Investigation of the Exam Question Types Attitude Scale for Secondary School Students: Development, Validity, and Reliability

Namudar İzzet KURBANOĞLU[a]   Mevlüde OLCAYTÜRK[b]

a: 🆔 0000-0002-5340-0855   🏛 Sakarya University, Turkiye   ✉ kurbanoglu@sakarya.edu.tr
b: 🆔 0000-0001-9032-7696   🏛 Sakarya University, Turkiye   ✉ mevludeolcay35@hotmail.com

Abstract

Developing students' positive attitudes toward a situation is extremely important. Many attitude scales aim to measure students' general attitudes toward a situation, but there is a scarcity of tools that can be easily employed to determine the type of questions students prefer in an exam. This paper reports the development and testing of an instrument like this for secondary school students, and 781 secondary school students participated in this study. The Exam Question Types Attitude Scale (EQTA-S) consisting of 10 items was developed in this study, and its validity and reliability were analyzed. Exploratory factor analysis was performed for the construct validity of the EQTA-S and an EQTA-S consisting of 10 items was obtained. The result of the confirmatory factor analysis showed that the scale is a good fit for the proposed one-factor structure. Additionally, the reliability of the scale was established by acceptable Cronbach alpha values that indicated reliability coefficients (Cronbach's alpha) of .81, .85, .84, and .82, respectively. Therefore, the findings of the study revealed that secondary school students' attitude scale toward exam question types provides valid and reliable results. As a result, this scale can be used to measure students' attitudes towards five different types of exam questions.

Keywords

Attitude, exam question types, test format, validity, reliability.

INTRODUCTION

Education is a concept that is designed to enhance and develop human potential by guiding the learning desires of students. Education is a comprehensive, diverse, and complex undertaking, not only in terms of acquiring gains but also in terms of tools that teach and measure them. Thus, education plays an important role in improving human quality (Astalini et al., 2019; Darmaji et al., 2019). In this context, education is also the measure of a nation's progress in science and technology (Kurniawan et al., 2019; Maison et al., 2019). Therefore, the education system is vital for every country in the world, and economically developed countries are distinguished from each other by the excellence of their education systems. Measurement and evaluation are pivotal in determining the quality of an education system.

In the education and training process, measurement and evaluation are used to measure the effectiveness of the educational methods and strategies, as well as the success of the students, revealing the strengths and weaknesses of the students in the learning process (Güler & Gelbal, 2010). For this reason, measurement and evaluation are accepted as the last stage in the evaluation of the education and training process (Delice & Ergene, 2015; Hlebowitsh, 2005; Oliva, 2005). It is the process of researching, recording, interpreting, and using information about the students (Kızılcık & Tan, 2011). Measurement and evaluation can be defined as the studies carried out to determine the quality of education and make the observed deficiencies progressively more successful (Ekinci & Köksal, 2011). Additionally, it is also defined as the process of gathering information and evidence about what students have been taught (Chen, 2003). Measurement and evaluation are different concepts that are closely related to each other (Ayaydın, 2010). Measuring is observing any attribute and expressing the result of the observation with numbers or adjectives (Turgut & Baykul, 2012). Whereas evaluation is to reach a decision resultant of comparing the results obtained through measurement with a criterion (Bahar et al., 2012). Examinations are used as measurement and evaluation tools in revealing the national and international educational status of countries (Kumandaş & Kutlu, 2014), evaluating the knowledge and skills of students in academic institutions (Dodeen, 2009), making decisions about passing a grade or graduating from a school (Kubiszyn & Borich, 2003; Popham, 1999). In this respect, Zeidner (2004) stated that exams are stimulants that create varied effects on students. Exams significantly affect most aspects of life as attempting exams is a frequent phenomenon in our lives (Hembree, 1988).

Today, as in the past, the most commonly used assessment tools to measure student learning are open-ended, multiple-choice, true-false, short-answer, and matching exams. In exam question types (open-ended, multiple-choice, true-false, short-answer, and matching) are generally developed, administered, and scored by teachers. Exams represent a particular measurement technique and consist of a series of questions, each with a correct answer, which test-takers usually answer orally or in writing. Exam questions are different from questions used to measure attitude, interest, or some other affective aspects of the individual. It can be said that each exam question type mentioned above has different advantages and disadvantages. Generally, students' perceptions, attitudes, preferences, and concerns toward these exam question types developed by teachers can be varied (Benjamin et al., 1981; Birenbaum & Feldman, 1998; Huxham et al., 1976; Kılıç & Çetin, 2018; Reteguiz, 2006; Önder, 2008; Zeidner, 1987).

In the literature, the preference for one exam question type over another has been discussed regarding the preparation of a question paper (Birenbaum& Feldman, 1998; Hudson & Treagust, 2013; Kılıç &

Çetin, 2018; Oosterhof, 2001; Tozoglu et al., 2004; Zeidner, 1993; Zoller & Ben-Chaim, 2007). However, while preparing the exam questions, the perspectives of the students who attempted the exam were generally ignored (Zeidner, 1987). In this context, educators have been interested in student attitudes toward a situation owing to its possible effects on learning. Attitude can be defined as the tendency of giving a positive or negative response to an object, person, institution, or incident (Ajzen, 2005). According to another definition, attitude is the readiness of an individual for giving cognitive, emotional, and behavioral reactions towards his own self, an object, or an incident that developed around him based on his knowledge and experience gained throughout his lifetime (Inceoglu, 1993). Although the figural definitions of attitude vary, most contemporary social psychologists accept that the evaluative aspect of attitude is its typical characteristic (Ajzen, 2005). Based on these definitions attitudes are composed of three components as follows: cognitive, emotional, and behavioral (McGuire, 1985).

The cognitive component of the attitude consists of the knowledge and beliefs that the person has about the attitude object. The emotional component of attitude is the positive or negative feelings of an individual about the attitude object (Koklu, 1995). The behavioral component of attitude shows the tendency of an individual to display an act about an attitude fact (Ozcan & Koca, 2020; Tavsancil, 2014). According to Dusic (1998), attitude is related to the emotional state of the student.

Student attitudes can tell a teacher a lot about the impact of emotional situations on the learning process. Therefore, many general attitude (for instance: Bhardwaj & Kaushik, 2014; Kind et al., 2007; Pell & Jarvis, 2001; Tezbaşaran & Yiğit, 2015) and test attitude (Dodeen, 2008; Zeidner, 1987) tools have been developed to measure the emotional state of students. However, a tool measuring secondary school students' attitudes toward exam question types has not been developed. Measuring secondary school students' attitudes toward question types may be useful to determine their preferences for types of questions in exams. There are a limited number of tools available to measure students' attitudes toward exam question types. Birenbaum and Feldman (1998) used a scale consisting of 16-word pairs to measure university students' attitudes toward open-ended and multiple-choice exam questions.

Since the education system in Turkey is based on exams, all students have to take exams and answer questions. As an imposition of the exam-oriented system, the evaluation of each student with the same exam question type eliminates individual differences and prevents differences in exam question type preferences among students. In this case, it is not the individual differences of the students that are taken into consideration, but the number of questions they answer correctly. In fact, learning is a unique practice that includes different methods for each individual. In cases where a certain subject is expected to be grasped in the same way for each individual, individual differences are ignored (Çolak & Fer, 2007). This prevents students from experiencing their individuality, revealing their individual differences, and prioritizing their interests and abilities. It is known that prioritizing individual differences in educational activities is effective in students' choice of exam question type (Birenbaum & Rosenau, 2006; Doğan & Kutlu, 2011). For this reason, it is important to determine the exam question type preferences of the students. Also, the types of questions may have a possible effect on students' test anxiety and attitudes (Birenbaum & Feldman, 1998). This can be the first step in understanding several related phenomena like the poor performance of some students during exams. In this sense, the study promises to develop a tool measuring students' attitudes toward exam question types. In this regard, the aim of the present study was to develop a valid and reliable tool for measuring secondary school students' attitudes toward exam question types. For this purpose, the following

question was investigated. Does the "Attitude scale toward exam question types" under development within the current study provide valid and reliable results?

METHOD

Sample of Study

The sample was 781 secondary school students (435 female and 346 male) at five different public secondary schools in Sakarya of Türkiye. Their ages ranged from 11 to 15 years, with a mean age of 13 years. Three random samples (347, 309, and 125 students) participated in this study by responding to different versions (containing the four different types of exam questions) of the exam question types attitude scale. These samples represented the actual percentage of both genders at the secondary school. Sample 1 consisted of 347 students (203 females [59%] and 144 males [41%]). Sample 2 had 309 students (159 females [51%] and 150 males [49%]). Finally, Sample 3 consisted of 125 students (73 females [58%] and 52 males [42%]).

Scale Development

The aim of the present study was to determine the validity and reliability of the secondary school students' attitude scale toward exam question types. To develop the EQTA-S researcher has examined the literature for similar scales and brainstormed about the students' question preferences in the exams. Items expressing these question preferences were noted down. In this way, a preliminary question pool including 15 items was developed based on the students' exam question types' preferences. All items on the scale were rated on a 5-point Likert scale: 1 (strongly disagree), 2 (disagree), 3 (neutral), 4 (agree), and 5 (strongly agree). The four types of exam questions (open-ended, multiple-choice, true-false, and short-answer) were included to assess the EQTA-S. For instance, to assess the same item four different question types were included to assess (i.e.,"I am happy when I am asked open-ended questions","I am happy when I am asked multiple-choice questions", "I am happy when I am asked true-false questions", "I am happy when I am asked short-answer questions"). The scale was administered in this format. Some items (3, 4, 5, 6, 8, 9, and 10) on the scale were positively worded to indicate an increased attitude toward exam question types (if students choose the appropriate question type, they would get a higher score). Conversely, other items (1, 2, and 7) on the scale were stated in the negative direction. A high score on this scale suggests a positive attitude toward exam question type. The development of the students' attitude scale toward exam question types and estimating its psychometric indices was conducted in the following manner: determining the validity of the EQTA-S, and testing the reliability of the EQTA-S.

Measures

Exam Question Types Attitude Scale (EQTA-S). The EQTA-S was designed to measure students' attitudes toward the four types of exam questions (open-ended, multiple-choice, true-false, and short-answer). Attitudes toward each exam question type were measured by 10-item a scale each which included different exam question type. Items were rated on a 5-point Likert scale: 1 (strongly disagree), 2 (disagree), 3 (neutral), 4 (agree), and 5 (strongly agree).The EQTA-S included 10 items: some items (3, 4, 5, 6, 8, 9, and 10) on scale were positively worded, and conversely, other items (1, 2, and 7) were stated in the negative (see, appendix). This scale was administered in a form containing the four different types of exam questions (i.e., open-ended, multiple-choice, true-false, and short-answer) at the same time.

Test Attitude: In order to determine test attitude levels of the secondary school students, the Test attitude Scale (TAS) was used. The TAS is a 20-item scale that measures the test attitudes of the secondary school students. The TAS was developed by Spielberger (1980) and translated into Turkish by Öner (1990). The TAS is a four-point Likert rating scale, ranging from 1 (never) to 4 (always). The scale Cronbach's alpha coefficient was calculated as .73.

Test Anxiety: In order to determine test anxiety levels of the secondary school students the Revised Test Anxiety Scale (RTA) was used. It was developed by Benson and El-Zahar (1994), and translated into Turkish by Akın and Demirci (2012). The RTA is a 20-item, fourpoint Likert rating scale, ranging from 1 (almost never) to 4 (almost always). The scale consists of 20-items and the Cronbach's alpha coefficient was calculated as .88.

Data Analysis

In this study, three different samples were used. Sample 1 included 347 secondary school students (203 females and 144 males). The draft tool was applied to this sample to determine the construct validity of the tool. In addition, this sample was administered the EQTA-S to examine the internal reliability and discriminatory power of the items by item-total correlations. Sample 2 involved 309 secondary school students (159 females and 150 males). This sample was administered the EQTA-S to establish the construct validity of the instrument by confirmatory factor analysis (CFA). Sample 3 consisted of 125 secondary school students (73 females and 52 males). To examine the concurrent validity of EQTA-S, this sample was administered the EQTA-S to establish the relationship between students' attitudes toward exam question types, test attitudes, and test anxiety.

Procedure

Participants were asked to complete the scales, consisting of a series of questions taken from the Exam Question Types Attitude Scale (EQTA-S), Test Attitude Scale, and Test Anxiety Scale. All participants were informed about the purpose of the study before completing the scales. The scales were administered to secondary school students in the first and second semesters of the 2020-2022 academic years. The application was carried out as follows.

In the second semester during the 2021-2022 academic year was selected randomly the secondary school students at two different public schools. Prior to the administration of the scales, permission for student participation was obtained and students voluntarily participated in the research. Later, this scale was administered in a form containing the four different types of exam questions (i.e., 40 items on paper) at the same time to 347 students (Sample 1). In the first semester during the 2022-2023 academic year was selected randomly the secondary school students at two different public schools. The scale was administered in a form containing the four different types of exam questions (i.e., 40 items on paper) at the same time to 309 students (Sample 2). In the second semester of the 2022-2023 academic year was selected randomly 125 secondary school students at one public school. Later, this sample was administered a test attitude scale, test anxiety scale, and a form containing the four different types of exam questions (i.e., 40 items on paper) at the same time (Sample 3). Completion of the scale was anonymous and there was a guarantee of confidentiality. The scale was administered to the students in groups in the classrooms. All administration typically required 20 to 30 minutes.

Ethical Principles

Ethics committee permission for this study was obtained from Sakarya University Ethics Committee with the decision dated 10.11.2021 and numbered 01/02.

FINDINGS

The following section describes the procedures that were employed in the development of the secondary school students' attitude scale toward exam question types.

Validity and Reliability of the Scale

Validity is one of the most important criteria for the development and assessment of a scale. The validity is an indication of how well an instrument actually measures what it is claimed to measure and helps to ensure that there are no logical errors in drawing conclusions from the data (Garson, 1998). To validate a scale, several pieces of evidence of validity are usually assessed. The most widely used aspects are (a) content, (b) construct, and (c) concurrent validity (Crocker & Algina, 1986). These three types were assessed for the present scale.

Reliability is defined as consistency in results from repeated measurements and this consistency is related to the homogeneity of the results (Taylor, 1999). The reliability analysis provides information about the relationships between individual items in the scale. In addition to, scale reliability allows you to study the properties of scale and calculates a number of commonly used measures of scale reliability (Chapman & van Auken, 2001). Starting from these considerations, the reliability of the EQTA-S was (d) calculated internal reliability and corrected item-total correlations.

Content Validity of Scale

To determine the content validity of the scale, 15-item the scale was examined by three experts from measurement and evaluation in education departments, and 10 teachers. The five irrelevant items with students' question type preferences were excluded from the scale and left a 10-item draft EQTA-S.

Construct Validity of Scale (Sample 1 and Sample 2)

At this stage, exploratory factor analysis (EFA) was deemed appropriate to determine the factor structure of the scale. The EFA (Sample 1) was performed to determine whether the items in the scale were significant as well as whether other subscales would emerge within the items. Starting from these opinions, to include an item with in a factor, the factor loadings must be at least 0.35 (Tabachnick & Fidel 2014). The EFA results of EQTA-S were given in Table 1.

Table 1

Factor loadings of EQTA-S items for each question type (Samle 1)

NO	ITEMS	Factor loadings for each question type			
		Open-ended	multiple-choice	true-false	short-answer
1	I get excited when faced with Q-TYP questions.	.36	.42	.38	.37
2	I panic when faced with Q-TYP questions.	.51	.49	.46	.40

3	I am happy when I am faced with Q-TYP questions.	.77	.81	.81	.76
4	Q-TYP questions intrigue me more than other types of questions.	.72	.73	.72	.69
5	Q-TYP questions make it easier for me to learn the subject.	.57	.57	.49	.50
6	I would like Q-TYP questions to be asked in all exams.	.71	.68	.76	.68
7	I have difficulty in understanding Q-TYP questions.	.48	.52	.37	.42
8	I give myself more while solving Q-TYP questions.	.42	.61	.50	.59
9	I feel comfortable and secure when solving Q-TYP questions.	.76	.65	.80	.71
10	I am better at solving Q-TYP questions.	.75	.77	.79	.77

QUESTION TYPES (Q-TYP) = Open-ended, Multiple-choice, True-false, and Short-answer

Findings from the EFA in Table 1, indicated a single factor with an eigenvalue > 1, which accounted for 35.57%, 39.7%, 39.0%, and 35.5% of the variance, respectively, and was characterized by an adequate sample size (Kaiser–Meyer–Olkin value = .87, .89, .70, and .85 respectively), lack of singularity (Bartlett's test of sphericity, χ2 = 1279.56, 1302.18, 1443.73, and 1184.12, respectively; df = 45; p<.001), and robust factor loadings (λ_{range} =.37-.77, .42-.81, .38-.81, .37-.77, respectively). The result of this analysis indicated that a one-factor structure was the best-fitting model. Therefore, another analysis was not explored. The structural validity of responses to the EQTA-S was further investigated by conducting a confirmatory factor analysis (CFA) to affirm the measurement model identified with sample 2 (N=309). The independent CFA conducted with data from sample 2 demonstrated a good model fit for the data, confirming the validity of the 10-item one-factor solution (χ^2/df = 2.77, CFI= .93, NNFI= .91, SRMR= .057, and RMSEA= .076).

Concurrent Validity of Scale (Sample 3)

There may be a positive or negative relationship between students' exam question type preferences and their test attitude and anxiety. Therefore, to examine the concurrent validity of EQTA-S, sample 3 was administered the EQTA-S, test attitude, and test anxiety scales. The relation between test attitude and test anxiety of EQTA-S was given in Table 2.

Table 2

EQTA-S Correlations with concurrent validity (test attitude and test anxiety) scales

Variables	Test attitude	Test Anxiety	Open-ended	Multiple-choice	True-false	Short-answer
Test attitude	1					
Test Anxiety	-.87**	1				
Open-ended	-.31**	.38**	1			
Multiple-choice	-,29**	.24**	-.03	1		
True-false	-.13	.16	.20*	.41**	1	
Short-answer	-.14	.16	.16	.19*	.25**	1

**p< .01; *p< .05

Results in Table 2 indicated that scores of EQTA-S involving open-ended and multiple-choice questions demonstrated associations with concurrent validity measures, and yielded small-to-moderate positive and negative. According to findings in Table 2, attitudes toward open-ended and multiple-choice exam questions of students were negatively correlated with test attitude and positively correlated with test anxiety. This result shows that exam anxiety increases and exam attitude decrease when students are not asked the types of questions, they prefer.

Internal Reliability and Corrected Item-Total Correlations (Sample 1)

This scale measures the attitudes toward exam question types of the secondary school students. The first step in selecting items on the scale was the discriminatory power of the items. The discriminatory power indicates whether the items have been well structured and meaningful (Boateng et al., 2018). In this context, good items will have discriminatory power that can distinguish the subject's characteristics based on the construct measured by the measuring instrument. The discriminatory power of items can observe through the corrected item-total correlations score. Items with a score of 0.30 are considered usable (Azwar, 2019).

Table 3

Cronbach's alpha and Corrected item-total correlations of EQTA-S

Q-TYP	Cronbach's alpha (α)	Corrected item-total correlations
Open-ended	.81	.35 - .68
Multiple-choice	.85	.50 - .74

True-false	.84	.42 - .74
Short-answer	.82	.40 - .70

When was examined the results in Table 3, the item-total correlation values of the EQTA-S were found to range between .35 and .74. Also, the internal reliability of EQTA-S was calculated using Cronbach alphas. The results of Cronbach alphas in Table 3 were calculated as .81, .85, .84, and .82, respectively. The internal reliability, as measured by Cronbach's alpha coefficients was acceptable for all measures ($\alpha > .70$).

RESULTS, DISCUSSIONS, AND SUGGESTIONS

Exams are widely used in the evaluation of student learning in the education and training process. Exams comprise varied question types like open-ended, multiple choice, true-false, matching, and short-answer questions; and each of these can induce different emotions in students. One of these feelings is students' attitudes toward exam question types. It is important for teachers to measure this attitude in order to determine students' exam question type preferences. For this reason, the present study investigated the validity and reliability of the EQTA-S for secondary school students.

Validity and reliability tests were conducted to determine the psychometric properties of the measurement tool developed in this study. In this context, three stages of validity tests were conducted: content, construct, and concurrent validities. Content validity was conducted with experts and teachers, exploratory factor analysis was carried out to fulfill construct validity, and exploratory factor analysis gives the best item description. Results from EFA and descriptive analyses actualized with Sample 1 provided initial evidence in favor of structural validity. Additionally, the findings from CFA and descriptive analyses conducted with Sample 2 supported the structural validity of responses to the EQTA-S, confirming the one-factor measurement model. Bivariant correlations conducted with Sample 3 provided inceptive support for convergent validity as well, showing that scores obtained from the EQTA-S had relations in the expected directions with scores from test attitude and test anxiety scales.

The reliability (Sample 1) of EQTA-S was assessed with internal reliability, and the internal reliability was calculated by Cronbach alphas. The Cronbach alpha values of EQTA-S for all question types (open-ended, multiple-choice, true-false, and short-answer) were calculated as .81, .85, .84, and .82, respectively, and they demonstrated good internal reliability. However, future researchers using the EQTA-S should always calculate internal reliability with their own data. Item-total correlation values of the EQTA-S were also found to range between .35 and .74. The results indicated that responses to EQTA-S items that were characterized by a one-factor measurement model were reliable. EQTA-S items have shown good psychometric properties via principal components factor analysis and adequate internal reliability. Thus, based on the analysis results on all items in the scale, ten items have quality and the ten-item scale has the lowest 10 points and the highest 50 points. Finally, results from the present study indicated that the EQTA-S is contextually appropriate, technically adequate, and practically usable to measure secondary school students' attitudes toward exam question types.

Studies conducted on student attitudes have received great attention in the literature. The attitude studies have been conducted at various grade levels, on various main themes, and in various cultures (Fraser, 1982; Schreiner & Sjoberg, 2004). However, in the studies conducted, a limited number of scales were developed to measure students' attitudes toward exam question types in the literature. Therefore, an attitude scale for exam question types has been developed to fill the gap in the field, as different types of questions will differ in student performance. That is, students who are evaluated with one type of exam question will perform differently if they are evaluated with another exam question type (Danili & Reid, 2005).

When research studies were reviewed, it was found that there were similarities between the findings of the current study and that of prior research studies. In the study of Birenbaum and Feldman (1998), attitudes toward open-ended and multiple-choice exam formats of students were measured by 16 semantic differential scales with 7 points each. The Cronbach alpha coefficients for open-ended and the multiple-choice exam formats in this scale were calculated as .76 and .86, respectively. The Cronbach alpha coefficient for the open-ended and the multiple-choice exam format in the EQTA-S was calculated as .81 and .85, respectively. This result showed that the EQTA-S for the open-ended exam format has high reliability.

As a result, this study reports the development and testing of an attitude scale that can be easily employed to determine the type of questions students prefer in an exam. We believe that EQTA-S makes a significant contribution to the literature because there is a scarcity of tools for this purpose and the findings of the study reveal that secondary school students' attitude scale toward exam question types has validity and reliability.

REFERENCES

Ajzen, I. (2005). *Attitudes, personality, and behavior*. UK: McGraw-Hill Education.

Akın, A., & Demirci, İ. (2012). Revised test anxiety scale: A validity and reliability study. *Educational Sciences and Practice, 11*(21), 103-118. Retrieved from https://www.researchgate.net/publication/264550506_Revize_Edilmis_Sinav_Kaygisi_Gecerlik_ve_guvenirlik_calismasi

Astalini A., Kurniawan, D, A., Darmaji, D., Sholihah, L, R., & Perdana, R. (2019). Characteristics of students' attitude to physics in muaro jambi high school. *Humanities & Social Sciences Reviews (HSSR), 7*(2), 91- 99. https://doi.org/10.18510/hssr.2019.7210

Ayaydın, A. (2010). A research on measurement and evaluation in design education. *Journal of Ahi Evran University Faculty of Education, 11*(2*)*, 159-172. Retrieved from https://dergipark.org.tr/en/pub/kefad/issue/59503/855357

Azwar, S. (2019). *Penyusunan skala psikologi [Making psychological scale]* (2nd ed.). Pustaka Pelajar.

Bahar, M., Nartgün, Z., Durmuş, S., & Bıçak, B. (2012). *Traditional complementary assessment and evaluation techniques* (5th Edition). Ankara: Pegem Akademi.

Benjamin, M., McKeachie, W.J., Lin, Y.G., & Holinger, D. P. (1981). Test anxiety: Deficits in information processing. *Journal of Educational Psychology, 73*(6), 816-824. https://psycnet.apa.org/doi/10.1037/0022-0663.73.6.816

Benson, L., & El-Zahar, N. (1994). Further refinement and validation of the Revised Test Anxiety Scale. *Structural Equation Modeling*, *1*, 203-221. https://doi.org/10.1080/10705519409539975

Bhardwaj, J. S., & Kaushik, V. (2014). Developing an attitude scale toward science at middle school level. *International Journal of Education and Science Research Review, 1*(1), 6-13. Retrieved from https://ijesrr.org/archive_issue.php?j_id=1

Birenbaum, M., & Feldman, R. A. (1998). Relationships between learning patterns and attitudes toward two assessment formats. *Educational Research, 40*(1), 90-98. https://doi.org/10.1080/0013188980400109

Birenbaum, M., & Rosenau, S. (2006). Assessment preferences, learning orientations, and learning strategies of pre-service and in-service teachers. *Journal of Education for Teaching, 32*(2), 213-225. https://doi.org/10.1080/02607470600655300

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research. A primer. *Frontiers in Public Health*, *6, 149*. https://doi.org/10.3389/fpubh.2018.00149

Chapman, K. J., & van Auken, S. (2001). Creating positive group project experiences: An examination of the role of the instructor on students' perceptions of group projects. *Journal of Marketing Education, 23*, 117–127. https://doi.org/10.1177/0273475301232005

Chen, H. (2003). *A study of primary school english teachers' beliefs and practices in multiple assessments: a case study in taipei city* (Unpublished master theses). Taipei: National Taipei Teachers College.

Crocker, L., & J. Algina. 1986. *Introduction to classical and modern test theory.* Orlando, FL: Harcourt Brace Jovanovich.

Çolak, E., & Fer, S. (2007). Linguistic equivalence reliability and validity study of learning approaches inventory. *Çukurova University Journal of Social Sciences Institute, 16*(1), 197-211. Retrieved from https://dergipark.org.tr/tr/pub/cusosbil/issue/4376/59944

Danili, E., & Reid, N. (2005). Assessment Formats: do they make a difference? *Chemistry Education Research and Practice, 6*, 204–212. https://doi.org/10.1039/B5RP90011E

Darmaji, D., Kurniawan, D. A., & Suryani, A. (2019). Effectiveness of basic physics II practicum guidelines based on science process skills. *Jurnal Ilmu Pendidikan Fisika* (JIPF), *4*(1), 1-7. https://dx.doi.org/10.26737/jipf.v4i1.693

Delice, A., & Ergene, Ö. (2015). Ölçek geliştirme ve uyarlama çalışmalarının incelenmesi: Matematik eğitimi makaleleri örneği [Examination of scale development and adaptation studies: The case of mathematics education articles]. *Karaelmas Journal of Educational Sciences, 3,* 60-75*.* Retrieved from https://dergipark.org.tr/en/pub/kebd/issue/67216/1049114

Dodeen, H. (2008). Assessing test-taking strategies of University students: Developing a scale and estimating its psychometrics indices. *Assessment and Evaluation in Higher Education, 33*(4), 409-419. https://doi.org/10.1080/02602930701562874

Dodeen, H. (2009). Test-related characteristics of uaeu students: test-anxiety, test-taking skills, guessing, attitudes toward tests, and cheating. *Journal of Faculty of Education, 26*, 31-66. Retrieved

from http://search.shamaa.org/PDF/Articles/TSIjre/JfeNo26Y2009/jfe_2009-n26_031-066_eng.pdf

Doğan, C. D., & Kutlu, Ö. (2011). The learning-related features that are effective in pre-service teachers' choice of new situation determination methods. *Kastamonu Journal of Education, 19*(2), 459-474. Retrieved from https://dergipark.org.tr/en/pub/kefdergi/issue/49052/625761

Dusic, D. M. (1998). What social cognitive factors influence faculty members' use of computers for teaching? A literature review. *Journal of Research on Computing in Education, 31*, 123-137. https://doi.org/10.1080/08886504.1998.10782246

Ekinci, H. Y., & Köksal, E. A. (2011). Developing an assessment and evaluation competency scale for primary school science and mathematics teachers. *Kastamonu Journal of Education, 19*(1),167-184. Retrieved from https://dergipark.org.tr/en/pub/kefdergi/issue/49053/625830

Fraser, B. J. (1978). Development of a test of science-related attitudes. *Science Education, 62*(4), 509-515. https://doi.org/10.1002/sce.3730620411

Garson, D. 1998. *Quantitative research in public administration.* http://www2.chass.ncsu.edu/garson/pa765/validity.htm

Güler, N., & Gelbal, S. (2010). Examining the reliability of open-ended mathematics questions according to classical test theory and generalizability theory. *Educational Sciences in Theory and Practice, 10*(2), 991-1019. Retrieved from https://eric.ed.gov/?id=EJ889199

Hembree, R. (1988). Correlates, causes, effecs and treatment of test anxiety. *Review of Educational Research, 58*(1), 47-77. https://doi.org/10.3102/00346543058001047

Hlebowitsh, P. S. (2005). *Desinging the school curriculum*. USA: Pearson Education.

Hudson, R. D., & Treagust, D. F. (2013). Which form of assessment provides the best information about student performance in chemistry examinations? *Research in Science & Technological Education, 31*(1), 49–65. https://doi.org/10.1080/02635143.2013.764516

Huxham, G. J., Lipton A., & Cummins, R. A. (1976). Student test type preference and its relation to personality and achievement. *Medical Education*, *10*(2), 90–96. https://doi.org/10.1111/j.1365-2923.1976.tb00538.x

Inceoglu, M. (1993). *Attitude perception communication.* Ankara: V Publishing.

Kılıç, Z., & Çetin, S. (2018). Investigation of students' examination type preferences in terms of some variables. *Elementary Education Online*, *17*(2), 1051-1065. https://doi.org/10.17051/ilkonline.2018.419353

Kind, P., Jones, K., & Barmby, P. (2007). Developing attitudes toward science measures. *International Journal of Science Education*, *29*, 871–893. https://doi.org/10.1080/09500690600909091

Kızılcık, H. Ş., & Tan, M. (2011). Developing a multiple-choice test on propulsion and momentum. *Kastamonu Journal of Education, 19*(1), 185-198. Retrieved from https://dergipark.org.tr/en/pub/kefdergi/issue/49053/625831

Koklu, N. (1995). Measurement of attitudes and options used in Likert type scales. *Ankara University Journal of Faculty of Educational Sciences, 28*(2), 81–93. https://doi.org/10.1501/Egifak_0000000299

Kubiszyn, T., & Borich, G. (2003). *Educational testing and measurement: classroom application and practice*. Hobokon NJ, USA: John Wiley and Sons.

Kumandaş, H., & Kutlu, Ö. (2014). The effects of the risk factors created by the exams used in student selection and placement in higher education on school success. *Turkish Journal of Psychology, 29*(74), 15-31. Retrieved from https://www.psikolog.org.tr/en/publications/yayinlar/10.31828/tpd1300443320140000m000046.pdf

Kurniawan, D. A., Perdana, R., & Kurniawan, W. (2019). Identification Attitudes of Learners on Physics Subjects. *Journal of Educational Science and Technology (EST), 5*(1), 56-63. Retrieved from https://ojs.unm.ac.id/JEST/article/view/8231/5282

Maison, M., Darmaji, D., Astalini, A., Kurniawan, D, A., & Indrawati, P, S. (2019). Science Process Skills and Motivation. *Humanities & Social Sciences Reviews (HSSR), 7*(5), 48- 56. https://doi.org/10.18510/hssr.2019.756

McGuire, J., & Priestley, P. (1985). *Offending behavior: Skills and stratagems for going straight.* London: Batsford Academic and Educational.

Oliva, P. F. (2005). *Developing the curriculum* (6.Edition). Newyork: Pearson Education.

Ozcan, H., & Koca, E. (2020). Development of the attitude towards science scale: A validity and reliability study. *Eurasian Journal of Educational Research, 85*, 109-134. Retrieved from https://dergipark.org.tr/en/pub/ejer/issue/52308/685271

Önder, O. (2008). *The effect of preparation with multiple choice and classical type questions on mathematics achievement and test anxiety levels.* Unpublished Master Thesis. Marmara University, İstanbul.

Öner, N. (1990). *Test anxiety inventory handbook.* Foundation for Disseminating Guidance in Higher Education. No:1, İstanbul.

Oosterhof, A. (2001). *Classroom Applications of Educational Measurement*. Upper Saddle River, New Jersey, Columbus, Ohio: Merill Prentice Hall, Inc.

Pell, T., & Jarvis, T. (2001). Developing attitude to science scales for use with children of ages from five to eleven years. *International Journal of Science Education, 23*, 847–862**.** https://doi.org/10.1080/09500690010016111

Popham, W. J. (1999). Why standardized test scores don't measure educational quality. *Educational Leadership*, *56*(6), 8-15. Retrieved from http://cepd410101.pbworks.com/w/file/fetch/48464721/measure.doc

Reteguiz, J. (2006). Relationship between anxiety and standardized patient test performance in the medicine clerkship. *Journal of General Internal Medicine, 21*, 415-418. https://doi.org/10.1111/j.1525-1497.2006.00419.x

Schreiner, C., & Sjoberg, S. (2004). *Sowing the seeds of ROSE. Background, Rationale, Questionnaire Development and Data Collection for ROSE (The Relevance of Science Education)*: A comparative study of students' views of science and science education. Oslo: Unipub AS.

Spielberger, C.D. (1980). *Test anxiety inventory: Preliminary professional manual*. Palo Alto, CA: Consulting Psychologist Press.

Tabachnick, B. G., & Fidell, L. S. (2014) *Using Multivariate Statistics*. 6th edn. Harlow: Pearson.

Tavsancil, E. (2014). *Measurement of attitudes and data analysis with SPSS.* Nobel Publishing, Ankara.

Taylor, J. R. (1999). *An introduction to error analysis: The study of uncertainties in physical measurements.* University Science Books. pp. 128–129. ISBN 0-935702-75-X.

Tezbaşaran, E., & Yiğit, R. (2015). A study on developing attitude scale toward nurses. *Hasan Ali Yücel Eğitim Fakültesi Dergisi,12*(24), 69-80. Retrieved from https://dergipark.org.tr/en/pub/iuhayefd/issue/8803/110063

Tozoglu, D., Tozoglu, M. D., Gurses, A., & Dogar, C. (2004). The students' perceptions: essay versus multıple-choıce type exams. *Journal of Baltic Science Education, 2*(6*),* 52-59*.* Retrieved from https://www.researchgate.net/profile/Cetin-Dogar/publication/291325299_The_students'_perceptions_essay_versus_multiple-choice_type_exams/links/5d072dcd458515b055d8e65d/The-students-perceptions-essay-versus-multiple-choice-type-exams.pdf

Turgut, M. F., & Baykul, Y. (2012). *Measurement and evaluation in education* (4th Edition). Ankara:Pegem Akademi.

Zeidner, M. (1987). Essay versus multiple-choice type classroom exams: The student's perspective. *Journal of Educational Research, 80*(6), 352-58. https://doi.org/10.1080/00220671.1987.10885782

Zeidner, M. (1993). *Essay versus multiple-choice-type classroom exams: The student's perspective.* In B. Nevo and R. S. Jäger (Eds.), *Educational and psychological testing: The test taker's outlook* (pp.85-111). Stuttgart: Hogrefe and Huber Publishers.

Zeidner, M. (2004). Test anxiety. in: spielberger CD (Ed.), *Encyclopedia of Applied Psychology* (pp. 545-556). New York: Elsevier.

Zoller, U., & Ben-Chain, D. (2007). Interaction between examination type, anxiety state and academic achievement in college science: An action oriented research. *Journal of Research in Science Teaching*, *26*(1), 65-77. https://doi.org/10.1002/tea.3660260107

APPENDIX

Appendix-A

		Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1*	I get excited when faced with Q-TYP questions.	5	4	3	2	1
2*	I panic when faced with Q-TYP questions.	5	4	3	2	1
3	I am happy when I am faced with Q-TYP questions.	1	2	3	4	5
4	Q-TYP questions intrigue me more than other types of questions.	1	2	3	4	5
5	Q-TYP questions make it easier for me to learn the subject.	1	2	3	4	5
6	I would like Q-TYP questions to be asked in all exams.	1	2	3	4	5
7*	I have difficulty in understanding Q-TYP questions.	5	4	3	2	1
8	I give myself more while solving Q-TYP questions.	1	2	3	4	5
9	I feel comfortable and secure when solving Q-TYP questions.	1	2	3	4	5
10	I am better at solving Q-TYP questions.	1	2	3	4	5

QUESTION TYPES (Q-TYP) = Open-ended, Multiple-choice, True-false, and Short-answer

Appendix B

		Kesinlikle katılmıyorum	Katılmıyorum	Kararsızım	Katılıyorum	Tamamen katılıyorum
1*	SST sorularla karşılaştığımda heyecanlanırım.	5	4	3	2	1
2*	SST sorularla karşılaştığımda paniğe kapılırım.	5	4	3	2	1
3	SST sorularla karşılaştığımda sevinirim.	1	2	3	4	5
4	SST sorular, diğer soru türlerine göre ilgimi daha çok çeker.	1	2	3	4	5
5	SST sorular konuyu öğrenmemi kolaylaştırır.	1	2	3	4	5
6	Bütün sınavlarda SST sorular sorulmasını isterim.	1	2	3	4	5
7*	SST soruları anlamakta güçlük çekerim.	5	4	3	2	1
8	SST soruları çözerken kendimi daha çok veririm.	1	2	3	4	5
9	SST soruları çözerken rahat ve güvenli hissederim.	1	2	3	4	5
10	SST soruları çözerken daha başarılıyım.	1	2	3	4	5

SINAV SORU TÜRLERİ (SST) = Açık-uçlu, Çoktan-seçmeli, Doğru-yanlış ve Kısa-cevaplı

Plagiarism Statement

Similarity rates of this article was scanned by the iThenticate software. No plagiarism detected.

Availability of Data and Materials

Not applicable.