International Journal of Assessment Tools in Education
2024, Vol. 11, No. 1, 126–147
https://doi.org/10.21449/ijate.1351510
journal homepage: https://dergipark.org.tr/en/pub/ijate
Research Article

Development of a knowledge test about scientists for secondary school students

Pelin Yıldırım[1*], **Gonca Keçeci**[2], **Fikriye Kırbağ Zengin**[2]

[1]Fırat University, Institute of Education Sciences, Elazığ, Türkiye
[2]Fırat University, Faculty of Education, Department of Mathematics and Science Education, Elazığ, Türkiye

ARTICLE HISTORY
Received: Aug. 28, 2023
Accepted: Feb. 11, 2024

Keywords:
Scientists,
Knowledge test,
Secondary school students,
Scale development,
Reliability.

Abstract: This study aims to develop a knowledge test to assess secondary school students' awareness of scientists. After determining the topics that the questions in the test would measure, 50 multiple-choice questions were prepared and presented to six experts in the field. Following the experts' feedback, the number of questions in the test was reduced to 44 due to issues related to question clarity, alignment with the curriculum, and cognitive appropriateness. The 44-question test underwent item analysis. The prepared test was subjected to a preliminary pilot application with 222 secondary school students. As a result of the preliminary pilot application, two questions with discrimination indices of 0.19 or lower were removed, and the final version of the test was arranged to consist of 42 questions for the pilot application. The pilot application was conducted with 211 secondary school students. In this application, secondary school students were administered the Knowledge Test of Scientists consisting of 42 questions, and item analysis was performed again. Two questions with discrimination indices of 0.19 or lower were removed from the pilot application. As a result of the pilot application, the Knowledge Test of Scientists had a KR-20 reliability coefficient of 0.94, and the Spearman Brown two-half test correlation coefficient was calculated as 0.92. The final version of the scale was arranged to consist of 40 items. These results demonstrate that the test is highly reliable, meaning that the test results accurately reflect the measured attribute and are replicable.

1. INTRODUCTION

Education is a process aimed at changing individuals' behaviors in the desired direction (Ertürk, 1972). This process encompasses various goals such as acquiring new knowledge, correcting erroneous behaviors, or altering undesirable behaviors. To achieve the intended changes in education, careful planning and control mechanisms are required (Turgut & Baykul, 2012).

The achievements and challenges of the education system play a significant role in improving this system. To overcome the issues encountered in education and to make future educational activities more effective, it is important to assess educational processes and analyze their outcomes (Ülger, 2021). Therefore, the assessment process in education holds great importance. Assessment is a process of gathering information, analyzing it, and making decisions based on criteria. These decisions are used to enhance educational processes and improve student achievement (Turgut & Baykul, 2012).

*CONTACT: Pelin YILDIRIM ✉ yildirim.pelin92@gmail.com Fırat University, Institute of Education Sciences, Elazığ, Türkiye

Education plays a crucial role in contributing significantly to the development of society, and decisions made in this context typically revolve around student success. However, making accurate and fair decisions becomes challenging without having a solid foundation in student achievement. Assessing student success forms the basis of educational decisions, and the reliability and validity of these assessments are of critical importance (Sarıbaş & Babadağ, 2015; Turgut & Baykul, 2012). Alongside these assessments, evaluations based on the appropriateness of criteria are also essential. Unfortunately, there is often a lack of awareness and understanding about measurement and evaluation in our country. This deficiency can lead to complications in measurement results, criteria, and evaluation outcomes, resulting in incorrect decisions. Decisions may be based on unreliable measurements, valid criteria may be chosen or misinterpreted, and in some cases, decisions may be made without utilizing any evaluation criteria at all. Such issues jeopardize the reliability and effectiveness of educational decisions (Turgut & Baykul, 2012). Therefore, more effort should be invested in assessing student success and allowing these assessments to shape educational decisions. Additionally, the focus should be on reliable and valid tools for measuring student achievement. In this way, educational decisions can become fairer and more evidence based. Overcoming these problems can contribute to the increased effectiveness of our education system. Multiple-choice tests, which are among the traditional measurement and evaluation techniques, can still be considered to play a significant role in determining students' learning levels today. This measurement method offers advantages such as minimizing subjective effects during scoring, providing greater content validity compared to other measurement methods, and making measurement and evaluation easier, especially in large groups. Therefore, it is considered a preferred measurement and evaluation technique by teachers (Çardak & Selvi, 2018; Doğan, 2009).

It is possible to observe that multiple-choice tests are frequently used in scientific studies related to education. In such studies, it is essential to meticulously follow the standard test development steps when creating achievement tests. These types of studies can contribute significantly to educational sciences by providing important information about groups' learning levels and achievements (Çardak & Selvi, 2018).

One of the fundamental goals of today's education system is to increase students' awareness of historical and cultural values (Meydan & Akkuş, 2014). However, the level of knowledge among students regarding scientists and their contributions can often be insufficient (Bıçak & Bilir, 2023). Educational programs often do not emphasize these important figures enough, leaving students with a limited understanding of the lives and scientific contributions of scientists. This limitation in knowledge can limit the awareness of the younger generation towards historical and cultural heritage, potentially causing them to miss out on valuable sources of inspiration (Arabacı & Dönel Akgül, 2022; Özdemir, 2022; Yıldırım & Keçeci, 2022). In addition, a detailed literature review has been conducted to determine the general perception and thoughts about scientists in the field of education. Within the scope of these studies, various methods such as open-ended questions, interviews, and drawings have been observed (Ağgül Yalçın, 2012; Bıçak & Bilir, 2023; Bilir et al., 2020; Duran & Bayar, 2019; Durukan, 2017; Karaçam et al., 2014; Özdemir, 2017; Özdeş & Aslan, 2019). As a result of the comprehensive literature review, it has been identified that there is only one achievement test related to scientists (Karasu, 2019). The achievement test developed by Karasu (2019) consists of 20 questions and includes questions about both Turkish and foreign scientists, focusing entirely on their inventions. On the other hand, the test developed within the scope of this study has a broader perspective. This 40-question test contains only questions about Turkish scientists, covering fundamental information about scientists to provide a comprehensive perspective. Additionally, the questions focus on the general contributions of scientists, providing a different emphasis from the previous study. This study particularly focuses on Turkish scientists and differentiates itself from previous studies by increasing the number of

questions. In this way, it aims to more detailedly assess students' knowledge about Turkish scientists and provide a more comprehensive measurement tool compared to previous studies.

At this point, it becomes evident that there is a lack of knowledge among secondary school students regarding scientists. The question of whether there is a deficiency in students' knowledge about who scientists are, what they do, and their contributions to society forms the basis of the problem. Simultaneously, the absence of an appropriate measurement tool to assess students' awareness of scientists poses a problem. In this context, the main problem addressed by this study is to tackle the lack of knowledge among secondary school students regarding scientists within an educational system that aims to increase awareness of historical and cultural values. This problem highlights the need to address the deficiency in knowledge about scientists and the absence of a suitable measurement tool to assess awareness. The study aims to develop a reliable and valid test to determine students' awareness of scientists. In line with this purpose, the fundamental study question of the study is, "How can the reliability of an achievement test developed to assess secondary school students' awareness of scientists be evaluated, and to what extent can the results be replicated?"

This test can serve the purpose of evaluating students' levels of knowledge about scientists, identifying deficiencies in science education, and providing support to students in these areas. This situation brings the potential to enhance science instruction. Furthermore, students' interest in scientists and science can be influenced by their own awareness. The developed test has the potential to increase students' interest in science by helping them better understand the importance and contributions of scientists. In addition to all these aspects, this study can contribute to the field's knowledge by introducing a new study and assessment tool to science education and assessment literature.

2. METHOD

This study focused on the development of an achievement test aimed at assessing secondary school students' awareness of scientists. This measurement tool is a type of achievement test used to evaluate secondary school students' knowledge about scientists. The test was designed to measure students' knowledge levels on topics such as who scientists are, the type of work they engage in, and their contributions to society.

The development process of the test followed the test development steps outlined by Crocker and Algina (1986). Initially, the objectives of the test were determined. Subsequently, a pool of 50 test items was created during this phase, and expert opinions were sought to gather necessary feedback. Based on the evaluations provided by experts, necessary corrections were made, and six items were excluded, reducing the number of questions to 44. In the process of developing the achievement test, a two-stage implementation was carried out, including the preliminary pilot and subsequent pilot. The preliminary pilot application and the subsequent pilot application were conducted to evaluate the effectiveness, discriminability, difficulty, and overall effectiveness of the test items. Based on the results of the preliminary pilot application, two items with discrimination indices below 0.19 (items 2 and 8) were removed from the test, some items were revised, and the number of items was reduced from 44 to 42 (see Table 4). Subsequently, the test consisting of 42 items underwent a pilot application. The pilot application was conducted to assess the impact of the corrections and to observe the overall performance of the test. This stage provided additional information on the reliability and validity of the test, examining the consistency of items among students and the accurate measurement of the intended topics. According to the results of the pilot application, two items with discrimination indices below 0.19 (items 1 and 40) were removed, reducing the number of items from 42 to 40 (see Table 6). Finally, the final test was composed of 40 distinct items, and test statistics were determined based on the results of the pilot application. A set of guidelines was created for evaluating test scores.

2.1. Study Group

The sample for the test consisted of a total of 433 students (222 for the preliminary pilot and 221 for the pilot application) who were enrolled in three different secondary schools in the city center of Elazig during the 2021-2022 academic year. These students were from 5th, 6th, 7th, and 8th grades. Yıldırım and Şimşek (2008) suggested maximizing diversity in the sample selection by conducting the study in schools with different levels of achievement and socio-economic statuses. Therefore, special attention was paid to schools with varying levels of achievement and socio-economic status in this study. Additionally, the selection of schools was influenced by their central location and easy accessibility.

As highlighted by Nunnally (1978), the sample size for testing exam items in a small student group should be at least five times the number of items in the test. The objective at this stage is to assess student responses and observe elements such as exam duration, test difficulty level, and confusion related to the items. In other words, the sample size is determined to evaluate the initial applications of the items and comprehend the overall structure of the test. This pertains to an initial pilot study aiming to assess the success of the test by evaluating students' reactions to the exam items. In the preliminary pilot application, findings obtained using Crocker and Algina's (1986) test development stages demonstrate the successful steps taken to ensure the reliability and validity of the test results. Consequently, the test has become a reliable measurement tool suitable for its intended purposes. The distribution of students by grade levels who participated in the preliminary pilot and pilot applications is shown in Table 1.

Table 1. *Distribution of students by grade levels.*

Type of application	Grade level	*f*
Preliminary pilot	5th grade	27
	6th grade	56
	7th grade	102
	8th grade	37
Pilot	5th grade	40
	6th grade	55
	7th grade	70
	8th grade	46

Table 1 shows the number of students from different grade levels who participated in various stages of the study. In the preliminary pilot application, 222 students participated. The majority of participants were from the 7th grade (102 students), followed by the 6th grade (56 students), 8th grade (37 students), and 5th grade (27 students). In the pilot application, 211 students participated. Again, the majority of participants were from the 7th grade (70 students), followed by the 6th grade (55 students), 5th grade (40 students), and 8th grade (46 students).

During the course of the study, necessary permissions were obtained from the Elazig Provincial Directorate of National Education. These permissions ensured that the study was conducted in compliance with legal and administrative requirements. Data confidentiality was rigorously maintained, and various measures were taken for this purpose. Participant selection was based on a voluntary basis, and Informed Consent Forms were signed by the parents of students who wished to participate in the study. This step ensured that participation was entirely voluntary and informed. During the data collection process, the personal information of participants was kept confidential and used only for the purpose of the study. This is a critical step in protecting privacy. Secure databases and encryption methods were used for data storage and analysis, ensuring protection against unauthorized access to the data. In addition, access to the obtained data was restricted to the authors of the study and authorized personnel only. This ensures that the data is used under the supervision of authorized individuals only. In conclusion, full

compliance with privacy policies ensured the highest level of data privacy and security. These steps ensured that the study was conducted in accordance with ethical standards and legal requirements.

2.2. Test Development Process

In accordance with the purpose of the study, the test development process was carried out in steps, and the test development steps proposed by Crocker and Algina (1986) were carefully followed. The main stages of creating the test can be listed as follows:

Goal Setting: The main objective of the study is to develop a valid and reliable knowledge test for scientists.

Defining Behaviors: The behaviors or topics that the test aims to measure were determined.

Development of Table of Specification: A specification table was created, indicating which behaviors would be included in the test and how each behavior would be measured.

Development of Item Pool: Based on the specification table, a pool of multiple-choice questions was prepared to measure each behavior.

Examination and Arrangement of Items: Prepared questions were examined and arranged by experts. The clarity of questions, the quality of distractors, and alignment with the curriculum were evaluated.

Creating a Trial Form: A trial form consisting of reviewed and edited questions was prepared.

Pilot Test: The prepared trial form was administered to a selected group of students for a pilot test. This stage aimed to understand how the test worked under field conditions and to identify possible problems.

Statistical Analysis of Items: The data obtained during the pilot test were used for the statistical analysis of the items. Correct answers were scored as 1 point, incorrect answers, questions with multiple markings, and unanswered questions were scored as 0 points. Item analysis was conducted to calculate item statistics, and items that did not meet the criteria or performed poorly were identified and removed from the test. This analysis was conducted to assess how the items performed in terms of the validity and reliability of the test.

Determining Statistical Properties: Statistical properties of the items such as difficulty levels and discrimination indices were evaluated. This provides information about the reliability and validity of the test.

Calculation of Test Scores: Based on the obtained data, participants' test scores were calculated and the results were analyzed.

In summary, after determining the topics to be measured in the test (basic information about scientists and their contributions), 50 multiple-choice questions were prepared. These questions were presented to six experts in the field, including four academic staff members (two professors, one associate professor, one Ph.D.) and two science teachers. Test item evaluations, focusing on the diverse academic titles and experience levels among experts, aimed to provide a more comprehensive and balanced perspective on the study. These evaluations aimed to develop questions that accurately measure students' knowledge, ensuring grammatical correctness and clarity in meaning. The expert panel, consisting of a professor, associate professor, and Ph.D. with extensive knowledge and experience in the field, allowed for a thorough assessment of the scientific content of the test. Additionally, an academic staff member from a state university specializing in Turkish education assessed the grammatical and semantic accuracy of the test questions. By focusing on grammatical rules and clarity, this expert contributed to the test's effectiveness in terms of language. The two science teachers in the expert panel provided an evaluation from a student perspective, being directly involved with students and having mastery over the curriculum. The teachers' suggestions aimed at revising

test questions to make them suitable, understandable, and aligned with learning objectives for students. The revisions proposed by experts generally aimed to improve the clarity of questions, correct grammatical errors, and ensure students could answer correctly. Furthermore, experts provided suggestions for reviewing cognitive levels and difficulty levels to enhance the test's relevance to its purpose and measurement of student achievements. In this way, experts combined different perspectives to offer a comprehensive assessment to enhance the overall quality of the test. As a result of expert opinions, the number of questions in the test was reduced from 50 to 44. The test, reduced to 44 questions, underwent an item analysis process. This analysis was employed to ensure the structural validity of the test and to enhance the content, clarity, and difficulty levels of the questions (Turgut, 1992).

The prepared test was administered to a total of 222 secondary school students, with 107 being girls and 115 boys (27 students at the 5th-grade level, 56 students at the 6th-grade level, 102 students at the 7th-grade level, and 37 students at the 8th-grade level) in the preliminary pilot application. The test scores of this group of students were arranged in order of achievement, and sixty students, equivalent to 27% of the total number of students, were selected to form the lower and upper groups. To assess the structural validity of the test items, item analysis was conducted, and the discrimination indices of the test items were examined. According to the evaluation conducted by Crocker and Algina (1986), items with a discrimination index greater than 0.40 were considered to be excellent, while items with an index value between 0.30 and 0.39 were considered good. On the other hand, items with an index value between 0.20 and 0.29 were identified as items that needed correction and improvement, and items with a discrimination index of 0.19 or lower were deemed to be removed from the test. As a result of the preliminary pilot application, two items with a discrimination index of 0.19 or lower were removed from the test, reducing the number of questions to 42. The test, consisting of 42 questions, was administered to 211 secondary school students (40 students at the 5th-grade level, 55 students at the 6th-grade level, 70 students at the 7th-grade level, and 46 students at the 8th-grade level) through the pilot application, and item analysis was conducted once again. As a result of the item analysis, two items with a discrimination index of 0.19 or lower were removed, reducing the number of questions to 40. The final version of the scale was arranged to include 40 items.

2.3. Data Analysis

The data collected were subjected to analysis through SPSS 22 software, and assessments of the validity and reliability of the test were carried out. Methods used to assess the reliability of a test can be categorized into two main categories: single-administration methods and two-administration methods. The various methods used to assess measurement reliability include test-retest, alternate form, split-half, interrater, Cronbach's alpha coefficient, Kuder-Richardson 20 (KR-20), and Hoyt's variance analysis. These methods are typically evaluated using the Pearson product-moment correlation coefficient (Crocker & Algina, 1986; Linn & Gronlund, 2000; Mehrens & Lehmann, 1991). Among the single-administration methods, which are based on the scores obtained by administering a test to a group only once, are KR-20 and 21, Cronbach's alpha, Hoyt's Variance Analysis, and the Test Split-Half methods (Karip, 2012). In this study, the reliability of the measurement results was calculated using split-half tests and KR-20 reliability types.

3. RESULTS

In this section, data related to the validity and reliability studies of the knowledge test developed for scientists are presented. These data have been collected to assess the measurement-related validity and reliability of the test. Validity studies aim to determine whether the test accurately reflects the concept or skill it intends to measure. Reliability studies, on the other hand, assess whether the test provides consistent and reliable results. The data analysis methods and results used in this study provide valuable information about the scientific validity and reliability of

the test. Therefore, these data will assist in understanding how reliable and valid the use and interpretation of the test scores are.

Within the scope of the study, eight science textbooks used in secondary schools affiliated with the Ministry of National Education (MoNE) during the 2020-2021 academic year were examined. The distribution of Turkish/Turkish-origin scientists mentioned in the textbooks according to grade levels is presented in Table 2.

Table 2. *Distribution of Turkish/Turkish-origin scientists mentioned in secondary school science textbooks by grade levels.*

Grade level	The mentioned Turkish/Turkish-origin scientists	*f*
5th grade (MEB Publications and SDR Vertical Publishing)	Ali Kuscu, Fatih Sultan Mehmet, Hezarfen Ahmet Celebi, Ibnu'l-Heysem, Vecihi Hurkus	5
6th grade (Sevgi Publications, MEB Publications-1, MEB Publications-2)	Ali Kuscu, Aziz Sancar, Canan Dagdeviren, Associate Professor Dr. Ozgur Sahin, El-Memun, Ibnu'l-Heysem, Ibn-i Sina, Mimar Sinan, Ulug Bey	9
7th grade (Tutku Publishing, MEB Publications)	Ali Kuscu, Aziz Sancar, Ibnu'l-Heysem, Ulug Bey	4
8th grade (SDR Vertical Publishing)		-

Table 2 reflects which scientists are presented to students in secondary school Science textbooks and in which grade levels. According to this table, there are the names of five scientists in the 5th-grade textbooks. The 6th-grade textbooks feature nine different scientists, making it the grade level with the highest number of scientist mentions compared to other grade levels. In 7th-grade textbooks, the names of four scientists are mentioned, while in 8th-grade textbooks, there is no mention of a Turkish/Turkish-origin scientist.

The developed knowledge test is focused exclusively on Turkish-Islamic scholars. In this context, there are no items related to contemporary influential scientists, such as Aziz Sancar and Canan Dagdeviren. This choice reflects the test's emphasis on a more historical and cultural perspective. The selection of scientists based on specific criteria in organizing the test aims to enable students to acquire knowledge at levels appropriate to their grade levels. While determining the number of items represented by each scientist, their impact on the history of science and the general knowledge level of students have been taken into account. Certain scientists are represented by more items due to their historical significance and contributions. For instance, prominent figures like Ebu Bekir er-Razi, Ibnu'l-Heysem, and el-Cezeri are represented by more items because of their extensive influence in various fields. This approach provides students with an opportunity to gain in-depth knowledge about the work of specific scientists and explore different aspects of their contributions. Furthermore, to enhance diversity in the test, a fair distribution among different scientific fields has been ensured. An attempt has been made to establish a balance among scientists focusing on various fields such as physics, mathematics, and astronomy. This diversity offers students the chance to explore different disciplines of science and enhance their overall knowledge levels. In this context, the distribution of items in the knowledge test according to scientists in the preliminary pilot application is presented in Table 3.

Table 3 displays which scientists the items in the knowledge test are based on and numerically represents the representation of these scientists in the test content. In total, there are items related to 16 different scientists in the knowledge test. Ebu Bekir er-Razi is the most represented scientist in the table with six different items. Following him are el-Cezeri and Abdurrahman el-Hazini, both represented with five different items.

Table 3. *Distribution of items in the knowledge test according to scientists in the preliminary pilot application.*

Scientists	Items
Ebu Bekir er-Razi	8-10-21-25-37-43
el-Cezeri	1-3-24-36-42
Abdurrahman el-Hazini	4-17-38-41
Ibnu'l-Heysem	5-11-29-44
Cabir ibn Hayyan	9-20-34
Farabi	12-19-28
Fergani	16-35-40
Ibn Sina	15-18-23
Ibnu'n-Nefis	6-13-32
Ulug Bey	22-26-33
Ali Kuscu	7-27
Aksemseddin	30
Biruni	14
Hezarfen Ahmet Celebi	31
John Dalton	2
Kindi	39

As a result of the item analysis conducted, the difficulty levels (p) and discrimination indices (d) of the items in the knowledge test within the scope of the preliminary pilot application are presented in Table 4.

Table 4. *Difficulty levels (p) and discrimination indices (d) of the items in the preliminary pilot application knowledge test.*

Questions	Groups	A	B	C	D	Blank	Filled	*p*	*d*
1	Upper	56	-	2	2	0	60	0.61	0.63
	Lower	18	14	17	11	0	60		
2	Upper	-	-	60	-	0	60	0.92	0.15
	Lower	3	4	51	2	0	60		
3	Upper	1	11	33	15	0	60	0.36	0.36
	Lower	6	17	11	26	0	60		
4	Upper	1	13	5	41	0	60	0.43	0.50
	Lower	9	22	18	11	0	60		
5	Upper	-	49	10	1	0	60	0.52	0.58
	Lower	5	14	22	19	0	60		
6	Upper	43	3	12	2	0	60	0.45	0.51
	Lower	12	16	27	5	0	60		
7	Upper	-	3	57	-	0	60	0.74	0.41
	Lower	8	8	32	12	0	60		
8	Upper	25	16	3	16	0	60	0.25	0.01
	Lower	14	18	13	15	0	60		
9	Upper	6	46	6	2	0	60	0.52	0.48
	Lower	26	17	13	4	0	60		
10	Upper	8	34	16	2	0	60	0.40	0.33
	Lower	16	14	23	7	0	60		
11	Upper	1	43	9	7	0	60	0.47	0.48
	Lower	8	14	21	17	0	60		

12	Upper	1	7	36	16	0	60	0.41	0.36
	Lower	9	20	14	17	0	60		
13	Upper	4	1	51	4	0	60	0.61	0.46
	Lower	9	12	23	16	0	60		
14	Upper	16	37	5	2	0	60	0.40	0.43
	Lower	29	11	15	5	0	60		
15	Upper	3	7	10	40	0	60	0.40	0.51
	Lower	17	13	21	9	0	60		
16	Upper	3	41	5	11	0	60	0.50	0.35
	Lower	13	20	14	13	0	60		
17	Upper	39	5	11	5	0	60	0.49	0.31
	Lower	20	10	10	20	0	60		
18	Upper	3	43	7	7	0	60	0.43	0.56
	Lower	21	9	19	11	0	60		
19	Upper	8	47	3	2	0	60	0.52	0.51
	Lower	10	16	14	20	0	60		
20	Upper	10	4	42	4	0	60	0.48	0.43
	Lower	18	13	16	13	0	60		
21	Upper	40	11	9	-	0	60	0.46	0.40
	Lower	16	13	13	18	0	60		
22	Upper	50	6	3	1	0	60	0.50	0.65
	Lower	11	20	16	13	0	60		
23	Upper	4	55	-	1	0	60	0.62	0.58
	Lower	17	20	15	8	0	60		
24	Upper	4	45	7	4	0	60	0.46	0.56
	Lower	19	11	13	17	0	60		
25	Upper	5	42	10	3	0	60	0.47	0.45
	Lower	9	15	17	19	0	60		
26	Upper	4	37	15	4	0	60	0.47	0.28
	Lower	8	20	25	7	0	60		
27	Upper	6	39	4	11	0	60	0.42	0.45
	Lower	14	12	22	12	0	60		
28	Upper	1	43	11	5	0	60	0.49	0.45
	Lower	17	16	25	2	0	60		
29	Upper	2	3	7	48	0	60	0.54	0.51
	Lower	6	20	17	17	0	60		
30	Upper	54	3	3	-	0	60	0.52	0.75
	Lower	9	18	14	19	0	60		
31	Upper	-	55	2	3	0	60	0.65	0.53
	Lower	9	23	20	8	0	60		
32	Upper	39	5	11	5	0	60	0.40	0.48
	Lower	10	18	22	10	0	60		
33	Upper	4	5	49	2	0	60	0.53	0.56
	Lower	16	15	15	14	0	60		
34	Upper	-	12	45	3	0	60	0.45	0.58
	Lower	14	26	10	10	0	60		
35	Upper	6	10	40	4	0	60	0.42	0.48
	Lower	17	19	11	13	0	60		
36	Upper	7	23	24	6	0	60	0.30	0.20

	Lower	14	12	12	22	0	60		
37	Upper	14	3	40	3	0	60	0.45	0.41
	Lower	15	20	15	10	0	60		
38	Upper	7	14	37	2	0	60	0.46	0.30
	Lower	10	20	19	11	0	60		
39	Upper	5	8	30	17	0	60	0.38	0.23
	Lower	17	15	16	12	0	60		
40	Upper	2	24	27	7	0	60	0.30	0.28
	Lower	16	20	19	14	0	60		
41	Upper	6	6	41	7	0	60	0.40	0.55
	Lower	17	18	8	17	0	60		
42	Upper	7	5	13	35	0	60	0.37	0.41
	Lower	16	12	22	10	0	60		
43	Upper	3	2	25	3	0	60	0.48	0.60
	Lower	11	18	20	11	0	60		
44	Upper	21	3	4	12	0	60	0.25	0.20
	Lower	9	17	17	17	0	60		

p: difficulty
d: discrimination

When examining Table 4, it can be seen that the difficulty indices of the test items range from 0.25 to 0.92. Items within this range indicate different levels of difficulty. Additionally, the discrimination indices of the items also vary between 0.01 and 0.75. These values indicate how well the items differentiate between different ability groups. A total of 31 items with a discrimination index of 0.40 or higher were used in the test, and it can be said that these items better measure the differences between different ability groups. Similarly, six items with a discrimination index between 0.30 and 0.39 were also included in the test, indicating an acceptable level of discriminative ability. Five items with a discrimination index between 0.20 and 0.29 (items 26, 36, 39, 40, and 44) were included in the test after necessary adjustments were made. These adjustments were found to enhance the discriminative power of the items. However, two items with a discrimination index of 0.19 or lower, related to John Dalton (a general question) and Ebu Bekir er-Razi, were removed from the test. As a result of these steps, following the preliminary pilot application, the reliability coefficient of the test was calculated as 0.87 for KR-20 and as 0.78 for the Spearman-Brown two-half test correlation value. This high reliability value indicates that the test's internal consistency has been achieved. Consequently, the final version of the scale has been organized to consist of 42 items for the pilot application. The distribution of items in the knowledge test according to scientists in the pilot application is presented in Table 5.

Table 5 indicates the basis of the knowledge test used in the pilot application and numerically represents the representation of various scientists in the test content. In total, there are items related to 15 different scientists in the knowledge test. Ebu Bekir er-Razi and el-Cezeri are the most represented scientists in the table with five different items each. Following them are Abdurrahman el-Hazini and Ibnu'l-Heysem, both represented with four different items.

The Science Knowledge Test consisting of 42 questions was administered to secondary school students through a pilot application, and item analysis was conducted again. The results of the item analysis, including item difficulty levels, discrimination indices, item-total correlations, and t-values, are presented in Table 6.

Table 5. *Distribution of items in the knowledge test according to scientists in the pilot application.*

Scientists	Items
Ebu Bekir er-Razi	10-21-25-37-43
el-Cezeri	1-3-24-36-42
Abdurrahman el-Hazini	4-17-38-41
Ibnu'l-Heysem	5-11-29-44
Cabir ibn Hayyan	9-20-34
Farabi	12-19-28
Fergani	16-35-40
Ibn Sina	15-18-23
Ibnu'n-Nefis	6-13-32
Ulug Bey	22-26-33
Ali Kuscu	7-27
Aksemseddin	30
Biruni	14
Hezarfen Ahmet Celebi	31
Kindi	39

Table 6. *Difficulty levels (p), discrimination indices (d), item-total correlations, and t-values of the items in the knowledge test in the pilot application.*

Questions	Groups	A	B	C	D	Blank	Filled	*p*	*d*	*r*	t^2
1	Upper	54*	-	1	2	0	57	0.99	0.08	0.16	1.59
	Lower	49	2	3	3	0	57				
2	Upper	11	8	32*	6	0	57	0.37	0.36	0.43	4.34***
	Lower	23	5	11	8	0	57				
3	Upper	4	9	7	37*	0	57	0.40	0.49	0.47	6.12***
	Lower	13	14	21	9	0	57				
4	Upper	3	45*	2	7	0	57	0.56	0.45	0.44	5.47***
	Lower	6	19	14	18	0	57				
5	Upper	40*	8	3	6	0	57	0.45	0.49	0.47	5.99***
	Lower	12	28	7	10	0	57				
6	Upper	-	-	54*	3	0	57	0.75	0.38	0.39	5.30***
	Lower	3	9	32	13	0	57				
7	Upper	4	42*	3	8	0	57	0.52	0.42	0.36	4.92***
	Lower	19	18	11	9	0	57				
8	Upper	12	31*	4	10	0	57	0.39	0.29	0.52	3.39***
	Lower	7	14	24	12	0	57				
9	Upper	3	42*	11	1	0	57	0.46	0.54	0.52	6.88***
	Lower	12	11	25	9	0	57				
10	Upper	17	2	34*	4	0	57	0.44	0.29	0.31	3.32***
	Lower	27	15	17	8	0	57				
11	Upper	3	5	46*	3	0	57	0.54	0.52	0.51	6.58***
	Lower	3	23	16	15	0	57				
12	Upper	10	34*	3	10	0	57	0.37	0.43	0.50	5.36***
	Lower	27	9	16	5	0	57				
13	Upper	6	5	10	36*	0	57	0.36	0.52	0.55	6.88***
	Lower	14	18	19	6	0	57				
14	Upper	8	39*	7	3	0	57	0.47	0.42	0.39	4.92***
	Lower	20	15	19	3	0	57				
15	Upper	38*	6	4	9	0	57	0.47	0.38	0.40	4.43***
	Lower	16	13	18	10	0	57				
16	Upper	4	42*	7	2	0	57	0.45	0.56	0.58	7.22***
	Lower	20	10	12	15	0	57				

17	Upper	2	45*	6	5	0	57	0.52	0.52	0.50	6.56***
	Lower	7	15	22	13	0	57				
18	Upper	6	4	41*	6	0	57	0.50	0.42	0.36	4.91***
	Lower	15	22	17	3	0	57				
19	Upper	38*	5	10	4	0	57	0.45	0.42	0.36	4.93***
	Lower	14	21	17	5	0	57				
20	Upper	46*	3	5	3	0	57	0.49	0.63	0.63	8.62***
	Lower	10	25	17	5	0	57				
21	Upper	3	52*	1	1	0	57	0.63	0.56	0.54	7.57***
	Lower	9	20	14	14	0	57				
22	Upper	5	42*	2	8	0	57	0.46	0.54	0.57	6.88***
	Lower	12	11	19	15	0	57				
23	Upper	9	41*	4	3	0	57	0.47	0.49	0.44	5.72***
	Lower	21	13	11	12	0	57				
24	Upper	10	34*	6	7	0	57	0.44	0.29	0.36	3.20***
	Lower	5	17	21	14	0	57				
25	Upper	4	35*	10	8	0	57	0.40	0.42	0.50	5.02***
	Lower	17	11	19	10	0	57				
26	Upper	10	38*	5	4	0	57	0.47	0.38	0.40	4.43***
	Lower	16	16	14	11	0	57				
27	Upper	8	3	1	45*	0	57	0.56	0.45	0.46	5.47***
	Lower	5	10	23	19	0	57				
28	Upper	50*	4	3	-	0	57	0.54	0.66	0.63	9.53***
	Lower	12	24	15	6	0	57				
29	Upper	1	52*	3	1	0	57	0.65	0.50	0.56	6.72***
	Lower	7	23	18	9	0	57				
30	Upper	37*	7	8	5	0	57	0.42	0.45	0.46	5.51***
	Lower	11	21	14	11	0	57				
31	Upper	3	2	48*	4	0	57	0.53	0.61	0.51	8.26***
	Lower	17	15	13	12	0	57				
32	Upper	8	4	41*	4	0	57	0.46	0.50	0.51	6.27***
	Lower	14	25	12	6	0	57				
33	Upper	12	5	37*	3	0	57	0.42	0.43	0.21	5.22***
	Lower	17	13	12	15	0	57				
34	Upper	10	13	24*	10	0	57	0.28	0.28	0.28	3.48***
	Lower	17	18	8	15	0	57				
35	Upper	6	5	39*	7	0	57	0.44	0.47	0.45	5.73***
	Lower	10	19	12	16	0	57				
36	Upper	4	6	34*	13	0	57	0.42	0.33	0.34	3.78***
	Lower	10	21	15	11	0	57				
37	Upper	10	7	29*	11	0	57	0.40	0.21	0.22	2.32**
	Lower	6	13	17	21	0	57				
38	Upper	12	13	25*	7	0	57	0.28	0.29	0.30	3.68***
	Lower	14	17	8	18	0	57				
39	Upper	6	8	38*	5	0	57	0.38	0.56	0.56	7.46***
	Lower	17	23	6	11	0	57				
40	Upper	6	5	6	40*	0	57	0.64	0.10	0.09	1.17
	Lower	12	9	12	34	0	57				
41	Upper	4	7	4	42*	0	57	0.44	0.57	0.50	7.57***
	Lower	6	27	15	9	0	57				
42	Upper	36*	6	7	8	0	57	0.38	0.63	0.59	6.18***
	Lower	8	21	17	11	0	57				

p: difficulty, *d*: discrimination, *r*: item-total correlation, *t*: value between lower 27% - upper 27%
[1]n=211 [2]$n_1=n_2=57$ *correct answer **p<.05 ***p<.001

Based on the results presented in Table 6, it is observed that there are 29 items with a discrimination index of 0.40 or higher, five items with a discrimination index ranging from 0.30 to 0.39, and six items with a discrimination index between 0.20 and 0.29. Additionally, there are two items with a discrimination index of 0.19 or lower related to el-Cezeri and Fergani, which have been removed from the test. After removing these two items, the item-total correlations for all remaining items in the test range from 0.21 to 0.63. A t-value of $p < .001$ is significant for 39 items, and a t-value of $p < .05$ is significant for one item.

Based on the pilot study results, the KR-20 reliability coefficient for the Scientist Knowledge Test is 0.94, and the Spearman-Brown two-half test correlation value is 0.92. In light of these findings, the final version of the scale has been arranged to include 40 items. The Scientists Knowledge Test is included in the Appendix.

The final version of "The Science Scientists Knowledge Test" consists of items specific to a total of 15 different scientists. The distribution of the items related to the scientists in the final version of the test is presented in detail in Table 7.

Table 7. *Relationship of scientists with items in the Science Scientists Knowledge Test.*

Scientists	Related items
Ebu Bekir er-Razi	10-21-25-37-43
Abdurrahman el-Hazini	4-17-38-41
el-Cezeri	3-24-36-42
Ibnu'l-Heysem	5-11-29-44
Cabir ibn Hayyan	9-20-34
Farabi	12-19-28
Ibn Sina	15-18-23
Ibnu'n-Nefis	6-13-32
Ulug Bey	22-26-33
Ali Kuscu	7-27
Fergani	16-35
Aksemseddin	30
Biruni	14
Hezarfen Ahmet Celebi	31
Kindi	39

When Table 7 is examined, it is observed that Ebu Bekir er-Razi, Ibnu'l-Heysem, and el-Cezeri are the focal points of more items in the test. In the developed knowledge test, some scientists like Aksemseddin, Hezarfen Ahmet Celebi, Biruni are represented by only one item.

4. DISCUSSION and CONCLUSION

This study aims to develop a knowledge test that meets the validity and reliability requirements to assess secondary school students' awareness of scientists. In line with this objective, the stages of test development have been followed to create a valid and reliable knowledge test. Upon reviewing the existing literature, it was found that there is only one achievement test available for assessing the awareness of scientists. However, similar test development procedures have been observed to be used in scales designed for science courses (Akbulut & Çepni, 2013; Keçeci et al., 2019; Yazıcı et al., 2022).

The results of the item analysis conducted within the scope of the study indicate that the developed test ensures structural validity. While the majority of items in the test have item discrimination index values ranging from 0.21 to 0.66, most of the items in the developed knowledge test have item discrimination index values of 0.40 and above. Additionally, the mean item discrimination index value for the test is 0.45. These findings demonstrate that the

developed test has been designed to be valid and aligned with its intended measurement domain. Structural validity is a critical indicator of measurement quality that assesses the internal consistency of a measurement instrument and the harmony among its items. The high item discrimination index values for the test items signify that these items effectively represent the measured construct and are in line with the test's intended purpose (Tekin, 2010). Therefore, establishing the structural validity of the test aids in confirming the reliability and validity of the test results. Study suggests that tests designed for science courses generally have mean item discrimination index values of 0.40 and above (Açıkgöz & Karslı, 2015; Demir et al., 2016). These findings indicate that the developed knowledge test is consistent with the mean item discrimination index value. Furthermore, all items in the test exhibit item-total correlations ranging from 0.21 to 0.63. There is one item with a significant T-value at the 0.001 level and another at the 0.005 level. Variations in item-total correlations provide further insights into the extent to which items contribute to the overall internal consistency and the measurement objective. High correlation values indicate that the item aligns with the test and enhances the overall reliability of the measurement, whereas low correlation values may raise questions about the consistency between the item and the measurement. These results collectively demonstrate that the test is a reliable and valid measurement instrument in its entirety.

The difficulty levels of the test were assessed using item difficulty index (p) values. These values range from 0.28 to 0.75, indicating that some items are challenging for students while others are relatively easy. The most difficult items are item 33 and item 37, whereas the easiest item is item 5. It can be expressed that the test is generally easy for students, and the questions are appropriate for their comprehension levels, with an average item difficulty index of 0.45. These index values are crucial for evaluating the difficulty levels of the items (Kan, 2011). These results are consistent with similar studies, which have shown that tests developed for science education generally have a moderate to high difficulty level. Additionally, the item difficulty index value of this knowledge test is similar to that of other tests designed for science education, indicating its comparability with these tests (Aymen Peker & Taş, 2019; Bolat & Karamustafaoğlu, 2019).

The reliability of the developed Science Scientists Knowledge Test was evaluated using the KR-20 coefficient, which was calculated as 0.94. This value falls within the range of $0.60 \leq \alpha < 0.90$, indicating that the test scores are highly reliable (Can, 2014). These results provide strong support for the high reliability of the students' test scores. Additionally, in line with the reliability levels of similar tests in the literature, the reliability of the test has been robustly demonstrated (Açıkgöz & Karslı; Saraç, 2018). However, despite the achieved high reliability, there are certain limitations regarding the generalizability of the test. Specifically, the focus of the test on Turkish-Islamic scholars may restrict the potential student groups to whom the test can be applied within a certain cultural and religious context. This situation could limit the overall validity of the test and may impact its applicability to student groups from different cultures or disciplines.

In conclusion, this study makes a significant contribution to the development of a reliable test aimed at assessing middle school students' awareness of scientists in the field of science education. The study can guide science education practitioners and researchers in providing an effective tool for measuring students' perceptions of science. Furthermore, future studies may assess the applicability of the test in a broader student population and contribute further to understanding learning efforts in science education. The analysis of items and the assessment of the test's reliability provide a substantial contribution to test development methodology. This study demonstrates that the steps in the knowledge test development process were followed accurately, and the test is suitable for its intended measurement purpose. Consequently, it can guide other researchers and educational professionals in test development. These aspects

indicate that the study enriches the knowledge in the field of science education and has the potential to offer an effective tool for measuring students' awareness of scientists.

5. RECOMMENDATIONS

Based on the results of this study, the following recommendations can be made:

- This study involved the development of a knowledge test to assess secondary school students' awareness of scientists. It is recommended that similar knowledge tests be developed for a wider range of topics and fields. This would allow for the evaluation of scientific awareness among students at different educational levels and across different disciplines.
- The Science Scientists Knowledge Test that has been developed has proven to be a powerful tool for measuring students' awareness of scientists. This test may encourage students to show more interest in science-related career opportunities and scientific study. Therefore, it is recommended that this test be widely used in schools and educational institutions.
- This study focused on assessing awareness of scientists. Future study could explore other aspects of students in the field of science education, leading to a more comprehensive understanding. This could provide more opportunities to enhance science education.

Declaration of Conflicting Interests and Ethics

The authors declare no conflict of interest. This research study complies with research publishing ethics. The scientific and legal responsibility for manuscripts published in IJATE belongs to the authors. **Ethics Committee Number**: Firat University Social Sciences and Humanities Research Ethics Committee, 23/11/2020-425157.

Contribution of Authors

Pelin Yıldırım: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing-original draft preparation, Writing – review & editing. **Gonca Keçeci**: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing-original draft preparation. **Fikriye Kırbağ Zengin**: Conceptualization, Project administration, Resources, Validation, Visualization.

Orcid

Pelin Yıldırım https://orcid.org/0000-0003-4425-2472
Gonca Keçeci https://orcid.org/0000-0002-2582-3850
Fikriye Kırbağ Zengin https://orcid.org/0000-0002-0547-8746

REFERENCES

Açıkgöz, M., & Karslı, F. (2015). Alternatif ölçme-değerlendirme yaklaşımları kullanılarak iş ve enerji konusunda geliştirilen başarı testinin geçerlilik ve güvenirlik analizi [Validity and reliability analysis of achievement test which developed for work and energy, using alternative measure-evaluation techniques]. *Amasya Education Journal*, *4*(1), 1-25. Erişim adresi: https://dergipark.org.tr/en/download/article-file/19638

Ağgül Yalçın, F. (2012). Öğretmen adaylarının bilim insanı imajlarının bazı değişkenlere göre incelenmesi [Investigation of prospective teachers' image of scientist with respect to some variables]. *Elementary Education Online*, *11*(3), 611-628. https://dergipark.org.tr/tr/download/article-file/90555

Akbulut, H.İ., & Çepni, S. (2013). Bir üniteye yönelik başarı testi nasıl geliştirilir?: İlköğretim 7. sınıf kuvvet ve hareket ünitesine yönelik bir çalışma [How to develop an achievement

test for a unit?: A study for grade 7 force and motion unit]. *Amasya Education Journal*, *2*(1), 18-44. https://dergipark.org.tr/en/download/article-file/19600

Arabacı, S., & Dönel Akgül, G. (2022). Fen Bilimleri ders kitaplarına bir bakış: Türk İslam bilim insanları [A overview of science textbooks: Turkish Islamic scientists]. *International Journal of Social Sciences Academy*, *4*(9), 386-404. https://doi.org/10.47994/usbad.1068516

Aymen Peker, E., & Taş, E. (2019). Üç aşamalı kavram tanı testi geliştirme ve geçerlik çalışması: 5. sınıf canlılar dünyasını gezelim tanıyalım ünitesi [Three tier concept diagnostic test development and validity study: Let's travel and learn about the living world unit of fifth grade]. *Journal of Computer and Education Research*, *7*(14), 515–539. https://doi.org/10.18009/jcer.602223

Bıçak, C.Y., & Bilir, V. (2023). Fen Bilimleri ders kitaplarında yer alan bilim insanları ile ilgili bilgilerin öğretmen gözünden değerlendirilmesi [Evaluation of the information regarding scientists in the science textbooks from the teacher's perspective]. *Mehmet Akif Ersoy University Journal of Education Faculty*, *65*, 487-511. https://doi.org/10.21764/maeuefd.1107010

Bilir, V., Tatlı, A., Yıldız, C., Emiroğlu, B.B., Ertuğrul, D., & Sakmen, G. (2020). Argümantasyon tabanlı öğrenme yaklaşımında kullanılan argümantasyon tekniklerinin ortaokul sekizinci sınıf öğrencilerinin bilim insanı imajları üzerine etkisi [The effect of argumentation techniques used in argumentation based learning approach on 8th grade secondary school students' image of scientist]. *Journal of Gazi University Gazi Education Faculty*, *40*(2), 481-510. https://dergipark.org.tr/en/download/article-file/722283

Bolat, A., & Karamustafaoğlu, S. (2019). "Vücudumuzdaki sistemler" ünitesi başarı testi geliştirme: Geçerlik ve güvenirlik [Development achievement of test on the unit "systems in our body": Validity and reability]. *Gazi Journal of Education Sciences*, *5*(2), 131-159. https://doi.org/10.30855/gjes.2019.05.02.008

Can, A. (2014). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi [Quantitative data analysis in the scientific research process with SPSS]* (2nd ed.). Pegem Akademi Yayıncılık.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart and Winston.

Çardak, Ç.S., & Selvi, K. (2018). Öğretim ilke ve yöntemleri dersi için bir başarı testi geliştirilmesi [Developing an achievement test for the teaching principles and methods course]. *Mediterranean Educational Research Journal*, *12*(26), 379-406. https://doi.org/10.29329/mjer.2018.172.19

Demir, N., Kızılay, E., & Bektaş, O. (2016). 7. sınıf çözeltiler konusunda başarı testi geliştirme: Geçerlik ve güvenirlik çalışması [Development of an achievement test about solutions for 7th graders: A validity and reliability study]. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education*, *10*(1). https://doi.org/10.17522/nefefmed.52947

Doğan, N. (2009). *Çoktan seçmeli testler [Multiple choice tests]*. H. Atılgan (Eds.), Eğitimde ölçme ve değerlendirme [Measurement and evaluation in education] içinde (3rd ed.) (pp. 223-268). Anı Yayıncılık.

Duran, E., & Bayar, A. (2019). İlkokul öğrencilerinin bilim ve bilim insanına ilişkin algıları [The perceptions of primary school students on science and scientist]. *Social Sciences Research Journal*, *8*(4), 14-29. https://dergipark.org.tr/en/download/article-file/870418

Durukan, Ü. G. (2017). Fen Bilimleri ders kitaplarında yer alan bilim insanları [Scientists in science textbooks]. *Karadeniz Sosyal Bilimler Dergisi*, *9(Kadın özel sayısı)*, 321-344. https://search.trdizin.gov.tr/tr/yayin/detay/265362/

Ertürk, S. (1972). *Eğitimde "program" geliştirme [Developing a "program" in education]*. Hacettepe Üniversitesi Basımevi.

Karaçam, S., Aydın, F., & Digilli, A. (2014). Fen ders kitaplarında sunulan bilim insanlarının basmakalıp bilim insanı imajı açısından değerlendirilmesi [Evaluation of scientists represented in science textbooks in terms of stereotype scientist image]. *Ondokuz Mayis University Journal of Faculty of Education*, *33*(2), 606-627. https://doi.org/10.7822/omuefd.33.2.19

Karasu, B. (2019). *Ortaokul Fen Bilimleri derslerinde işlenen bilim insanlarına yönelik formüllü tişört yönteminin akademik başarıya etkisi [The effect on academic success of formulated t-shirt method for teaching scientists' studies in middle school science course]* [Master's thesis]. Ağrı İbrahim Çeçen University.

Karip, E. (Ed.). (2012). *Ölçme ve değerlendirme [Measurement and evaluation]* (5th ed.). Pegem Akademi.

Keçeci, G., Yıldırım, P., & Kırbağ Zengin, F. (2019). Sistemler akademik başarı testi: Geçerlik ve güvenirlik çalışması [Systems academic achievement test: Validity and reliability study]. *National Journal of Education Academy*, *3*(1), 96-114. https://doi.org/10.32960/uead.514188

Linn, R.L., & Gronlund, N.E. (2000). *Measurement and assessment in teaching* (8th ed.). Merrill/Prentice-Hall.

Mehrens, W.A., & Lehmann I.J. (1991). *Measurement and evaluation in education and psychology* (4th ed.). Harcourt Brace College.

Meydan, A., & Akkuş, A. (2014). Sosyal bilgiler öğretiminde müze gezilerinin tarihi ve kültürel değerlerin kazandırılmasındaki önemi [The importance of getting students adopt the historical and cultural values through museum visits in social studies teaching]. *Marmara Coğrafya Dergisi*, *29*, 402-422. https://doi.org/10.14781/mcd.92390

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.

Özdemir, E. (2017). *Ortaokul öğrencilerinin bilim insanı imajı hakkındaki görüşleri [Image of secondary school students' views of scientist]* [Unpublished master's thesis]. Necmettin Erbakan University.

Özdemir, E. (2022). Ortaokul Fen Bilimleri ders kitaplarının bilim insanları bağlamında incelenmesi [Analysis of secondary school Science textbooks in the context of scientists]. *Journal of Research in Education and Society*, *9*(2), 403-429. https://doi.org/10.51725/etad.1098860

Özdeş, S., & Aslan, O. (2019). Kız öğrencilerin bilim insanı cinsiyetine yönelik algılarını ve bilim insanı olma isteklerini etkileyen faktörler [The factors affecting the girl students' perceptions about the gender of scientist and their willingness to become a scientist]. *Ondokuz Mayis University Journal of Education Faculty*, *38*(1), 190-214. https://doi.org/10.7822/omuefd.501944

Saraç, H. (2018). Fen Bilimleri dersi 'maddenin değişimi' ünitesi ile ilgili başarı testi geliştirme: Geçerlik ve güvenirlik çalışması [Development of achievement test about science lesson 'change of matter' unit: Validity and reliability study]. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, *18*(1), 416-445. https://doi.org/10.17240/aibuefd.2018..-388815

Sarıbaş, S., & Babadağ, G. (2015). Temel eğitimin temel sorunları [Basic problems of primary education]. *Anatolian Journal of Educational Leadership and Instruction*, *3*(1), 18-34.

Tekin, H. (2010). *Eğitimde ölçme ve değerlendirme [Measurement and evaluation in education]* (20th ed.). Yargı Yayınevi.

Turgut, M.F. (1992). *Eğitimde ölçme ve değerlendirme [Measurement and evaluation in education]*. Saydam Matbaacılık.

Turgut, M.F. (1997). *Eğitimde ölçme ve değerlendirme metotları [Measurement and evaluation methods in education]*. Gül Yayınevi.

Turgut, M.F., & Baykul, Y. (2012). *Eğitimde ölçme ve değerlendirme [Measurement and evaluation in education]* (4th ed.). Pegem Akademi.

Ülger, K. (2021). Uzaktan eğitim modelinde karşılaşılan sorunlar-fırsatlar ve çözüm önerileri [Problems-opportunities and solution suggestions encountered in the distance education model]. *International Journal of Contemporary Educational Studies*, *7*(1), 393-412.

Yazıcı, F., Gül, Ş., & Sözbilir, M. (2022). "Bitki ve hayvanlarda üreme, büyüme ve gelişme" konusuna yönelik bir başarı testinin geliştirilmesi [The development of an achievement test towards "reproduction, growth and development in plants and animals" subject]. *Journal of Atatürk University Kazım Karabekir Education Faculty*, *46*, 11-19. https://doi.org/10.5152/AUJKKEF.2022.837043

Yıldırım, A., Şimşek, H. (2008). *Sosyal bilimlerde nitel araştırma yöntemleri [Qualitative research methods in the social sciences]*. Seçkin Yayıncılık.

Yıldırım, P., & Keçeci, G. (2022). Türk-İslam alimlerinin Fen Bilimleri ders kitaplarında ve öğretim programında yeri [The place of Turkish-Islamic scholars in Science textbooks and curriculum]. *Van Yüzüncü Yıl University Journal of Education*, *19*(2), 429-445. https://doi.org/10.33711/yyuefd.1064898

APPENDIX

Turkish version of Scientists Knowledge Test

BİLİM İNSANLARI BİLGİ TESTİ

Sevgili Öğrencim;

Bu çalışma ile fen alanında bilime yön vermiş Türk-İslam alimlerine ilişkin bilgi testinin geliştirilmesi, ortaokul öğrencilerinin bilgi düzeyinin belirlenmesi amaçlanmaktadır. Çalışmanın sonuçları yalnızca bilimsel amaçla kullanılacak olup, kişisel bilgileriniz kimseyle paylaşılmayacaktır.

Cinsiyetiniz : ( ) Kız ( ) Erkek

Sınıf Düzeyiniz : ( ) 5. sınıf ( ) 6. sınıf ( ) 7. sınıf ( ) 8. sınıf

1. Aşağıdakilerden hangisi filli su saatinde yer alan unsurlardan değildir?

A) Fil B) Zümrüdüanka Kuşu C) Fıskiye D) Hükümdar

2. Türkistan'da yetişen, yer çekimi ve terazilerle alakalı çalışmalar yapan fizik, astronomi ve matematik alimi aşağıdakilerden hangisidir?

A) Fatih Sultan Mehmed B) Ali kuşçu

C) İbn Sina D) Abdurrahman el-Hazini

3. Batı'da "*Alhazen*" ve "*Alhacen*" gibi isimlerle tanınan ve Kitabü'l-Menazir (Görüntüler Kitabı) isimli eseri olan Türk-İslam alimi aşağıdakilerden hangisidir?

A) Ali Kuşçu B) İbnü'l-Heysem C) Biruni D) Piri Reis

4. Yaşadığı dönemde "*İkinci İbn Sînâ*" olarak anılan Türk-İslam alimi aşağıdakilerden hangisidir?

A) İbnü'n-Nefis B) Uluğ Bey C) İbnü'l-Heysem D) el-Cezeri

5. "*Kuşçu*" veya "*Kuşi*" lakabıyla anılan, Uluğ Bey'den matematik ve astronomi dersleri alan Türk-İslam alimi aşağıdakilerden hangisidir?

A) İbnü'n-Nefis B) Farabi C) Ali Kuşçu D) Mimar Sinan

6. Şaraptan, saf alkol elde eden ilk bilim insanı kimdir?

A) Biruni B) Cabir İbn Hayyan C) Evliya Çelebi D) el-Cezeri

7. Kızıl ve kızamık hastalıklarının iki ayrı hastalık olduğunu ortaya koyan bilim insanı aşağıdakilerden hangisidir?

A) Harezmi B) Ebu Bekir er-Razi C) Abdurrahman el-Hazini D) el-Cezeri

8. Aşağıdakilerden hangisi, İbnü'l-Heysem'in optikle ilgili en önemli eseridir?

A) Mizanü'l Hikme (Hikmetin Terazisi)

B) Kitabü'l-Menazir (Görüntüler Kitabı)

C) el-Kanun Fi't-Tıbb (Tıbbın Kanunu)

D) Risaletü'l-Fethiyye

9. Hava titreşimlerinden ibaret olan ses olayının ilk mantıklı açıklamasını hangi bilim insanı yapmıştır?

A) İbn Sina B) Evliya Çelebi C) Farabi D) Uluğ Bey

10. Ciltte oluşan beyaz lekelerden söz eden ve bunun deride yer yer renk kaybı şeklinde ortaya çıkan bir hastalık olduğunu belirten Türk -İslam alimi kimdir?

A) Uluğ Bey B) Ali Kuşçu C) İbnü'n-NefisD) Biruni

11. "*Piknometre*" denilen bir aletle, cisimlerin özgül ağırlıklarını ilk defa ölçen Türk-İslam alimi aşağıdakilerden hangisidir?

A) Akşemseddin B) Biruni C) Piri Reis D) Ali Kuşçu

12. Menenjit hastalığını ve türlerini ilk defa tespit eden bilim insanı aşağıdakilerden hangisidir?

A) el-Cezeri B) Cabir İbn Hayyan
C) Musaoğulları (Benî Musa) D) İbn Sina

13. Batı'da "*Alfraganus*" ismiyle çok meşhur olan Türk-İslam alimi aşağıdakilerden hangisidir?

A) İbnü'n-Nefis B) Fergani C) İbn Sina D) Akşemseddin

14. Aşağıdakilerden hangisi, çeşitli maddelerin özgül ağırlıklarını tespit eden Türk-İslam alimlerinden biridir?

A) Abdurrahman el-Hazini B) Piri Reis
C) İbnü'l-Heysem D) Mimar Sinan

15. Batı'da "*Avicenna*" ismiyle şöhret yapan Türk-İslam alimi aşağıdakilerden hangisidir?

A) Fatih Sultan Mehmed B) İbn Sina
C) Piri Reis D) Mimar Sinan

16. Batı'da Alfarabius, Abunazar gibi isimlerle meşhur olan Türk-İslam alimi aşağıdakilerden hangisidir?

A) el-Cezeri B) Farabi C) İbn Sina D) Mimar Sinan

17. Batı dünyasında "*Geber*" ismiyle tanınan, Doğu'da ve Batı'da kimyanın kurucusu kabul edilen Türk-İslam alimi aşağıdakilerden hangisidir?

A) el-Cezeri B) Mimar Sinan C) Cabir İbn Hayyan D) Uluğ Bey

18. Hastanelere klinik sistemini ilk yerleştiren kişi aşağıdakilerden hangisidir?

A) Ebu Bekir er-Razi B) Farabi
C) İbnü'l-Heysem D) el-Cezeri

19. Aşağıdakilerden hangisi 15. yüzyılın en büyük gökbilimcilerinden biridir?

A) Uluğ Bey B) Akşemseddin C) Ebu Bekir er-Razi D) Fatih Sultan Mehmed

20. Aşağıdakilerden hangisi Türk-İslam alimidir?

A) Galileo B) Avicenna C) Newton D) Archimedes

21. el-Cezeri, en önemli bilimsel buluşlarını hangi alanda yapmıştır?

A) Matematik B) Fizik C) Kimya D) Astronomi

22. Gırtlak sinirini keşfeden ve bu sinirin bazen sağ tarafta olduğunu da ifade eden Türk-İslam alimi aşağıdakilerden hangisidir?

A) Mimar Sinan B) Ebu Bekir er-Razi
C) Abdurrahman el-Hazini D) el-Cezeri

23. Uluğ Bey'in, kurduğu rasathanede gerçekleştirilen gözlemlerin sonuçlarının toplandığı dünyaca meşhur eserinin adı aşağıdakilerden hangisidir?

A) Mizanü'l Hikme (Hikmetin Terazisi)
B) Zic-i Uluğ Bey
C) Kitabü'l-Menazir (Görüntüler Kitabı)
D) el-Kanun Fi't-Tıbb (Tıbbın Kanunu)

24. Astronomiyle ilgili Risâle fi'l-Heyʾe isimli Farsça eserine yeni ilaveler yaparak Risaletü'l-Fethiyye adıyla Arapça'ya çeviren ve Fatih Sultan Mehmed'e sunan Türk-İslam alimi kimdir?

A) el-Cezeri B) Ali Kuşçu C) Farabi D) İbnü'n-Nefis

25. Müzikte sesleri notalarken ve sesleri bölümlerken logaritmayı icat eden Türk-İslam alimi aşağıdakilerden hangisidir?

A) İbn Sina B) Farabi C) Akşemseddin D) el-Cezeri

26. Gelişmiş kamerayı tasarlayan Leonardo da Vinci'den çok daha önce karanlık odayı keşfeden Türk- İslam alimi aşağıdakilerden hangisidir?

A) Ali kuşçu B) Farabi C) İbn Sina D) İbnü'l-Heysem

27. Tıp tarihinde ilk defa mikrop meselesini ortaya atan ve hastalıkların bu yolla bulaştığı fikrini öne süren Türk-İslam alimi kimdir?

A) Akşemseddin B) Fatih Sultan Mehmed
C) Ali Kuşçu D) Mimar Sinan

28. Yaptığı özel bir aletle İstanbul'da Galata Kulesi'nden havalanarak Boğaz'ı geçip Üsküdar'a indiği rivayet edilen Türk-İslam alimi aşağıdakilerden hangisidir?

A) el-Cezeri B) Hezarfen Ahmed Çelebi C) Mimar Sinan D) Uluğ Bey

29. I. İdrar incelemesiyle şeker hastalığını teşhis ve tespit etmiştir.
II. Döneminde "*İkinci İbn Sînâ*" olarak anılmıştır.
III. Küçük kan dolaşımını keşfetmiştir.
IV. Filli su saati, en önemli buluşlarından biridir.

İbnü'n-Nefis ilgili, yukarıdaki bilgilerden hangileri yanlıştır?

A) I ve IV B) Yalnız III
C) I, II ve III D) Hepsi

30. Uluğ Bey'in 1421 yılında Semerkant'ta açtığı, dönemin en gelişmiş rasathanesi aşağıdakilerden hangisidir?

A) İstanbul Rasathanesi
B) Meraga Rasathanesi
C) Semerkant Rasathanesi
D) Hemedan Rasathanesi

31. Geliştirdiği kimyasal fırınlar sayesinde, bugün bildiğimiz kostik soda, sülfirik asit, arsenik asit ve nitrik asit gibi temel asitleri kimyasal yollar ile elde eden Türk-İslam alimi aşağıdakilerden hangisidir?

A) Uluğ Bey B) Akşemseddin C) Cabir İbn Hayyan D) Farabi

32. Gezegenler gibi Güneş'in de kendi yörüngesinde Batı'dan Doğu'ya bir hareketi olduğunu ilk defa açıklayan Türk-İslam alimi aşağıdakilerden hangisidir?

A) Hezarfen Ahmed Çelebi B) Farabi C) Fergani D) İbn Sina

33. I. En önemli bilimsel buluşlarını kimya alanında yapmıştır.
II. Doğu'da ve Batı'da kimyanın kurucusu kabul edilmektedir.
III. Filli su saati ve değişken şekilli fıskiye, Cezeri'nin en önemli buluşlarındandır.
IV. Fil, zümrüdüanka kuşu, seyis, hükümdar ve halı, el-Cezeri'nin filli su saatinde yer alan unsurlardandır.

el-Cezeri ile ilgili, yukarıdaki bilgilerden hangileri doğrudur?

A) I ve II B) Yalnız III
C) III ve IV D) Hepsi

34. Simyayı tıbbın hizmetine sunan ilk kişi kimdir?

A) İbnü'l-Heysem B) Mimar Sinan C) Ebu Bekir er-Razi D) el-Cezeri

35. Aşağıdakilerden hangisi, Abdurrahman el-Hazini'nin en önemli eserlerinden biridir?

A) el-Kanun Fi't-Tıbb (Tıbbın Kanunu)
B) Risaletü'l-Fethiyye
C) Mizanü'l Hikme (Hikmetin Terazisi)
D) Kitabü'l-Menazir (Görüntüler Kitabı)

36. Gökyüzünün ve denizlerin mavi görünmelerinin nedenini ilk defa doğru bir şekilde açıklayan Türk-İslam alimi aşağıdakilerden hangisidir?

A) el-Cezeri B) Abdurrahman el-Hazini
C) Kindi D) Musaoğulları (Benî Musa)

37. Yapmış olduğu deneyler sonucunda bütün cisimlerin yerin merkezine doğru, bir kütle çekim kuvveti ile çekildiklerini gösteren ve yer çekimini keşfeden Türk-İslam alimi aşağıdakilerden hangisidir?

A) Cabir İbn Hayyan B) İbn Sina C) Abdurrahman el-Hazini D) el- Cezeri

38. Aşağıdakilerden hangisi, el-Cezeri'nin en önemli bilimsel buluşlarından biridir?

A) Piknometre B) Hikmet Terazisi C) Usturlap D) Filli su saati

39. Allerji üzerine ilk eser yazan alim aşağıdakilerden hangisidir?

A) Fatih Sultan Mehmed B) Uluğ Bey C) Piri Reis D) Ebu Bekir er-Razi

40. Aşağıdakilerden hangisi önemli bir fizikçi ve optikçidir?

A) İbnü'l-Heysem B) İbn Sina C) Fatih Sultan Mehmed D) Harezmi

Bilim İnsanları Bilgi Testi Cevap Anahtarı

1. C	2. D	3. B	4. A	5. C	6. B	7. B	8. B	9. C	10. C
11. B	12. D	13. B	14. A	15. B	16. B	17. C	18. A	19. A	20. B
21. B	22. B	23. B	24. B	25. B	26. D	27. A	28. B	29. A	30. C
31. C	32. C	33. C	34. C	35. C	36. C	37. C	38. D	39. D	40. A