Teacher Anxiety Scale for Organizing Trips to Out-of-School Learning Environments: Development and Validity of the Scale

Selçuk ARIK[*]
Department of Mathematics and Science Education, Tokat Gaziosmanpaşa University, Tokat, Turkey
ORCID: 0000-0003-4496-8104

Aykut Emre BOZDOĞAN
Department of Mathematics and Science Education, Tokat Gaziosmanpaşa University, Tokat, Turkey
ORCID: 0000-0002-5781-9960

This study aims to develop a valid and reliable "Teacher Anxiety Scale for Organizing Trips to Out of School Learning Environments" to measure teachers' anxiety about out-of-school learning environments. This research focuses on the methodological validity and reliability of the scale. The study group of the research consisted of 394 teachers working in public and private schools affiliated to the Ministry of National Education in different cities of Turkey. In the development of the scale, the processes of literature review, creating an item pool, obtaining expert opinion, pilot test, creating the trial form, data collection, data analysis (exploratory factor analysis, repeated exploratory factor analysis, confirmatory factor analysis, item analysis, and reliability analysis), and creating the final form were followed. As a result of the analysis, a valid and reliable scale consisting of 28 items and four factors was created. The factors of the scale were named "Bureaucracy-Related (BR) Anxiety", "Safety Risks-Related (SRR) Anxiety", "Harm-Related (HR) Anxiety" and "Pedagogy-Related (PR) Anxiety", respectively. While the Cronbach's Alpha reliability coefficient for the total scale was determined as .944; the reliability coefficients for the factors were determined as BR Anxiety Cr α=.868, SRR Anxiety Cr α=.922, HR Anxiety Cr α=.903, and PR Anxiety Cr α=.952, respectively. The results indicate that the scale is quite reliable.

Introduction

Today, as twenty-first-century skills is gaining great importance, what is expected from education systems is to raise individuals who are self-sufficient, self-regulated, self-critical, self-confident, searchers, questioners, critical thinkers with creative thinking skills and can solve their problems using scientific processes (Arık, 2019, 2021; Bozdoğan, 2016). For

---

[*] Correspondency: selcuk.arik@gop.edu.tr

individuals who will adapt to today's world to gain the skills and experiences mentioned above, the experiences acquired outside the school are as important as the skills and experiences gained inside the school. In this context, it has been seen that education and training have become a life-long process under each circumstance, everywhere: at home, at school, at work, in the garden, in the park, in the museum (Bozdoğan, 2016).

In today's world, where lifelong learning and learners are also gaining importance, learning environments should not be considered only in the classroom. Learning should also be experienced outside the classroom (Sen, Ertaş-Kılıç, Oktay, Ekinci & Kadirhan, 2021). Because it is stated that teaching activities in out-of-school learning environments support and enrich the teaching activities performed in the classroom (Berberoğlu & Uygun, 2013). Since planning trips to these environments provide the opportunity to learn by practicing and experiencing, they will save the courses from being addicted to the book and the classroom environment (Demir, 2007a; Özgen, 2011). Out-of-school learning environments known as informal learning environments include natural environments (streams, lakes, caves, etc.), mass media (radio, television, newspaper, etc.), science, sports and art centers, schoolyards, various parks, and gardens (botanical parks, arboretums, zoos, etc.). They can also consist of many areas such as museums, forested areas, national parks, observatories, factories (Bozdoğan, 2016; Eshach, 2007; Hannu, 1993; Howe & Disinger, 1988). It can be said that the education given in these environments is more fun than the formal environments, provides the opportunity to act flexibly, provides the opportunity to have natural experiences, and offers different individuals the chance to gain different experiences with different activities (Taylor & Caldarelli, 2004). In addition, studies have shown that research conducted outside the classroom or out of school enriches the learning environment (Ramey-Gassert, 1997), increases students' self-confidence, that is, their courage to learn (Bozdoğan, 2007, 2016; Melber & Abraham, 1999), increases academic achievement by supporting learning experiences at school (Bozdoğan, 2008; Gerber, Cavallo & Marek, 2001; Hannu, 1993), and provides the development of scientific process skills (Erten & Taşçi, 2016).

The positive effects of out-of-school learning environments on education and training are not limited to these. These effects can be increased by the diversity of experimental research (e.g., the diversity of the out-of-school learning environment, the diversity in the research to be conducted in these environments, the diversity between learning groups, etc.) (Arık & Yılmaz, 2020). Teachers have great responsibilities in this context. In the literature, the planned trips that are well associated with the subjects and achievements in the curriculum are thought to yield positive results (Bowker & Tearle, 2007; Kisiel, 2005; Tal, Bamberger & Morag, 2005). However, studies show that teachers do not much prefer out-of-school environments as a learning environment (Bowker & Tearle, 2007; Kisiel, 2005; Tal, Bamberger & Morag, 2005). Similarly, some other studies also support the same results that teachers do not much prefer out-of-school environments as learning environments (Bozdoğan, 2007; Carrier, 2009; Güven, Gazel & Sever, 2010; Moseley, Reinke & Bookout, 2002; Orion, Hofstein, Tamir & Giddings, 1997; Pekin & Bozdoğan, 2021; Simmons, 1998; Tatar & Bağrıyanık, 2012). This situation can be explained by the difficulties faced by teachers especially in organizing field trips (Bozdoğan, 2007; Pekin & Bozdoğan, 2021; Tal & Morag, 2009). Bozdoğan (2016) stated that the difficulties experienced by the teachers during the field trip are classified as (1) administrative, guidance, and pedagogical difficulties, (2) not being able to take an active role in the field trip, (3) not having any idea about the planning of the field trips, and (4) not being sufficient in guiding the students.

These challenges that teachers may encounter in out-of-school learning environments cause

negative attitudes and thoughts, anxiety, worries, and even fear for these environments in teachers (Bozdoğan, 2018; Pekin & Bozdoğan, 2021; Tatar & Bağrıyanık, 2012). Anxiety can be defined as the feeling of tension and fear against the possibility of a danger, which can originate from both the inner and outer worlds of individuals. Spielberger (1972) defines anxiety as unpleasant, emotional, and observable reactions such as sadness, perception, and tension caused by stressful situations (as cited in Büyüköztürk, 1997). It can be said that there are not many studies investigating the concerns about the teachers not using the out-of-school environments in their courses. A study conducted in this context determined that prospective science teachers have various concerns about ensuring the safety and control of students during the field trip, not being able to meet the guidance needs, and not keeping the motivation of the students at a high level (Bozdoğan, 2018). Another study also determined that the majority of secondary school teachers (Science, Social Studies, Turkish, and Mathematics teachers) were concerned about the safety of students during the field trip (Pekin & Bozdoğan, 2021). However, different studies reveal that teachers cannot use out-of-school learning environments for their courses much due to financial difficulties, lack of time, the intensity of the curriculum, and various concerns about safety (Demir, 2007b; Kılıç, 2018).

When the effect of out-of-school learning environments on the lifelong learning process is observed, it is seen that increasing the number of studies about these learning environments has great importance. This is possible by eliminating the worries, anxieties, or fears of teachers who plan education and training programs in out-of-school learning environments. When the literature is examined, it is seen that various measurement tools have been developed to be used to measure various characteristics or situations in out-of-school learning environments. The first of these is the 30-item scale developed by Bozdoğan (2016) to measure the self-efficacy belief of organizing trips to out-of-school environments. The first scale is the 30-item scale developed by Bozdoğan (2016) to measure the self-efficacy belief of organizing trips to out-of-school environments. Again, in the literature, there is a 24-item scale developed by Balkan-Kıyıcı and Yavuz-Topaloğlu in 2016 to determine the attitudes, behaviors, activities, and competencies of science teachers about using out-of-school learning environments that support classroom education activities. Another scale is the "out-of-school learning regulation scale". The scale developed by Bolat and Köroğlu (2020) consists of 29 items and four factors. Adıyaman and Ünal (2020), on the other hand, developed an informal learning environment scale consisting of 10 items to determine the views of prospective teachers on environments such as science centers, botanical gardens, camps, industrial establishments, zoos, and museums. Another scale development study in the literature is the teacher self-efficacy beliefs scale for out-of-school learning activities and it consists of 29 items (Göloğlu-Demir & Çetin, 2021a). The other scale developed by Göloğlu-Demir and Çetin (2021b) is the teacher attitude scale towards out-of-school learning activities, which consists of 25 items. Another scale is the perception scale for out-of-school environments and consists of 16 items (Şen et al., 2021). When the literature on teacher anxiety in out-of-school learning environments is examined, there appears only one scale developed to measure the anxiety levels of science teachers (Üner, 2019). In this study, a scale was developed to investigate the concerns of teachers teaching in different branches in the preschool, primary, secondary, and high schools about organizing field trips to out-of-school learning environments. In addition, the items of the prepared measurement tool were prepared by focusing on teacher concerns before, during, and after the trip.

In this context, the aim of this study was determined as the development of a teacher anxiety scale about organizing trips to out-of-school learning environments. It is thought that this developed scale will contribute to the determination of teachers' concerns about organizing

trips to out-of-school learning environments and to take the necessary steps to eliminate their deficiencies in this context, by filling this gap in the literature.

Method

Research Model

This research is a scale development study designed according to the descriptive survey model. In descriptive survey studies, researchers ask a series of questions to explore the opinions of a large audience on a particular topic or subject. They encode and analyze their answers in standardized categories and generalize the analysis results to similar student groups (Fraenkel, Wallen, & Hyun, 2011). In this study, a descriptive survey model was used because a large sample was needed to investigate teachers' anxieties about organizing trips to out-of-school learning environments (Cohen, Manion, & Morrison, 2007). In this direction, validity, reliability, and item analyzes were conducted to test the psychometric properties of the Teacher Anxiety Scale for Organizing Trips to Out-of-School Learning Environments, which was prepared by the researchers in line with expert opinions.

Study Group

The data of the study were collected from volunteer participants in the fall semester of the 2021-2022 academic year. The study group consisted of 394 teachers (307 females, 77.9%; 87 males, 22.1%) who were teaching at the pre-school, primary, secondary, and high school levels of public and private education institutions affiliated to the Ministry of National Education in different provinces of Turkey. The mean age of the study group is 34.2. The study group was determined through convenience sampling and snowball sampling methods. For factor analysis, the study group is acceptable to have a sample in the range of 100-200 (MacCallum, Widaman, Zhang, & Hong, S., 1999). At the same time, it is expressed in the literature that the number of items in the sample should be five times (Child, 2006; Tavşancıl, 2002). In this study, the sample was tried to be determined by taking the situations expressed in the literature into account.

Development Stages of the Scale

Validity, reliability, and item analyzes were conducted in order to test the psychometric properties of the measurement tool, which was prepared to determine the teachers' anxieties about trips to out-of-school learning environments. In the development of the measurement tool, the process shown in Figure 1 was followed (Cohen & Swerdlik, 2010; Crocker & Algina, 1986; DeVellis, 2003; Fraenkel, et al., 2011; Kan, 2009; Özcan & Arık, 2019; Tezbaşaran, 2008).

```
┌─────────────────────────────────────────────────────────────────┐
│                        Literature review                          │
├──────────────────┬────────────────────────┬──────────────────────┤
│     Anxiety      │ Out-of-school learning  │   Measurement tools  │
│                  │      environments       │                      │
└──────────────────┴────────────────────────┴──────────────────────┘
                              ▼
┌─────────────────────────────────────────────────────────────────┐
│                     Creating the item pool                        │
└─────────────────────────────────────────────────────────────────┘
                              ▼
┌─────────────────────────────────────────────────────────────────┐
│            Expert opinions (Content and face validity)            │
├──────────────────┬────────────────────────┬──────────────────────┤
│   Field expert   │ Assessment and          │ Turkish language     │
│                  │ evaluation expert       │ expert               │
└──────────────────┴────────────────────────┴──────────────────────┘
                              ▼
┌─────────────────────────────────────────────────────────────────┐
│                        Pilot application                          │
└─────────────────────────────────────────────────────────────────┘
                              ▼
┌─────────────────────────────────────────────────────────────────┐
│               Final arrangements for trial form of scale          │
└─────────────────────────────────────────────────────────────────┘
                              ▼
┌─────────────────────────────────────────────────────────────────┐
│                         Data collection                           │
└─────────────────────────────────────────────────────────────────┘
                              ▼
┌─────────────────────────────────────────────────────────────────┐
│                          Data analysis                            │
├──────────────────────┬─────────────────┬──────────────────────────┤
│ Structural validity  │  Item analysis  │  Reliability analysis    │
│ (EFA ve CFA)         │                 │                          │
└──────────────────────┴─────────────────┴──────────────────────────┘
                              ▼
┌─────────────────────────────────────────────────────────────────┐
│              Final arrangements for final form of scale           │
└─────────────────────────────────────────────────────────────────┘
```
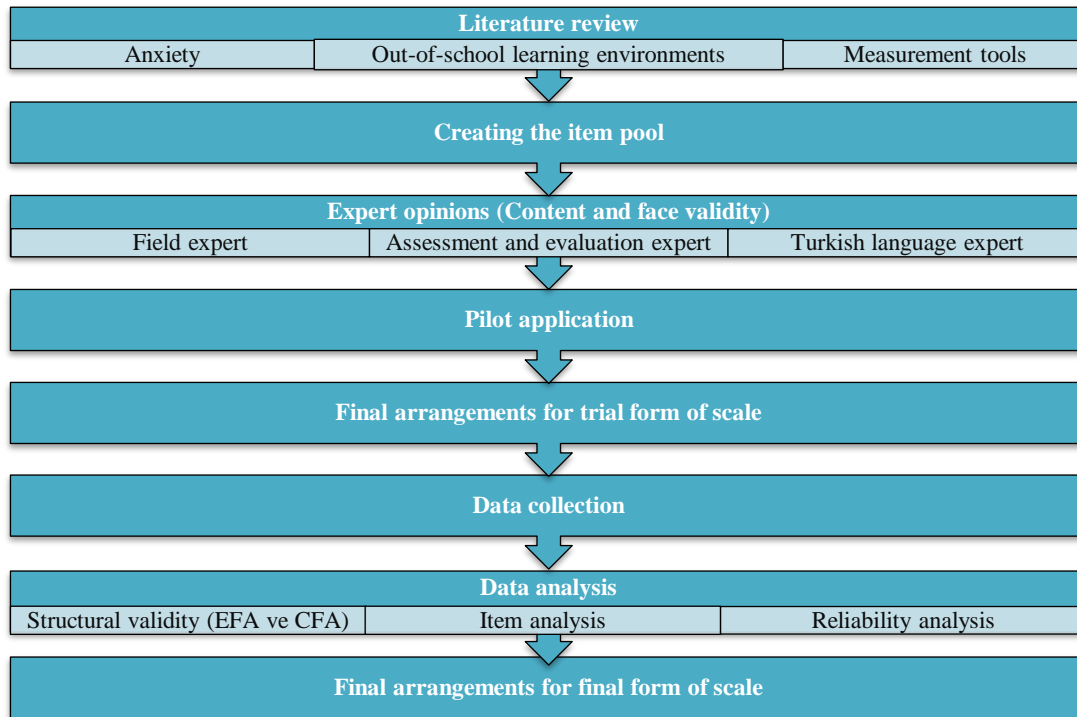
Figure 1. Scale development process

In the first stage of the research, the literature was reviewed. First of all, the literature on the structure of anxiety-related scales, indicators of anxiety, measurement, and characteristics of anxiety, and measurement tools related to anxiety were reviewed. While creating the scale items, Spielberger's (1966) anxiety theory was considered. Spielberger (1966) examined anxiety under two sub-factors as state and trait anxiety in his theory. While state anxiety is an emotional reaction state that occurs as a result of individuals perceiving a particular situation as a threat; on the other hand, trait anxiety is the tendency of the individual to have a state of an anxious life. The duration and severity of state anxiety vary according to the individual's perception of danger and the severity of the perceived threat. However, in the case of trait anxiety, individuals perceive ordinary events as worrisome situations. They feel themselves in a constant state of threat and danger. A high level of trait anxiety causes state anxiety to occur more frequently and intensely (Spielberger, 1966). This study tried to measure the state anxiety of the teachers about the trips to be organized to out-of-school learning environments. For this purpose, teachers were asked to evaluate how they felt (tension, uneasiness, nervousness, panic, etc.) during the trip. Then, literature was reviewed about the out-of-school learning environments and their features, the field trips to out-of-school learning environments and the problem situations encountered during these trips, and the measurement tools related to out-of-school learning environments (Bozdoğan, 2007; Bozdoğan, 2016; Eshach, 2007; Hannu, 1993; Howe & Disinger, 1988; Taylor & Caldarelli, 2004; Ramey-Gassert, 1997; Melber & Abraham, 1999). When the literature on out-of-school learning environments is examined, it is observed that there are measurement tools related to self-efficacy beliefs, perceptions, and attitudes towards out-of-school learning environments (Adıyaman & Ünal, 2020; Balkan Kıyıcı & Yavuz Topaloğlu, 2016; Bolat & Köroğlu, 2020; Bozdoğan, 2016; Göloğlu Demir & Çetin, 2021a, 2021b; Şen et al., 2021). In a study similar to this research, a measurement tool was developed to measure the anxiety of science teachers about out-of-school learning environments (Üner, 2019).

In the second stage, the items of the measurement tool were written in light of the information obtained from the literature (Adıyaman & Ünal, 2020; Balkan Kıyıcı & Yavuz Topaloğlu, 2016; Bolat & Köroğlu, 2020; Bozdoğan, 2007, 2016, 2018; Göloğlu Demir & Çetin, 2021a, 2021b; Pekin & Bozdoğan, 2021; Şen et al., 2021; Tal & Morag, 2009; Tatar & Bağrıyanık, 2012; Üner, 2019). While preparing the items of the measurement tool, the following has been considered: The items are written in a plain and understandable language, the items examine the relevant behaviour and are not narrow-scoped, each item contains only a judgment, opinion, emotion, or thought, unnecessary explanations and expressions are avoided in the item's origin, no expressions leading to options are used in the item's origin, and the items are appropriate for the student level (Crocker & Algina, 1986; Cohen & Swerdlik, 2010; DeVellis, 2003). The items of the measurement tool were prepared in a way to consist of a total of 56 items including situations that cause anxiety before the trip (18 items), situations that cause anxiety during the trip (28 items), and situations that cause anxiety after the trip (10 items) in the process of organizing a trip to out-of-school learning environments. For ranging the measurement tool, the five-point Likert type method was preferred since it provides direct and easier measurement and is sensitive and useful. Accordingly, each positive item in the scale was ranked as; "Strongly agree (5 points)", "Agree (4 points)", "Neutral (Neither agree nor disagree) (3 points)", "Disagree (2 points)" and "Strongly disagree (1 point)". The data collection tool consists of two subsections. In the first section, which aims to determine demographic information, there are multiple-choice and open-ended items. In the second section, there are five-point Likert-type items to determine the anxiety levels of teachers about out-of-school learning environments.

In the third stage, the content and face validity of the measurement tool was examined in line with expert opinions. The 56-item measurement tool, which was prepared by the researchers in a five-point Likert type, was examined by a field expert (out-of-school learning field expert), an assessment and evaluation expert, and a Turkish language expert. Then, the items were ranked according to their validity (sufficient, insufficient, should be changed). The experts examined whether the items measured teacher concerns about out-of-school learning environments and the comprehensibility of expressions and statements. As a result of expert opinion, the title of the measurement tool "Teacher Anxiety Scale for Out-of-School Learning Environments" was changed to "Teacher Anxiety Scale for Organizing Trips to Out-of-School Learning Environments". In addition, the item stating "I am concerned about the fact that all of the misconceptions are determined after the trip, which is acquired in out-of-school learning environments" was found insufficient as a result of expert opinion and was excluded from the measurement tool because it was out of the research content. Three new items were added to the measurement tool in line with expert opinions. In addition, the scale items were listed in line with the expert opinions, the instruction regarding the measurement tool (the purpose of the study, how to do the coding) was arranged, and the face validity of the measurement tool was ensured by preparing its pilot form.

In the fourth stage, as a result of expert opinions, the five-point Likert type measurement tool consisting of 58 items was applied to eight science teachers who were continuing their graduate education in the field of science education. The teachers were then asked to score and evaluate the items. As a result of the pilot application, it was determined that the average scores of the items ranged between 1.625 and 4.375 (mean=2.84, SD=.64). Accordingly, items with an average score higher than .80 were accepted as having sufficient face validity. The incomprehensibility of the items was determined within the scope of teachers' opinions and necessary arrangements were made. In addition, an item recommended by the teachers was included in the measurement tool.

Data Collection

In the fifth stage, the 59-item "Teacher Anxiety Scale for Organizing Trips to Out-of-School Learning Environments" was prepared for application by making final arrangements before the application. An "informed consent form" explaining that participation in the research is on a voluntary basis and that participants can withdraw from the research at any time, and an introduction stating the purpose of the research was prepared and added to the measurement tool.

In the sixth stage, the data of the research were collected through the measurement tool prepared for the implementation. Applications of the measurement tool take an average of 20-25 minutes. The data of the study were collected in a two-week period.

Data Analysis

In the seventh stage, the data obtained from the research were analyzed in order to ensure the validity and reliability of the measurement tool. At this stage, structural validity, reliability, and item analyzes were carried out.

In order to determine the extent to which the measurement tool measures the structure that is intended to be measured, structural validity was examined (Crocker & Algina, 1986; Field, 2013). It was carried out through factor analysis, which is called a 'data reduction technique' (Pallant, 2015). Factor analysis is a statistical test to explore whether variables in a single data set form relatively independent consistent subsets with other variables in the same data set. There are two main types of factor analysis: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) (Tabachnick, & Fidell, 2013). In this study, in order to ensure the structural validity of the measurement tool; Data were analyzed using the 'hybrid model' in which EFA and CFA were used together (Aydın, Yerdelen, Yalmancı & Göksu, 2014; Matsunaga, 2010).

Through EFA, data are defined and summarized by grouping the interrelated variables together. At this stage, variables are usually combined, and hypotheses related to the basic process are obtained. In this study, factor extraction techniques of principal component analysis (PCA) and principal axis analysis (PAF) were used together (Tabachnick, & Fidell, 2013). The reason for using two different factor extraction techniques in the study is that there are some opinions in the literature that PCA is not a factor analysis (Costello, & Osborne, 2005; Field, 2013; Matsunaga, 2010). In this study, EFA was conducted on two study groups. In the first part, EFA was performed through PCA in order to present an empirical summary of the data set; In the second part, repeated EFA was performed through PAF in order to free the data set from error variance and provide a theoretical solution (Tabachnick, & Fidell, 2013). The parallelism between the first EFA and the repeated EFA will ensure the consistency of the data obtained and the support of the structure obtained as a result of factor analysis (Aydın, et al., 2014). Statistical Package for the Social Sciences (SPSS) (v. 21.0) program was used to analyze the data in EFA (Tabachnick, & Fidell, 2013).

After the EFAs were completed, the CFA was performed to determine the suitability of the obtained factor structure. CFA is a statistical technique based on testing (confirming) theories and hypotheses about the structure formed by a series of variables (Pallant, 2015). In CFA, the suitability of the factors determined by EFA to the factor structures determined by hypotheses is tested. In order to determine the model fit obtained as a result of DFA, $\chi 2$ /df, GFI, CFI, and RMSEA values are examined. However, IFI, RMR, NFI, and AGFI values can also be checked

(Karagöz, 2019). In this study, the values of χ2 /df, GFI, CFI, NFI, IFI, TLI, RMSEA, and 90% confidence interval, which were suggested in the literature, were examined to determine the model fit. SPSS Analysis of Moment Structures (AMOS) (v. 23.0) program was used to analyze the data in DFA.

Reliability analysis, which is defined as "the degree to which the measurement tool does not contain random errors", was determined by calculating the Cronbach's α internal consistency coefficient and Split-Half test reliability coefficient, which are among the methods that require a single test application (Crocker & Algina, 1986).

In the analysis of the item and test parameters of the measurement tool, the independent t-test results of the differences between the item-total test correlation (ITC) and the item average scores of the lower 27% and upper 27% groups, which were created according to the total score of the test, were examined (Büyüköztürk, 2020; Karagöz, 2019; Pallant, 2015; Tabachnick, & Fidell, 2013).

Ethical Committee

The scientific and ethical appropriateness of the study was evaluated by the Social and Human Sciences Ethics Committee of Tokat Gaziosmanpaşa University within the scope of the meeting dated 19.11.2021 and numbered 23. Session 01-23 and approved (Decision 23.08).

Findings

Study 1

The structural validity of the "Teacher Anxiety Scale for Organizing Trips to Out-of-School Learning Environments" was examined through EFA.

Sample 1

The sample of the first study in which EFA will be applied consists of 171 teachers (137 females, 80.1%; 34 males, 19.88%). The mean age of the study group is 33.9.

Procedure 1

In the factor analysis of the 59-item measurement tool (Field, 2013; Thompson, 2004), the Principal Component Analysis (PCA) technique, which is one of the most used factor extraction methods in social sciences (Warner, 2012; Matsunaga, 2010), and Promax (kappa=4), one of the oblique axis rotation methods, and EFA were conducted. Having a factor load value of at least .40 in the EFA (Field, 2013; Stevens, 2009), the difference between the two highest load values for an item should be at least .10 (Costello, & Osborne, 2005; Büyüköztürk, 2020) and have at least three items in a factor (MacCallum, et al., 1999; Comrey, & Lee, 1992).

Based on the data obtained as a result of the EFA, in the first stage, nine items (e.g.: "That out-of-school learning environments are boring makes me worried.") were removed because the difference between the two highest load values is lower than .10; 15 items (e.g.: "I refrain from that parents do not trust me about activities to be done in out-of-school environments.") were removed because factor load value was below .40. And three items were excluded because they were loaded on a single factor (e.g.: "Sharing the photos and videos taken during the trip on social media environments makes me nervous."), and four items were excluded because of

having similar meanings in the same factor (e.g.: The items of "I have some drawbacks whether the activities would not meet the goals during the trip," and "I have some concerns whether the trip would not meet its goals." have similar meanings, so, the first item with a lower load was removed from the scale).

After extracting the items, EFA was repeated with 28 items. As a result of the analysis performed to determine the suitability of the data structure for EFA, the Kaiser-Meyer-Olkin (KMO) value was found as .908. The chi-square value ($\chi 2$ (df=378) = 3801.843) which was calculated as a result of the Bartlett sphericity test was determined to show a statistically significant difference (p < .001). KMO value above .60 is an acceptable value (Field, 2013; Pallant, 2015). KMO value above .90 is considered perfect (Kaiser, 1974). The significant difference in the chi-square value calculated as a result of the Bartlett sphericity test indicates that the correlation matrix is suitable for the multivariate normal distribution (Field, 2013; Raykov, & Marcoulides, 2008; Tabachnick, & Fidell, 2013). As a result, the obtained data show that the correlation matrix provides the hypothesis of multivariate normality and is suitable for factor analysis. In this direction, the data obtained as a result of the EFA, reliability analysis, and item analysis conducted to ensure the validity and reliability of the "Teacher Anxiety Scale for Organizing Trips to Out-of-School Learning Environments" are given in Table 1.

Table 1. The result of reliability and validity analysis of "Teacher Anxiety Scale for Organizing Trips to Out-of-School Learning Environments"

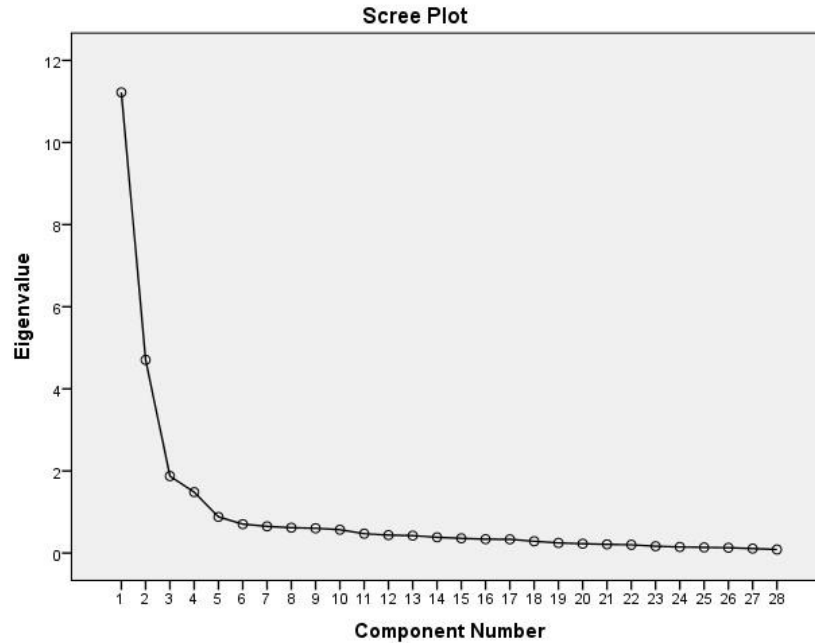| Item numbers	t L%27U%27[2]	Factors (FL/ITC/Crα*)[1]				Whole scale
		F1 (PR[3])	F2 (BR[4])	F3(SRR[5])	F4 (HR[6])	
I43	11.05**	.913/.685/.940				
I44	8.16**	.874/.607/.941				
I40	10.91**	.865/.720/.940				
I41	12.04**	.862/.764/.939				
I39	9.57**	.835/.646/.941				
I42	10.38**	.824/.665/.939				
I33	9.38**	.822/.639/.941				
I13	6.20**	.761/.552/.942				
I59	6.47**	.739/.564/.941				
I55	9.95**	.716/.599/.941				
I38	10.16**	.706/.664/.940				
I32	6.61**	.672/.569/.941				
I35	10.70**	.650/.654/.940				
I49.	11.39**	.495/.661/.940				
I5.	11.54**		.918/.587/.941			
I3	10.20**		.857/.521/.942			
I6	10.98**		.779/.605/.941			
I2	9.40**		.763/.592/.941			
I1	8.92**		.698/.514/.942			
I4	7.52**		.684/.483/.942			
I19	5.62**			.925/.467/.942		
I20	5.31**			.922/.470/.942		
I24	7.03**			.795/.530/.942		
I21	6.68**			.783/.521/.942		
I29	9.86**				.887/.583/.941	
I30	9.72**				.814/.559/.941	
I28	9.26**				.814/.586/.941	
I27	10.31**				.728/.583/.941	
Eigenvalue		11.222	4.705	1.874	1.483	
Ratio of variance explanation (%)		40.079	16.805	6.693	5.297	68.873
Number of items		14	6	4	4	28
Mean ($\bar{X}$)		33.67	20.80	17.03	14.56	86.05
Standard Deviation (s)		13.72	6.50	3.79	4.44	22.04
Skewness		.80	-.38	-1.29	-.56	.27
Kurtosis		.18	-.68	1.21	-.45	.25

Figure 2. Scree Plot

When the graph in Figure 2 is examined, it is observed that the factor structure is four. A highly accelerated decline after the first and second breakpoints on the graph is observed and after the third and fourth breakpoints, a significant decrease is observed, although not highly accelerated. At the breakpoints after the fourth factor, while the graph continues almost horizontal, any significant decrease is not observed. It is seen that the factor structure determined by considering Kaiser's eigenvalue and the scree plot are consistent.

The Pearson correlation coefficients between the factor scores of the "Teacher Anxiety Scale for Organizing Trips to Out-of-School Learning Environments" and the total score are given in Table 2.

Table 2. The correlation coefficient between the scores of factors and the whole scale

Factors	The correlation between the scores of the factors			
	F1	F2	F3	Whole scale
F1				.87*
F2	.41**			.74**
F3	.24**	.52**		.59**
F4	.41**	.51**	.58**	.71**

**p<.001

When Table 2 is examined, it is seen that the correlation coefficient between the factors varied between .24 (between F1 and F3) and .58 (between F3 and F4) and the t value was significant (p<.001); the correlation between whole scale and the factors varied between .59 (Total scale and F3) and .87 (Total scale and F1) and the t value was significant (p<.001).

The items were examined to name the factors. The items in the first factor refer to the anxiety related to the teacher's pedagogy, the items in the second factor are about the anxiety related to obtaining permission procedures before the trip, the items in the third factor are for the anxiety related to the safety risks that students may encounter during the trip. And the items in the

fourth factor refers to the anxiety about the possibility of the students' giving harm to themselves, others, and the materials and works during the trip. In line with this information, the first factor was named Pedagogy-Related (PR) Anxiety (M13, M32, M33, M35, M38, M39, M40, M41, M42, M43, M44, M49, M55, and M59), the second factor was named Bureaucracy-Related (BR) Anxiety (M1, M2, M3, M4, M5, and M6), the third factor was named Safety Risks-Related (SRR) Anxiety (M19, M20, M21, and M24), and the fourth factor was named Harm-Related (HR) Anxiety (M27, M28, M29, and M30).

Study 2

EFA was repeated in a different sample to determine how the four-factor structure obtained as a result of EFA was distributed in a different sample.

Sample 2

The sample of the study, in which repeated EFA will be applied, consists of 81 teachers (64 females, 79.01%; 17 males, 20.99%). The mean age of the study group was 33.17.

Procedure 2

In order to determine how the 28-item-four factor structure formed as a result of the first EFA analysis was distributed in a different sample, the EFA was repeated using Principal axis factor analysis (PAF), a factor extraction technique, and Promax (kappa=4) method, one of the oblique axis rotation methods (Field, 2013; Thompson, 2004). Factor load criteria in the first study were used for the repeated EFA.

The KMO value (.866; KMO>.60), obtained as a result of repeated EFA, and the chi-square value, calculated as a result of the Bartlett sphericity test ($\chi 2$ (df=378) = 1991.613) showed a statistically significant difference (p < .001). This indicates that the correlation matrix provides multivariate normality assumption and is suitable for factor analysis (Field, 2013; Pallant, 2015; Kaiser, 1974; Raykov, & Marcoulides, 2008; Tabachnick, & Fidell, 2013).

According to the repeated EFA results, the 28 ($\bar{X}$=83.51) item four-factor structure obtained in the first EFA was supported. Similar to the first EFA result, the items in this data set have ranged from as; fourteen ($\bar{X}$=31.65) items in the first factor, six ($\bar{X}$=20.38) items in the second factor, four ($\bar{X}$=16.96) items in the third factor, and four ($\bar{X}$=14.51) items in the fourth factor. This four-factor structure obtained from this data set explains 66.741% of the total variance. Each sub-factor explains 40.298%, 15.855%, 6.206 and 4.382% of the total variance, respectively. The factor loadings of the items vary between .507 and .948 (factor loading>.40). Item-total correlations ranged from .44 to.75. It was determined that the t-test results showing the significance of the differences between the item averages of the lower 27% and upper 27% groups also showed statistically significant differences for all items. In order to determine the distribution of the scores of the measurement tool, skewness and kurtosis values were examined. The skewness value for the whole measurement tool was determined as .35 and the kurtosis value as .28. The skewness values were .98, -.56, -1.39, and -.58 for the four factors, respectively, and the kurtosis value was determined as .48, -.42, 1.38, and -.57.

The Cronbach's (α) reliability coefficient related to the reliability of the measurement tool was determined as .945 for the total of the measurement tool. Cronbach's (α) reliability coefficient was determined as .957, .884, .927, and .895 for four factors, respectively. In addition, the

Cronbach's α reliability coefficient obtained when the item was removed varied between .942 and .945.
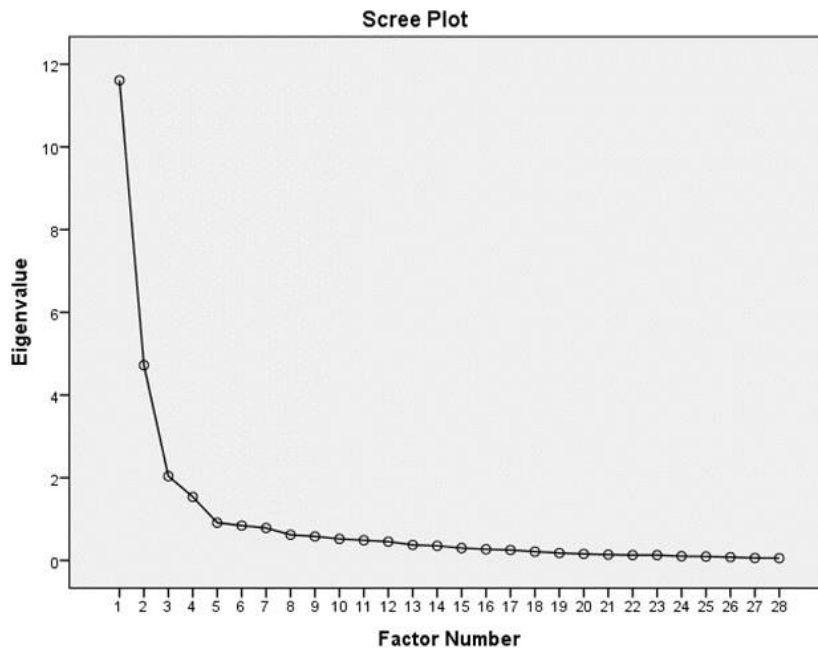

Figure 2. Scree plot for repeated EFA

When the scree plot graph (Figure 3) was examined to determine the most appropriate factor structure and number, it was determined that the first EFA and the second EFA graphs were similar. Similar to the graph in the first analysis, a highly accelerated decline after the first and second breakpoints was observed; after the third and fourth breakpoints, a significant decline was observed.

It was determined that the correlation coefficient between the factors ranged between .23 (between F1 and F3) and .57 (between F3 and F4) and it showed t value significant ($p < .001$). In addition, it was determined that the correlation between the whole measurement tool and the factor varied between .57 (Total scale and F3) and .87 (Total scale and F1), and the t value was significant ($p < .001$).

The validity, reliability, and item analysis results obtained from the second study show parallelism with the results obtained from the first study. As a result of the second study, the 28-item four-factor structure obtained in the first study was supported.

Study 3

In the third stage, confirmatory factor analysis (CFA) was performed to confirm the factor structure obtained as a result of the EFA analysis.

Sample 3

The sample of the study, in which CFA will be applied, consists of 223 teachers (170 females, 76.23%; 53 males, 23.77%). The mean age of the study group was 34.42.

Procedure 3

Before CFA, the data set was examined in terms of assumptions. First, it was examined whether there was any missing data in the data set. After the analysis made by SPSS, no missing data was found in the data set. Secondly, the extreme values in the data set were examined. As a result of the extreme value analysis, it was determined that there was no data that should be removed from the data set. Thirdly, the normality of univariate and multivariate of the new data set was examined. The kurtosis value of the data set was found to vary between -1.18 and 2.59; the skewness value was determined to vary between -1.78 and .91. These values show that the data set provides the univariate normality. After the analysis with the AMOS program to ensure multivariate normality, the data of a person with a high Mahalanobis distance value and tending to disrupt the normal distribution were removed from the data set.

The 28 items and the four-factor structure obtained as a result of the EFA were examined by performing CFA through the SPSS AMOS (v. 23.0) program. Maximum likelihood technique was used in DFA. For DFA fit indices of the 222 person data set, "Chi-Square/Degree of Freedom" (χ^2 /df), "Comparative Fit Index" (CFI), "Goodness of Fit Index" (GFI), "Normed Fit Index" (NFI), "Incremental Fit Index" (IFI), "Tucker-Lewis Index" (TLI) and "Root Mean Square Error of Approximation" (RMSEA) values were examined. The fit indices obtained as a result of CFA to verify the four-factor structure of the 28-item measurement tool are as follows: χ^2 (df=329) = 539.125; χ^2 /df = 1.64; p < .001; CFI = .96; GFI = .86; NFI = .90; IFI=.96; TLI=.95; RMSEA=.05. When the goodness of fit indexes of the model are examined, it was determined that the ratio of Chi-Square to degrees of freedom showed a good fit (χ^2 /df<2) (Tabachnick, & Fidell, 2013); RMSEA value was below the threshold value (RMSEA≤.06) (Hu & Bentler, 1999) and was in close fit (RMSEA 05) (Browne & Cudeck, 1993); CFI, IFI, and TLI values showed perfect agreement (≥.90) (Hu & Bentler, 1999); GFI value showed an acceptable fit (≥.85) (Browne & Cudeck, 1993); and, NFI value showed an acceptable fit (≥.90) (Browne & Cudeck, 1993; Sümer, 2000). The estimated standardized regression coefficients for each item range from .57 to .95. In Table 3, correlations between factors (Phi values) that were obtained as a result of CFA are given.

Table 3. Correlations between Factors (Phi values)

| Factors | The correlation coefficients between the scores of the factors | | |
	PR Anxiety	BR Anxiety	SRR Anxiety
PR Anxiety			
BR Anxiety	.46[**]		
SRR Anxiety	.36[**]	.64[**]	
HR Anxiety	.37[**]	.22[**]	.50[**]

[**]p<.001

When Table 3 is examined, it was determined that the correlations between the factors obtained as a result of CFA varied between .22 and .64 and showed a statistically significant difference.

As a result, it can be said that the four-factor structure obtained as a result of EFA is compatible with the model obtained as a result of CFA. It can be stated that the factor structure of the 28-item four-factor measurement tool obtained as a result of EFA and repeated EFA was confirmed as a result of CFA. The item numbers, item averages, and standard deviations, and skewness, and kurtosis values regarding the final version of the measurement tool are given in Table 4.

Table 4. Descriptive statistics of the items in the final form of the scale

Factors	Item number (Trial form)	Item number (Final form)	$\overline{X}$	s	Skewness Statistics	Kurtosis Statistics
BR Anxiety	I6	I1	3.694	1.213	-.587	-.635
	I5	I2	3.203	1.352	-.151	-1.144
	I4	I3	3.680	1.336	-.543	-1.013
	I3	I4	3.198	1.367	-.244	-1.142
	I2	I5	3.833	1.253	-.749	-.618
	I1	I6	3.680	1.223	-.489	-.794
SRR Anxiety	I19	I7	4.360	.982	-1.574	1.779
	I20	I8	4.441	.924	-1.776	2.588
	I21	I9	4.383	.952	-1.615	2.211
	I24	I10	4.347	.933	-1.544	2.070
HR Anxiety	I30	I11	3.680	1.326	-.625	-.792
	I29	I12	3.626	1.297	-.603	-.746
	I28	I13	3.883	1.249	-.880	-.250
	I27	I14	3.869	1.224	-.847	-.276
PR Anxiety	I13	I15	2.189	1.284	.853	-.400
	I32	I16	2.266	1.334	.835	-.462
	I33	I17	2.423	1.322	.649	-.720
	I35	I18	2.793	1.329	.233	-1.109
	I38	I19	2.577	1.411	.381	-1.127
	I39	I20	2.757	1.231	.179	-.886
	I40	I21	2.414	1.184	.631	-.394
	I41	I22	2.469	1.224	.518	-.653
	I42	I23	2.586	1.280	.393	-.868
	I43	I24	2.324	1.212	.704	-.325
	I44	I25	2.234	1.240	.905	-.110
	I49	I26	2.829	1.321	.140	-1.099
	I55	I27	2.797	1.375	.221	-1.176
	I59	I28	2.451	1.271	.582	-.700

Reliability of the Final Form of the Measurement Tool

In order to ensure the reliability of the final form of the measurement tool, the reliability coefficient Crα was calculated. The Crα reliability coefficient calculated for the total measurement tool was determined as .944. The Crα reliability coefficients calculated for each sub-factor, respectively, were determined as .868 (BR Anxiety), .922 (SRR Anxiety), .903 (HR Anxiety), and .952 (PR Anxiety) respectively. The Split-Half reliability coefficient, which was made to ensure the two-half test reliability of the measurement tool, was determined as .893 for the first half and .948 for the second half. These reliability coefficient values obtained show that the measurement tool has sufficient reliability (Bland & Altman, 1997; DeVellis, 2003).

Discussion and Conclusions

This study tried to develop a valid and reliable measurement tool to determine the anxiety levels of teachers about organizing trips to out-of-school learning environments. The sub-dimension of "Spielberger's (1966) Anxiety Theory" called "state anxiety" and the information expressed in the literature about out-of-school learning environments were used to create the items of the measurement tool. In the development of the measurement tool, while the trips to the out-of-school learning environments were determined as a situation, the processes before, during, and after the trip were taken into consideration. As a result of factor and item analyzes, a measurement tool consisting of 28 items and four sub-factors were obtained. The sub-factors of the measurement tool were determined as "Bureaucracy-Related Anxiety", "Safety Risks-Related Anxiety", "Harm-Related Anxiety" and "Pedagogy-Related Anxiety". In the first factor called "Bureaucracy-Related Anxiety", there are six items that aim to determine teachers' concerns about the process of getting permission before the trip. In the

second factor, named "Safety Risks-Related Anxiety", there are four items that aim to determine teachers' concerns about negative situations that may happen to students (children) during the trip. In the third factor, named "Harm-Related Anxiety", there are 4 items that aim to determine the teachers' concerns about students' (children's) harming each other, others, or objects in the environment during the field trip. In the last factor called "Pedagogy-Related Anxiety", there are 14 items that include teachers' pedagogical concerns such as learning-teaching process, classroom management, measurement, and evaluation.

The final stage confirmatory factor analysis was conducted regarding the structural validity of the measurement tool. As a result of CFA, it was concluded that 28 items and a four-factor structure were supported. The Cronbach's alpha reliability coefficient calculated to determine the reliability of the measurement tool was determined as .944 for the total measurement tool. The Cronbach's alpha reliability coefficient calculated for the sub-factors of the measurement tool was determined as .868 for "Bureaucracy-Related Anxiety", .922 for "Safety Risks-Related Anxiety", .903 for "Harm-Related Anxiety", and .952 for "Pedagogy-Related Anxiety". At the same time, the Split-Half reliability coefficient values of the measurement tool, which were made to provide two-half test reliability, were determined as .893 for the first half and .948 for the second half, respectively. All the data obtained show that the measurement tool is a valid and reliable measurement tool to determine teachers' anxiety about trips to out-of-school learning environments.

It is known that teachers do not prefer out-of-school learning environments (Bozdoğan, 2007; Carrier, 2009; Güven, et al., 2010; Moseley, et al., 2002; Orion, et al., 1997; Pekin & Bozdoğan, 2021; Simmons, 1998; Tatar & Bağrıyanık , 2012) and they have negative attitudes and thoughts, worry, anxiety and even fears for these environments (Bozdoğan, 2018; Pekin & Bozdoğan, 2021; Tatar & Bağrıyanık, 2012). However, when the national literature is examined, it is seen that the number of studies investigating teachers' concerns about out-of-school learning environments is quite limited (Bozdoğan, 2018; Pekin & Bozdoğan, 2021; Demir, 2007b; Kılıç, 2018). At the same time, it is seen that the number of measurement tools to be used to determine teacher anxiety against out-of-school learning environments and field trips to these environments is quite limited (Üner, 2019). In this scale development study, Üner (2019) developed a scale to determine the anxiety levels of science teachers about out-of-school learning environments. In this study, EFA, item analysis, and reliability analysis were performed for the scale development analysis. Our study differs from other studies by including the teachers teaching in different branches of preschool, primary school, secondary school, and high school in the study group. During the development of the measurement tool, repeated EFA and CFA were performed. While the structure of the measurement tool obtained as a result of EFA is supported by repeated EFA; This structure obtained by CFA was confirmed (Crocker & Algina, 1986; Tabachnick, & Fidell, 2013). At the same time, focusing on the sub-dimension of "Spielberger's (1966) Anxiety Theory" called "state anxiety" in the preparation of the items of the measurement tool and focusing on the processes before, during, and after the trip to the out-of-school learning environments indicates the originality of the assessment tool.

References

Adıyaman, A., & Ünal, F. (2020). Pre-service teachers' views on the informal learning environments. *Journal of Humanities and Tourism Research, 10*(1), 95-114. doi: https://doi.org/10.14230/johut784.

Arık, S. (2019). *Acid-base rocket: Argumentation based inquiry STEM activity*. *II. International 19 May* Innovative Scientific Approaches Congress, 242–255, Samsun.

Arık, S. (2021). STEM teachers at work for a sustainable future. H. Özcan (Ed.), *STEM Educational Applications-II* (pp. 132–156). İstanbul: Pusula.

Arık, S., & Yılmaz, M. (2020). Systematic review of experimental studies on environmental academic achievement and attitudes toward environment. *Cumhuriyet International Journal of Education*, *9*(2), 494–535. doi: https://doi.org/10.30703/cije.624562.

Aslan, S., & Kan, A. (2021). A validity and reliability study on the scale for the attitude toward Syrian students. *The Journal of Turkish Social Research, 25*(1), 193-206.

Aydın, S., Yerdelen, S., Yalmancı, S. G., & Göksu, V. (2014). Academic motivation scale for learning biology: A scale development study. *Education and Science*, *39*(176), 425–435.

Balkan-Kıyıcı, F., & Yavuz-Topaloğlu, M. (2016). A Scale development study for the teachers on out of school learning environments. *Malaysian Online Journal of Educational Sciences, 4*(4), 1-13.

Berberoğlu, E., & Uygun, S. (2013). Examining of outdoor education development in the world and in Turkey. *Mersin University Journal of The Faculty of Education, 9*(2), 32-42.

Bland, J. and Altman, D. (1997) Statistics Notes: Cronbach's Alpha. *BMJ,* 314, 572. doi: http://dx.doi.org/10.1136/bmj.314.7080.572.

Bolat, Y., & Köroğlu, M. (2020). Out-of-school learning and scale of regulating out-of-school learning: Validity and reliability study, *International Journal of Education Technology and Scientific Researches, 5*(13), 1630-1663.

Bowker, R., & Tearle, P. (2007). Gardening as a learning environment: A study of children's perceptions and understanding of school gardens as part of an international project. *Learning Environments Research, 10*(2), 83-100.

Bozdoğan, A. E. (2007). *Role and Importance of Science and Technology Museum in Education*. (Unpublished doctoral dissertation). Gazi University, Institute of Educational Science. Ankara.

Bozdoğan, A. E. (2008). The assessment of the science centres of the pre-service science teachers in terms of science education: The case of Feza Gürsey Science Center. *The Journal of Uludağ University Faculty of Education, 21*(1), 19–41.

Bozdoğan, A. E. (2016). Development of self-efficacy belief scale for planning and organizing educational trips to out of school settings. *Journal of Theoretical Educational Science, 9*(1), 111-129. doi: http://dx.doi.org/10.5578/keg.9475

Bozdoğan, A. E. (2018). The effects of elective course designed with different contents on pre-service science teachers' self-efficacy beliefs and knowledge about organizing curriculum based field trips. *Journal of Baltic Science Education*, *17*(3), 446–461. doi: https://doi.org/10.33225/jbse/18.17.446.

Browne, M. W., & Cudeck, R. (1993). *Alternative ways of assessing model fit*. In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models (pp. 136-162). Newbury Park, CA: Sage.

Büyüköztürk, Ş. (1997). Development of anxiety scale for research. *Educational Administration: Theory and Practice, 3*(4), 453-464.

Büyüköztürk, Ş. (2020). *Manual of data analysis for social sciences.* (28th Ed.). Ankara: Pegem.

Carrier, S. J. (2009). The effects of outdoor science lessons with elementary school students on preservice teachers' self-efficacy. *Journal of Elementary Science Education*, *21*(2), 35–48. doi: https://doi.org/10.1007/BF03173683.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245–276. doi: https://doi.org/10.1207/s15327906mbr0102_10.

Child, D. (2006). *The Essentials of factor analysis* (3th.). Continuum.

Clayton, S., & Karazsia, B. T. (2020). Development and validation of a measure of climate change anxiety. *Journal of Environmental Psychology*, *69*, 101434. https://doi.org/https://doi.org/10.1016/j.jenvp.2020.101434.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159. doi: https://doi.org/10.1037/0033-2909.112.1.155.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education* (6th ed.). London: Routhledge.

Cohen, R., & Swerdlik, M. (2010). *Psychological testing and assessment: An introduction to tests and measurement* (7th ed.). New York: McGraw Hill.

Comrey, A.L., & Lee, H.B. (1992). *A first course in factor analysis* (2nd ed.). New York: Psychology Press. doi: https://doi.org/10.4324/97813158275.

Costello, A. B. ve Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*(7), 1-9.

Crocker, L., & Algina, J. (1986) *Introduction to classical and modern test theory*. New York, Harcourt Brace Jovanovich.

Çokluk, Ö., Şekercioğlu, G. & Büyüköztürk, Ş. (2010). *Multivariate statistical SPSS and LISREL applications*. Ankara: Pegem.

Demir, M. K. (2007a). The study of classroom teacher candidates' perspectives towards field trip method. *Gazi University Journal of Gazi Educational Faculty, 27*(3), 83-98.

Demir, M. K. (2007b). Field trip application characteristics of classroom teachers in life study and social studies courses. *The Journal of Turkish Educational Sciences, 5*(2), 323-341.

DeVellis, R. F. (2003). *Scale development: Theory and application* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Erten, S., & Taşçi, G. (2016). Developing activities of out of the school learning environments for science classes, and analysing their effects on students' scientific process skills. *Erzincan University Journal of Education Faculty (EUJEF)*, *18*(2), 638–657. doi: https://doi.org/10.17556/jef.41328

Eshach, H. (2007). Bridging in-school and out-of-school learning: Formal, non-formal, and informal education. *Journal of Science Education and Technology*, *16*(2), 171–190. doi: https://doi.org/10.1007/s10956-006-9027-1.

Field, A. (2013). *Discovering statistics using IBM SPSS Statistics* (4th ed.). London: Sage Publications.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2011). *How to design and evaluate research in education.* (8th ed.). New York: McGraw-Hill.

Gerber, B. L., Cavallo, A. M. L., & Marek, E. A. (2001). Relationships among informal learning environments, teaching procedures and scientific reasoning ability. *International Journal of Science Education*, *23*(5), 535–549. doi: https://doi.org/10.1080/09500690116971.

Göloğlu-Demir, C., & Çetin, F. (2021a). Development of a scale to measure teachers' self-efficacy beliefs toward out-of-school learning (OoSL) activities. *The Journal of Turkish Educational Science*, *19*(1), 613-634. doi: https://doi.org/10.37217/tebd.901426.

Göloğlu-Demir, C., & Çetin, F. (2021b). The development of an attitude scale to measure teachers' attitudes to out-of-school learning activities (OOSLA): A validity and reliability studies. *Third Sector Social Economic Review, 56*(2), 895-910.

Güven, A., Gazel, A. A., & Sever, R. (2004). Problems faced by history teachers in trip-observation practices. *Atatürk University Journal of Kazım Karabekir Education Faculty*, *9*, 225–235.

Hannu, S. (1993). *Science centre education, motivation and learning in informal education*. (Unpublished Doctoral Dissertation). Helsinki University, Finland.

Howe, R. W., & Disinger, J. F. (1988). *Teaching environmental education using out of school settings and mass media*. ERIC Clearinghouse for Science Mathematics and Environmental Education Columbus OH. doi: https://www.ericdigests.org/pre-9215/mass.htm.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. doi: https://doi.org/10.1080/10705519909540118.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement,* 20, 141–151. doi: https://doi.org/10.1177/001316446002000116.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31–36. doi: https://doi.org/10.1007/BF02291575.

Kan, A. (2009). *Qualities that must be present in measurement tools.* H. Atılgan (Ed.) In Measurement and evaluation in education (pp. 23-80). Ankara: Anı.

Karagöz, Y. (2019). *SPSS-AMOS-META Applied statistical analysis*. (2nd ed.). Ankara: Nobel Academy.

Kılıç, A. (2018). *Evaluation of oral and local history activities regarding "yesterday, today, tomorrow" theme in life studies course*. (Unpublished master thesis). İstanbul University Institute of Educational Sciences, İstanbul.

Kisiel, J. (2005). Understanding elementary teacher motivations for science fieldtrips. *Science Education, 86*(6), 936–955.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84–99. doi: https://doi.org/10.1037/1082-989X.4.1.84.

Matsunaga, M. (2010). How to factor-analyze your data right: do's, don'ts, and how-to's. *International Journal of Psychological Research*, *3*(1), 97–110. doi: https://doi.org/10.21500/20112084.854.

Melber, L. M., & Abraham, L. M. (1999). Beyond the classroom: Linking with Informal Education. *Science Activities*,*36*(1), 3–4. doi: https://doi.org/10.1080/00368129909601027.

Moseley, C., Reinke, K., & Bookout, V. (2002). The effect of teaching outdoor environmental education on preservice teachers' attitudes toward self-efficacy and outcome expectancy. *The Journal of Environmental Education*, *34*(1), 9–15. doi: https://doi.org/10.1080/00958960209603476.

Orion, N., Hofstein, A., Tamir, P., & Giddings, G. J. (1997). Development and validation of an instrument for assessing the learning environment of outdoor science activities. *Science Education*, *81*(2), 161–171.

Özcan, H., & Arik, S. (2019). Development of the attitude scale towards environmental pollution: Validity and reliability study. *Iğdır University Journal of Social Sciences,* 17, 425-456.

Özgen, N. (2011). A different approach to physical geography teaching method: Trip-observation supported education. *International Journal of Geography and Geography Education, 23*, 373-388.

Pallant, J. (2015). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS.* Maidenhead, Berkshire: McGraw Hill.

Pekin, M., & Bozdoğan, A. E. (2021). Examining secondary school teachers' self-efficacy in organizing trips to out-of-school environments in terms of different variables: Sample of Tokat province *International Journal of Turkish Educational Studies, 17*,114–133. doi: https://doi.org/10.46778/goputeb.956719

Ramey-Gassert, L. (1997). Learning science beyond the classroom. *The Elementary School Journal*, *97*(4), 433–450. doi: https://doi.org/10.1086/461875.

Raykov, T., & Marcoulides, G.A. (2008). *An introduction to applied multivariate analysis* (1st ed.). Routledge. doi: https://doi.org/10.4324/9780203809532.

Scherer, R. F., Wiebe, F. A., Luther, D. C., & Adams, J. S. (1988). Dimensionality of coping: factor stability using the Ways of Coping Questionnaire. *Psychological reports, 62*(3), 763–770. doi: https://doi.org/10.2466/pr0.1988.62.3.763.

Simmons, D. (1998). Using natural settings for environmental education: Perceived benefits and barriers. *The Journal of Environmental Education*, *29*(3), 23–31. doi: https://doi.org/10.1080/00958969809599115.

Spielberger, C. D. (1972). *Anxiety: Current trend in theory and research*. New York: Academic Press.

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences,* (5th ed.). Newyork: Routledge/Taylor & Francis Group.

Sümer, N. (2000). Structural equation models: Basic concepts and example practices. *Turkish Psychological Articles, 3*(6), 49-74.

Şen, A.İ, Ertaş-Kılıç, H., Oktay, Ö., Ekinci, S., & Kadirhan, Z. (2021). Learning science outside the classroom: Development and validation of the out-of-school learning environments perception scale. *Journal of Outdoor and Environmental Education, 24:*19–36. doi: https://doi.org/10.1007/s42322-020-00070-7.

Tabachnick, B.G., & Fidell, L.S. (2013). *Using multivariate statistics.* (6th Edition). Boston: Pearson Education.

Tal, R., Bamberger, Y., & Morag, O. (2005). Guided school visits to Natural History Museums in Israel: Teachers' roles. *Science Education, 89*(6), 920-935.

Tal, T., & Morag, O. (2009). Reflective practice as a means for preparing to teach outdoors in an ecological garden. *Journal of Science Teacher Education*, *20*(3), 245–262. doi: https://doi.org/10.1007/s10972-009-9131-1.

Tatar, N., & Bağrıyanık, K. E. (2012). Opinions of science and technology teachers about outdoor education. *Elementary Education Online, 11*(4), 882–896.

Tavsancil, E. (2002). *Measurement of attitudes and data analysis with SPSS*. Ankara: Nobel.

Taylor, E. W., & Caldarelli, M. (2004). Teaching beliefs of non-formal environmental educators: a perspective from state and local parks in the United States. *Environmental Education Research*, *10*(4), 451–469. doi: https://doi.org/10.1080/1350462042000291001.

Tezbaşaran, A. (2008). *The likert-type scale development guide*. (3th Edition). Mersin: Turkish Psychological Association.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications. In Exploratory and confirmatory factor analysis: Understanding concepts and applications.* Washington, DC: American Psychological Association. doi: https://doi.org/10.1037/10694-000.

Üner, S. (2019). *A study of an evaluation scale for science teachers' anxiety levels towards out-of-school learning environments.* (Unpublished master thesis). Hacettepe University Institute of Educational Sciences, Ankara.