it above the average time. These findings showed us that there was a design problem about the speaker icon on the page. As a solution to it, we redesigned the speaker icon and added a directive text on the use of the icon to the page (Figure 7). In the new design, the sound stopped if the user clicked on the speaker icon a second time.

- In the old design of the Body Mass Index (BMI), a user should have clicked the warning "Hey, you have a new task!" to read and complete a mission (Figure 5). Also, the case texts were plain. However, we observed some design challenges during the usability test. While performing task 4, participants requested help from researchers according to the observations during implementation. This was the first challenge because we expected participants to complete the task without asking for help. The second challenge is that three of the participants couldn't complete the task. It was observed that instead of clicking on the case text at the top, participants went to the bottom of the page with the mouse and searched for the relevant case in the body mass index section. In other means, they couldn't notice the case texts. According to the result obtained from interviews and observations, participants stated the reason for this challenge was as a visually appealing BMI tool and misunderstanding of the directive. They stated that the directives in the case texts were not clear and understandable enough. A participant view is as follows:

  I didn't notice "Hey, you have a new mission!" warning at the top of the page and we should have clicked on it to open. But I looked below to find the case. It could have been better if we could see the details on the page below instead of clicking to open it."

  As a solution to these specified design challenges in the new design, the case texts could be seen and done on the page without any clicking. Moreover, a directive text was placed over the cases, and it was written in bold. In addition, the directive texts were re-written in a more intimate and motivating language (Figure 5).

- While participants performed task 6, they had difficulty completing the task. There was one participant who could not finish the task and one participant who needed help from the researcher. According to the observations and interviews, the challenge with this task was that the participants clicked the "Sexual Transmitted Diseases Crossword" menu instead of the "Sexual Transmitted Diseases Word Hunt" menu. There were two different crossword activities starting with the same words on the page. The reason for this confusion is to start both pages' titles with "Sexually Transmitted Diseases …". As a solution to this design challenge on the page, we redesigned it. After the revisions, a menu titled "Sexually Transmitted Diseases Activities" was added, and also "Crossword" and "Word Hunt" activities were added as a sub-menu under this menu (Figure 6). Also, the text "Solve" button in the Crossword activity was changed to "I couldn't find it."

- In the old design, the user was to correctly match male reproductive organs with the related numbers on the right side of the figure (Figure 3). However, this design had caused navigating problems on the page. While performing task 7, participants could not notice that they needed to move words to numbers, because the number field was not marked visually and there was not a directive showing how to do the activity. While performing task 7, one participant requested help from researchers. In the interviews and observations made, participants stated that the directive was insufficient, the list was not remarkable, and therefore they had difficulties while performing task 7. As a solution to these challenges, we redesigned the number field and added a directive text. In the new design, the numbers, which indicated the organ, were divided with a line (Figure 6).

- In the interviews, some participants suggested an intro video as a solution to navigating problems. In parallel with this suggestion, an informative intro video indicating the general use of the training was developed. The video was about the navigation in the menus and the use of the tools.

- In the old design of the website, menus were opened when clicking. In the new design, an arrow sign indicating submenus was designed because participants had difficulty finding submenus (Figure 6).

- While the participants were performing tasks, it was observed that some participants were not sure which page they were on. As a solution to this challenge, we designed breadcrumb navigation on a section under the main menu to indicate which page the user is on (Figure 6).

- In the old design, there was no button on the website for the user to logout. The web page was designed as session-based, that is when a user closed the page in the browser, the user was also logging out of the page. However, this design was contrary to the frequently viewed user behavior in the online learning environments. As a solution to this, a "Logout" button was added to the website in order to ensure that the user felt safe and to adapt to the users' behavior in the new design.

- In the old design, there was no symbol showing the tasks completed when a task was completed, so some participants were hesitant whether the task was completed or not. As a solution to this design challenge, we designed a figure that indicates completed activities on the page. It was added to the right of the menu titles to enable students to understand if activities were completed (Figure 6).

- In the old design, the user was not informed about how to use the buttons on the page. It was expected from the user. But this design has caused navigating problems on the page. As a solution to this challenge, we added directive texts showing how to use buttons on each page. Also, if the user wants to pass directly without reading the directives, a design has been made on the page so that they can exit from it by clicking the "Skip to explanations" button on the top right (Figure 8).

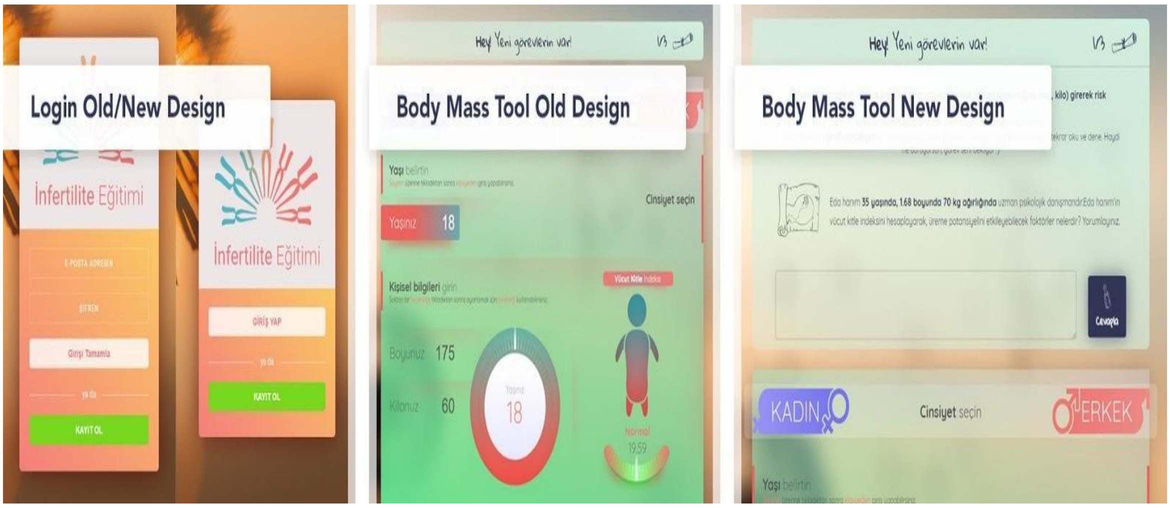

Figure 5. Screenshots of logging in and registering, and a mission for Body Mass Index

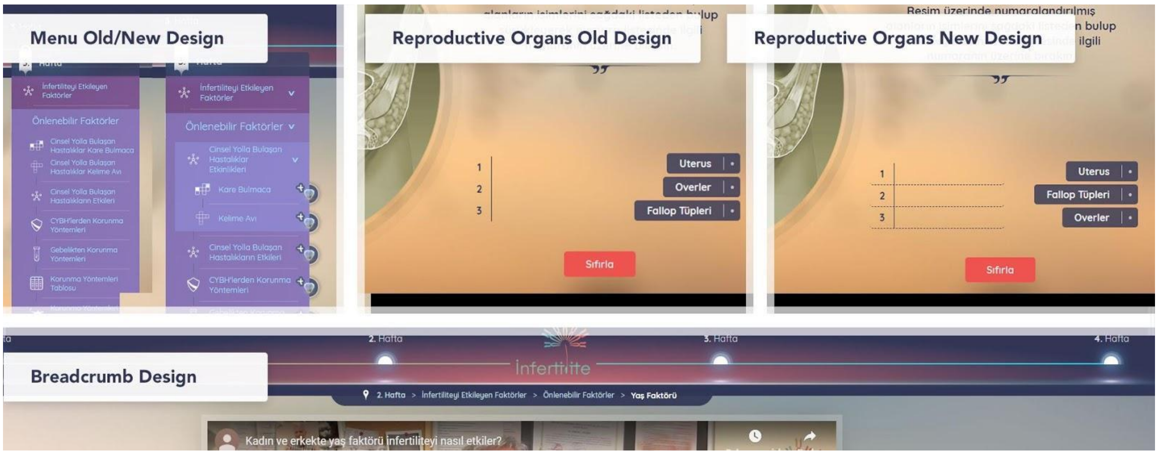

Figure 6. Old and new design of menus and submenus, old and new design of the reproductive organs matching game, and breadcrumb navigation

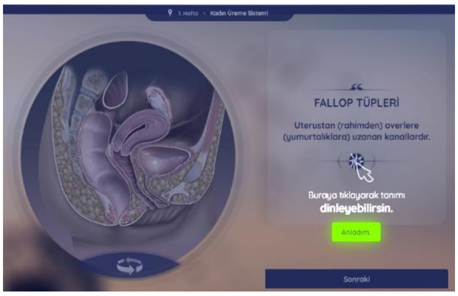

Figure 7. New design of the speaker icon

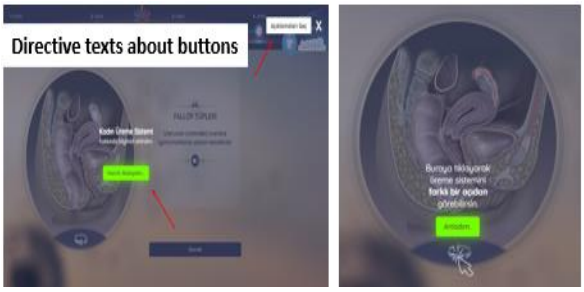

Figure 8. Adding of directive texts on the page

Iteration II. Results

The participants identified the main positive aspects of the education as interactive, rich in terms of multimedia elements and entertaining. In addition, division of content on an unit basis, separation of weekly content into titles (menus), and use of the website on different devices (smart board, computer) were the other positive features of the education. It was also stated that users had problems related to the interface design. These included: 1) navigation problems on the website, 2) inability to find some submenus under the main menu, and 3) confusion as to how to finish the education. As a solution to these design problems, a redesign was made on the relevant tools (Figure 5, 6, 7 and 8).

Iteration III. Implementation in a class

Research Question 3

How usable is the online environment as a system for a real implementation in a class?

Method of Iteration III

Sixteen undergraduate students (13 female and 3 male) participated in the implementation. None of the students had previous experience with online education. The students voluntarily participated in the study within the scope of an undergraduate course on health knowledge for 4 weeks.

The system usability interview form which aims to reveal the participants' experiences and suggestions about the online education and the system usability scale (SUS), which was used in Iteration II were used after implementation. Database records were also used after completing the implementation. The results of the observation and interview were interpreted together.

Iteration III. Challenges and Results

Design of the directions and feedback

The participants described the directions on the website as "sufficient and explanatory". They found the feedback "sincere, understandable and entertaining". Some of the participants suggested that feedback about ending the activity could be added to the end of each week.

> The directives and feedbacks were entertaining. The directions were facilitating everything. They were checking whether I am ready or not and asking in the form of "Are you ready? Here we go!". I progress easier in this way.
>
> A question was asked before watching the video. Then a text was shown in the form of "Let me watch!". I closed it but I liked it, I laughed and thought that "what a lovely idea". I thought that they thought well of it. It came to me very intimate, I felt I was chatting with my friend and felt like he/she was teaching something to me. The feedbacks have reduced such formality and made me feel sincere.
>
> The first part was over, but it didn't seem something like "it was complete". I was confused. So, I did the activities twice. It may be added something like "You've done it correctly, it is completed.

Page design

The participants evaluated the page designs as "simple, understandable, interesting and consistent". They stated as a negative opinion that circles indicating completion of weekly activities were not working on the website. As this feature was not working on the site, the participants frequently consulted the teacher or repeated the activities over and over again. This was an important design challenge while the implementation process. After that, we solved this problem.

> The design is understandable. The location of different buttons is different, and the same color is used for the same function. The design doesn't distract attention.

The content

The participants classified their opinions related to the content as "important", "interesting", "a reliable information source", "rich in multimedia elements", and "classification is guiding".

> The importance of the subject is vital to be willing to follow the whole online lesson. We know from the media/ news that sexual intercourse age is decreasing. We are 21 years old now and this is the first time to take such a course. I think this course should have been given before.

System Usability Scale

The SUS was applied in the class and the highest score was 100 and the lowest score was 65. The means SUS score of the website is 85.47 (SD = 11.26). That is, the participants found the website to be highly usable.

Iteration IV. Comparison of Iteration II and III, and Decision

Research Question 4

Is there any significant increase in system usability between the usability test and real implementation in a class?

Method of Iteration IV

This iteration compared the results of system usability scores of previous iterations as a component of design. One of the nonparametric statistical methods, Mann-Whitney U test, was used to compare the system usability between the usability test and real implementation.

Iteration IV Results

A Mann-Whitney U test was conducted to compare the system usability SUS scores between the usability test and real implementation. The test indicated that there was a significant difference between the SUS scores of the participants in the usability test and the implementation in a class, U=23.50, p< .05. (Table 4); the SUS scores were higher in the implementation class.

Table 4: Results of the Mann-Whitney U test

Periods	n	Mean rank	Sum of ranks	U	p
System Usability in the implementation in a class	16	15.03	240.50	23.50	.013
System Usability in the usability test	8	7.44	59.50		

Discussion, Conclusion and Suggestions

The design-based research method enabled us to design and develop an "Online Infertility Prevention Training" in an effective way that produced positive and promising evidence regarding the use of online education about health sciences for university students in a blended learning environment. User-based approaches of design-based research in this study also facilitated the researchers to determine significant user behaviors while using an online educational website. Çağıltay (2001) stated that usability studies were very beneficial when developing a digital system and could be used as a part of a formative and summative evaluation. As a whole, the four iterations and their outputs positively contributed to the development of an online educational website. However, some prior work indicated that researchers preferred to use only one iteration during design-based studies and did not report revisions due to time and resource limitations (Zheng, 2015). Conducting multiple iterations in the current study led to an increase in system usability scores. Therefore, two main implementations contributed to the improvement of the design used in the current study. All in all, conducting usability testing made a great contribution to determine design challenges and to observe users' behaviors while they are using the online learning environment. Also, starting the design process conducting need analysis ensured us to reveal users' expectations and to create usable knowledge relating to the content of the environment.

Prior research found the Internet is one of the important sources of health information for patients receiving infertility treatment or healthy individuals (Zelkowitz et al., 2016; Zillien et al., 2011). Unlike other online infertility education studies where the usability scores were low (Marriott et al., 2008), the usability in this study gradually increased as a result of utilizing design-based research. Therefore, the evidence-based positive results indicate that online education can enable people to investigate the source of infertility-related knowledge without usability problems and social labeling. In addition, improving system usability contributed to practical knowledge for researchers working on message design in the development of e-learning environments in the field of health sciences.

The first iteration started with the needs analysis which aimed to reveal university students' available infertility knowledge and their preferences about multimedia design. Tripp and Bichelmeyer (1990) stated that instructional designers should begin the software development process by examining the current context and determining the needs because a computer-based system is centered on the human cognition system. It is remarkable that many studies noticed the importance of identifying learners' previous knowledge about the subject not only in the field of health sciences but also in other fields (Hickey et al., 2017). A successful website should involve substantial careful consideration of users' preferences in addition to the accuracy of information provided (Djamasbi et al., 2007). As a result, the obtained detailed information about the users facilitated studying the processes used by teams of instructional technologists and provided them with a road map before starting the design of the education in the current study. The practitioners were able to make design decisions more readily because they knew the expectations of the users. Moreover, data from a large sample based on users' knowledge and multimedia preferences led to increased student satisfaction and effectiveness of the system, which we evaluated in the other iterations. The needs analysis results indicated that more than %80 of students would like to see frequently asked questions and answers about infertility, expert (doctor) videos, and animations in an online training about infertility. Thus, the main design decision was to use doctor videos on topics or questions in which learners had limited knowledge. The most important benefit of using doctor videos was that learners were satisfied with the accuracy of the knowledge relayed. The usability test and main implementation results indicated that this decision increased student satisfaction.

The Iteration I result showed that there were significant differences between the preferences of female and male in the development of the online training platform, with the exception of content information. More specifically, female students preferred to see "questions and answers" and "expert videos", but men students preferred animations and different designs. There were significant differences between the preferences of female and male in most studies examining gender differences in the development of websites (Cyr & Bonanni, 2005; Djamasbi et al., 2007; Huang & Yuan, 2017; Simon, 2001). The current study considered gender-related design preferences as a whole because the ratio of both female and male involved in a design decision was comparable. Therefore, this study used of expert videos, subject matter related with the other sex, or animations in the design. We recommend that other practitioners consider both gender preferences and expectations when designing.

Before designing instruction, determination of students' prior knowledge enables designers to identify learning problems, find the appropriate teaching level, and select successful feedback (Hailikari et al., 2008). This study highlighted the importance of prior knowledge and started teaching with basic concepts of the reproductive systems and then presented biological and preventable factors relevant to infertility. In online education, the reproductive system was the favorite and the most popular section among the students, and they repeatedly used this part during their education. This result also indicated that prior knowledge must be considered in the design. Furthermore, the nature of the subject is complex because it includes many concepts that university students have not previously learned. Bruner, who made two important contributions to teaching through concept teaching and discovery learning, argued that concept teaching includes teaching the definition, its features, and its examples (Ozmen, 2004). Therefore, the design required use of the concept teaching method including first teaching spelling and then the function of a concept. The fact that the concept teaching in the online environment was realized through a drag and drop game was also a favorite element for the students.

Iteration II indicated that the participants had some navigation problems in the online environment. For example, not being able to guess what issues were under a menu, having difficulty in understanding the instructions, increasing the time to perform tasks, and not closing the web site. These challenges caused navigation problems but especially increased cognitive load for the human cognitive system (Altun, 2000). There is a significant effect of navigation design types on recall and retention in e-learning environments (Dikbas Torun & Altun, 2014). Therefore, this study indicated that it is important to fix these various navigation problems before implementation. After redesigning the navigation problems of the online environment, the system usability scores increased. In addition, Iteration II indicated that adding a limitation of skipping videos while watching doctor videos increased the rate of video viewing. Fiorella and Mayer (2018) stated that breaking the video into pieces and asking students for explanations in video breaks can help learners control essential processing or cognitive load. Similar to this study, users tended to skip videos in e-learning environments (Bayazıt & Akçapınar, 2018; Kim et al., 2014). In our daily life, students may skip videos on social media, especially when the video duration is long. Therefore, adding a limitation of skipping in videos or giving feedback about it may be suggested to other designers who want to use educational videos in e-learning designs. In addition, segmenting videos into short sections and presenting essential content knowledge can be helpful to prevent the skipping behaviors of students. Lastly, the observations and the interviews indicated that users couldn't clearly follow the instructions on the page. The instructions were revised according to Mayer's (2009) multimedia principles (signaling and personalization principles) and then were presented to the users' experience (Iteration III). As a result, they didn't have any problems understanding the instruction during the implementation in a class.

As a conclusion, the results obtained from this study show that design and user-based research approaches in multimedia learning environments about health sciences is useful to determine target users' behaviors explicitly and design challenges. Four iterations and their outputs of the current study was made a great contribution to the designing process of the online education website and usability scores of the website positively. The results of this investigation also show that while teaching concrete parts of the human body, interactive animations and visuals may be useful to motivate students to use the content again and again. Therefore, their memorization of labels, definitions and the place of an organ can increase. Taken together, this research has shown significant findings related to students' multimedia learning preferences in health sciences, how can implement usability tests on online sexual health education environments, design principles of these environments etc.

Scientific Ethics Declaration

The authors declare that the scientific ethical and legal responsibility of this article published in JESEH journal belongs to the authors.

Acknowledgements or Notes

This study was supported by The Scientific and Technological Research Council of Turkey – TUBITAK [grant number 215K001].

References

Adam, M., McMahon, S. A., Prober, C., & Bärnighausen, T. (2019). Human-centered design of video-based health education: an iterative, collaborative, community-based approach. *Journal of Medical Internet research*, *21*(1), e12128, DOI: https://doi.org/10.2196/12128

Altun, A. (2000). Patterns in cognitive processes and strategies in hypertext reading: A case study of two experienced computer users. *Journal of Educational Multimedia and Hypermedia*, *9*(1), 35-55.

Alwashmi, M., Hawboldt, J., Davis, E., Marra, C., Gamble, J. M., & Ashour, W. A. (2016). The effect of smartphone interventions on patients with chronic obstructive pulmonary disease exacerbations: A systematic review and meta-analysis. *JMIR mHealth and uHealth*, *4*(3), e105, DOI: https://doi.org/10.2196/mhealth.5921

Baddeley, A. D. (1999). *Human memory*. Boston, MA: Allyn & Bacon.

Bailey, B. (2006). *Getting the complete picture with usability testing*. Retrieved 20.03.2018, from https://www.usability.gov/get-involved/blog/2006/03/complete-picture-with-usability-testing.html.

Bangor, A., Kortum, P., & Miller, J. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, *24*(6), 574–594. https://doi.org/10.1080/10447310802205776

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, *13*(1), 1-14. https://doi.org/10.1207/s15327809jls1301_1

Baran, B. Siyez, D. M., Kagnici, Y. D., Seymenler, S., Esen, E. Ozturk, B. & Siyez, E. (2017, August 7-8). *Determination of the needs for development of infertility psycho-educational program and the design of a website about infertility for university students* [Paper presentation]. 20th Euro Congress on Psychiatrists and Psychologists, Rome, Italy. DOI: 10.4172/2378-5756-C1-023 https://www.walshmedicalmedia.com/conference-abstracts-files/2378-5756-C1-023-020.pdf

Bayazıt, A., & Akçapınar, G. (2018). Çevrimiçi dersler için video analitik aracının tasarlanması ve geliştirilmesi. *İlköğretim Online*, *17*(1). https://doi.org/10.17051/ilkonline.2018.413719

Boren, T., & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. *IEEE Transactions on Professional Communication*, *43*(3), 261-278. DOI: 10.1109/47.867942.

Borzekowski, D. L., Fobil, J. N., & Asante, K. O. (2006). Online access by adolescents in Accra: Ghanaian teens' use of the internet for health information. *Developmental psychology*, *42*(3), 450-458. https://doi.org/10.1037/0012-1649.42.3.450

Chi, X., Hawk, S.T., Winter, S., Meeus, W. (2015). The effect of comprehensive sexual education program on sexual health knowledge and sexual attitude among college students in Southwest China. *Asia Pacific Journal of Public Health, 27*(2), 2049-2066. https://doi.org/10.1177/1010539513475655

Conceição, C., Pedro, J., & Martins, M. V. (2017). Effectiveness of a video intervention on fertility knowledge among university students: A randomised pre-test/post-test study. *The European Journal of Contraception & Reproductive Health Care*, *22*(2), 107-113. https://doi.org/10.1080/13625187.2017.1288903

Cyr, D.& Bonanni, C. (2005). Gender and website design in e-business. *International Journal of Electronic Business*, *3*(6), 565–582. https://doi.org/10.1504/IJEB.2005.008536

Çağıltay, K. (2011). *İnsan bilgisayar etkileşimi ve kullanılabilirlik mühendisliği: Teoriden pratiğe*. Ankara: ODTÜ Geliştirme Vakfı Yayıncılık.

Davids, M. R., Harvey, J., Halperin, M., & Chikte, U. M. (2015). Determining the number of participants needed for the usability evaluation of e-learning resources: A Monte Carlo simulation. *British Journal of Educational Technology*, *46*(5), 1051-1055. https://doi.org/10.1111/bjet.12336.

Deng, M., Kuriansky, J., Yang, W., Lao, S., & Lao, Y. (2017). Discussion on the application of sexuality education's flipped education in Chinese college students. *The Journal of Sexual Medicine*, *14*(5), e272. https://doi.org/10.1016/j.jsxm.2017.04.686

Dikbas Torun, E., & Altun, A. (2014). The effect of levels of processing with navigation design types on recall and retention in e-learning environments. *Behaviour & Information Technology*, *33*(10), 1039-1047. https://doi.org/10.1080/0144929X.2014.945963

Djamasbi, S., Tullis, T., Hsu, J., Mazuera, E., Osberg, K., & Bosch, J. (2007). Gender preferences in web design: usability testing through eye tracking. *AMCIS 2007 Proceedings*, 133.

El Said, G. R. (2017). Understanding how learners use massive open online courses and why they drop out: Thematic analysis of an interview study in a developing country. *Journal of Educational Computing Research*, *55*(5), 724-752. https://doi.org/10.1177/0735633116681302

Epstein, Y. M., Rosenberg, H. S., Grant, T. V., & Hemenway, N. (2002). Use of the Internet as the only outlet for talking about infertility. *Fertility and Sterility*, *78*(3), 507-514. https://doi.org/10.1016/S0015-0282(02)03270-3

Fiorella, L., & Mayer, R. E. (2018). What works and doesn't work with instructional video. *Computers in Human Behavior*, 89, 465–470. *https://doi.org/10.1016/j.chb.2018.07.015*

Gungor, I., Rathfisch, G., Kizilkaya Beji, N., Yarar, M., & Karamanoglu, F. (2013). Risk‐taking behaviours and beliefs about fertility in university students. *Journal of clinical nursing*, *22*(23-24), 3418-3427. https://doi.org/10.1111/jocn.12097

Gülbahar, Y., Kalelioğlu, F., & Madran, O. (2008). Öğretim ve değerlendirme yöntemi olarak web macerası'nın kullanışlılık açısından değerlendirilmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, *41*(2), 209-236.

Hailikari, T., Katajavuori, N., & Lindblom-Ylanne, S. (2008). The relevance of prior knowledge in learning and instructional design. *American Journal of Pharmaceutical Education*, 72(5), 113.

Harvey, S., Memon, A., Khan, R., & Yasin, F. (2017). Parent's use of the Internet in the search for healthcare information and subsequent impact on the doctor–patient relationship. *Irish Journal of Medical Science (1971-)*, *186*(4), 821-826. https://doi.org/10.1007/s11845-017-1555-6

Hattink, B., Droes, R. M., Sikkes, S., Oostra, E., & Lemstra, A. W. (2016). Evaluation of the Digital Alzheimer Center: Testing usability and usefulness of an online portal for patients with dementia and their carers. *JMIR research protocols*, *5*(3), e5040, DOI: https://doi.org/10.2196/resprot.5040

Hickey, C., McAleer, S., & Khalili, D. (2017, November). Design, implementation and evaluation of a blended intensive short-term dynamic psychotherapy course. In *Medical Education Scholarship Forum Proceedings* (Vol. 3).

Huang Z., Yuan L. (2017) Gender Differences in Tourism Website Usability: An Empirical Study. In: Marcus A., Wang W. (eds) *Design, User Experience, and Usability: Understanding Users and Contexts*. DUXU 2017. Lecture Notes in Computer Science, vol 10290. Springer, Cham. https://doi.org/10.1007/978-3-319-58640-3_32

ISO 9241 (1998). Ergonomic requirements for office work with visual display terminals-Part 11: Guidance on usability.

Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014, March). Understanding in-video dropouts and interaction peaks in online lecture videos. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 31-40). ACM.

Marriott, J. V., Stec, P., El-Toukhy, T., Khalaf, Y., Braude, P., & Coomarasamy, A. (2008). Infertility information on the World Wide Web: A cross sectional survey of quality of infertility information on the internet in the UK. *Human Reproduction*, *23*(7), 1520-1525. https://doi.org/10.1093/humrep/den072

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge: University of Cambridge Press.

Nielsen, J. (2012, 4 June). *How many test users in a usability study?*. Retrieved 14.02.2018 from https://www.nngroup.com/articles/how-many-test-users/

Paivio, A. (2001). *Mind and its evolution: A dual-coding approach*. Mahwah, NJ: Erlbaum.

Poelmans, S., Wessa, P., Milis, K., Bloemen, E., & Doom, C. (2008, November). *Usability and acceptance of e-learning in statistics education, based on the compendium platform,* In Proceedings of the International Conference of Education, Research and Innovation, 1-10.

Rieber, L. P. (2017). Participation patterns in a massive open online course (MOOC) about statistics. *British Journal of Educational Technology*, *48*(6), 1295-1304. https://doi.org/10.1111/bjet.12504

Seymenler, S. (2017). *Investigating infertility knowledge and attitudes of university students* (Thesis number: 486455) [Master's thesis, Dokuz Eylül University]. Council of Higher Education Thesis Center. https://tez.yok.gov.tr/

Simon, S.J. (2001). The impact of culture and gender on web sites: an empirical study, *The Data Base for Advances in Information Systems*, 32(1), pp.18–37. https://doi.org/10.1145/506740.506744

Siyez, D.M., Seymenler, S., Esen, E., Siyez, E., Kağnıcı, Y., Baran, B., and Öztürk, B. (2018). Investigating knowledge levels of university students about infertility. *Turkish Journal of Urology, 44*(2), 153-161. DOI: 10.5152/tud.2018.78861

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 19–30). New York: Cambridge University Press.

Şahin Y. T. (2012). İnfertil Çiftlerin İnfertilite Stresiyle Başa Çıkma Yöntemleri. Unpublished PhD Thesis, İstanbul University Health Sciences Institute, İstanbul.

Thato, R., & Penrose, J. (2013). A brief, peer-led HIV prevention program for college students in Bangkok, Thailand. *Journal of Pediatric and Adolescent Gynecology*, 26(1), 58-65. https://doi.org/10.1016/j.jpag.2012.09.011

Travers, J. F., Elliott, S. N., & Kratochwill, T. R. (1993). Educational psychology: Effective teaching, effective learning. Madison, WI, England: Brown & Benchmark/Wm. C. Brown Publ.

Tripp, S. D., & Bichelmeyer, B. (1990). Rapid prototyping: An alternative instructional design strategy. *Educational Technology Research and Development*, *38*(1), 31-44.

Trotter, M. I., & Morgan, D. W. (2008). Patients' use of the Internet for health related matters: a study of Internet usage in 2000 and 2006. *Health Informatics Journal*, *14*(3), 175-181. https://doi.org/10.1177/1081180X08092828

TUIK (2018). *Hanehalkı Bilişim Teknolojileri (BT) Kullanım Araştırması*, Ankara. www.tuik.gov.tr

Virvou, M., & Katsionis, G. (2008). On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE. *Computers & Education*, *50*(1), 154-178. https://doi.org/10.1016/j.compedu.2006.04.004

von Rosen, A. J., von Rosen, F. T., Tinnemann, P., & Müller-Riemenschneider, F. (2017). Sexual health and the internet: Cross-Sectional study of online preferences among adolescents. *Journal of Medical Internet Research*, *19*(11), e379. doi:10.2196/jmir.7068. https://doi.org/10.2196/jmir.7068

Williams, K. M., Stafford, R. E., Corliss, S. B., & Reilly, E. D. (2018). Examining student characteristics, goals, and engagement in Massive Open Online Courses. *Computers & Education*, *126*, 433-442. https://doi.org/10.1016/j.compedu.2018.08.014

Wojcieszek, A. M., & Thompson, R. (2013). Conceiving of change: a brief intervention increases young adults' knowledge of fertility and the effectiveness of in vitro fertilization. *Fertility and Sterility*, *100*(2), 523-529. https://doi.org/10.1016/j.fertnstert.2013.03.050

Zegers-Hochschild, F., Adamson, G. D., de Mouzon, J., Ishihara, O., Mansour, R., Nygren, K., ... & Van der Poel, S. (2009). The international committee for monitoring assisted reproductive technology (ICMART) and the world health organization (WHO) revised glossary on ART terminology, 2009. *Human reproduction*, *24*(11), 2683-2687.

Zelkowitz, P., Robins, S., & Grunberg, P. (2016). Searching for Infertility Information Online: Differences Between Men and Women. *Iproceedings*, *2*(1): e37. DOI: 10.2196/iproc.6243

Zheng, L. (2015). A systematic literature review of design-based research from 2004 to 2013. *Journal of Computers in Education*, *2*(4), 399–420. https://doi.org/10.1007/s40692-015-0036-z

Zillien, N., Haake, G., Fröhlich, G., Bense, T., & Souren, D. (2011). Internet use of fertility patients: a systemic review of the literature. *Journal of Reproductive Medicine and Endocrinology*, *8*(4), 281-287. https://www.kup.at/kup/pdf/10199.pdf

Author(s) Information

Bahar Baran
Dokuz Eylül University, Department of Computer Education and Instructional Technologies, Buca Faculty of Education, Izmir/TURKEY
Contact e-mail: bahar.baran@deu.edu.tr
ORCID iD: 0000-0002-9179-3469

Şirin Nur Yacı
Dokuz Eylül University, Distance Education Application and Research Center, Izmir/TURKEY
ORCID iD: 0000-0001-8212-221X

Murat Öcal
Dokuz Eylül University, Department of Computer Education and Instructional Technologies, Buca Faculty of Education, Izmir/TURKEY

Diğdem Müge Siyez
Dokuz Eylül University, Department of Guidance and Psychological Counseling, Buca Faculty of Education, Izmir/TURKEY
ORCID iD: 0000- 0003-4724-3387

Journal of Education in Science, Environment and Health
Volume 8, Issue 1, 2022 https://doi.org/10.21891/jeseh.1059283

The Effect of Augmented Reality Applications in Biology Lesson on Academic Achievement and Motivation

Esra Omurtak, Gulcin Zeybek

Article Info

Article History

Published:
01 January 2022

Received:
04 June 2021

Accepted:
20 November 2021

Keywords

Augmented reality,
Instructional technologies,
Biology course,
Academic achievement,
Motivation

Abstract

In this study, it was aimed to determine the effect of activities based on augmented reality applications in 9th grade biology lesson on students' achievement and motivation towards biology lesson and to get students' opinions about the applications. In the study, mixed method was used. The study group of the research consisted of students of two different classes at the 9th grade level of a private high school located in the city center of Karaman in the 2018-2019 academic year. When the research findings were examined, it was seen that the academic achievement post-test mean scores of the experimental group students were found to be significantly higher than the mean scores of the control group students. According to another finding of the study, there was no significant difference between the experimental group and the control group students in terms of the total mean scores obtained from the motivation questionnaire for biology lesson. However, a significant difference was found in favor of the experimental group in terms of "anxiety in exams" dimension. As a result of the research, the students stated that they liked the biology lessons in which augmented reality applications were carried out, that augmented reality applications made abstract concepts concrete and made them more permanent, and that the lessons conducted in this way were fun.

Introduction

Today, there is a need to train individuals who can acquire increasingly complex knowledge and skills and keep up with the rapid change of this information. In addition to the rapid increase in knowledge, since 21st century skills are technology-oriented, the question of how information technologies can be used to support and enrich education, has led studies in the field of education and educational technology in recent years (Kozma & Anderson, 2002; Wang & Hannafin, 2005; Watson, 2001). In addition to these, it is thought that current curricula are insufficient to respond to the differing expectations of today's youth, which is called the digital generation or generation z. 21st century students, called digital natives, differ from previous generations due to their distinctive features such as their desire to access information very quickly, games instead of serious studies, visual and graphic elements instead of long texts, having a parallel cognitive structure, and being able to do more than one job simultaneously (Bilgiç et al.,, 2011). These innate features of digital natives have also led to innovations in the world of education and a new learning culture has emerged. In this sense, while teaching processes are being structured, new technological approaches have started to be preferred. For these reasons, it has become imperative to know the characteristics of the new generation and to configure the education-teaching process with teaching methods suitable for their characteristics, reshape curricula and environments to allow the use of innovative technologies (Somyürek, 2014; İzgi Onbaşılı, 2018).

Considering the developmental periods, the emergence of computers first, then the internet, and then the development and spread of mobile devices, and the growing technological transformation had important reflections on education and training environments (Hayes et al., 2004). Printed materials, which make learning more tangible and permanent, but have controversial dimensions in terms of cost, portability, usability and accessibility, have begun to move to digital media in recent years (Smaldino et al., 2019). Now, it has become inevitable to use digital tools and materials as materials in addition to teaching lessons with more than one teaching method (Seferoğlu, 2006). Integrating education with new technologies has allowed teachers to transform learning materials consisting of fixed text and graphics into more interactive multimedia materials (Huang et al., 2016). With the widespread use of the internet in the digital environment, hypermedia (video, sound, animation, picture, etc.) and web 2.0 tools have begun to be used more in learning and teaching environments (Andersen, 2007; Greenhow et al., 2009). Today, in addition to these technological tools, new applications are developed through mobile and wearable technologies and used in learning environments. One of these innovative digital technologies used is Augmented Reality technology.

Augmented Reality

Augmented Reality (AR) technology, which was developed in the 1960s, inspired by the work of Ivan Sutherland and his students at Harvard and Utah universities on computer graphics, came into existence fully in the 1970s. While it was officially first used by the United States Air Force and NASA, this technology became widespread after the 1990s and reached wider masses (Feiner, 2002). According to Azuma (1997), AR is a virtuality-based technology that enables us to interact with the real world in real time. According to Cabero and Barroso (2016), AR is a technology that allows the user to see the real world by combining real elements with virtual attachments. In other words, AR is a virtual reality application in which users interact with virtual objects while interacting with the real world without affecting the real world (Zhu et al., 2004). It is the creation of an interactive environment between the virtual world and the real world that can accommodate the characteristics of both environments. AR technology is used to create this interactive environment (Bronack, 2011; Klopfer & Squire, 2008). In other words, AR enriches the existing reality with virtual objects it adds to the real environment and makes it more dynamic (Cheng & Tsai, 2013; Kerawalla et al., 2006). AR provides a seamless interface and natural interaction for users by combining the real and virtual worlds (Cai et al., 2014; Kaufmann, 2003; Matcha & Rambli, 2013). Azuma (1997) lists the characteristic features of AR technology as follows in his study to determine the limits of AR technology:

• Combines the real with the virtual,
• It is real-time interactive,
• It is 4 dimensional.

AR technology is the placement of four different elements in the real world in three dimensions: the camera, the computer infrastructure, a marker and the real world. AR is a technology based on the appearance of a four-dimensional object designed on the target picture, output or materials determined by a developer and creating the effect of the object being on the target picture, output or material (Augment, 2020). In addition to being used in various fields such as AR technology, science, industry, trade, health, space and aviation, it has recently had an important and widespread use in the field of education.

Use of Augmented Reality in Education

The development in technology has transformed AR from a technology that only works with special equipment, to a technology that can be easily used on personal computers or mobile devices. Today, it is seen that AR is used in many different sectors such as tourism, military, health, advertising and education (Buluş Kırıkkaya & Şentürk, 2018; Gümüş & Boydaş, 2021; İçten & Bal, 2017; İzgi Onbaşılı, 2018; Yen et al., 2013). Especially the widespread use of mobile devices has paved the way for the use of AR in educational environments (Wu et al., 2013). In the reports of institutions and organizations that direct educational technologies in the international context, it is predicted that AR is among the potential technologies at the K-12 and university level, and that it can make a remarkable contribution to learning with its immersive and developable structure in the future (Johnson, et al., 2014; Johnson. et al., 2016).

AR technology is used in teaching inaccessible or invisible objects in educational environments, in the application of dangerous experiments, and in embodying abstract concepts. In addition, it is known that it has advantages for cognitive and affective development such as increasing permanence in learning, making lessons more enjoyable, and increasing students' interest (Lee, 2012; Wu et al., 2013). At the middle and high school level, science is one of the fields that students have difficulty in (Timur & Özdemir, 2018). The reason for this is suggested that science subjects are rich in abstract concepts. In order to understand abstract scientific concepts, students need to construct mental models (Ibáñez et al., 2014). AR makes it possible to teach situations that cannot be observed in the classroom environment by combining real and virtual (Kerawalla et al., 2006; Shelton & Hedley, 2002). Thus, science subjects that cannot be observed in real life can be taught by doing and experiencing. In addition, AR enables abstract concepts to be learned more easily by concretizing them (Abdüsselam, 2014; Kamarainen et al., 2013; Núñez et al., 2008; Shelton & Hedley, 2002; Shelton & Stevens, 2004; Wu et al., 2013). The use of AR applications has become even more important, especially in the teaching of applied sciences, where the use of technological tools and equipment such as mathematics and science are more needed; As the idea that AR can be used in educational environments has become widespread, many applications that can be used in science education have been designed recently.

The field of science is of great importance in the development of countries. However there are still problems related to the learning-teaching process in Turkey, such as the lack of integration of learning materials and technology in science education, and the predominance of teacher-centered methods in teaching. It is thought that the fact that students' science achievement is not at the desired level in the national and international context may be related to these problems (Fidan & Tuncel, 2018; Ültay & Ültay, 2020). The fact that AR provides the opportunity to have real world assets and objects in digital environments in the same environment (Azuma et al., 2001) adds some exciting features to the learning of students learning in traditional learning environments. In addition, thanks to AR, some features that cannot be seen in the real world can be modeled three-dimensionally and numerically, and this information can be presented together with real world assets, allowing some abstract concepts to be concretized (Timur & Özdemir, 2018), it increases academic success by transforming the learning process from abstract to concrete, especially in teaching science subjects (Abdüsselam & Karal, 2012).

Research in recent years indicates that AR makes many contributions to educational environments. It is among the results obtained in studies that the use of AR increases academic achievement, supports learning by doing and experiencing and learning by questioning, increases class participation and motivation, makes lessons more enjoyable, facilitates concept teaching and reduces misconceptions (Abdüsselam, 2014; Bressler & Bodzin, 2013; Cai et al., 2014; Chiang et al., 2014; Chen & Wang, 2015; Enyedy et al., 2012; Furió et al., 2015; Hsiao et al., 2012; Hwang et al., 2016; Ibáñez et al., 2014; Kerawalla et al., 2006; Lin et al., 2013; Matcha & Rambli, 2013; Perez-Lopez & Contero, 2013; Rosenbaum et al., 2006; Sırakaya, 2015; Singhal et al., 2012; Shelton & Hedley, 2002; Solak & Cakir, 2015; Squire & Jan, 2007; Tian et al., 2014; Vilkoniene, 2009; Wang et al., 2014; Zhang et al., 2014).

Like cognitive factors, affective factors also play an important role in students' achievement (Tuan et al., 2005). Motivation, which is among these affective factors, is defined as the initiation of the necessary process to meet a need (Waterman, 2005). It is known that motivation is a factor that plays an important role in the learning process and that highly motivated students participate more in the lesson (Dede & Yaman, 2008; Wolters & Rosenthal, 2000). In this context, AR may contribute to the positive development of students' attitudes and motivations towards learning with its features. Today, the use of technology to increase the desire to learn has become a very popular research topic (Di Serio et al., 2013). The fact that this research investigates the effect of AR applications on motivation as well as the effect of academic achievement and at the same time including student opinions makes the research important in terms of contributing to the field. Greene (2007) defined the mixed method as "more than one way of seeing". In this study, it was considered important to determine the academic success and motivation levels of the students, as well as their experiences and opinions about the practice, which were determined quantitatively. In addition, when we look at AR studies in education, it can be said that most of them are at secondary school level and there are relatively few studies at high school level. However, high school students love to learn by exploring and having fun, they prefer fast access to information, thus using technology while researching and learning, they prefer graphics over texts. It is mentioned that the increasing use of augmented reality technologies in educational environments is due to some unique characteristics of young individuals. Kapil and Roy (2014) defined the digital generation as individuals who want a technology-oriented life, are productive social media users, live in a communication environment equipped with advanced technologies, and can actively use technology to overcome the problems they encounter. In order for individuals with these characteristics to participate actively in their learning processes and to increase their interest and gain effective permanent learning experiences, multimedia materials should be created for them and used in teaching environments (Wang et al., 2013).

Based on the idea that technology-based approaches can be useful in different school levels and in different courses, in this study, it was aimed to determine the effects of activities based on AR applications on the ninth grade biology lesson "cell" unit "cellular structures and their functions" subject on the academic achievement of students and their motivation for the biology lesson, and to get students' opinions on the applications carried out. The subject of the 9th grade cell unit "cellular structures and their functions" includes sub-headings such as parts of prokaryotic cells, the structure of eukaryotic cells and the parts that make up this structure, the functions that organelles take in the cell, different cell examples, and intracellular organization. These subjects and concepts are quite abstract and it is important to teach them concretely through various visuals, three-dimensional models, videos and simulations. In addition in the Biology curriculum of the Ministry of National Education, while comparing different cell samples visual elements (photos, pictures, drawings, cartoons, etc.), graphic organizers (concept maps, mind maps, diagrams, etc.), e-learning objects and applications (animation, video, simulation, infographic, augmented and virtual reality applications, etc.) is recommended. Considering all these, it is thought that AR applications will be an effective way to teach the subject of "cellular structures and functions" of the cell unit.

The hypotheses of the experimental research carried out in this direction are presented below:

1. The posttest point averages of the experimental group in which the activities based on AR applications were applied were significantly higher than the posttest point averages of the control group in which the activities based on the current program were applied.
2. The motivation score averages of the experimental group in which the activities based on AR applications were applied were significantly higher than the motivation point averages of the control group in which the activities based on the current program were applied.

The sub-problem regarding the qualitative dimension of the research is given below:

- What are the opinions of the experimental group students about the AR activities applied within the scope of the research?

Method

Research Design

The design of the research is the explanatory sequential mixed design. The purpose of this design is to use qualitative data to explain quantitative findings in more detail. The first stage in this process includes the collection of quantitative data and the analysis of the data. Afterwards, it is aimed to help explain the answers given in the quantitative dimension with the interviews made at the qualitative stage (Creswell, 2014). The explanatory design is probably the most understandable of the mixed method designs (Creswell & Plano Clark, 2011). One of the reasons for choosing this design is that the researcher and the main problem of the research are quantitative oriented. In addition, its two-stage structure makes this design easy to apply (Creswell & Plano Clark, 2011). The researcher preferred this design in order to carry out the two methods in separate stages, to collect only one type of data at a time, and to support the quantitative data using qualitative data and to explain it in more depth and detail. The static-group pretest-posttest design was applied in the quantitative dimension of the study. In this design, the experimental group and the control group are selected without random assignment. Pretest and posttest are applied to both groups. Experimental process is done only to the experimental group (Cresswell, 2014). While experimental studies are carried out with students in schools, it is often possible to work with ready-made groups. It is generally not possible to make changes in groups and to perform random assignment. In this research, it was necessary to work with previously formed groups, since the experimental study was started in the middle of the academic term. However the experimental and control groups were randomly assigned from these groups. The case study was applied in the qualitative dimension of the study. The case study delves into and analyzes the factors that explain the current situation or influence change and development (Best & Kahn, 2014). With the case study design, the researcher aimed to facilitate the explanation and understanding of the possible change in the scores obtained from the academic achievement test and motivation questionnaire before and after the experimental application. In the context of the case study, the interview method was used to collect the qualitative data of the research. The purpose of the interview is to learn what is in the mind of the individual and to reach his point of view (Patton, 1990). Interviews are used to obtain information about the knowledge and experiences of the individual (Best & Kahn, 2014). Interviews allow to obtain information about what cannot be seen directly and to make alternative explanations about what is seen (Glesne, 2016). In this study, this method was used to obtain the experiences, opinions and suggestions of the participants regarding the experimental process. Within the scope of the research, it was aimed to reach a rich and a variety of data that can confirm each other by employing more than one data collection method.

Study Group

The students of two different classes, who were studying at the ninth grade level of a private high school in Karaman, were the study groups of the research. One of these groups was determined randomly as the experimental group and the other as the control group. The experimental group consisted of 17 students and the control group consisted of 21 students. The average age of both groups is 15. Students in both the experimental and control groups do not have any knowledge or experience of AR applications. While determining the study group of the experimental research, previously formed ready groups were used. Considering the qualitative dimension of the research, this method can be expressed as easily accessible case sampling. Easily accessible case sampling is often used when the researcher is not able to use other sampling methods. This sampling method brings speed and practicality to the research (Yıldırım & Şimşek, 2018). It was envisaged that the students in the experimental group of the research would be the participants of the qualitative dimension of the research. It was envisaged that the students in the experimental group of the research would be the participants of the qualitative dimension of the research. However, the researcher ended the data collection process at the point where the concepts and processes that emerged while the interview data were collected and analyzed began to repeat each other. At this stage, interviews were conducted with 11 participants. These participants participated in the interviews voluntarily.

In order to determine whether the experimental and control group students were equivalent in terms of academic achievement levels in the biology course cell unit, the independent samples t test was applied to determine whether there was a significant difference between the academic achievement pretest score averages. The t test results of the students' pretest mean scores according to the group are shown in Table 1.

Table 1. The *t* test results of the students' academic achievement pretest mean scores according to the group

Group	n	x̄	Sd	df	*t*	*p*
Experiment	17	10.24	2.71	36	0.757	0.415
Control	21	11.00	3.38			

$p \leq .05$.

When Table 1 is examined, it is seen that there is no significant difference between the academic achievement pretest mean scores of the experimental and control group students ($t_{(36)=}0.757$, $p \leq 0.05$). Based on this finding, it can be said that the experimental and control group students are equivalent in terms of academic achievement pretest mean scores in the biology course cell unit. In order to determine whether the experimental and control group students were equal in terms of their motivation levels for the biology lesson, the independent samples *t* test was applied to determine whether there was a significant difference between the motivation questionnaire pretest score averages. The *t* test results of the students' motivation pretest mean scores according to the group are shown in Table 2.

Table 2. The *t* test results of the students' motivation pretest mean scores according to the group

Factor	Group	n	x̄	Sd	df	*t*	*p*
Intrinsic motivation	Experiment	17	20.18	2.94	36	0.585	0.413
	Control	21	19.05	4.96			
Extrinsic motivation	Experiment	17	21.12	3.02	36	0.457	0.470
	Control	21	20.33	3.50			
Interest in learning	Experiment	17	19.29	3.67	36	1.748	0.091
	Control	21	16.76	5.03			
Responsibility for learning	Experiment	17	20.12	3.24	36	1.628	0.482
	Control	21	19.19	4.51			
Confidence in learning	Experiment	17	19.47	4.93	36	1.134	0.264
	Control	21	17.62	5.06			
Anxiety in exams	Experiment	17	16.18	4.71	36	1.084	0.286
	Control	21	17.71	4.04			
Total	Experiment	17	116.35	12.10	36	0.919	0.364
	Control	21	110.67	23,03			

p≤..05.

When Table 2 is examined, it is seen that there is no significant difference between the motivation pretest mean scores of the experimental and control group students in terms of total motivation level and in terms of each sub-dimension ($t_{(36)=}0.919$, $p \leq 0.05$). Based on this finding, it can be said that the experimental and control group students are equivalent in terms of their motivation pretest mean scores.

Data Collection Tools

The experimental design of the research is the static group pretest-posttest design, and an achievement test consisting of 43 questions was developed by the researcher to be applied to the experimental and control groups as pretest and posttest. While creating the items for the test, cooperation was made with biology experts and language experts, the relevant literature was scanned to ensure content validity, and opinions were received from three experts working in the relevant field. As a result of expert evaluations, corrections were made on the test and the agreement value between the evaluations of the experts was calculated (Miles & Huberman, 1994); It was determined as 0.81. The test was applied to students studying in two different high schools in the city of Karaman in order to calculate its reliability. The reliability of the test as a result of the application was calculated by the Kuder-Richardson (KR-20) method as r=0.95. Item difficulties (pj) were determined based on the data obtained from the pre-application and the average difficulty of the test was calculated as 0.44. Accordingly, it is possible to say that the developed test is a medium difficulty test. In the study, item discrimination values were calculated. As a result of the obtained values, item 9 (pj=0.36, qj=0.64, rjx= 0.26) and item 17 (pj=0.53, qj= 0.47, rjx=0.29) were not included in the final test due to their low efficacy. After these changes, 41 questions remained in the test. After the changes, the reliability of the test was recalculated with the KR-20 method and the reliability coefficient was found to be r=0.95. The high values of this coefficient, close to 1.00, indicate that reliability in terms of internal consistency is provided (Demirel, 2008). This value is quite good for an achievement test.

In order to determine the effect of AR applications on the motivation of students towards biology lesson, "Motivation Questionnaire for Biology Lesson" (BDME), which was developed by Glynn and Koballa (2006) and adapted into Turkish by Ekici (2009), was used. The original questionnaire developed by Glynn and Koballa (2006) consists of six dimensions. These are *intrinsic motivation, extrinsic motivation, interest in learning biology, responsibility in learning biology, confidence in learning biology, and anxiety in biology exams*. It is stated that the Cronbach Alpha reliability coefficient for the overall scale was determined as .93. The questionnaire translated into Turkish was applied to 646 high school students, and its validity and reliability were determined. According to the results of the analysis, the number of items, which was 30 in the original questionnaire, was also preserved in the Turkish questionnaire. According to the results of the factor analysis, it was determined that the Turkish questionnaire had six dimensions, just like the original questionnaire. There are 5 items in each dimension of the Turkish questionnaire. While the Cronbach Alpha reliability coefficient was found to be .87 for the overall questionnaire, it was determined that it ranged from .80 to .89 in its dimensions. These values are similar to previous studies (Glynn & Koballa, 2006; Glynn et al., 2007; Glynn et al., 2009). Questionnaire items were arranged in a 5-point Likert type. Positive items are never: 1 point, rarely: 2 points, sometimes: 3 points, usually: 4 points, and always: 5 points, while negative items are evaluated in the opposite way. In addition, statistically significant and positive correlations were determined between the six factors in the questionnaire. It states that these relationships between the dimensions of the scale are moderate and positive, and that the scale consists of factors independent of each other (Büyüköztürk, 2006). As a result; it has been determined that the Turkish form of the biology lesson motivation questionnaire is suitable for Turkish, valid and reliable in terms of language.

At the end of the experimental process, standardized open-ended interviews were conducted with the volunteer participants. For this purpose, an interview form consisting of open-ended questions was prepared by the researchers. The interview form was evaluated by the experts in the field of curriculum development and measurement-evaluation in terms of the suitability of the questions, and by the language experts in terms of clarity. As a result the form was found appropriate to apply to the participants. As examples of the questions in the interview form; *"Do you think AR applications contributed to your learning, remembering and your achievement in the course? (If so, how did it contribute?), Do you think AR applications contributed to your interest and motivation towards the lesson? (If so, how did it contribute?), Have AR applications changed your study method? (If so, how did it change?), Do you think AR applications contributed to the classroom environment, teacher-student and student-student relations? (If so, how did it contribute?), What do you think about the software you use for AR applications?, What changes would you like to make if you were to make the application again?, Do you think that education should be supported with AR applications in the future? (If so, why it should be supported?)"* can be given.

Implementation Process and Data Collection

In the study, the experimental application was made in the *cell* unit of the ninth grade biology course. At the beginning of the application, the achievement test prepared by the researcher and the biology lesson motivation questionnaire were given to the experimental and control group students as a pre-test. The experimental process was started with the subject of *cell and organelles* subject of the *cell* unit. The biology course, in which the experimental application is made, is a 3-hour course per week and is mostly taught in the classroom environment. The experimental group students were informed by the researcher about the definition of *AR* and mobile and computer-assisted AR applications before the application. AR mobile application was installed on the phones of the experimental group students and personal accounts were created for each of them. The students acquired visuals on the subject *cells and organelles*, and shot short videos describing the functions and properties of *cells and organelles*. In order to combine images and videos, they opened the relevant application on computers in the computer laboratory and logged into the program with their user accounts. Then they combined the pictures and videos they uploaded to the application. Finally, they introduced the pictures using the application on their mobile devices and played the videos they had uploaded in this way. While these procedures were carried out for a total of 18 lesson hours for 6 weeks, 3 lessons per week, in the experimental group; in the control group, the lessons continued in line with the current program, and the subjects were explained by the teacher.

As soon as the experimental application was completed, the achievement test and the biology lesson motivation questionnaire were applied to the experimental group and the control group as a post-test. On the other hand interviews were held with 11 students in the experimental group; at the beginning of the interviews, the purpose, importance, scope of the research and where the data obtained from the research will be used were explained to the participants; in order to facilitate the breakdown and analysis of the interview data and to prevent data loss, it was explained to the participants that the interview could be recorded with audio, and their verbal consent was obtained in this regard.

Analysis of Data

The quantitative data of the study were analyzed through IBM SPSS Statistics 24 program. For the statistical techniques to be applied in the study, the Shapiro-Wilk test was applied in order to determine whether the distribution of the measurements of the dependent variables is normal or not, taking into account the group size smaller than 50, and the test results are given in Table 3.

Table 3. Tests of normality

		Shapiro-Wilk		
	Group	Statistic	df	p
Academic achievement pre-test	Experiment	.905	17	.083
	Control	.898	21	*.032*
Academic achievement post-test	Experiment	.883	17	*.036*
	Control	.917	21	.077
Intrinsic motivation pre-test	Experiment	.952	17	.496
	Control	.886	21	*.019*
Intrinsic motivation post-test	Experiment	.887	17	*.041*
	Control	.950	21	.345
Extrinsic motivation pre-test	Experiment	.929	17	.210
	Control	.939	21	.209
Extrinsic motivation post-test	Experiment	.930	17	.219
	Control	.915	21	.070
Interest in learning biology pre-test	Experiment	.958	17	.587
	Control	.975	21	.842
Interest in learning biology post-test	Experiment	.939	17	.312
	Control	.936	21	.183
Responsibility in learning biology pre-test	Experiment	.928	17	.199
	Control	.921	21	.090
Responsibility in learning biology post-test	Experiment	.906	17	.085
	Control	.845	21	*.004*
Confidence in learning biology pre-test	Experiment	.896	17	.058
	Control	.951	21	.348
Confidence in learning biology post-test	Experiment	.929	17	.211
	Control	.899	21	*.033*
Anxiety in biology exams pre-test	Experiment	.930	17	.219
	Control	.927	21	.118
Anxiety in biology exams post-test	Experiment	.899	17	.065
	Control	.931	21	.147
Total motivation pre-test	Experiment	.937	17	.286
	Control	.943	21	.247
Total motivation post-test	Experiment	.938	17	.298
	Control	.938	21	.195

In the analysis, the statistical (null) hypothesis was established as "the distribution of scores does not differ significantly from the normal distribution", so the calculated p value was higher than α = .05, it was interpreted as that the scores at this significance level did not show excessive deviation from the normal distribution and are appropriate (Büyüköztürk, 2007). Looking at Table 3; it can be said that most of the pretest and posttest measurements show a normal distribution. Based on this data, independent samples *t* test from parametric statistical techniques was used to determine whether there was a significant difference between the scores of the experimental group students and the control group students from the academic achievement test and the motivation questionnaire. Six different points less than α = .05 were neglected considering the group size greater than 15.

The qualitative data obtained by the interview method were analyzed with the descriptive analysis technique. Descriptive analysis is the lowest level and simplest form of analysis in which data is shown, described, illustrated and explained as it is (Sönmez & Alacapınar, 2011). Descriptive analysis is mostly used in research where the conceptual structure of the research is clearly determined beforehand (Yıldırım & Şimşek, 2006). Since the conceptual framework and themes were determined at the beginning of the research process, the descriptive analysis method was preferred. The questions in the interview form were also effective in determining the themes. In this study, the interview data were recorded audibly and these data were listened to by the researcher on the day of the interview and recorded in the word processing program. Relevant themes were determined for descriptive analysis, the data obtained were processed according to the thematic framework and presented with direct quotations. The following validity and reliability studies were conducted in the quantitative and qualitative dimensions of the research:

1. In the study, in which quantitative and qualitative data collection methods were used together, the researchers became a natural part of the research process by spending time in the field and conducting direct interviews with the participants, both in the experimental process and the interview process. Therefore, in this study, researchers play a participant role. The closeness of the researcher to the participants increases the validity (Creswell, 2017).
2. The processes followed in preparing the achievement test and interview form used in the research are explained under the heading of data collection tools.
3. The research process was enriched with qualitative and quantitative data collection techniques. Different data collection techniques were used together to increase the credibility of the research.
4. The qualitative findings of the research were clearly presented, and participant confirmation was obtained for the data collected at the end of the one-by-one interviews with the students. The findings obtained from the interviews were presented through direct quotations to increase the transferability, and the opinions of the researchers were reflected in the interpretation stage after the data were collected and analyzed.
5. The analysis results of the collected data were examined by other researchers and experts.
6. The research process was explained in detail.

Findings

Findings Regarding the First Hypothesis of the Study

Table 3 shows the independent samples t-test results, which were conducted to determine whether the posttest mean scores of the experimental group students in which the AR applications were applied and the control group students in which the current program was applied showed a significant difference.

Table 3. The *t* test results of the students' academic achievement posttest mean scores according to the group

Group	n	x̄	Sd	df	*t*	*p*
Experiment	17	28.06	2.11	36	17.922	0.000*
Control	21	14.86	2.37			

*p≤..05.

When Table 3 is examined, it is seen that the posttest mean scores of the experimental group students are significantly higher than the posttest mean scores of the control group students ($t_{(36)}$=17,922; p≤0.05). Based on this finding, it can be said that AR applications are effective in achieving academic achievement in the ninth grade biology cell unit compared to the current application.

Findings Regarding the Second Hypothesis of the Study

Table 4 shows the independent samples *t* test results of the mean scores obtained from the biology lesson motivation questionnaire of the experimental group students in which AR applications were applied and the control group students in which the current application was applied. When Table 4 is examined, there is no significant difference between the mean scores of the experimental group students and the control group students in the motivation questionnaire for the biology lesson in terms of "intrinsic motivation, extrinsic motivation, interest in learning, responsibility in learning, trust in learning" dimensions and total score. However, a significant difference was found in favor of the experimental group in terms of "anxiety in exams" ($t(36)=2.593$; $p\leq 0.05$). Based on this finding, it can be said that AR applications performed in the ninth grade biology lesson cell unit increase students' anxiety about exams compared to the current curriculum.

Table 4. The *t* test results of the students' motivation posttest mean scores according to the group

Factor	Group	n	$\bar{x}$	Sd	df	*t*	*p*
Intrinsic motivation	Experiment	17	20.77	3.21	36	1.272	0.211
	Control	21	19.24	4.01			
Extrinsic motivation	Experiment	17	22.29	1.57	36	1.344	0.187
	Control	21	21.14	3.23			
Interest in learning	Experiment	17	19.00	4.05	36	0.733	0.468
	Control	21	17.86	5.29			
Responsibility for learning	Experiment	17	20.65	2.47	36	0.648	0.521
	Control	21	19.86	4.50			
Confidence in learning	Experiment	17	19.65	2.62	36	0.730	0.470
	Control	21	18.67	5.00			
Anxiety in exams	Experiment	17	20.06	2.93	36	2.593	0.014*
	Control	21	16.86	4.35			
Total	Experiment	17	122.41	11.45	36	1.688	0.100
	Control	21	113.62	18.82			

* $p\leq..05$.

Findings regarding the Sub-problem related to the Qualitative Dimension of the Research

While the research findings related to the interviews with the participants were given, the common opinions of the participants were collected and presented together and supported by direct quotations.

Table 5. Qualitative findings of the research

Themes	Codes
Positive aspects (S1, S2, S4, S7)	It includes visuality, reflects reality, is practical, enjoyable and entertaining
Negative aspects (S4, S6, S7)	Taking time Not taking notes
Contribution to learning (S1, S2, S3, S4, S6, S7, S8, S9, S10, S11)	Ensuring learning by doing and experiencing, being permanent, increasing level of learning and course achievement, improving academic self-confidence
Effect on study methods	Changing study method, working with visual elements, seeing the subject more detailed, visuals draw attention, no change
Contribution to the interest and motivation for the course (S1, S2, S3, S4, S8)	Destroying the prejudices about the course, increasing interest and motivation, providing excitement and happiness, getting bored, decreasing motivation
Its effect on the classroom environment, teacher-student and student-student relations (S2, S3, S6, S7, S8, S11)	Helping, sharing, exchanging ideas, teamwork, a friendly environment
Opinions on the software used (S1, S3, S10, S11)	Being different and entertaining, providing access to information in a different way,
(S2, S6, S8)	Difficult to use, slow
Suggestions (S1, S3, S4, S7, S8, S9, S11)	Application of biology course in different subjects, physics, chemistry and geography courses.
(S4, S6)	Giving more time
(S2, S3, S7, S11)	Similar practices in other schools, supporting learning by doing and experiencing

Students participating in the research, regarding the positive aspects of AR applications; they stated that it reflects the reality, that it is enjoyable and entertaining. Examples of student views on this issue are presented below:

> S1; I liked it, it was pretty good, I liked that it was hands-on and visual.
> S2; I liked that the app is visual and fun.
> S4; It was a great practice for me to not get bored in the lesson.
> S7; I really liked that the app reflects reality.

Students, regarding the negative aspects of AR applications; they expressed the opinion that not taking notes and taking time. Examples of student views on this issue are presented below:

> S4; Using the app was taking time.
> S6; Applications took a long time.
> S7; It was fun, but it was better when we took notes, and we didn't take notes here.

Regarding the contribution of AR applications to learning; they stated that they learn by doing and living, applications contain visuals, and they provide easy and permanent learning. Examples of student views on this issue are presented below:

> S1; The class was more fun. We learned the subject better. We don't like to memorize. I think I am and will be more successful with AR applications.
> S2; I remembered what I learned more easily because my visual intelligence was better.
> S3; It was more permanent and fun to practice rather than reading and writing.
> S4; The fact that it was hands-on rather than narrative helped us recognize and visually recall organelles.
> S6; Because we did it ourselves, we learned better by seeing all the finer details.
> S7; Using and applying technology has been more effective.
> S8; I learned better thanks to the practical teaching of the course subject rather than being taught by writing it down.
> S9; It had an effect on me learning better and remembering easily.
> S10; It contributed to my achievement in the biology course, my self-confidence increased.
> S11; We started to believe that we could do better on the exams.

A few of the students stated that AR applications had an effect on their study methods and they understood better visually. Some students stated that although the application attracted attention with its visual side, it still did not change their study methods. Examples of student opinions are presented below:

> S8; It changed the way I study, now I started working with visual elements.
> S10; My study activities have changed.
> S11; While I used to perceive the cell as simple, with this application I changed my way of working because I saw that the cell and its organelles are more detailed.
> S4; My study method hasn't changed, but the visuals and pictures have caught my attention.

The students who participated in the research stated that AR applications increased their interest and motivation towards the lesson and reduced their prejudices about the biology lesson. One participant stated that his motivation decreased due to the long duration of AR applications. Examples of student opinions are presented below:

> S2; I started coming to biology classes more excited and happy.
> S3; The studies really caught my attention and it changed my perspective on biology.
> S4; With the app, my interest in biology grew.
> S8; With this study, my prejudice to biology was destroyed and my motivation increased.
> S1; I got bored with the applications because it took so long and my motivation dropped.

The students stated that the AR application increased the sharing and exchange of information, and made the classroom environment more friendly. Student opinions regarding this are given below:

> S2; There was a more friendly atmosphere with our teacher and our friends.
> S3; Our teacher helped us a lot because it was the first time we did such a practice, and sometimes teamwork was effective.

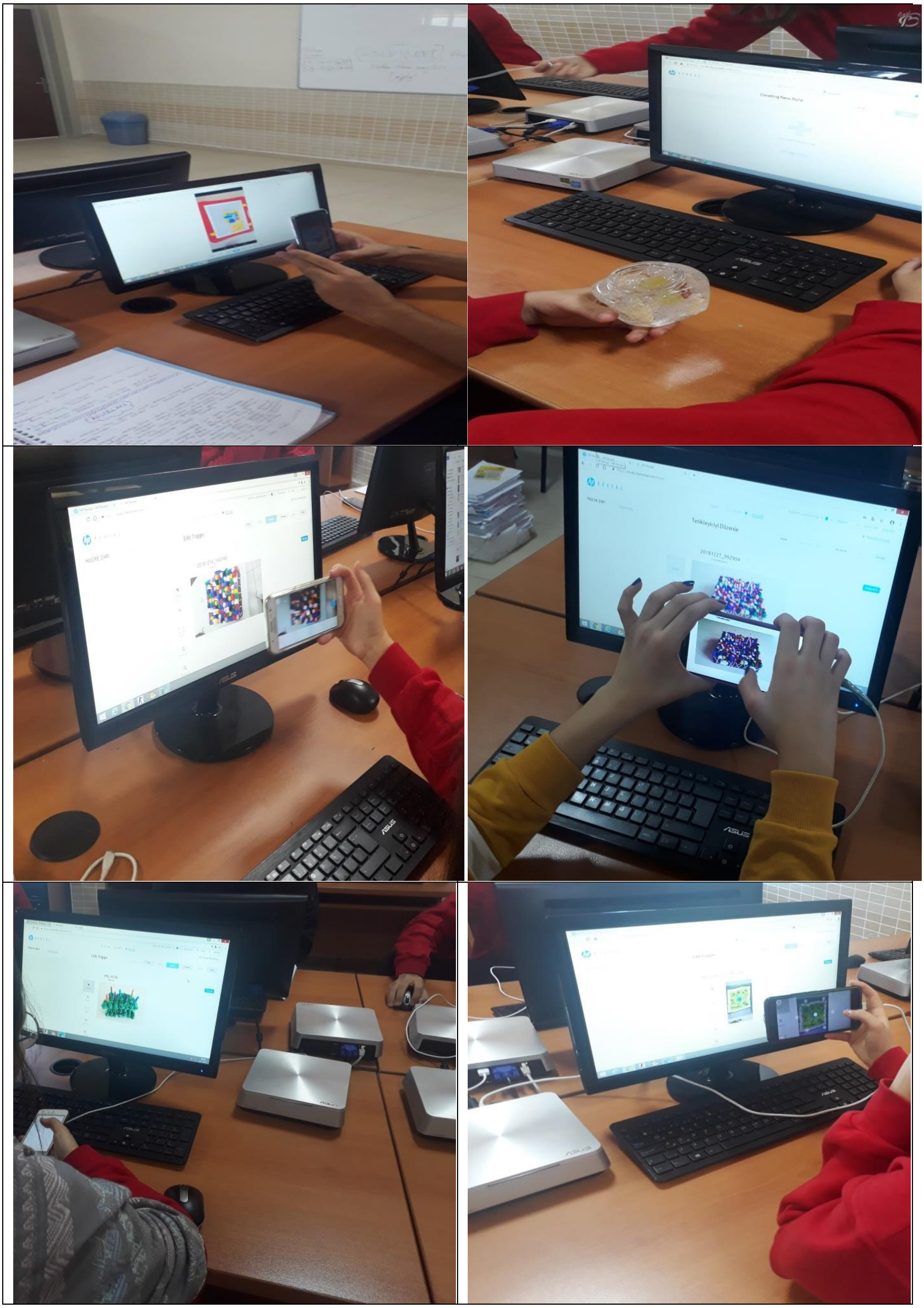

Figure 1. Visuals of AR applications

S6; We became more sincere with our teacher and our sharing with our friends increased.
S7; We exchanged ideas with our friends.
S8; We learned by discussing organelles with our friends.
S11; We got help from our teacher and helped and socialized with our friends.

While some of the students participating in the research emphasized that AR software is fun, reflects reality and has easy access to information; others stated that they had difficulties in installing and using the application and that this application should be developed. Examples of student opinions are presented below:

S1; Although I struggled a bit, it was a good application.
S3; It was a different application, more reflective of the virtual environment to us.
S10; I liked it very much because it is an application where I can access information very easily.
S11; I found the app fun and very beautiful.
S2; I had a very difficult time installing and using the software.
S6; I had a hard time introducing pictures and uploading videos. It would be better if the application was developed.
S8; It would be nice if the app was faster.

The students participating in the research stated that it would be good to make AR applications in other units of the biology course, in physics, chemistry and geography courses. Student opinions regarding this are given below:

S1; Because it is 3D, we remembered it. I would like to do it in other lessons as well.
S3; I learned the cell unit very well. If we did it in physics and chemistry classes, I would learn very well.
S4; It would be great if we could do in geography class. I would like it to be embodied in this way, as other biology subjects also remain abstract.
S7; I cannot understand many subjects in physics class, I would like to do this application in physics class as well.
S8; I would love to do it in chemistry and geography classes.
S9; This technology should be used in all courses.
S11; I would like to do it in different units as well. It should also be applied in physics and chemistry classes.

Students made various suggestions regarding AR applications. Some of these recommendations are presented below:

S2; "I think that students should be active with these and similar practices and should be supported for learning by doing."
S3; "It would be more comfortable if the application was done in more class hours. It is an application where we will be more successful by learning through the concretization of abstract concepts."
S7; "This practice should definitely be done in other schools as well."
S11; "AG studies should be supported in education and in our country."

The students' views on AR applications are stated above. In Figure 1 above, visuals of AR applications are given.

Discussion and Conclusion

When the findings related to the first hypothesis of the research are examined, it was seen that the achievement posttest point averages of the experimental group students were significantly higher than the posttest averages of the control group students. Based on this finding, it can be said that AR applications are effective in achieving academic achievement in the ninth grade biology cell unit compared to the current curriculum. This finding can be interpreted as the use of AR supported teaching materials in science education will provide positive learning outcomes. This finding shows parallelism with the qualitative findings of the study. The participants of the study stated that AR applications enable learning by doing and experiencing, increase the level of learning and academic success, that what they learn in this way is permanent and their academic self-confidence improves. In addition, the students evaluated the applications positively in terms of including visuality, reflecting the truth, being enjoyable and entertaining, increasing interest and motivation, providing excitement and happiness, and destroying the prejudices about the lesson. All these factors can be seen as factors that have positive effects on academic achievement. When the relevant literature is examined, it has been seen that similar results have been reached in many studies (Abdüsselam & Karal, 2012; Alınlı & Yazıcı, 2020; Bacca et al., 2014; Eroğlu, 2018; Ersoy et al., 2016; Girgin, 2018; Gün & Atasoy, 2017; Ibáñez et al., 2014; İbili, 2013; Kırıkkaya & Şentürk, 2018; Küçük et al., 2014; Özbek & Ak, 2020; Özdemir & Özçakır, 2017).

When the findings related to the second hypothesis of the research are examined, no significant difference was found between the motivation total score averages of the experimental group and the control group students and the sub-dimension averages of intrinsic motivation, extrinsic motivation, interest in learning, trust in learning and responsibility in learning. Affective variables such as attitude, motivation, academic success, etc. can change over a longer period of time compared to other variables. The 6-week experimental period may not be enough time to see a significant difference in motivation scores. On the other hand, the experimental group students stated that AR applications increased their interest and motivation towards the lesson, and thus they came to the lessons more enthusiastically and excitedly. When the studies on AR applications are examined, it can be seen that the effective use of new technologies such as AR in education can affect students' imaginations, improve their creativity, and increase their motivation and attitudes towards the lesson (Arıcı, 2013; Ersoy, Duman & Öncü, 2016; Uluyol & Eryılmaz, 2014). Many studies on the use of technology in learning activities suggest that learning environments using AR can increase learning motivation and effectiveness (Huang et al., 2016). Since AR environments have the ability to attract students' attention easily, students can adapt to the subject they are interested in more easily (Winkler et al., 2002). In this context, AR technology can offer new opportunities to promote learning.

In this study, it was determined that the motivation questionnaire "exam anxiety" sub-dimension score averages were higher in the experimental group, in other words, students who performed AR applications approached exams with anxiety. Exam anxiety is defined as intense anxiety that prevents the previously learned knowledge from being used effectively during the exam and leads to a decrease in success (Sarıkaya & Gemalmaz, 2021). The fact that the student stands out with her study feature, keeps her grades high and fulfills what is asked of her in the lessons causes her to be considered academically successful (Kaya et al., 2012). The students' encounter with a new application may have caused them to think that they could face a new situation in the exams. Or, because they have studied in a different way than before, they may not be aware of how much they have learned and may have doubts about using the information they have learned in the exam. Although the students generally liked this application and stated that it facilitated their learning and increased the permanence of what was learned, it can be said that this application did not relieve their worries about the exams.

When the findings regarding the sub-problem related to the qualitative dimension of the research are examined, students participating in the research; stated that they like biology lessons in which AR applications are carried out; they found the applications different, enjoyable and entertaining, and the applications helped to embody abstract concepts and make the subjects more memorable; They stated that cooperation between the students and the teacher increased in the classroom and learning by experience took place. Students, who stated that their interest and motivation towards the lesson increased thanks to the applications, suggested that these and similar applications be used in different units and subjects of the biology lesson, in other fields of science such as physics and chemistry, and in the geography lesson. The results of similar studies in the literature largely support the findings obtained from the interviews with the students in this study. The findings of the researchers, who determined that AR applications embody abstract concepts, increase attention and motivation, increase student participation by creating a learning environment by doing, provide learning with fun, improve interaction and cooperation in the classroom, are in line with the results of this study (Huang et al., 2016; Ivanova & Ivanov, 2011; O'Brien & Toms, 2005; Squire et al., 2008; Walczak et al., 2006; Wei et al., 2015; Wojciechowski & Cellary, 2013; Yoon et al., 2012).

Today, many researches on educational technologies show a tendency towards emerging technologies. Educational applications developed for use on desktop computers, tablets and mobile phones enrich the imagination of students with the help of various multimedia elements brought together, embody abstract concepts and add excitement to learning (Timur & Özdemir, 2018). Using the ever-evolving technology, and presenting the lessons by making use of technological devices, is of great importance for students to better understand the subject and to be interested in the lesson. In this way, students can come to the lesson more willingly, concentrate on the lesson and lose their interest later than in a normal lesson. In addition, such practices, in which contemporary techniques are used instead of classical methods to prepare a course material, can provide new perspectives on education (Avcı & Taşdemir, 2019). In line with these views, it can be said that an innovative technology such as AR can bring a new breath to the teaching of subjects that are difficult in the field of science. Nowadays, when we can see a wide variety of applications of AR in education, the ability to connect reality and digital content is constantly evolving and offers more options for teachers and students day by day (Girgin, 2018).

If another result of the research is mentioned; some of the students stated that the applications take time, they have difficulties in some stages and the software needs to be developed. The inadequacy of students' knowledge and experience about technological applications and instructional software may cause these opinions and misconceptions about the effectiveness of the application. In addition to these, the fact that students who meet a new application are more active by getting rid of the traditional ways they are used to may have caused this result and some difficulties.

Developments in information and communication technologies offer us new opportunities in the field of education. AR technology can make classrooms more engaging and knowledge more applicable. A large proportion of teenagers today own smartphones, and many of them are active smartphone users who use these tools to play games and connect with friends. A much smaller proportion of teens use their phones to do homework, gather information on a topic, etc. uses for educational purposes. When smartphones and AR are combined for education, students can be given extra digital information about any subject in various ways and complex information can be made easier to understand (Girgin, 2018).

Recommendations

It has been determined that biology lessons taught with AR activities are more effective in increasing academic achievement than the lessons taught with the activities of the current program. For this reason, it is thought that including AR applications in curricula and lesson plans, focusing on the use of activities with AR applications in the lessons will increase the academic achievement of the students and a high level of efficiency will be obtained from the teaching activities carried out in this way. The important thing here is to ensure that AR applications are suitable for the level of students. It can be said that the implementation of activities supported by AR in educational institutions will bring a different perspective in terms of the integration of technology into education. It was concluded that there was a significant difference in favor of the experimental group in the "anxiety in exams" sub-dimension of the motivation questionnaire, and the AR activities applied to the experimental group students caused the students to worry about the exams. It is normal for students who are introduced to a new application to worry about what they will do and what they will encounter in the exams. It is thought that such anxieties of students will decrease with the use of similar technological applications more in lessons.

When the findings about the challenging aspects of these activities of the experimental group students, in which AR applications were made, were examined, it was concluded that some students had problems with insufficient lesson hours and long applications. In this case, it can be suggested that the time allocated to AR applications should be arranged in line with student opinions. In addition to the fact that the mobile application used in the research is largely useful, it has been revealed that it perceives the pictures difficult and there are problems in loading them. It is thought that the mobile application should be developed in order to avoid these problems in new applications to be made in the future. Students who like to learn with activities based on AR applications stated that they want to use this application in physics, chemistry and geography lessons. In this direction, it can be suggested that AR applications should be used more in other lessons, especially in lessons that students find abstract and have difficulties. In this study, the effect of activities based on AR applications on student achievement and motivation in biology lesson was tried to be measured. Researchers who will study on this subject may be advised to investigate the effects of similar practices on variables such as interest, attitude, and permanence of what has been learned.

Scientific Ethics Declaration

The authors declare that the scientific ethical and legal responsibility of this article published in JESEH journal belongs to the authors.

Acknowledgements or Notes

This article was produced from the master's thesis presented by Esra Omurtak at Karamanoğlu Mehmetbey University Social Sciences Institute.

References

Abdüsselam, M. S. (2014). A study of art and craft teaching programs according to their function of generating "ındustrial design education" qualification for teacher candidates. *Pegem Journal of Education and Instruction, 4*(1), 59–74. http://doi.org/10.14527/pegegog.2014.004

Abdüsselam, M. S., & Karal, H. (2012). Fizik öğretiminde artırılmış gerçeklik ortamlarının öğrenci akademik başarısı üzerine etkisi: 11. Sınıf manyetizma konusu örneği. *Eğitim ve Öğretim Araştırmaları Dergisi, 1*(4), 170-181. http://jret.org/FileUpload/ks281142/File/20a.abdusselam.pdf

Alınlı, C., & Yazıcı, F. (2020). The effect of augmented reality practices on students' academic achievement in grade 8th T.R. The history of revolution and Kemalism course and attitudes of students towards augmented reality practice. *International Journal of Turkish Educational Studies, 8*(15), 99-113. https://doi.org/10.46778/goputeb.780344

Andersen, P. (2007). *What is Web 2.0?: ideas, technologies and implications for education* (Vol. 1). JISC Bristol.

Arıcı, V. A. (2013). A study on 3d-virtual reality in science education programs: "Solar system and beyond: Space puzzle" Unit sample (Unpublished master's thesis). Adnan Menderes University.

Augment. (2014). http://augmentedev.com/

Avcı, A. F., & Taşdemir, Ş. (2019). Periodic table teaching with augmented and virtual reality. *Journal of Selcuk-Technic, 18*(2), 68-83. http://sutod.selcuk.edu.tr/sutod/article/view/468

Azuma, R. (1997). A survey of augmented reality. *Presence-Teleoperators and Virtual Environments, 6*(4), 355- 385. https://doi.org/10.1162/pres.1997.6.4.355

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE Computer Graphics and Applications, 21*(6), 34–47. http://doi.org/10.1109/38.963459

Bacca, J., Baldiris, S., Fabregat, R., & Graf, S. (2014). Augmented reality trends in education: a systematic review of research and applications. *Journal of Educational Technology & Society, 17*(4), 133. https://dugi-doc.udg.edu/bitstream/handle/10256/17763/AugmentedRealityByNcNd.pdf?sequence=1&isAllowed=y

Best, J. W., & Kahn, J. V. (2014). *Research in education* (10th Edt.). Pearson.

Bilgiç, H. G., Duman, D., & Seferoğlu, S. S. (2011). The characteristics of digital natives' and their effects of on the design of online environments. *Academic Informatics, 2*(4), 1-7.

Bressler, D. M. ve Bodzin, A. M. (2013). A mixed methods assessment of students' flow experiences during a mobile augmented reality science game. *Journal of Computer Assisted Learning, 29*(6), 505–517. https://doi.org/10.1111/jcal.12008

Bronack, S. C. (2011), The role of immersive media in online education. *Journal of Continuing Higher Education, 59*(2), 113–117. https://doi.org/10.1080/07377363.2011.583186

Buluş Kırıkkaya, E., & Şentürk, M. (2018). The impact of using augmented reality technology in the solar system and beyond unit on the academic achievement of the students. *Kastamonu Education Journal, 26*(1), 181-189. https://doi.org/10.24106/kefdergi.375861

Büyüköztürk, Ş. (2006). *Manual of data analysis for social sciences [Sosyal bilimler için veri analizi el kitabı].* Pegem Academy Publishing.

Cabero, J., & Barroso, J. (2016). The educational possibilities of augmented reality. *Journal of New Approaches in Educational Research, 5*(1), 44-50. https://doi.org/10.7821/naer.2016.1.140

Cai, S., Wang, X., & Chiang, F.-K. (2014), A case study of augmented reality simulation system application in a chemistry course. *Computers in Human Behavior, 37,* 31–40. https://doi.org/10.1016/j.chb.2014.04.018

Chen, C., & Wang, C.-H. (2015). Employing augmented-reality-embedded instruction to disperse the imparities of individual differences in earth science learning. *Journal of Science Education and Technology, 24*(6), 835–847. https://link.springer.com/content/pdf/10.1007/s10956-015-9567-3.pdf

Cheng, K.-H., & Tsai, C.-C. (2013). Affordances of augmented reality in science learning: Suggestions for future research. *Journal of Science Education and Technology, 22*(4), 449–462. https://link.springer.com/content/pdf/10.1007/s10956-012-9405-9.pdf

Chiang, T. H. C., Yang, S. J. H., & Hwang, G.-J. (2014). An augmented reality-based mobile learning system to improve students' learning achievements and motivations in natural science inquiry activities. *Educational Technology and Society, 17*(4), 352–365. https://www.jstor.org/stable/10.2307/jeductechsoci.17.4.352

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th edt.). Sage Publications.

Creswell, J. W. (2017). *Qualitative inquiry and research design: Choosing among five approaches* (4th edt.). Sage Publications.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd Edt.). Sage Publications.

Dede, Y., & Yaman, S. (2008). Fen öğrenmeye yönelik motivasyon ölçeği: Geçerlik ve güvenirlik çalışması. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi, 2*(1), 19–37. https://dergipark.org.tr/en/pub/balikesirnef/issue/3366/46483

Demirel, Ö. (2008). *Curriculum development in education (11th edt.). [Eğitimde program geliştirme].* Pegem Academy Publishing.

Di Serio, Á., Ibáñez, M. B., & Kloos, C. D. (2013). Impact of an augmented reality system on students' motivation for a visual art course. *Computers & Education, 68,* 586-596. https://doi.org/10.1016/j.compedu.2012.03.002

Ekici, G. (2009). Adaptation of the biology motivation questionnaire to Turkish. *Çağdaş Eğitim Dergisi, 34*(365), 6-15. https://toad.halileksi.net/sites/default/files/pdf/biyoloji-dersi-motivasyon-anketi-toad.pdf

Enyedy, N., Danish, J. A., Delacruz, G., & Kumar, M. (2012). Learning physics through play in an augmented reality environment. *International Journal of Computer-Supported Collaborative Learning, 7*(3), 347–378. https://10.1007/s11412-012-9150-3

Eroğlu, B. (2018). Evaluation of teaching astronomy concept with augmented reality application for secondary student (Unpublished master's tehesis) Karadeniz Technical University, Trabzon. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp ID: 494395

Ersoy, H., Duman, E., & Öncü, S. (2016). Artırılmış gerçeklik ile motivasyon ve başarı: Deneysel bir çalışma. *Journal of Instructional Technologies and Teacher Education, 5* (1), 39-44. https://dergipark.org.tr/en/pub/jitte/issue/25089/264804

Feiner, S. (2002), Augmented reality: A new way of seeing. *Scientific American, 286*(4), 48–55. https://www.jstor.org/stable/26059641

Fidan, M., & Tuncel, M. (2018, October). *Ortaokul fen eğitimine yönelik artırılmış gerçeklik uygulamalarının tasarımı.* Paper presented at the meeting of The 6th International Congress on Curriculum and Instruction, Kars, Turkey. https://www.kafkas.edu.tr/dosyalar/egitimYeni/file/epokkongresi.pdf

Furió, D., Juan, M.-C., Segui, I., & Vivó, R. (2015). Mobile learning vs. traditional classroom lessons: a comparative study. *Journal of Computer Assisted Learning, 31*(3), 189–201. https://doi.org/10.1111/jcal.12071

Girgin, S. (2018). The effect of augmented reality technology on the academic success of science course. *Journal of Social And Humanities Sciences Research (JSHSR), 5*(30) 4283-4290. https://www.researchgate.net/profile/Soenmez-Girgin/publication/330323405

Glesne, C. (2016). *Becoming qualitative researchers: An introduction* (5th edt.). Pearson.

Glynn, S. M., & Koballa, T. R. Jr. (2006). Motivation to learn in college science. In J. Mintzes and W. H. Leonard (Eds). *Handbook of college science teaching* (pp. 25-32). Arlington, VA: National Science Teachers Association Pres.

Glynn, S. M., Taasoobshirazi, G., & Brickman, P. (2007). Nonscience major learning science; a theoretical model of motivational. *Journal of Research in Science Teaching. 44,* 1088-1107. https://doi.org/10.1002/tea.20181

Glynn, S. M, Taasoobshirazi, G., & Brickman, P. (2009). Science motivation questionnaire: Construct validation with nonscience major. *Journal of Research in Science Teaching, 46*(2), 127-146. https://doi.org/10.1002/tea.20267

Greene, J. C. (2007). *Mixed methods in social inquiry.* Jossey-Bass.

Greenhow, C., Robelia, B., & Hughes, J. E. (2009). Learning, teaching, and scholarship in a digital age: Web 2.0 and classroom research: What path should we take now? *Educational Researcher, 38*(4), 246-259. https://doi.org/10.3102/0013189X09336671

Gümüş, K., & Boydaş, O. (2021). Examining the place of augmented reality applications in education. *The Journal of International Social Research, 14*(76), 321-330. https://eds.s.ebscohost.com/eds/pdfviewer/pdfviewer?vid=0&sid=01f2e2c4-ba5e-4d54-ab1d-8953e89861c7%40redis

Gün, E. T., & Atasoy, B. (2017). Artırılmış gerçeklik uygulamalarının ilköğretim öğrencilerinin uzamsal yeteneklerine ve akademik başarılarına etkisi. *Eğitim ve Bilim, 42*(191), 31-51.

Hayes, P., Joyce, D., & Pathak, P. (2004). Ubiquitous learning - an application of mobile technology in education. In L. Cantoni & C. McLoughlin (Eds.), *Proceedings of ED-MEDIA 2004--World Conference on Educational Multimedia, Hypermedia & Telecommunications* (pp. 1811-1816). Lugano, Switzerland: Association for the Advancement of Computing in Education (AACE). m https://www.learntechlib.org/primary/p/12724/.

Hsiao, K.-F., Chen, N.-S., & Huang, S.-Y. (2012). Learning while exercising for science education in augmented reality among adolescents. *Interactive Learning Environments, 20*(4), 331–349. https://doi.org/10.1080/10494820.2010.486682

Huang, T-C., Chen, C.-C., & Chou, Y.-W. (2016). Animating eco-education: To see, feel, and discover in an augmented reality-based experiential learning environment. *Computers & Education, 96,* 72-82. https://doi.org/10.1016/j.compedu.2016.02.008

Hwang, G.-J., Wu, P.-H., Chen, C.-C., & Tu, N.-T. (2016). Effects of an augmented reality-based educational game on students’ learning achievements and attitudes in real-world observations. *Interactive Learning Environments, 24*(8), 1895–1906. https://doi.org/10.1080/10494820.2015.1057747

Ibáñez, M. B., Di Serio, A., Villarán, D., & Kloos, C. D. (2014). Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness. *Computers & Education, 71,* 1-13. https://doi.org/10.1016/j.compedu.2013.09.004

İçten, T., & Bal, G. (2017). Review of recent developments and applications in augmented reality. *Gazi University Journal of Science, 5*(2), 111-136. https://dergipark.org.tr/en/download/article-file/840617

İzgi Onbaşılı, Ü. (2018). The effect of augmented reality practices on the attitudes of primary school students towards augmented reality practices and science motivation. Ege Journal of Education, 19(1), 320-337. https://doi.org 10.12984/egeefd.390018

Ivanova, M., & Ivanov, G. (2011). Enhancement of learning and teaching in computer graphics through marker augmented reality technology. *International Journal on New Computer Architectures and Their Applications, 1*(1), 176-184. https://www.researchgate.net/profile/Malinka-Ivanova/publication/230772782

İbili, E. (2013). *Development, implementation and assessment of the effect augmented reality on geometry teaching materials for geometry classes* (Unpublished doctoral dissertation). Gazi University. Ankara. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp ID: 349085

Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2014). *NMC Horizon Report: 2014 Higher Education Edition.* Austin, Texas: The New Media Consortium. https://cdc.qc.ca/pdf/2014-Horizon-Report-creative-commons-copy.pdf

Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., & Hall, C. (2016). *NMC Horizon Report: 2016 Higher Education Edition.* Austin, Texas: The New Media Consortium. https://www.learntechlib.org/p/171478/

Kamarainen, A. M., Metcalf, S., Grotzer, T., Browne, A., Mazzuca, D., Tutwiler, M. S., & Dede, C. (2013). EcoMOBILE: Integrating augmented reality and probeware with environmental education field trips. *Computers ve Education, 68,* 545–556. http://doi.org/10.1016/j.compedu.2013.02.018

Kapil, Y., & Roy, A. (2014). A critical evaluation of generation Z at workplaces. *International Journal of Social Relevance & Concern, 2*(1), 10-14. https://ijournals.in/wp-content/uploads/2017/07/IJOURNAL_CAMERA_READY.pdf

Kaufmann, H. (2003, February). *Collaborative augmented reality in education.* Paper presented at the meeting of the IMAGINA. Monte Carlo, Monaco. https://www.researchgate.net/profile/Hannes_Kaufmann/publication/2555518_Collaborative_Augmented_Reality_in_Education/links/0912f508031dc45254000000/Collaborative-Augmented-Reality-in-Education.pdf

Kaya A, Bozaslan H, & Genç G. The effect of university students' parental attitudes on their problem solving skills, social anxiety levels and academic achievement. *Dicle University Journal of Ziya Gökalp Education Faculty, 18,* 208-225. https://dergipark.org.tr/tr/pub/zgefd/issue/47947/606639

Kerawalla, L., Luckin, R., Seljeflot, S. ve Woolard, A. (2006). Making it real: exploring the potential of augmented reality for teaching primary school science. *Virtual Reality, 10*(3–4), 163–174. https://doi.org/10.1007/s10055-006-0036-4

Kırıkkaya, E. B., & Şentürk, M. (2018). Güneş sistemi ve ötesi ünitesinde artırılmış gerçeklik teknolojisi kullanılmasının öğrenci akademik başarısına etkisi. *Kastamonu Eğitim Dergisi, 26*(1), 181-189. https://doi.org/10.24106/kefdergi.375861

Klopfer, E., & Squire, K. (2008), Environmental detectives: the development of an augmented reality platform for environmental simulations. *Educational Technology Research and Development, 56*(2), 203–228. https://link.springer.com/content/pdf/10.1007/s11423-007-9037-6.pdf

Kozma, R., & Anderson, R. (2002). Qualitative case studies of innovative pedagogical practices using ICT. *Journal of Computer Assisted Learning, 18*(4), 387-394. https://doi.org/10.1046/j.0266-4909.2002.00250.doc.x

Küçük, S., Yılmaz, R. M., Baydaş, Ö., & Göktaş, Y. (2014). Ortaokullarda artırılmış gerçeklik uygulamaları tutum ölçeği: geçerlik ve güvenirlik çalışması. *Eğitim ve Bilim, 39*(176), 383-392. http://eb.ted.org.tr/index.php/EB/article/view/3590

Lee, K. (2012). Augmented reality in education and training. *TechTrends, 56*(2), 13-21. https://link.springer.com/content/pdf/10.1007/s11528-012-0559-3.pdf

Lin, T.-J., Duh, H. B.-L., Li, N., Wang, H.-Y., & Tsai, C.-C. (2013). An investigation of learners' collaborative knowledge construction performances and behavior patterns in an augmented reality simulation system. *Computers & Education*, *68,* 314–321. https://doi.org/10.1016/j.compedu.2013.05.011

Matcha, W., & Rambli, D. R. A. (2013). Exploratory study on collaborative interaction through the use of augmented reality in science learning. *Procedia Computer Science, 25,* 144–153. https://doi.org/10.1016/j.procs.2013.11.018

Miles, M, B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook.* (2nd edt.). Sage Publications.

Núñez, M., Quirós, R., Núñez, I., Carda, J. B., Camahort, E., & Mauri, J. L., (2008, July). *Collaborative augmented reality for inorganic chemistry education.* Paper presented at the meeting of the International Conference of Mathematics and Computers in Science and Engineering, WSEAS. https://www.researchgate.net/profile/Ricardo-Quiros/publication/234777537_Collaborative_augmented_reality_for_inorganic_chemistry_education/links/02e7e51836a411d74f000000/Collaborative-augmented-reality-for-inorganic-chemistry-education.pdf

Oblinger, D., & Oblinger, J. (2005). Is it age or IT: First steps toward understanding the net generation. *Educating the Net Generation*, *2*(1–2), 20. https://studentconduct.tamucc.edu/assets/IsItAge.pdf

O'Brien, H. L. & Toms, E. G. (2005, November). *Engagement as process in computer mediated environments.* Poster presented at ASIS&T, Charlotte, North Carolina.

Özbek, F., & Ak, Ş. (2020). The augmented reality in 4th grade Turkish lesson: Effect on achievement and motivation. *Kastamonu Education Journal*, *28*(4), 1668-1679. http://doi.org/ 10.24106/kefdergi.4003

Özdemir, D., & Özçekır, B. (2019). An analysis of the effects of augmented reality activities in teaching fractions on 5th grade students' math achievement and attitudes. *Adıyaman University Journal of Educational Sciences, 9*(1), 21-41. http://dx.doi.org/10.17984/adyuebd.495731

Patton, M. Q. (1990). *Qualitative research and evaluation methods* (3rd edt.). Sage Publications.

Perez-Lopez, D., & Contero, M. (2013). Delivering educational multimedia contents through an augmented reality application: A case study on its impact on knowledge acquisition and retention. *Turkish Online Journal of Educational Technology - TOJET, 12*(4), 19–28. https://files.eric.ed.gov/fulltext/EJ1018026.pdf

Prensky, M. (2001). Digital natives, digital immigrants part 1. *On the Horizon, 9*(5), 1-6. https://www.marcprensky.com/writing/Prensky%20-%20Digital%20Natives,%20Digital%20Immigrants%20-%20Part1.pdf

Rosenbaum, E., Klopfer, E., & Perry, J. (2006). On location learning: Authentic applied science with networked augmented realities. *Journal of Science Education and Technology, 16*(1), 31–45. http://doi.org/10.1007/s10956-006-9036-0

Sarıkaya, S., & Gemalmaz, A. (2021). Factors affecting test anxiety. *Jour Turk Fam Phy, 12*(2), 99-107. https://10.15511/tjtfp.21.00297.

Seferoğlu, S. S. (2006). *Instructional technologies and material design [Öğretim teknolojileri ve materyal tasarımı].* Pegem Academy Publishing.

Sırakaya, M. (2015). *Effects of augmented reality applications on students' achievement, misconceptions and course engagement* (Unpublished doctoral dissertation). Gazi University, Ankara. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp ID 419423

Singhal, S., Bagga, S., Goyal, P., & Saxena, V. (2012). Augmented chemistry: Interactive education system. *International Journal of Computer Applications*, *49*(15), 1–5. https//doi.org/10.5120/7700-1041

Shelton, B. E., & Hedley, N. R. (2002, September). *Using augmented reality for teaching earth-sun relationships to undergraduate geography students.* Paper presented at the meeting of the IEEE International Workshop, Darmstadt, Germany.

Shelton, B. E., & Stevens, R. (2004, June). *Using coordination classes to interpret conceptual change in astronomical thinking.* Paper presented at the meeting of the 6th International Conference for the Learning Sciences. Lawrence Erlbaum and Associates, Mahweh, NJ. https://doi.dx.org/10.22318/icls2004.634

Smaldino, S. E., Lowther, D. L., Russell, J. D., & Mims, C. (2019). *Instructional technology and media for learning* (12th edt.). Pearson.

Solak, E., & Cakir, R. (2015). Exploring the Effect of Materials Designed with Augmented Reality on Language Learners' Vocabulary Learning. *Journal of Educators Online, 12*(2), 50–72. https://files.eric.ed.gov/fulltext/EJ1068381.pdf

Somyürek, S. (2014). Gaining the attention of generation z in learning process: augmented reality. *Educational Technology Theory and Practice, 4*(1), 63-80. https://doi.org/10.17943/etku.88319

Sönmez, V., Alacapınar, F. (2011). *Illustrated scientific research methods [Örneklendirilmiş bilimsel araştırma yöntemleri].* Anı Publishing.

Squire, K. D., & Jan, M. (2007). Mad city mystery: Developing scientific argumentation skills with a place-based augmented reality game on handheld computers. *Journal of Science Education and Technology*, *16*(1), 5–29. http://doi.org/10.1007/s10956-006-9037-z

Squire, K. D., Jan, M., Matthews, J., Wagler, M., Martin, J., & Devane, B. (2008). *Wherever you go, there you are: Pace-based augmented reality games for learning.* http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=C82FB26748C7B6A2B0709358AB08AD7B?doi=10.1.1.173.821andrep=rep1andtype=pdf

Tian, K., Endo, M., Urata, M., Mouri, K. ve Yasuda, T. (2014). Multi-viewpoint smartphone AR-based learning system for astronomical observation. *International Journal of Computer Theory and Engineering*, *6*(5), 396–400. http://doi.org/10.7763/IJCTE.2014.V6.897

Timur, B., & Özdemir, M. (2018). Teachers' views on the use of augmented reality environments in science education. *International Journal of Turkish Education Sciences*, *6*(10), 62-75. https://dergipark.org.tr/tr/pub/fmgted/issue/61510/882023

Tuan, H.-L., Chin, C.-C., Shiehh, S.-H. (2011). The development of a questionnaire to measure students' motivation towards science learning. International Journal of Science Education, 27(6), 639-654. https://doi.org/10.1080/0950069042000323737

Uluyol, Ç. ve Eryılmaz, S. (2014). Examining pre-service teachers' opinions regarding to augmented reality learning. *Gazi University Gazi Faculty of Education Journal*, *34*(3), 403-413. https://dergipark.org.tr/en/download/article-file/77514

Ültay, N., & Ültay, E. (2020). A comparative investigation of the views of pre-school teachers and teacher candidates about STEM. *Journal of Science Learning*, *3*(2), 67-78. https://doi.org/ 10.17509/jsl.v3i2.20796

Vilkoniene, M. (2009). Influence of augmented reality technology upon pupils' knowledge about human digestive system: The results of the experiment. *Online Submission*, *6*(1), 36–43. https://files.eric.ed.gov/fulltext/ED503889.pdf

Walczak, K., Wojciechowski, R., & Cellary, W. (2006, November). *Dynamic interactive VR network services for education.* Paper presented at the meeting of the ACM symposium on virtual reality software and technology. https://doi.org/10.1145/1180495.1180552

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, *53*(4), 5- 23. https://link.springer.com/content/pdf/10.1007/BF02504682.pdf

Wang, X., Kim, M. J., Love, P. E., & Kang, S. C. (2013). Augmented reality in built environment: classification and implications for future research. *Automation in Construction*, *32*, 1-13 https://doi.org/10.1016/j.autcon.2012.11.021

Wang, H.-Y., Duh, H. B.-L., Li, N., Lin, T.-J., & Tsai, C.-C. (2014). An investigation of university students' collaborative inquiry learning behaviors in an augmented reality simulation and a traditional simulation. *Journal of Science Education and Technology*, *23*(5), 682–691. http://doi.org/10.1007/s10956-014-9494-8

Waterman, A. S. (2005). When effort is enjoyed: Two studies of intrinsic motivation for personally salient activities. *Motivation and Emotion*, *29*(3), 165–188. https://link.springer.com/content/pdf/10.1007/s11031-005-9440-4.pdf

Watson, D. M. (2001). Pedagogy before technology: Re-thinking the relationship between ICT and teaching. *Education and Information Technologies*, *6*(4), 251-266. https://link.springer.com/content/pdf/10.1023/A:1012976702296.pdf

Wei, X., Weng, D., Liu, Y., & Wang, Y. (2015). Teaching based on augmented reality for a technical creative design course. *Computers & Education*, *81*, 221-234. https://doi.org/10.1016/j.compedu.2014.10.017

Winkler T., Kritzenberge H., & Herczeg, M. (2002, June). *Mixed reality environments as collabrative and constructive learning spaces for elementary school children.* Paper presented at the meeting of the World Conference on Educational Multimedia, Hypermedia & Telecommunications, Colorado.

Wojciechowski, R., & Cellary, W. (2013). Evaluation of learners' attitude toward learning in ARIES augmented reality environments. *Computers and Education*, *68*, 570–585. https://doi.org/10.1016/j.compedu.2013.02.014

Wolters, C. A., & Rosenthal, H. (2000). The relation between students' motivational beliefs and their use of motivational regulation strategies. *International Journal of Educational Research*, *33*(7), 801–820. https://doi.org/10.1016/S0883-0355(00)00051-3

Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., & Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers and Education*, *62*, 41–49. http://doi.org/http://dx.doi.org/10.1016/j.compedu.2012.10.024

Yen, J.-C., Tsai, C.-H., & Wu, M. (2013). Augmented reality in the higher education: Students' science concept learning and academic achievement in astronomy. *Procedia - Social and Behavioral Sciences, 103,* 165–173. https://doi.org/10.1016/j.sbspro.2013.10.322

Yıldırım, A., & Şimşek, H. (2018). *Qualitative research methods in the social sciences [Sosyal bilimlerde nitel araştırma yöntemleri]* (117h edt.). Seçkin Publishing.

Yoon, S. A., Elinich, K., Wang, J. Steinmeier, C., & Tucker, S. (2012). Using augmented reality and knowledge-building scaffolds to improve learning in a science museum. *International Journal of Computer-Supported Collaborative Learning, 7*(4), 519-541. https://link.springer.com/article/10.1007/s11412-012-9156-x

Zhang, J., Sung, Y.-T., Hou, H.-T. ve Chang, K.-E. (2014). The development and evaluation of an augmented reality-based armillary sphere for astronomical observation instruction. *Computers and Education, 73,* 178–188. http://doi.org/10.1016/j.compedu.2014.01.00

Zhu, W., Owen, C., Li, H. & Lee, J.-H. (2004), Personalized in-store e-commerce with PromoPad: an augmented reality shopping assistant. *Electronic Journal for E-commerce Tools and Applications, 1*(3), 1-19. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.83.8198&rep=rep1&type=pdf

Author(s) Information

Esra Omurtak
Ministry of Education, Babaoğlu High School
Karaman/Turkey
Contact e-mail: esraomurtak@outlook.com
ORCID iD: 0000-0002-1855-4824

Gülçin Zeybek
Karamanoğlu Mehmetbey University
Karaman/Turkey
Contact e-mail: gozkan80@hotmail.com
ORCID iD: 0000-0002-5509-5129

Journal of Education in Science, Environment and Health
Volume 8, Issue 1, 2022 https://doi.org/10.21891/jeseh.1056832

Critical Thinking Dispositions as a Predictor for High School Students' Environmental Attitudes

Ali Orhan

Article Info

Article History

Published:
01 January 2022

Received:
04 August 2021

Accepted:
13 November 2021

Keywords

Environmental attitudes
Critical thinking

Abstract

This non-experimental quantitative study aimed to determine if high school students' critical thinking (CT) dispositions and environmental attitudes differ by gender, father, and mother's educational background and their CT dispositions are significant predictors of environmental attitudes. UF/EMI Critical Thinking Disposition Instrument and Environmental Attitude Scale were used to collect data in this study carried out with 139 high school students. It was found out that students' CT dispositions did not differ by gender, father, and mother's educational background. Also, while students' environmental attitudes differed by gender in favor of females, they did not differ by mother and father's educational background. Besides, engagement, maturity, and innovativeness were significant predictors of the students' environmental attitudes. A positive relationship was found between students' CT dispositions and environmental attitude scores and engagement, maturity, and innovativeness together explained 31% of the total variance on their environmental attitudes.

Introduction

With the increasing population, the world is facing many critical environmental problems such as deforestation, scarcity of water, climate change, etc. (Steg et al., 2014a). One of the main reasons for these problems is human interaction with the environment (Cardinale et al., 2012; Vlek & Steg, 2007), and changing people's attitudes and behaviors towards the natural environment can be seen as an important solution for these problems (Keniger et al., 2013; Steg et al., 2014b). Students are expected not only to be able to protect the environment but also solve the environmental problems (Fua et al., 2018) and positive environmental attitudes are important for this. Schultz et al. (2004) defines the environmental attitude as beliefs that have an impact on individuals' behavioral intentions regarding the activities related to environment. According to Gifford and Sussman (2012), environmental attitude has three components which are cognitive (thoughts and knowledge about the environment), affective (feelings about the environment), and behavioral (environment-friendly behavioral intentions and actions).

It is a well-established belief that attitudes and real world behaviors are connected to each other (Fishbein & Ajzen, 1975; Ajzen 1991). When individuals have positive attitudes, they can show positive behaviors in the real world. Some previous studies showed a strong relationship between environmental attitudes and environmental-friendly behaviors (Poortinga et al., 2004; Iverson & Rundmo, 2001). Therefore, studying environmental attitudes is important because they may determine behaviors (Gifford & Sussman, 2012). The existing studies highlight that many factors such as gender, educational background, age, and income (Lee et al. 2015; Blocker & Eckberg, 1989; Dlamini et al., 2021; Mohai, 1992; Zhang, 1994; Marquart-Pyatt, 2008; Kemmelmeie et al., 2002; Franzen & Vogl, 2013a) significantly impact individuals' environmental attitudes. Based on these previous studies, the literature indicates that females, more educated, younger, and economically better situated individuals have more positive environmental attitudes. Literature also suggests that individuals' political beliefs and value orientations are significant predictors of environmental attitudes (Poortinga et al., 2004; Steg et al. 2014b). While biospheric values positively affect environmental attitudes, individuals with egoistic values have less concern about environmental issues (De Groot & Steg, 2008; Steg et al., 2011). Also, levels of trust (Meyer & Liebe, 2010), engaging in nature-related outdoor activities (Hausbeck et al., 1992), and levels of tolerance and understanding (Milfont & Gouveia, 2006) are closely related to environmental attitudes. Besides, the previous studies has also identified a set of variables that can be related to environmental attitudes such as environmental quality, urbanization, and population density of the countries (Franzen, 2003; Franzen & Meyer, 2010; Franzen & Vogl, 2013b). Based on the previous literature, people who live in wealthier countries display higher levels of environmental attitude (Franzen, 2003; Diekmann & Franzen, 1999; Dunlap & York, 2008).

In this study, it was expected that critical thinking (CT) is another important variable that can affect individuals' environmental attitudes. Individuals with high CT skills or dispositions do not use these skills only during the learning process. They also use them at any moment of their daily life. Thanks to CT, individuals can realize the condition of their environment better and develop their curious attitude towards environmental problems. They also understand the environmental problems and investigate the obtained data to find the best way to solve this problem. Therefore, this makes CT a necessary step to preserve environmental sustainability (Puspitasari et al., 2016; Amin et al., 2020). Indeed, previous literature emphasizes attitudes towards renewable energy sources are influenced by CT dispositions of individuals (Lee, 2016; Opitz, 2016; Sakschewski et al., 2014; Güven & Çakır, 2019). Also, some previous studies indicate that CT teaching improves not only students' CT skills or dispositions but also their environmental attitudes (Hofreiter et al., 2007, Orhan, 2021; Muhsilin et al., 2019; Mengi Us, 2019). Therefore, previous literature shows a positive relationship between CT and environmental attitudes. Although there are studies investigating the relationship between the attitudes towards renewable energy sources and CT dispositions (Lee, 2016; Opitz, 2016; Sakschewski et al., 2014; Güven & Çakır, 2019), very few studies were encountered aiming to investigate the relationship between students' general environmental attitudes and CT dispositions in the literature. Therefore, the aim of this study was to investigate if high school students' CT dispositions are a significant predictor of their environmental attitudes. To this end, the following questions were sought:

1. Do high school students' CT dispositions and environmental attitudes significantly differ by their gender, father, and mother's educational background?
2. Are high school students' scores on UF/EMI Critical Thinking Disposition Instrument (CTDI) subscales significant predictors of their environmental attitudes?

Method

In this non-experimental quantitative study, a cross-sectional survey design was employed. High school students' environmental attitude was determined as the dependent variable and their scores on engagement, maturity, and innovativeness sub-dimensions were determined as predictor variables.

Study Group

The study group consisted of 139 students (76 female, 63 male) studying in different high schools in northern Turkey in the academic year of 2020-2021. The age of the students ranged between 13 and 16 (M=14.46, Mdn=14, Mo=14, SD=0.52). While Stevens (1996) suggests 15 participants for each predictor variable, Pallant (2001) recommends 40 participants for each predictor variable in multiple regression analysis. In this current study with 3 predictor variables which are engagement, maturity and innovativeness, it can be said that the sample size of 139 is enough. I also computed a-priori power analysis for conducting a linear multiple regression analysis before the study. The minimal sample size of students needed for this analysis (alpha=0.05; power=0.95; 3 predictors) to detect a medium effect size (f^2=0.15) based on the previous literature would be 119, based on Faul et al.'s (2007) G*Power 3 software program. Therefore, the acquired sample size of 139 in this study was very good.

Data Collection Tools

UF/EMI Critical Thinking Disposition Instrument (CTDI)

CTDI was used to determine students' CT dispositions in this study. CTDI was developed by Irani et al. (2007) and is composed of three sub-dimensions which are engagement (11 items), maturity (7 items), and innovativeness (7 items). After the Turkish adaptation study which was conducted with 342 ninth- and tenth-grade students by Kılıç and Şen (2014), one item was omitted from the instrument, and the remaining items were found to be consistent with the construct in the original instrument. Cronbach's alphas of instrument's sub-dimensions ranged from 0.70 to 0.88. For the total scale, Cronbach's alpha was found to be 0.89. Cronbach's alpha values calculated for this study were 0.87 for the engagement sub-dimension, 0.64 for the maturity sub-dimension, and 0.67 for the innovativeness sub-dimension. Furthermore, the reliability estimate for the total scale was calculated as 0.86.

Environmental Attitude Scale (EAS)

EAS developed by Uzun et al. (2019) was used to determine students' environmental attitudes. The reliability and validity studies were conducted with 1687 students. The scale is composed of three subscales which are the environmental behavior subscale (13 items), the environmental opinion subscale (11 items), and the environmental emotion subscale (16 items). Cronbach's alphas of the subscales ranged from 0.82 to 0.94. For the total scale, Cronbach's alpha was found to be 0.94. Cronbach's alpha values calculated for this study were 0.85 for the behavior subscale, 0.75 for the opinion subscale, and 0.87 for the emotion subscale. Furthermore, the reliability estimate for the total scale was calculated as 0.90.

Data Collection

After getting ethical committee approval from Zonguldak Bülent Ecevit University (No. 44858 dated 03.06.2021), the data for the current study were gathered in the spring term of 2020-2021 academic year. Privacy and confidentiality issues were shared with all students and they were informed about their right to withdrawal anytime they want. The instruments took approximately 20-25 minutes to complete.

Data Analysis

The collected data were analyzed in SPSS 20 statistical software. No missing data were observed as a result of a review for each variable, and normality was tested with Skewness and Kurtosis values. Z transformation was used to determine outliers per variable and Z-scores with values greater than 3.29 were accepted as potential outliers (Tabachnick & Fidell, 2012). Z-scores indicated no influential outliers should be excluded. Also, multivariate outliers were checked by using Mahalanobis Distance (Mahalanobis D^2) and no influential outliers were found to be excluded from the dataset. Pearson correlation coefficients, CI, VIF, and tolerance values between the predictor variables were investigated to make sure there is no high correlation among them. For the assumptions of MANOVA, multivariate normality was tested with the Henze-Zirkler test, and it was seen multivariate normality was presented in the data. Also, Box's M test results indicated that covariance between the groups was equal (Box's M test=200.080; $p>0.01$). Besides, Levene test results made for each variable to examine whether or not the variance between independent variable groups are equal showed that variance was equal between groups for engagement ($F=1.439$; $p>0.05$), maturity ($F=0.737$; $p>0.05$), innovativeness ($F=1.230$; $p>0.05$), behavior ($F=1.287$; $p>0.05$), opinion ($F=1.346$; $p>0.05$), and emotion ($F=1.630$; $p>0.05$). Independent samples t-test, MANOVA test, Pearson correlation, and multiple linear regression with enter method were used in this study.

Results

Descriptive statistics including means, standard deviations of all measured subscales of the EAS and CTDI instruments, and independent samples t-test results by gender are presented in Table 1 below.

Table 1. Descriptive statistics for the EAS and CTDI scales

	Total Sample (n=139)		Female (n=76)		Male (n=63)				
	$\bar{X}$	sd	$\bar{X}$	sd	$\bar{X}$	sd	$t_{(139)}$	p	Cohen's d
CTDI subscales									
Engagement	3.87	0.60	3.93	0.63	3.80	0.56	1.292	0.19	-
Maturity	3.82	0.51	3.84	0.50	3.81	0.53	0.317	0.75	-
Innovativeness	3.85	0.54	3.90	0.56	3.79	0.51	1.234	0.21	-
Total	3.85	0.48	3.89	0.50	3.80	0.45	1.113	0.26	-
EAS subscales									
Behavior	2.76	0.68	2.96	0.64	2.52	0.65	3.944	0.00	0.67
Opinion	4.57	0.38	4.64	0.28	4.49	0.47	2.119	0.03	0.36
Emotion	4.08	0.55	4.17	0.51	3.97	0.57	2.147	0.03	0.36
Total	3.80	0.42	3.92	0.38	3.66	0.42	3.781	0.00	0.64

As it can be seen in the descriptive statistics in Table 1, high school students had a positive and strong inclination towards engagement ($\overline{X}$=3.87), maturity ($\overline{X}$=3.82), and innovativeness ($\overline{X}$=3.85) subscales. Students also had strong CT dispositions ($\overline{X}$=3.85). Besides, while students had highly positive attitudes towards opinion ($\overline{X}$=4.57) and emotion ($\overline{X}$=4.08) subscales, they had moderately positive attitudes towards behavior subscale ($\overline{X}$=2.76). Also, it was found out that students' engagement subscale (t_{139}=1.292, p>0.05), maturity subscale (t_{139}=0.317, p>0.05), innovativeness subscale (t_{139}=1.234, p>0.05), and total CT disposition (t_{139}=1.113, p>0.05) scores did not differ by gender although females revealed higher scores for each subscale and total scores. However, the difference was found to be significant for behavior subscale (t_{139}=3.944, p<0.05), opinion subscale (t_{139}=2.119, p<0.05), emotion subscale (t_{139}=2.147, p<0.05), and total environmental attitude (t_{139}=3.781, p<0.05) scores. Female students had higher scores than male students for each EAS subscale and total scale. While gender variable had a small effect on students' opinion (d=0.36) and emotion (d=0.36) subscales scores, it had a medium effect on students' behavior subscale (d=0.67) and total EAS scores (d=0.64) based on Cohen's (1988) classification.

Table 2. MANOVA test results of students' CTDI and EAS subscales according to the educational background of father and mother

Effect		Value	F	Hypothesis df	Error df	p	η^2
Intercept	Wilks' Lambda (λ)	0,010	1884.925	6.00	114.00	0.00	0.990
Mother's EB	Wilks' Lambda (λ)	0,775	1.258	24.00	398.90	0.18	0.064
Father's EB	Wilks' Lambda (λ)	0.767	1.316	24.00	398.90	0.14	0.062
Mother's EB* Father's EB	Wilks' Lambda (λ)	0.614	0.891	66.00	615.45	0.71	0.078

Note: Mother's EG means mother's educational background; Father's EG means father's educational background

As it can be seen in Table 2, according to results of MANOVA test which is made to determine whether students' CTDI and EAS subscales scores differ according to their mother and father's educational background, it was found out that CTDI and EAS subscales scores did not differ according to students' mother's educational background variable (λ=0.775; $F_{(139)}$=1.258; p>0.05), father's educational background variable (λ=0.767; $F_{(139)}$=1.316; p>0.05), and mother's educational background*father's educational background variable (λ=0.614; $F_{(139)}$=0.891; p>0.05).

Table 3. MANOVA test results of students' CTDI and EAS subscales according to the educational background of father and mother

Source	Dependent variable	Sum. of squares	df	Mean square	F	p	η^2
Mother's EB	Engagement	2.677	4	0.669	1.949	0.10	0.061
	Maturity	0.616	4	0.154	0.583	0.67	0.019
	Innovativeness	2.301	4	0.575	2.200	0.07	0.061
	Behavior	5.008	4	1.252	3.189	0.01	0.097
	Opinion	0.211	4	0.053	0.360	0.83	0.012
	Emotion	0.356	4	0.089	0.350	0.84	0.012
Father's EB	Engagement	0.374	4	0.093	0.272	0.89	0.009
	Maturity	0.969	4	0.242	0.918	0.45	0.030
	Innovativeness	1.189	4	0.297	1.137	0.34	0.037
	Behavior	0.095	4	0.024	0.061	0.99	0.002
	Opinion	0.694	4	0.174	1.183	0.32	0.038
	Emotion	2.980	4	0.745	2.927	0.02	0.090
Mother's EB* Father's EB	Engagement	2.710	11	0.501	0.718	0.72	0.062
	Maturity	0.883	11	0.054	0.304	0.98	0.027
	Innovativeness	2.212	11	0.345	0.769	0.67	0.066
	Behavior	5.511	11	0.246	1.276	0.24	0.106
	Opinion	0.591	11	0.080	0.366	0.96	0.033
	Emotion	3.800	11	0.201	1.357	0.20	0.111

Note: Mother's EG means mother's educational background; Father's EG means father's educational background

Table 3 shows that students' engagement scores ($F_{(139)}$=2.710; p>0.05), maturity scores ($F_{(139)}$=0.883; p>0.05), innovativeness scores ($F_{(139)}$=2.212; p>0.05), behavior scores ($F_{(139)}$=5.511; p>0.05), opinion scores ($F_{(139)}$=0.591; p>0.05), and emotion scores ($F_{(139)}$=3.800; p>0.05) did not significantly differ according to students' mother's educational background*father's educational background variable. Therefore, it can be said that students' mother's and father's educational background together were not a significant variable that affects high school student's CTDI and EAS subscales scores.

Table 4. Inter-correlations among the subscales of CTDI and EAS

	Maturity	Innovativeness	Total CTDI	Behavior	Opinion	Emotion	Total EAS
Engagement	0.561**	0.757**	0.900**	0.490**	0.208*	0.398**	0.501**
Maturity	-	0.587**	0.810**	0.447**	0.072	0.432**	0.451**
Innovativeness	-	-	0.899**	0.502**	0.181*	0.416**	0.507**
Total CTDI	-	-	-	0.552**	0.180*	0.476**	0.559**
Behavior	-	-	-	-	0.297**	0.563**	0.874**
Opinion	-	-	-	-	-	0.245**	0.573**
Emotion	-	-	-	-	-	-	0.813**

Note: **correlation is significant at p< 0.01; *correlation is significant at p< 0.05

As can be seen in Table 4, most of the subscales of CTDI and EAS were found to be moderately to highly correlated to each other. However, there was not a significant correlation between maturity and opinion subscales scores (r=0.072; p>0.05). Also, high school students' total CTDI scores were moderately correlated to their total EAS scores (r=0.559).

Table 5. Multiple linear regression results between environmental attitude and engagement, maturity, and innovativeness scores

	B	Std. Error	β	t	p
Constant	1.932	0.248	-	7.804	0.00
Engagement	0.154	0.078	0.221	1.976	0.05
Maturity	0.159	0.074	0.195	2.161	0.03
Innovativeness	0.174	0.089	0.225	1.964	0.04

R=0.559, R^2=0.313, $F_{(3, 138)}$= 20.466, p<0.01

As can be seen in Table 5, the established regression model was significant ($F_{(3,138)}$=20.466, p<0.01). As a result of the multiple linear regression analysis, it was seen that engagement (β=0.221, $t_{(138)}$=1.976, p<0.05), maturity (β=0.195, $t_{(138)}$=2.161, p<0.05) and innovativeness (β=0.225, $t_{(138)}$=1.964, p<0.05) scores were significant predictors of high school students' environmental attitudes (R=0.559, R^2=0.313, p<0.01). Engagement, maturity, and innovativeness together explained 31% of the total variance on their environmental attitudes. Besides, innovativeness (β=0.225) was the significant predictor with the largest effect on students' environmental attitudes. It was followed by engagement (β=0.221) and maturity (β=0.195). The regression equation related to the model can be seen below:

Environmental attitude= 1.932 + 0.154*engagement + 0.159*maturity + 0.174*innovativeness.

Conclusion and Discussion

In this current study, it was aimed to determine if high school students' CT dispositions and environmental attitudes significantly differ by their gender, father, and mother's educational background and their scores on CTDI subscales are significant predictors of environmental attitudes. It was found out that students had a positive and strong inclination towards all three CT disposition subscales. Also, gender, father, and mother's educational background did not significantly affect high school students' CT dispositions. In some studies in the literature, it was found that CT dispositions did not significantly differ by gender (Liu et al., 2019; Khandaghi et al., 2011; Mahmoud & Mohamed, 2017; Polat, 2017; Özdemir, 2005; Walsh & Hardy, 1999; Facione et al., 1995), and mother and father's educational background (Polat, 2017; Özdemir, 2005; Şahin, 2018; Aybek & Ekinci, 2010). In a meta-analysis study by Çolak et al. (2019), it was found out that the effect of gender, mother and father's educational background variables on CT dispositions was very small. However, there are also some contradictory studies which concluded that CT dispositions significantly differed by gender (Ricket & Rudd, 2004; Kökdemir, 2003; Ferah, 2000; Tümkaya et al., 2009; Walsh, 1996; Rodriquez, 2000).

Therefore, it can be said that most of the studies, including this study, concluded that CT dispositions were not significantly predicted by gender and hence, gender was a variable that should not be considered in the explanation of CT dispositions. However, some other studies revealed that gender was a critical variable considering individuals' CT dispositions and hence, should be considered. Due to this inconsistency between the results of previous studies, it can be said that further research and analysis regarding the effect of gender on CT dispositions are necessary.

Also, this study revealed that while students had highly positive inclination towards opinion and emotion subscales, they had relatively low positive attitudes towards behavior subscale. In other words, although they had a strongly positive opinion and emotion towards the environment, they seemed to perform a few positive actions towards it. This result is confirmed by many previous studies (Mifsud, 2011; Rebolj & Devetak, 2013; Liu et al., 2015; Saka, 2016). The results of this study indicated that the students' environmental attitudes significantly differed by gender. Female students had higher scores than male students for each EAS subscale and total scale. There are some studies in the literature which indicated that environmental attitudes significantly differed by gender in favor of female participants (Yüksel & Yıldız, 2019; Gökçe et al., 2007; Tikka et al., 2000; Ekici, 2005; Duarte et al., 2017; Arnocky & Stroink, 2010; Pauw & Petegem, 2010; Coertjens et al., 2010). Also, in his meta-analysis study, Gökmen (2021) found out that gender had a small effect on environmental attitudes and females had significantly more positive environmental attitudes. Therefore, this result is consistent with the existing literature. This result may be attributable to the expectation that females are more interested and concerned about environmental problems than males (Spellmann et al., 2003; Mohai, 1992; Blocker & Eckberg, 1989). This expectation derives from the claim that while females are socialized to be family nurturers and caregivers, males are socialized to be family breadwinners and economic providers from their childhood (Mohai, 1992; Blocker & Eckberg, 1989). While this situation leads to develop a motherhood mentality for females, it develops a marketplace mentality for males (McStay & Dunlap, 1983). Therefore, the nurturing attitudes and the motherhood mentality of females which are the results of this socialization make them to have concerns about many social issues such as environmental pollution, scarcity of water, and poverty. However, because of their marketplace mentality, economic growth is the priority for males and that can cause environmental pollution (Blocker & Eckberg, 1989; McStay & Dunlap, 1983). Therefore, based on this argument (Blocker & Eckberg, 1989; Mohai, 1992; McStay & Dunlap, 1983), it is expected that females have more positive environmental attitudes than males and the results of this study justified this expectation. However, it can be said that this difference is based on gender roles and socialization (Zelezny et al., 2000). In this study, it was found out that students' mother and father's educational background did not significantly affect their environmental attitudes. This result is confirmed by earlier studies (Gökçe et al., 2007; Yavuz, 2019; Şahin, 2015; Şama, 2003; Sadık & Çakan, 2010; Gürbüz et al., 2007). Therefore, it can be said that the educational background of mother and father is not an important factor on environmental attitudes. Actually, it can be expected that not educational backgrounds but environmental attitudes of mother and father can significantly affect children's environmental attitudes (Ballantyne et al., 1998; Istead, 2009). In other words, when the mother and father have positive environmental attitudes, we can expect that the children will also have positive environmental attitudes.

According to another result obtained from the study, it was found that engagement, maturity, and innovativeness were significant predictors of the students' environmental attitudes. There was a positive relationship between engagement, maturity, innovativeness, and students' environmental attitude scores and it was determined that engagement, maturity, and innovativeness together explained 31% of the total variance on their environmental attitudes. This result may be attributable to the fact that individuals with high CT dispositions are eager to learn new things, curious about what is going on around them, resourceful, flexible, actively engaged in their surroundings, self-confident, being aware of own predispositions and biases, and able to modify their ideas in line with situational requirements (Facione et al., 1995). Also, they see the incidents as multidimensional rather than one-dimensional (Girot, 2000) and they have greater skills at making mature judgments (Facione et al., 1995). Therefore, it can be said that when the individuals have higher CT dispositions, they also have more positive environmental attitudes. Saka (2016) found out that CT dispositions significantly predicted environmental ethics approaches in her study which aimed to investigate the effects of CT dispositions on environmental ethics approaches. Also, Güven and Çakır (2019) found a significant relationship between the students' attitudes towards renewable energy sources and CT dispositions. There are also other studies which concluded that CT can affect individuals' thoughts, decisions, and attitudes towards renewable energy sources (Lee, 2016; Opitz, 2016; Sakschewski et al., 2014). Therefore, it can be said that this finding is confirmed by many previous studies. Also, previous experimental research (Hofreiter et al., 2007; Mengi Us, 2019; Muhsilin et al., 2019) indicated that CT teaching improves not only students' CT skills and dispositions but also their environmental attitudes which revealed that higher CT dispositions result in higher environmental attitudes.

Besides, this study revealed that innovativeness (β=0.225) was the significant predictor with the largest effect on students' environmental attitudes. It was followed by engagement (β=0.221) and maturity (β=0.195). According to Irani et al. (2007), innovativeness is about measuring the disposition of being eager to look for new knowledge. Individuals with a positive inclination towards innovativeness subscale are intellectually curious about new challenges and want to know more about what is going on around them. They, who are also described as "hungry learners", actively seek to know more about their life, and their world through research, reading, and questioning. Therefore, individuals with high dispositions towards innovativeness are highly interested in their environment, life, and the world. They enjoy being occupied with new challenges and obstacles such as environmental problems. Therefore, the result which indicated innovativeness was the significant predictor with the largest effect can be explained by this. Also, individuals with high dispositions towards engagement subscale value the reasoning process, are self-confident, and eager to find situations which she/he can use reasoning skills, and solve problems (Irani et al., 2007). Besides, According to Irani et al. (2007), individuals with a positive inclination towards maturity subscale are open-minded and need enough evidence and data before making a decision. They are also aware that most of the problems can have more than one possible solution and they know that their decisions can be influenced by their own prejudices, tendencies, experiences or other people. Therefore, individuals with high dispositions towards engagement and maturity subscales are expected to have highly positive environmental attitudes because they make true decisions regarding daily problems such as environmental issues by using their reasoning skills. So, this can be shown as a possible explanation for the result which indicated engagement and maturity were the significant predictors of environmental attitudes.

In short, this study concluded that high school students had a positive and strong inclination towards all three CT disposition subscales and gender, father and mother's educational background did not significantly affect their CT dispositions. Another result of this study was that although high school students had strongly positive opinions and emotions towards the environment, they seemed to perform a few environmentally positive actions. This result is important because having a positive opinion and emotion towards the environment is not enough to solve today's challenging environmental problems. In addition to having positive opinions and emotions, individuals should also show positive actions towards the environment. Therefore, students should be encouraged to take action to preserve the environment. This study revealed that while the students' environmental attitudes significantly differed by gender in favor of female students, they did not differ by students' mother and father's educational background. Another important result of this study was that engagement, maturity, and innovativeness were significant predictors of the high school students' environmental attitudes and they together explained 31% of the total variance on their attitudes towards the environment. This result emphasized the importance of CT which is one of the 21^{st}-century skills to improve the students' environmental attitudes. Therefore, CT should be implemented during the effort of increasing environmental attitudes and improving students' CT dispositions should be one of the most important aims of today's educational systems because it will also have a positive impact on students' environmental attitudes.

Limitations and Recommendations

Although this study is important to shed light on the relationship between high school students' CT dispositions and environmental attitudes, it has several limitations. First, it has a limitation in terms of its target sample. This study was carried out with high school students and it can be a good idea to repeat this study with other samples from different educational levels. Second, it has a limitation in terms of data collection tools. The research data were only collected with quantitative tools in this study. Therefore, future studies with qualitative or mixed methods may help provide a further understanding of the relationship between CT dispositions and environmental attitudes.

Scientific Ethics Declaration

The author(s) declare that the scientific ethical and legal responsibility of this article published in JESEH journal belongs to the author(s).

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179–211.

Amin, S., Utaya, S., Bachri, S., Sumarmi, & Susilo, S., (2020). Effect of problem-based learning on critical thinking skills and environmental attitude. *Journal for the Education of Gifted Young Scientists, 8*(2), 743-755. doi:http://dx.doi.org/10.17478/jegys.650344.

Arnocky, S. & M. Stroink. (2010). Gender differences in environmentalism: The mediating role of emotional empathy. *Current Research in Social Psychology, 16*(9), 1–14.

Aybek, B. & Ekinci, Ö. (2010). Analysis of the empathy and the critical thinking disposition of the teacher candidates. *Elementary Education Online, 9*(2), 816-827.

Ballantyne, R., Connell, S., & Fien, J. (1998). Students as catalysts of environmental change: A framework for researching intergenerational influence through environmental education. *Environmental Education Research, 12*(3-4), 1-11.

Blocker, T. J. & Eckberg, D. L. (1989). Environmental issues as women’s issues: General concerns and local hazards. *Social Science Quarterly, 70*, 586–593.

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., & Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature, 486*(7401), 59–67.

Coertjens, L., Pauw, J. B., Maeyer, S., & Petegem, P. V. (2010). Do schools make a difference in their students’ environmental attitudes and awareness? Evidence from pisa 2006. *International Journal of Science and Mathematics Education, 8*(3), 497–522. doi:10.1007/s10763-010-9200-0.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lewrence Erlbaum Associates.

Çolak, İ., Türkkaş-Anasız, B., Yorulmaz, Y. İ., & Duman, A. (2019). The role of gender, grade level, and parents’ educational attainment variables on teacher candidates’ critical thinking dispositions: A meta-analysis. *e-International Journal of Educational Research, 10*(1), 67-86.

De Groot, J. I. M. & Steg, L. (2008). Value orientations to explain beliefs related to environmental significant behavior: How to measure egoistic, altruistic, and biospheric value orientations. *Environment and Behavior, 40*, 330–354.

Diekmann, A. & Franzen, A. (1999). The wealth of nations and environmental concern. *Environment and Behavior, 31*, 540–549.

Dlamini, S., Tesfamichael, S. G., & Mokhele, T. (2021). Socio-demographic determinants of environmental attitudes, perceptions, place attachment, and environmentally responsible behaviour in Gauteng province, South Africa. *Scientific African, 12*, 1-9.

Duarte, R., Escario, J. J., & Sanagustín, M. V. (2017). The influence of the family, the school, and the group on the environmental attitudes of European students. *Environmental Education Research, 23*(1), 23-42. doi:10.1080/13504622.2015.1074660

Dunlap, R. E. & York, R. (2008). The globalisation of environmental concern and the limits of the postmaterialist values explanation: evidence from four multinational surveys. *The Sociological Quarterly, 49*, 529–563.

Ekici, G. (2005). Investigation of high school students’ attitudes toward environmental education. *Eurasian Journal of Educational Research, 18*, 71-83.

Facione, P. A., Giancarlo, C. A., Facione, N. C., & Gainen, J. (1995). The disposition toward critical thinking. *Journal of General Education, 44*(1), 1-25.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Ferah, D. (2000). *Investigation of perception of problem solving skills and problem solving approaches of the Turkish army academy cadets in terms of gender, grade, academic success and being leader* [Unpublished master`s thesis]. Hacettepe University.

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

Franzen, A. (2003). Environmental attitudes in international comparison: An analysis of the ISSP Surveys 1993 and 2000. *Social Science Quarterly, 84*, 297–308.

Franzen, A. & Meyer, R. (2010). Environmental attitudes in cross-national perspective: A multilevel analysis of the ISSP 1993 and 2000. *European Sociological Review, 26*, 219–234.

Franzen, A. & Vogl, D. (2013a). Two decades of measuring environmental attitudes: a comparative analysis of countries. *Global Environmental Change, 23*, 1001–1008.

Franzen, A. & Vogl, D. (2013b). Acquiescence and the willingness to pay for environmental protection: A comparison of the ISSP, WVS, and EVS. *Social Science Quarterly, 94*(3), 637-659.

Fua, J. L., Wekke, I. S., Sabara, Z., & Nurlila, R. U. (2018). Development of environmental care attitude of students through religion education approach in Indonesia. *IOP Conf. Series: Earth and Environmental Science, 175*, 1-7.

Gifford, R. & Sussman, R. (2012). Environmental attitudes. Susan D. Clayton (Ed) In *The Oxford Handbook of Environmental and Conservation Psychology* (p. 65-80). Oxford University Press.

Girot, E. A. (2000). Graduate nurses: Critical thinkers or better decision makers? *Journal of Advanced Nursing, 31*(2), 288-297.

Gökçe, N., Kaya, E., Aktay, S., & Özden, M. (2007). Elementary students' attitudes towards environment. *Elementary Education Online, 6*(3), 452-468.

Gökmen, A. (2021). The effect of gender on environmental attitude: A meta-analysis study. *Journal of Pedagogical Research, 5*(1), 243-257.

Gürbüz, H., Kışoğlu, M., & Erkol, M. (2007). Assessment of the biology teacher candidates' environmental attitudes for informal and formal learning surroundings. *Bayburt Faculty of Education Journal, 2*(3),74-84.

Güven, G. & Çakır, N. K. (2019). The relation between teachers' attitudes towards renewable energy sources and critical thinking dispositions. *Journal of Baltic Science Education*, *18*(5), 717-731.

Hausbeck, K. W., Milbrath, L. W., & Enright, S. M. (1992). Environmental knowledge, awareness, and concern among 11th-grade students: New York state. *Journal of Environmental Education, 24*, 27-34.

Hofreiter, T. D., Monroe, M. C., & Stein, T. V. (2007). Teaching and evaluating critical thinking in an environmental context. *Applied Environmental Education & Communication, 6*(2), 149-157.

Irani, T., Rudd, R., Gallo, M., Ricketts, J., Friedel, C., & Rhoades, E. (2007). *Critical thinking instrumentation manual.* http://step.ufl.edu/resources/critical_thinking/ctmanual.pdf.

Istead, L. L. (2009). *Intergenerational eco-education: An exploration of child influence or parental environmental behavior* [Unpublished master's thesis]. Royal Roads University.

Iverson, H. & Rundmo, T. O. (2001). Environmental concern and environmental behavior among the Norwegian public. *Journal of Risk Research*, *5*, 265-279.

Kemmelmeier, M., Krol, G., & Kim, Y. H. (2002). Values, economics and pro-environmental attitudes in 22 societies. *Cross-Cultural Research, 36*, 256–285.

Keniger, L. E., Gaston, K. J., Irvine, K. N., & Fuller, R. A. (2013). What are the benefits of interacting with nature*? International Journal of Environmental Research and Public Health*, *10*(3), 913–935.

Khandaghi, M. A., Pakmehr, H., & Amiri, E. (2011). The status of college students' critical thinking disposition in humanities. *Procedia Social and Behavioral Sciences, 15*, 1866–1869.

Kılıç, H. E. & Şen, A. İ. (2014). Turkish adaptation study of UF/EMI Critical Thinking Disposition Instrument. *Education and Science, 39*(176), 1-12.

Kökdemir, D. (2003). *Decision making and problem solving under uncertainty* [Unpublished PhD thesis]. Ankara University.

Lee, P. L. (2016). Misconceptions and biases in German students' perception of multiple energy sources: Implications for science education. *International Journal of Science Education, 38*(6), 1306-1056.

Lee, T. M., Markowitz, E. M., Howe, P. D., Ko, C. Y., & Leiserowitz, A. A. (2015). Predictors of public climate change awareness and risk perception around the world. *Nature Climate Change, 5*, 1014–1020.

Liu, N. Y., Hsu, W. Y., Hung, C. A., Wu, P. L., & Pai, H. C. (2019). The effect of gender role orientation on student nurses' caring behaviour and critical thinking. *International Journal of Nursing Studies, 89*, 18-23.

Liu, S. Y., Yeh, S. S. C., Liang, S. W., Fang, W. T. & Tsai, H. M. (2015). A national investigation of teachers' environmental literacy as a reference for promoting environmental education in Taiwan. *The Journal of Enviromental Education*, *46*(2), 114- 132.

Mahmoud, A. S. & Mohamed, H. A. (2017). Critical thinking disposition among nurses working in puplic hospitals at Port-Said Governorate. *International Journal of Nursing Sciences, 4*, 128-134.

Marquart-Pyatt, S.T. (2008). Are there similar sources of environmental concern? Comparing industrial countries. *Social Science Quarterly, 89*, 1312–1335.

Mengi Us, F. (2019). *An action research to develop environmental awareness and critical thinking skills in secondary school students with environmental education for sustainable development* [Unpublished PhD thesis]. Çukurova University.

McStay, J. R. & Dunlap, R. E. (1983). Male-female differences in concern for environmental quality. *International Journal of Women's Studies, 6*(4), 291-301.

Meyer, R. & Liebe, U. (2010). Are the affluent prepared to pay for the planet? Explaining willingness to pay for public and quasi-private environmental goods in Switzerland. *Population and Environment 32*, 42–65.

Mifsud, M. C. (2011). An investigation on the environmental knowledge, attitudes and behavior of Maltese youth. *US-China Education Review, 3*, 413-422.

Milfont, T. L. & Gouveia, V. V. (2006). Time perspective and values: An exploratory study of their relations to environmental attitudes. *Journal of Environmental Psychology, 26*, 72-82.

Mohai, P. (1992). Men, women, and the environment: An examination of the gender gap in environmental concern and activism. *Society & Natural Resources, 5*, 1–19.

Muhsilin, M., Rosiana, I., Rahayuningsih, Y., & Suharyana, Y. (2019). The efforts to improve environmental behavior and critical thinking of students through guided inquiry-based learning on environmental education-based science. *Jurnal Penelitian dan Pembelajaran IPA, 5*(2), 202-218.

Opitz, S. (2016). *Students' progressing understanding of the energy concept: An analysis of learning in biological and crossdisciplinary contexts* [Unpublished PhD thesis]. Christian-Albrechts Universitat.

Orhan, A. (2021). *The effect of critical thinking teaching with different approaches on high school students' critical thinking skills, dispositions and environmental attitudes* [Unpublished PhD thesis]. Düzce University.

Özdemir, S. M. (2005). Assessing university students' critical thinking skills for some variables. *The Journal of Turkish Educational Sciences, 3*(3), 297-316.

Pallant, J. (2001). *SPSS survival manual*. Open University Press.

Pauw, J. B. & Petegem, P. V. (2010). A Cross-national perspective on youth environmental attitudes. *The Environmentalist, 30*(2), 133–144. doi:10.1007/s10669-009-9253-1.

Polat, M. (2017). *Examining primary school teachers'critical thinking disposition and creativity levels in terms of difference variable* [Unpublished master's thesis]. Adıyaman University.

Poortinga, W., Steg, L., & Vlek, C. (2004). Values, environmental concern, and environmental behavior: A study into household energy use. *Environment and Behavior, 36*, 70–93.

Puspitasari, E., Sumarmi, S., & Amirudin, A. (2016). Integrasi Berpikir Kritis dan Peduli Lingkungan melalui Pembelajaran Geografi dalam Membentuk Karakter Peserta Didik SMA. *Education Journal: Theory, Research, and Development, 1*(2), 122–126.

Rebolj, N. & Devetak, I. (2013). 12-year-old students' attitude towards and views about the environmental problems. European Science Education Research Association, ESERAe_Book_2013.

Ricket, J. & Rudd, R. (2004). Critical thinking skills of FFA leaders. *Journal of Southern Agricultural Education Research, 54*(1), 7-20.

Rodriquez, G. (2000). *Demographics and disposition as predictors of the application of critical thinking skills in nursing practice* [Unpublished PhD thesis]. Colorado State University.

Sadık, F., & Çakan, H. (2010). Biyoloji bölümü öğrencilerinin çevre bilgisi ve çevre sorunlarına yönelik tutum düzeyleri. *Journal of Çukurova University Institute of Social Sciences, 19*(1), 351-365.

Saka, M. (2016). A study of preservice teachers' critical thinking dispositions as predictors of environmental ethics approaches. *Sakarya University Journal of Education, 6*(3), 100-115.

Sakschewski, M., Eggert, S., Schneider, S., & Bogeholz, S. (2014). Students' socioscientific reasoning and decision-making on energy-related issues-development of a measurement instrument. *International Journal of Science Education, 36*(14), 2291-2313.

Schultz, P. W., Shriver, C., Tabanico, J. J., & Khazian, A. M. (2004). Implicit connections with nature. *Journal of Environmental Psychology, 24*, 31-42.

Spellmann, G., Field, K. & Sinclair, J. (2003). An investigation into UK higher education students' knowledge of global climatic change. *International Research in Geographical and Environmental Education, 12*(1), 6-17.

Steg, L., Bolderdijk, J. W., Keizer, K., & Perlaviciute, G. (2014a). An integrated framework for encouraging pro-environmental behavior: The role of values, situational factors and goals. *Journal of Environmental Psychology, 38*, 104–115.

Steg, L., De Groot, J. I. M., Dreijerink, L., Abrahamse, W., & Siero, F. (2011). General antecedents of personal norms, policy acceptability, and intentions: The role of values, worldviews, and environmental concern. *Society and Natural Resources, 24*(4), 349-36.

Steg, L., Perlaviciute, G., Van der Werff, E., & Lurvink, J. (2014b). The significance of hedonic values for environmentally-relevant attitudes, preferences and actions. *Environment and Behavior, 46*(2), 163-192.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences*. Lawrence Erlbaum.

Şahin, K. (2018b). *Examining the critical thinking and empathy tendencies of sports high school students* [Unpublished master's thesis]. Gazi University.

Şahin, M. (2015). *Investigation of secondary school students' environmental literacy levels* [Unpublished master's thesis]. Aksaray University.

Şama, E. (2003). Teacher candidates' attitudes toward environmental problems. *Gazi University Journal of Gazi Educational Faculty, 23*(2), 99-110.

Tabachnick, B. G. & Fidell, L. S. (2012). *Using multivariate statistics* (6th ed.). Pearson.

Tikka, P. M., Kuitunen, M. T., & Tynys, S. M. (2000). Effects of educational background on students' attitudes, activity levels, and knowledge concerning the environment. *The Journal of Environmental Education, 31*(3), 12–19.

Tümkaya S., Aybek, B., & Aldağ, H. (2009). An investigation of university students' critical thinking disposition and perceived problem solving skills. *Eurasian Journal of Educational Research, 36,* 57-74.

Uzun, N., Gilbertson, K. L., Keleş, Ö., & Ratinen, I. (2019). Environmental attitude scale for secondary school, high school and undergraduate students: validity and reliability study. *Journal of Education in Science, Environment and Health, 5*(1), 79-90.

Walsh, C. M. (1996). *Critical thinking disposition of university students in practice disciplines (nursing, education and business) and nonpractice disciplines (English, history and psychology): An exploratory study* [Unpublished PhD thesis)]. University of Maryland.

Walsh, C. M. & Hardy, R. C. (1999). Dispositional differences in critical thinking related to gender and academic major. *Journal of Nursing Education, 38*(4), 149-155.

Vlek, C. & Steg, L. (2007). Human behavior and environmental sustainability: Problems, driving forces and research topics. *Journal of Social Issues, 63*(1), 1-19.

Yavuz, E. (2019). *Investigation of environmental attitudes of secondary school students in terms of different variables* [Unpublished master's thesis]. Gazi University.

Yüksel, Y. & Yıldız, B. (2019). Investigating students' attitudes towards environment in regard to numerous variables. *International Journal of Field Education, 5*(1), 88-110.

Zhang, J. (1994). Environmental hazards in the Chinese public's eyes. *Risk Analysis*, *14*, 163-167.

Zelezny, L. C., Chua, P. P., & Aldrich, C. (2000). New ways of thinking about environmentalism: Elaborating on gender differences in environmentalism. *Journal of Social Issues*, *56*(3), 443–457. doi:10.1111/0022-4537.00177.

Author Information

Ali Orhan
Zonguldak Bülent Ecevit University
School of Foreign Languages, Zonguldak, Turkey
Contact e-mail: *ali.orhan@beun.edu.tr*
ORCID ID: 0000-0003-1234-3919

Journal of Education in Science, Environment and Health
Volume 8, Issue 1, 2022 https://doi.org/10.21891/jeseh.1029468

Applying an Ecology Metaphor in a Mixed Methods Analysis of High School Science Program Infrastructure

Niyazi Erdogan, Carol L. Stuessy

Article Info

Article History

Published:
01 January 2022

Received:
10 August 2021

Accepted:
21 December 2021

Keywords

Science program infrastructure,
Ecology,
Secondary analysis,
Rubric development

Abstract

A sequential exploratory mixed methods approach guided research on the science program infrastructures (SPI) of a sample of 28 high schools representing 1,370 high schools in a large southwestern state. Comparisons of SPI were made between and among schools representing highly successful schools of low (n=9) and high diversity (n=10) with less successful, high-diversity high schools (n=9). Content analyses of interview data from science program teacher liaisons revealed the following characteristics for typical high school SPIs: (a) shared leadership within a diverse community of actors, including teachers, school principals, special education teachers, department heads, content-area leaders, curriculum directors, and district representatives; (b) supports for shared, balanced decision making, including frequent meetings, established communication channels, and explicit department head responsibilities; and (c) actions extending beyond general tasks of disseminating information to include resolution of issues regarding state-mandated test scores, curriculum-related tasks, and choice of professional development topics. We include a discussion of the implications for including SPI as an important mediating layer within the ecology of high schools. The science program layer links administration with classrooms by employing efficient, equitable, and effective practices to advance the goals of science achievement and college readiness established at national and state levels for all high school graduates.

Introduction

Effective science programs within high school learning ecologies provide the connection between state, district, and school policies and their consequent decisions to science teachers. Teachers are critical actors contributing within the infrastructures of districts, schools, and classrooms who enact policies and decisions to advance national and state goals for K-12 science education (National Research Council [NRC], 1996). The *National Science Education Standards* (NRC, 1996) argue that the most important resource for an effective science program is the professional teachers who serve the community's children. Chapter 7 of the Standards elaborates features of effective science programs, asserting that science teachers are best served by programs that continually work to replace the typical norms of isolation, conformity, and competition amongst high school teachers with new norms of community building, collegiality, openness, and trust. The Standards explain that replacement of typical norms occurs with realignment and readjustment:

> Schedules must be realigned, time provided, and human resources deployed regularly to discuss individual student learning needs and to reflect and conduct research on practice. In a community of learners, teacher work together to design the curriculum and assessment, … [taking] part in other professional growth activities. Time must be available for teachers to … hold meetings during the school day. (p. 223)

Not knowing what to expect in our exploration of science programs within the 1,370 high schools in the state of Texas, we chose an inductive, mixed methods approach to describe the infrastructures of science programs currently operating in the state. We developed an understanding of the complex relationships existing within the science program infrastructures (SPI) of 28 schools representing three types of schools (i.e., highly successful, highly diverse; highly successful, less diverse; less successful, highly diverse) through a constant comparison method to identify the salient components of each science program. The method required the identification of infrastructural features of each school, one by one, to constantly compare all features existing within the sample of 28 schools. Features were then clustered and named to represent a component. The diversity existing within the component was described by the frequencies of occurrence for each of the features within the component.

Expressed in the form of a rubric, features and components were then compared across the three school types, thus providing a snapshot for comparison on the basis of science program infrastructure.

This study extends work initiated by a project on policy research in science education funded by the National Science Foundation (NSF). This project was initiated to study the Teacher Professional Continuum (TPC) of high school science teachers in the state of Texas. In 2011, our research group was awarded an extension to investigate the features of "achievement gap schools" residing within the state. Only a few (i.e., 28 out of 1,370) of these notable, highly successful, highly diverse schools were found among the state-maintained database providing quantitative data describing each of the high schools in the state. In the final year of the project, these highly successful, highly diverse schools became the focus of policy research questions about the qualities of these unique schools.

Although several research projects have been designed to explore different categories and groups of programs within high schools (Giles & Hargreaves, 2006; Goodman, 1995; Irwin & Farr, 2004), none exist to compare the infrastructure of science programs in highly successful and less successful high schools varying in diversity status. In the extension study of our research group, researchers specifically investigated policies and practices of science programs within schools and asked questions about science programs, including comparisons with less successful and less diverse schools existing in the state. Ultimately, knowledge from the project has provided researchers and practitioners with a deeper understanding of the infrastructures of high schools' science programs as well as the unique characteristics of the unique subset of schools identified as both highly successful and highly diverse.

This research addressed questions of interest to mixed methods researchers, policy makers in science education, and science education practitioners in schools. We specifically looked for the most common infrastructural elements of high school science programs existing within a large, demographically complex southwestern state, paying special attention to the unique qualities of science programs in highly successful, highly diverse schools. Using a two-phase sequential exploratory secondary analysis mixed methods design (Figure 1), we were able to analyze both qualitative and quantitative phases to describe and compare the infrastructural components found within 28 high school science programs.

Research Background

Literature Review

Content analysis is the common methodology used to generate rubrics from interview data (Krippendorff, 2004). Rubrics have been used in qualitative research for the categorization of themes or ideas (Denzin & Lincoln, 2011). Additionally, rubrics also have been used in quantitative research for counting the occurrences of specific elements (Hsieh & Shannon, 2005). As a mixed methods tool for measurement, rubrics combine the process of creating general categories from themes and ideas with the identification of specific elements within those categories. In the social sciences, rubrics also provide a means for conducting comparative analysis across key demographic groups found within larger populations.

Scholars have yet to use rubrics to identify categories and elements within science program infrastructure. In fact, much of the literature related to infrastructure in education research focuses on actors and contextual factors found in a school but not within a specific content program (Englert & Tarrant, 1995). Consequently, science program infrastructure has yet to be defined in the literature. Instead, most of the sources informing our work on science programs typically focused on issues related to: (1) comprehensive school reform (Waldron & McLeskey, 2009), (2) teacher-researcher community (Englert & Tarrant, 1995), and (3) sustainability of innovative schools (Giles & Hargreaves, 2006). While literature specific to science program infrastructure is rare, some articles on infrastructure in high schools in general were found.

Infrastructure in high schools has been examined from various perspectives ranging from the broad perspective of comprehensive school reform to a narrower perspective focusing on sustainability of innovative schools (Erdogan & Stuessy, 2015a; Erdogan & Stuessy, 2015b; Erdogan, 2014; Navruz et al., 2014). An analysis of infrastructure within high schools has, on occasion, included a comparison across different school types. For example, MacNeil, Prater, and Busch (2009) used multivariate analysis of variance to assess differences in infrastructure between schools of varying success levels. As the literature on high school infrastructure was studied, several common categories emerged as important for consideration. These categories included collaboration among actors, communication and leadership, and diversity of school populations. As mentioned previously, none of the studies reviewed included discussions about the infrastructure of science programs specifically. Based on the preponderance of the literature, studies about infrastructure typically focused on one of two areas. Qualitative studies typically focused on the elements related to actors found within school infrastructure. For example, Englert and Tarrant (1995) identified three important characteristics of actors related to school infrastructure: (a) actors are the primary components in school infrastructure; (b) actors can be selfish learners within the infrastructure; and (c) actors can be drivers of change.

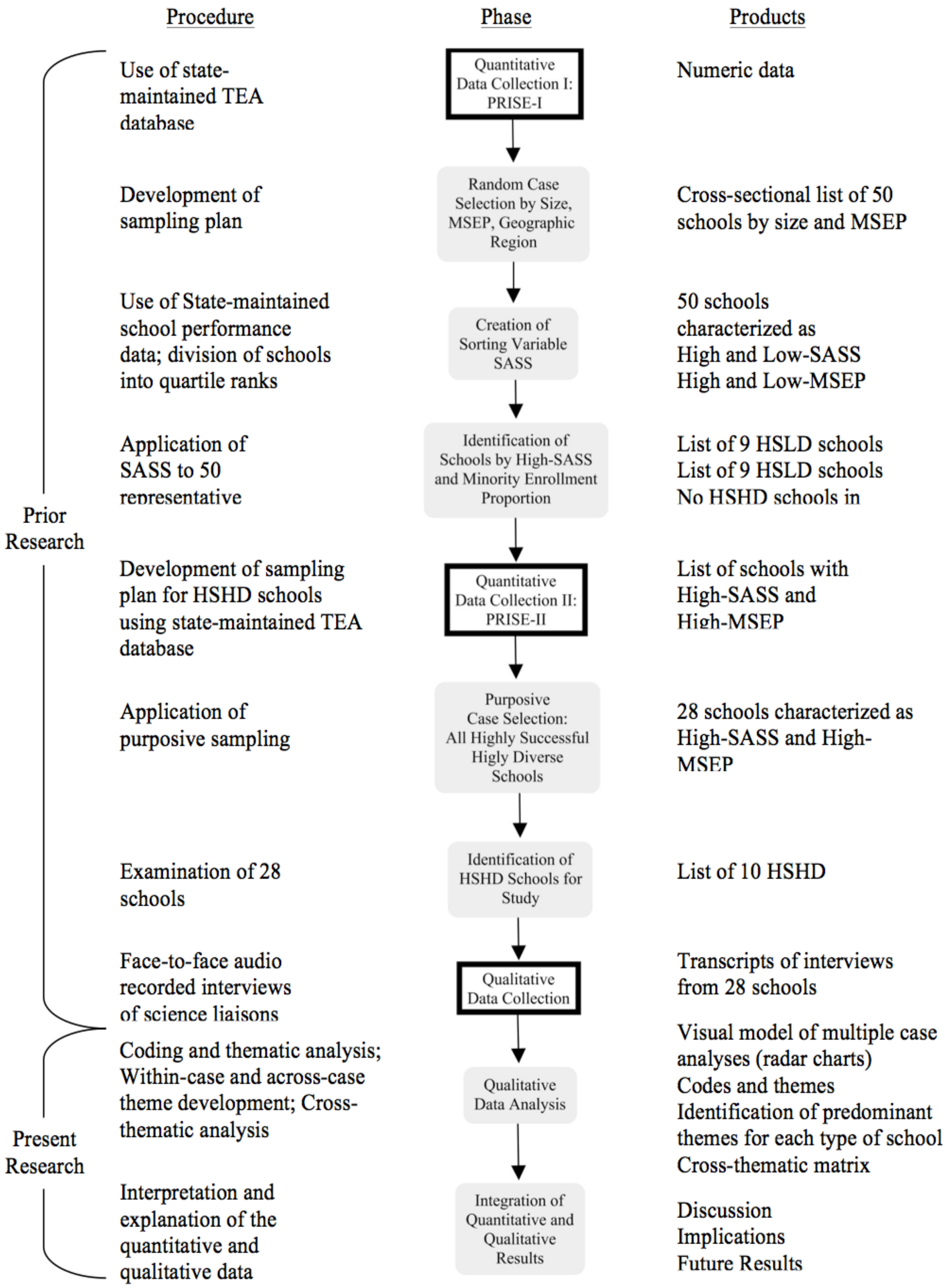

Figure 1. Research diagram.

Alternatively, quantitative studies in our review of literature usually focused on elements related to contextual factors within school infrastructure. MacNeil, Prater, and Busch (2009) in a quantitative analysis of 25,000 students and 1,700 teachers in 29 schools concluded that infrastructural features such as a school's success rating affected the student-learning environment. As a result of these findings, we chose a mixed methods design to blend the types of studies we reviewed in order to identify numerous types of factors that used to describe the science program infrastructures in high schools across Texas.

Conceptual Framework

An ecology metaphor proposed by Weaver-Hightower (2008) helped us conceptualize schools as complex, interdependent, and structural learning environments. As Weaver-Hightower (2008) presented in his metaphor, the ecology of schools consists of actors, actions, and contextual factors. For the purposes of our analysis, we identified actors, actions, and supports as the integral infrastructural components of the science program layer of the school ecology. Infrastructural components, even within a layer, cannot be viewed as deconstructed parts. Infrastructural components are viewed in terms of how they interact within and among the various organizational layers within the school ecology. Without an examination of connections between and among actors, actions, and supports at all layers, we would be unable to view school ecology as a whole, functioning system. An understanding of science program, therefore, is only one layer of the ecology, interconnected with other layers within the entire school system.

Actors in school ecologies interact in complex ways that parallel those found in ecosystems. Ecological terms describing relationships (e.g., competition, cooperation, predation, symbiosis) may be helpful in understanding relationships between and among actors within and among school layers. For example, actors (i.e., students, teachers, support staff, and administrators) may cooperate or even rely on one another to meet a common goal. How actors relate to one another is also affected by environmental boundaries and extant conditions. Schools differ in what these particular boundaries and extant conditions may be, but ultimately personal, support, and task factors intermix to affect how a given school approaches change (Weaver-Hightower, 2008). Taken together, the ecology metaphor provides a powerful conceptual model for understanding the complexity of schools.

Method

Mixed methods research is a pragmatic approach used to consider questions of interest through both qualitative and quantitative lenses. We conducted this sequential exploratory mixed methods secondary analysis of science program infrastructure in two phases. The initial qualitative phase using content analysis for rubric development was followed by a quantitative phase using comparative analysis. The following research questions were considered: (1) Regardless of schools' success and diversity, what elements exemplify science program infrastructure? (2) When comparing highly successful, highly diverse schools with two other types of schools (i.e., highly successful, less diverse schools and less successful, highly diverse high schools), what elements of science program infrastructure are unique to highly successful, highly diverse schools?

Sampling

Schools were identified as highly successful (HS) or less successful (LS) using the School Aggregate Science Score (SASS) created during prior research (Stuessy & Bozeman, 2011). Schools were identified as high diversity (HD) or low diversity (LD) using Minority Student Enrollment Proportion (MSEP) also created during prior research (Stuessy & Bozeman, 2011). Combining these two identifiers allowed us to describe each of 28 schools in our sample as HSHD, HSLD, or LSHD. This analysis used school and teacher level data to describe and model science program infrastructure within and across three high school types: HSHD (n=10), HSLD (n=9), and LSHD (n=9).

Sampling was conducted in two stages. In the first stage, sampling reflected the design used in the original study for which participating schools were selected. A stratified random sampling design was originally used to select the nine HSLD and nine LSHD schools in the current analysis. In contrast, a purposive sampling design was used to select the ten HSHD high schools. In this second stage of sampling, a purposive sampling design was used to ensure selection of representative highly successful and highly diverse high schools.

Phase I

The analysis for this study involved an examination of archived data from interviews with the science teacher leader from each science program within the selected 28 high schools. The first, qualitative, phase of analysis used content analysis of interview data to develop a coding rubric. During the content analysis, reliability was demonstrated when researchers working independently agreed on data elements to be coded, grouped, and categorized (Bogdan & Biklen, 1992). Initiative rubric development relied on a shared analysis of nine science program interviews to establish temporary codes, groups, and categories. Three interviews from each of the three school types (i.e., highly successful, high diversity, HSHD; highly successful, less diversity, HSLD; less successful, high diversity, LSHD) were analyzed. Over a three-week period of peer debriefing and refinement, we generated a rubric containing five major categories divided into 17 groups possessing 60 unique and mutually exclusive elements (see Appendix). Ultimately, this rubric was used to conduct content analyses for all 28 program interviews to enable the interpretation of differences in science program infrastructure across the three types of schools.

We used simple percent agreement to establish inter-coder reliability of the rubric, using methods described by Lombard, Snuder-Duch, & Bracken (2002). Specifically, each researcher used the final rubric to independently analyze program interviews of a subset of nine interviews. Upon completion of the independent analysis, we calculated percent agreement by dividing the total number of agreements for all three researchers by the total number of potential agreements. Using these percentages, final inter-coder reliability for the rubric was determined to be 87%, a value assessed as acceptable by most researchers when conducting exploratory research using content analysis (Lombard, Snyder-Duch, & Bracken, 2002).

Phase II

The second, quantitative, phase of our analysis used frequency and comparative analysis. During frequency analysis, measures of center, spread, and shape were determined by simple counts. We identified the most common elements of science program infrastructure, regardless of school success and diversity, which included information about categories within science program infrastructure identified through the rubric generated during Phase I. During the comparative analysis, differences in measures related to schools’ success and diversity were identified and/or tested. We compared the frequency scores for science program infrastructure in the three types of schools (i.e., HSHD, HSLD, and LSHD). In our comparative analysis, we concurred that a higher percentage in HSHD schools was to be used in establishing the criterion of "most likely to occur."

Results and Discussion

The integration of multiple data forms (i.e., qualitative and quantitative data) is a key feature of analysis in mixed methods research. Integration of data provided a way to advance our understanding of the complex, interdependent, and structural elements occurring within high school science program infrastructures. In this section, a summary of results for our analysis is organized by the two research questions.

Elements Exemplifying Science Program

Regardless of schools’ success and diversity, what elements exemplify science program infrastructure? The rubric in the Appendix consolidates qualitative elements emerging during the content analysis of interviews from teacher leaders representing 28 science programs. The rubric contains 60 elements within 14 groups and three components. In the content analysis, we achieved consensus that 59 of 60 elements occurred in at least one of the sampled science programs. These results suggest that elements emerging in our analysis would be likely to appear in analyses of infrastructures supporting other schools' science programs residing within the state of Texas. We organized the rubric to include Components (i.e., Actors, Supports, and Program Tasks), Groups (e.g., Teacher Numbers, Group Meetings, Other SP Members) and Elements within groups (e.g., within Teacher Numbers: Many, Several, Few). The frequencies of occurrence for elements are reported on the rubric for HSHD, HSLD, and LSHD schools. These numbers indicate the results of an analysis of interview transcripts in which we applied the rubric to confirm the existence of elements for each school within the sample, to identify the elements best exemplifying elements of science program infrastructures occurring within all three school types, to compare occurrences of elements within each of the school types, and to identify elements most frequently occurring in HSHD schools. We mention here that we report frequencies only to provide evidence of co-occurrence and not as evidence of causal proof that an increase in any of the occurrences of an element would lead to an increase in student success.

Table 1. Percentages of occurrence for 60 science program infrastructure elements in 28 science programs

Component	Element	n	%
	Actors		
Teacher Numbers	*Many (> 6)	**16**	**57**
	Few (1-3)	**8**	**29**
	Several (4-6)	3	11
Group Meetings	Whole group only	**13**	**46**
	*Whole group and subject-area	**7**	**25**
	Subject-area only	1	4
Other SP Members	School principal	**9**	32
	Special education	**8**	**29**
	In-school non science	6	21
	Central office	5	18
Department Head	Appointed	5	18
	Volunteered	**9**	**32**
Identified Leaders within the SP Community	*Department head	**23**	**82**
	*Content area leader	**11**	**39**
	District representative	**10**	**36**
	Curriculum director	**9**	**32**
	Instructional leader	6	21
	Veteran teacher	4	14
	Everyone a leader	3	11
	Supports		
Meeting Frequencies	As needed	**13**	**46**
	More than once a week	**8**	**29**
	Twice a month	4	14
	Once a month	4	14
	Once a week	2	7
Department head compensation	Stipend	**9**	**32**
	Flexibility /time off	4	14
	Reduced teaching load	3	11
Communication	Email, newsletter	**11**	**39**
	Use of opinion surveys	2	7
	With other programs	1	14
	With students	0	0
Decision Making	*Shared /balanced	**18**	**64**
	Top-down	5	21
	Bottom-up	2	7
Department Head Responsibilities	Budget	**12**	**43**
	*Provide professional development	**10**	**36**
	Paperwork	**9**	**32**
	Scheduling	**7**	**25**
	*Instructional coaching	5	18
	TAKS tasks	4	14
Department Head Compensation	Place substitute teachers	1	4
	Stipend	**9**	**32**
	Flexibility/time off	4	14
	Reduced teaching load	3	11
	Program Tasks		
General	*Disseminate information	**13**	**46**
	*Housekeeping	5	18
	*Participate in hiring/staffing	3	11
Assessment	Resolve TAKS issues	**11**	**39**
	*Develop common assessments	6	21
	Examine student work	6	21

Continued

	Program Tasks (continued)		
	Assess teaching with TAKS	4	14
Curriculum	*Plan/share lessons	**14**	**50**
	*Coordinate within content area	**9**	**32**
	*Share lab materials	**7**	**25**
	Align curriculum vertically	5	18
	Choose informal science opportunities	4	14
	*Work in interdisciplinary groups	2	7
	Design informal activities	0	0
	Develop tutorials	1	4
Professional Development	*Choose PD topics	**13**	**46**
	Read books together	4	14
	Conduct action research	2	7
	Adopt targeted strategies	2	7

*Note: *marks elements with highest frequencies and characteristic of HSHD schools*

Table 1 presents a summary of the content analyses of interview data from science program teacher liaisons. The following elements occurred in typical high school SPIs: (a) shared leadership within a diverse community of actors, including teachers, school principals, special education teachers, department heads, content-area leaders, curriculum directors, and district representatives; (b) supports for shared, balanced decision making, including frequent meetings, established communication channels, and explicit department head responsibilities; and (c) tasks extending beyond general actions of disseminating information to include more specific tasks of resolving issues regarding state-mandated test scores, performing curriculum-related tasks, and choosing professional development topics.

The most common elements were: (1) presence of a department head within the program leadership (n=23); (2) shared or balanced autonomy in decision making within the program (n=18); and (3) seven or more teachers within the science program (n=16). Less common, but still describing the typical science program infrastructure were four additional elements: (1) meeting frequencies as needed (n=15); (2) information dissemination as a program task (n=13); and (3) use of email and/or newsletter communication (n=11). Taken together, these nine elements exemplified typical high school science program infrastructure.

Elements of Science Program Infrastructure Most Frequently Occurring in HSHD Schools

When comparing highly successful, highly diverse schools with two other types of schools (i.e., highly successful, less diverse schools and less successful, highly diverse high schools), what elements of science program infrastructure are unique to highly successful, highly diverse schools? Table 1 also identifies the elements of science program infrastructure most frequently occurring in HSHD schools.

Actors within HSHD schools were most likely to have more than six members (57%); meet in both whole group and subject-area groups (25%); and share leadership responsibilities among department heads (82%), content-area leaders (39%) and/or curriculum directors (32%). Supports were most likely to include established communication channels by email and/or newsletters (39%), shared/balanced decision-making (64%), and department head responsibilities related to teacher learning, either by professional development (36%) and/or instructional coaching (18%). Program tasks were most likely to include general tasks of disseminating information (46%), housekeeping (18%), and hiring new science teachers (11%); and more specific curriculum tasks of planning/sharing lessons (50%), within-content area coordination (32%), sharing laboratory materials (25%), and working in interdisciplinary groups (7%).

In order to visually represent our findings, we created a spider diagram in which you can see the most common elements (Figure 2). As seen in the diagram, HSHD schools are more likely to have whole-group science program meetings, be involved in general tasks of information dissemination and curriculum tasks of lesson planning/sharing, have content area leaders, with department heads providing PD, and possessing shared/balanced decision making. (Also see the rubric in the Appendices displaying frequency counts for each of the characteristics displayed in Figure 2.).

Conclusion

The rubric designed in this study contained a large number of individual elements (n=60). However, only 20 elements of the 60 elements occurred at a frequency above 29%. (Refer to Table 1.) Most frequently (> 45%) identified elements included department heads as leaders (82%), shared/balanced decision making (64%), teacher numbers > 6 (57%), SP meetings occurring as needed (54%) in whole groups (46%), with program tasks including planning/sharing lessons (50%), information dissemination (46%), and choosing PD topics (46%). In reference to the ecology metaphor, Actors, Actions (SPI Tasks), and Supports are identified by these elements. In addition, our findings suggest that science program infrastructures in HSHD schools are different from the infrastructure of science programs in HSLD and LSHD schools. Specifically, infrastructure within HSHD schools differed in Actors (teacher numbers, both whole group and subject-area meetings), types of actor-leaders (principals and content-area leaders), shared decision making with department heads providing both professional development and instructional coaching, and SP tasks including more general and specific tasks, especially regarding curriculum, sharing materials, and working in interdisciplinary groups.

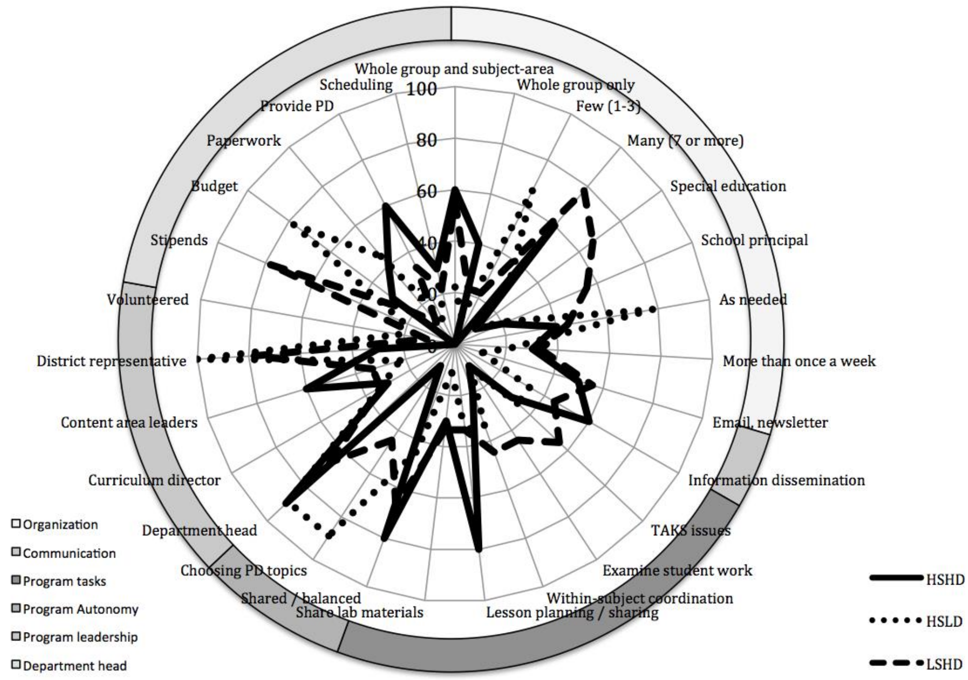

Figure 2. Most frequent elements of science program infrastructure occurring in HSHD schools.

Table 2 summarizes a qualitative cross-case comparison of the 16 elements identified from the analysis to most likely occur in HSHD schools. The cross-case comparison yielded two predominant categories of SP infrastructure elements most likely to occur in HSHD schools: Cross-Cutting School Community Structures (existing within the Whole SP, within Content-area Groups, and extending Outside the SP Community); and Shared Leadership (Distributed Leaders and Department Head Leadership in Professional Learning).

Cross-element comparisons revealed two infrastructural emphases for HSHD schools: Cross-Cutting Community Structures and Shared Leadership. These results suggest that the science program environment in HSHD schools, when compared to HSLD and LSHD schools, demonstrates differences in terms of their distributed leadership and professional learning roles of their science department heads.

Further evidence of exemplary practice is provided by the *National Science Education Standards* A and F, which explicitly describe the needs for effective science education in terms of community involvement and leadership. Standard A states the need for clearly defined leadership at the school and district levels "vested in a variety of people, including teachers, school administrators, and science coordinators" (p. 212) in order to provide opportunities to learn and teach science. Standard F identifies strong leadership as a quality of effective communities of learners, recommending a dramatic change from the "hierarchical and authoritarian leadership often in place in schools and in schools districts today," with a leadership structure "inevitably [requiring] that teachers and administrators rethink traditional roles and responsibilities and take on new ones" (p. 223). Such is the case with the SPI of HSHD schools. A shared leadership structure was identified more frequently in HSHD

(80%; 8 out of 10) schools than in their less diverse (49%; 4 out of 9) or less successful (67%; 6 out of 9) counterparts. Furthermore, department heads, curriculum directors, and content area leaders were most likely to be identified as leaders within HSHD schools. Furthermore, responsibilities for budget, scheduling, and paperwork for department heads were more likely to extend beyond these more traditional roles to include department heads as instructional coaches and professional development providers. (These frequencies of occurrence as evidence of diversified leadership roles can be reviewed in Table 2.)

Table 2. Elements of SP infrastructure most likely to occur in highly successful, highly diverse schools

	School Community Structures			Shared Leadership	
Elements	Whole SP	Content-area	Outside SP community Within- and outside-school	Distributed Leadership	Department Head Leads Professional Learning
Teacher numbers > 6	X				
Meetings	X	X			
Department head leaders	X			X	
Content-area leaders	X		X	X	
Curriculum director leaders	X		X	X	
Communication by email and newsletters	X		X	X	
Shared/balanced decision making	X			X	
Dept head responsibilities for professional development					X
Dept head responsibilities for instructional coaching					X
SP Task disseminate information	X			X	
SP Task housekeeping	X			X	
SP Task hiring new science teachers			X	X	
SP Task planning/ sharing lessons		X			
SP Task within-content area coordination		X			
Sharing laboratory materials	X	X			
Working in interdisciplinary groups		X	X		
Totals	10	4	4	8	2

The National Science Education Standard F also maintains that schools must work as communities "that encourage, support, and sustain teachers" (p. 222) as they implement an effective science program:

> Many previous reform efforts have failed because little attention was paid to the connection between teacher enhancement, curriculum development, and the school as a social and intellectual community. Teachers with new ideas, skills, and exemplary materials often worked in an environment where reform was not valued or supported. (p. 222)

Evidence of community structures such as those described in the Standards is provided in HSHD schools by the higher frequencies of occurrence for several SPI elements. HSHD schools were most likely to meet in whole groups and subject-area groups; these schools were also most likely to identify content-area leadership within the SPI. They identified shared/balanced decision making more frequently than the other two types of schools, and they identified general and specific tasks more frequently in terms of housekeeping, information dissemination, developing common assessments, planning/sharing lessons, and sharing laboratory materials. Overall, HSHD schools exhibited more of the Standards-based qualities of "community" and "leadership" than in the other two types of schools.

Recommendations

Our design of the rubric for identifying elements within SPI is a "first" in terms of identifying the elements of infrastructure currently existing within high school science programs. Development of the rubric allowed us to identify the most common elements composing science program infrastructures and to compare science program infrastructures across different school types. Random sampling led to our ability to generalize to all schools within the state of Texas, including the few HSHD schools purposively chosen to represent an important subgroup residing within the populations of all Texas high schools. We recommend that other schools desiring to strengthen their SPI use the rubric as a starting point to identify critical elements within their own infrastructures, particularly those involving community and leadership structures.

Our findings also support that identification of HSHD schools as potential leaders in developing stronger SPI for schools struggling with issues related to equitable and effective policies and practices supporting decision-making and task completion at the science program level. This study has also contributed to information for the previous research initiative striving to identify the unique qualities associated with the powerful group of "achievement gap" schools. As such, the findings of this study contribute to the much broader ecological perspective used by our research group to define an exemplary model for highly successful, highly diverse high schools in Texas.

Scientific Ethics Declaration

The authors declare that the scientific ethical and legal responsibility of this article published in JESEH journal belongs to the authors.

References

Bogdan, R. C., & Biklen, S. K. (1992). *Qualitative research for education* (2nd ed.). Boston: Allyn and Bacon.

Denzin, N. K., & Lincoln, Y. S. (2011). *The Sage handbook of qualitative research* (4th ed.). Thousand Oaks, CA: Sage.

Englert, C. S., & Tarrant, K. L. (1995). Creating collaborative cultures for educational change. *Remedial and Special Education, 16*, 325–336.

Erdogan, N. (2014). *Modeling successful inclusive STEM high schools: An analysis of students' college entry indicators in Texas* [Unpublished doctoral dissertation]. Texas A&M University.

Erdogan, N., & Stuessy, C. L. (2015a). Modeling successful STEM high schools in the United States: An ecology framework. *International Journal of Education in Mathematics, Science and Technology, 3*(1), 77-92.

Erdogan, N., & Stuessy, C. L. (2015b). Examining inclusive STEM schools' role in the college and career readiness of students in the United States: A multi-group analysis of students' achievement outcomes. *Educational Sciences: Theory & Practice, 15*(6), 1517-1529. DOI: 10.12738/estp.2016.1.0072

Giles, C., & Hargreaves, A. (2006). The sustainability of innovative schools as learning organizations and professional learning communities during standardized reform. *Educational Administration Quarterly, 42*(1), 124–156.

Goodman, J. (1995). Change without difference: School restructuring in historical perspectives. *Harvard Educational Review, 65*(1), 1–29.

Hsieh, H., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*, 1277-1288.

Irwin, J. W., & Farr, W. (2004). Collaborative school communities that support teaching and learning. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 20*, 343-363.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage.

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research, 28*(4), 587-604.

MacNeil, J. A., Prater, D. L., & Busch, S. (2009). The effects of school culture and climate on student achievement. *International Journal of Leadership in Education: Theory and Practice, 12*(1), 73-84.

National Research Council. (1996). Chapter 7, Science education program standards. In *National Committee on Science Education Standards and Assessment, National science education standards* (pp. 209-225). Washington DC: National Academies Press.

Navruz, B., Erdogan, N., Bicer, A., Capraro, R. M., & Capraro, M. M. (2014). Would a STEM school 'by any other name smell as sweet'? *International Journal of Contemporary Educational Research, 1*(2), 67-75.

PRISE Research Group. (2011). *Policy and research initiative in science program-II research manual: School, program, and teacher instruments*. Retrieved from http://prise.tamu.edu

Stuessy, C. L., & Bozeman, D. (2011). *Policy research initiative in science education—II: School, program and teacher instruments*. Retrieved from Texas A&M University, Policy Research Initiative in Science Education website: http://prise.tamu.edu/documents/PRISE_-_II_Instrument_Manual[1].pdf

Waldron, N. L., & McLeskey, J. (2009). Establishing a collaborative school culture through comprehensive school reform. *Journal of Educational and Psychological Consultation, 20*(1), 58–74.

Weaver-Hightower, M. B. (2008). An ecology metaphor for educational policy analysis: A call to complexity. *Educational Researcher*, *37*(3), 153-167.

Authors Information

Niyazi Erdogan
Texas A&M University
4232 TAMU, College Station, TX, USA 77843
Contact e-mail: *niyazierdogan84@gmail.com*
ORCID iD: 0000-0001-6373-0930

Carol L. Stuessy
Texas A&M University
4232 TAMU, College Station, TX, USA 77843

Appendix

Overall Inter-Rater Agreement 87.0%	**Organization 86.9**															**Communication 96.3**			
	Groups and Subgroups 87.7			**Teacher Numbers 100.0**			**Other SP Members 76.0**				**Meeting Frequency 83.7**					**Within SP 92.6**		**Outside SP 100.0**	
	Whole group and Subject-area	Whole-group only	Content/subject only	Few (1-3)	Several (4-6)	Many (7 or more)	Special education	In-school non-science people	Central Office	School principal	As needed	Once a month	Twice per month	One a week	> Once a week	Email, newsletters	Opinion surveys	Students	Other programs
HSHD	6	4	1	0	3	6	1	2	1	2	4	2	2	1	3	5	1	0	0
HSLD	2	1	0	6	0	3	1	1	1	2	7	1	1	0	2	1	1	0	1
LSHD	5	2	0	2	0	7	6	3	3	5	4	1	1	1	3	5	0	0	0

SP Tasks 86.0																**Leadership Within SP 80.9**						
General 89.2		**Review Student Progress 80.3**			**Curriculum 82.7**						**Program-Level PD and Improvement 95.6**					**School-Level 84.0**						
Housekeeping	Information dissemination	TAKS Issues	Common assessments	Examine student work	Within-subject coordination	Vertical alignment	Lesson planning/sharing	Sharing lab materials	Developing tutorials	Interdisciplinary grouping	Reading books together	Design informal sci activities	Action research	Adopt targeted strategies	Assess teaching with TAKS	Department head	Curriculum director	Instructional leader	Content area leaders	Veteran teachers	Everyone a leader	District-level Representative 77.8
3	6	3	3	1	2	0	8	3	0	1	1	0	0	0	1	9	3	2	6	0	1	3
1	3	3	1	1	3	4	3	1	0	1	1	0	0	0	3	8	3	1	2	2	0	0
1	4	5	2	4	4	1	3	3	1	0	2	0	2	2	0	6	3	3	3	2	2	7

Department Head 85.3												**SP Autonomy in Decision Making 86.4**					
Assignment 77.8		**Compensation 95.1**			**Duties 83.1**							**Overall 87.7**			**Operational Matters 85.2**		
Appointed	Volunteer	Stipends	Reduced Teaching Load	Flexibility/Time Off	Budget	Scheduling	Paperwork	TAKS Tasks	Instructional Coaching	Provide PD	Place Substitute Teachers	Bottom-Up	Shared/Balanced	Top-Down	Hiring/Staffing Involvement	Choosing Topics for PD	Informal Sci Opportunities
1	0	0	0	0	3	3	4	0	3	6	0	0	8	2	3	1	2
1	2	2	0	1	7	2	4	2	0	1	0	2	4	2	0	8	0
3	1	7	3	3	2	2	1	2	2	3	1	0	6	2	3	4	2