Cumhuriyet International Journal of Education, 11(1): 228-239, 2022
DOI: https://doi.org/10.30703/cije.1016641

Cumhuriyet International Journal of Education

| cije.cumhuriyet.edu.tr | Founded: 2011 Available online, ISSN: 2147-1606 Publisher: Sivas Cumhuriyet Üniversitesi

Developing Two-Tier Diagnostic Instrument to Determine Misconceptions on Socioscientific Issues

Fatma Güler Gürsel[1,a,*], Behiye Akçay[2,b]

[1]*Istanbul University-Cerrahpasa, Institute of Graduate Studies, Science Education PhD Program, Istanbul/Turkey*
[2]*Istanbul University-Cerrahpasa, Hasan Ali Yucel Faculty of Education, Department of Mathematics and Science Education, Istanbul/Turkey*
**Corresponding author*

Research Article

History

Received: 07/11/2021
Accepted: 31/01/2022

ABSTRACT

Socioscientific issues, which we encounter more and more frequently, will gain much more importance in the coming years. Socioscientific issues have many topics that individuals will experience decision-making processes in the future. Adequate knowledge of the concepts lies at the basis of the decision-making process. The aim of this study is to develop a two-tier diagnostic test to identify misconceptions in socioscientific issues and to study the validity and reliability of this test. In this scale, which was developed in the mixed method, there are questions about energy resources, global warming, genetic engineering and cloning. In order to develop the two-tier test, a process consisting of three basic stages was carried out: determining the content, obtaining information about students' misunderstandings, and developing the diagnostic test. The data of the study were collected from eighth and ninth grade students. In order to determine the final version of the scale; the reliability, difficulty and discrimination values were analyzed with the data of 385 students. A test-retest application was also made with the participation of 55 students. The KR-20 value of the 30-item two-tier socioscientific issues concept test developed is 0.85, the average difficulty value is 0.46987, the average discrimination value is 0.497364, and the Pearson Correlation coefficient for test-retest reliability was 0.776. As a result of the study, a valid and reliable two-tier diagnostic test consisting of 30 items that can be used to identify misconceptions in socioscientific issues has been developed.

Keywords: Socioscientific issues, Misconception, Two-tier diagnostic test, Science education, Scale development

Sosyobilimsel Konularda Kavram Yanılgılarını Belirlemeye Yönelik İki Aşamalı Teşhis Testi Geliştirme

Bilgi

**Sorumlu yazar*

Süreç
Geliş: 07/11/2021
Kabul: 31/01/2022

Bu çalışma ön inceleme sürecinde ve yayımlanmadan önce iThenticate yazılımı ile taranmıştır.

ÖZ

Günümüzde gittikçe daha sık karşımıza çıkan sosyobilimsel konular, ilerleyen yıllarda çok daha fazla önem kazanacaktır. Sosyobilimsel konular kapsamında pek çok konu yer almakta ve gelecekte bireylerin bu konular ile ilgili karar verme süreçleri yaşayacakları öngörülmektedir. Karar verme sürecinin temelinde de konu ile ilgili kavramların yeterli düzeyde bilinmesi bulunmaktadır. Bu çalışmanın amacı sosyobilimsel konulardaki kavram yanılgılarını belirlemeye yönelik iki aşamalı bir teşhis testinin geliştirilmesi ve bu testin geçerlik ve güvenirlik çalışmasının yapılmasıdır. Karma yöntemde geliştirilen bu ölçekte sosyobilimsel konulardan enerji kaynakları, küresel ısınma, genetik mühendisliği ve klonlamaya dair sorular yer almaktadır. İki aşamalı testin geliştirilmesi için içeriğin belirlenmesi, öğrencilerin yanlış anlamaları hakkında bilgi edinilmesi ve teşhis testinin geliştirilmesi şeklinde üç temel aşamadan oluşan bir süreç gerçekleştirilmiştir. Araştırmanın verileri sekizinci ve dokuzuncu sınıf öğrencilerinden toplanmıştır. Ölçeğin son hâlini belirlemek adına 385 öğrencinin verileri ile yapılan analizler sonunda ortalama güçlük ve ortalama ayırt edicilik değerleri, testin güvenirliği için K-20 değeri hesaplanmıştır. Analizler tamamlandıktan sonra 55 öğrencinin katılımı ile test tekrar test uygulaması yapılmıştır. Yapılan çalışma sonucunda geliştirilen 30 maddelik iki aşamalı sosyobilimsel konular kavram testinin KR-20 değeri 0.85, ortalama güçlük değeri 0.46987, ortalama ayırt edicilik değeri 0.497364 ve test tekrar test güvenirliği için Pearson Korelasyon katsayısı 0.776'dır. Çalışma sonucunda sosyobilimsel konularda kavram yanılgılarını belirlemeye yönelik kullanılabilecek olan 30 maddeden oluşan iki aşamalı geçerli ve güvenilir bir teşhis testi geliştirilmiştir.

Anahtar Kelimeler: Sosyobilimsel konular, Kavram yanılgısı, İki aşamalı teşhis testi, Fen eğitimi, Ölçek geliştirme

[a] *fgg2307@hotmail.com* *https://orcid.org/0000-0002-2892-7743* [b] *bbezir@gmail.com* *https://orcid.org/ 0000-0002-0546-8759*

How to Cite: Gürsel, F. G., Akçay, B. (2022) Developing Two-tier diagnostic instrument to determine misconceptions on socioscientific issues, Cumhuriyet International Journal of Education, 11(1): 228-239

Introduction

Science affects human life in many aspects (Jennings, 1982) and it is thought that the decisions on the selection of issues such as energy sources to be used in the future will be made by the public as well as experts and politicians (Lay, Khoo, Treagust, & Chandrasegaran, 2013). Topics like global warming, cloning, alternative fuels have been termed as "socioscientific issues (SSI)" because of central roles of both social and scientific factor in these dilemmas (Sadler, 2004). SSI is naturally controversial; there are also aspects of moral reasoning or ethical concern when making decisions (Zeidler & Nichos, 2009). Science education should respond to a variety of contemporary concerns and crises so there is a need to develop a curriculum that pays more attention to these topics (Hodson, 1994). Within the scope of the special aims of the Science course curriculum of the Ministry of National Education (MoNE) of the Republic of Turkey, the expression "to develop reasoning ability, scientific thinking habits and decision-making skills by using socioscientific issues" is included (MoNE, 2018a). In the 4th item of the special objectives section of the curriculum of the current Science Practices course of the MoNE, the expression "to develop reasoning, scientific thinking habits and decision-making skills by using socioscientific issues" is included and the expression "decides on socioscientific issues by logical reasoning" is included in the learning outcomes (MoNE, 2018b).

There is a general consensus about that students come to science class with some ideas which differs from what is accepted by science community and do not agree with scientific explanations, teachers and scientists' ideas and expressed as misconceptions, preconceptions, alternative frameworks, or children's science (Treagust, 1988; Anderson, Fisher, & Norman, 2002). Some of the methods to reveal students' misconceptions are concept maps, estimation-observation-explanation, interviews, drawings, phenomenography, V diagrams and word association (Karataş, Köse, & Coştu, 2003; Treagust, 1988).

In the past, individual interviews were generally used to identify misconceptions, since 1971 multiple-choice (MC) tests have also been used, allowing teachers to easily identify students' misconceptions in a defined area (Treagust, 1988). Because making interviews with students is time consuming and requires substantial training (Treagust, & Haslam, 1986) even though they give richer information about students' performance (Briggs, Alonzo, Schwab, & Wilson, 2006). Therefore, tests are tools widely used in education (Yaghmour, Obaidat, & Hamadnedh, 2016). The tests used in the detection of misconceptions can be grouped under five groups as short-answer tests, tests that require classification, multiple-choice tests, two-tier tests and open-ended tests (Demirci, & Efe, 2007). Recently, many studies have focused on defining and correcting concepts that differ from scientific knowledge, and it has become common to develop multiple-choice questions for this purpose but multiple-choice tests cannot distinguish whether the answers are due to lack of knowledge or misconceptions (Caleon, & Subramaniam, 2010). In these tests, which enable to identify misconceptions in a short way, students cannot explain the reason for their answer (Demirci, & Efe, 2007; Kenan, & Özmen, 2014). The student who does not know the correct answer may be able to give the correct answer, results may be affected by the power of reading comprehension, the expressions should be clear and understandable and sufficient time should be given (Karataş et al., 2003). To eliminate disadvantages, two-tier tests which developed by Treagust are used to determine misconceptions and the comprehension level of students (Sıbıç, Akçay & Arık, 2020). Two-tier tests' first tier consists of a content question, while the second tier requires a reasoning response (Chandrasegaran, Treagust, & Mocerino, 2007). The types of two-tier tests developed to eliminate the stated negativities are shown in Table 1 (Karataş et al, 2003):

The scale development process includes a number of steps. Treagust (1988) states that developing a two-tier test has 3 main stages and 10 steps (see Table 2).

Some of the data collection tools used in the literature to identify misconceptions about socioscientific issues for secondary school students are shown in Table 3.

The aim of this study is to develop a two-tier diagnostic test for secondary school curriculum that can be used to identify students' misconceptions about socioscientific issues. To make choices about socioscientific issues consciously in the future, it is extremely important for education stakeholders to understand the nature of SSI and its teaching and effectively implement SSI-related activities (Bayram, 2021). Recent studies showed that students have some misconceptions about socioscientific issues. Aubrecht (2018) argued that students' misconceptions about global warming includes pollution, the hole in the ozone layer are causes of the global warming, lack of understanding of relationship between human actions and climate change; greenhouse gases, and greenhouse mechanism. Most students believe that food produced by genetic engineering like as genetically modified organisms (GMO) are harmful and caused variety of health problems (Topaloğlu, & Kıyıcı, 2018). Students have misconceptions about energy sources including nuclear, fossil and renewable energy sources as well (Kaplan, 2019; Bahar, & Aydın, 2002). Looking at the literature, it is seen that open-ended questions and multiple-choice tests are used to identify misconceptions about socioscientific issues. However, these scales mostly address a specific socioscientific issue like as global warming (Kılınç, Stanisstreet, & Boyes, 2008), adaptation and natural Selection (Bakırcı, & Çalık, 2013), greenhouse effect (Bozkurt, & Cansüngü-Koray, 2002). The lack of a comprehensive socioscientific issues concept test for the secondary school curriculum makes this study important. The Socioscientific Issues Concept Test developed in this study is an important scale that can be used by teachers as it includes socioscientific subjects in the Science curriculum.

Table 1. Types and contents of two-tier tests

Types of Two-Tier Tests	First Tier	Second Tier
Multiple choice two-tier tests	Multiple choice	Multiple choice (+Open ended)
Two-tier tests that require classification	True-False	Multiple choice (+Open ended)
Open-ended two-tier tests	Multiple choice	Open ended

Table 2. According to Treagust (1988), two-tier test development steps

Stage	Step	Process
Defining The Content	Identifying propositional knowledge statements	It enables the researcher to examine the nature of the content.
	Developing a concept map	
	Relating propositional knowledge to the concept map	It provides the internal consistency of the test. It ensures the reliability that the propositions and concepts are in the same field.
	Validating the content	After the content is reviewed, if there is anything that needs to be changed, corrections are made. Thus, the development of any problem that is not related to the concepts to be taught is prevented. The content and concepts to be developed must be scientific.
Obtaining Information About Students' Misconceptions	Examining related literature	The misconceptions in the literature on the subject are examined.
	Conducting unstructured student interviews	Interviews are held with students who are knowledgeable about related concepts. In this way, it helps to identify misunderstandings and/or misconceptions while preparing multiple-choice questions and provides ideas for the next steps.
	Developing multiple choice content items with free response	Multiple choice items are written according to the taught topic. Each item is based on a limited number of propositions and addresses encountered misconceptions. Under each multiple-choice item, a space is left for the student to write the reason for their answer. Misconceptions become more evident when the written items are applied to students in a class.
Developing a Diagnostic Test	Developing the two tier diagnostic tests	In the first part of each item, there is usually a question with 2 or 3 options. The second part of each question provides 4 possible reasons for the answer given in the first part. The options include a correct answer, a defined misconception, a misconception, and, if necessary, a simple wrong answer. The options here are shaped by the data collected from the literature, preliminary practice and interviews. If more than one misconception is expressed by the students, they can be put in the options.
	Designing a specification grid	It is designed to ensure that the expressions in the test cover all concepts fairly.
	Continuing refinements	Improving the two-tier diagnostic test by applying it in different classes ensures that the test as a whole can be used to examine students' misconceptions. Each application to different groups helps to separate the misconceptions for each item.

Method

In this study, exploratory sequential design, one of the mixed method typologies, was used since two-tier test development steps of Treagust were used and these steps are both qualitative and quantitative (see Figure 1). Exploratory sequential design mostly used in studies like as theory and instrument development studies. In this method, researcher starts with qualitative methods to collect data including literature review, interviews, and document analysis. After analyzing qualitative data to explore the results of the first phase, quantitative method used to collect data (Creswell, 2016; Creswell, & Plano Clark, 2018).

While developing the Socioscientific Issues Concept Test, the following steps determined by Treagust (1988) were followed:

Defining the content: Energy sources, global warming, cloning and genetic engineering were determined as the content of the scale. A concept map was prepared about the determined topics and the concepts were associated (see Figure 2).

Table 3. Developed tests on socioscientific issues

Researcher	Subject of the Test	Type of the Test	Participants
Boyes, Stanisstreet & Papantoniou (1999)	Ozone layer	Five-point scale 36 items	11-16 years
Bozkurt, & Cansüngü-Koray (2002)	Greenhouse effect	Three-point scale 16 items	6-7. grades
Kılınç, Stanisstreet, & Boyes (2008)	Global warming	Five-point scale 18 items	15-16 years
Ayvacı, & Şenel-Çoruhlu (2009)	Environmental problems	5 open ended questions	4. 5. 6. 7. 8. 10.11. grades
Demirbaş, & Pektaş (2009)	Environment	17 open ended questions	6-7-8 grades
Aydın (2011)	Heredity	14 open ended questions	8. grade
Bodzin (2012)	Energy Sources	39 MC	8. grade
Bakırcı, & Çalık (2013)	Adaptation and Natural Selection	Two-tier 6 questions	8. grade
Kılınç, Boyes, & Stainsstreet (2013)	Nuclear energy	Five-point scale	6, 7, 8 and 9. grades
Artun, & Okur (2015)	Environment	3 open ended	6, 7 and 8. grades
Dawson (2015)	Climate change	Open ended Semi-structured interview (20 participants)	14-15 years
Bakırcı, & Yıldırım (2017)	Greenhouse effect	Two-tier 5 questions 13 MC	7. grade

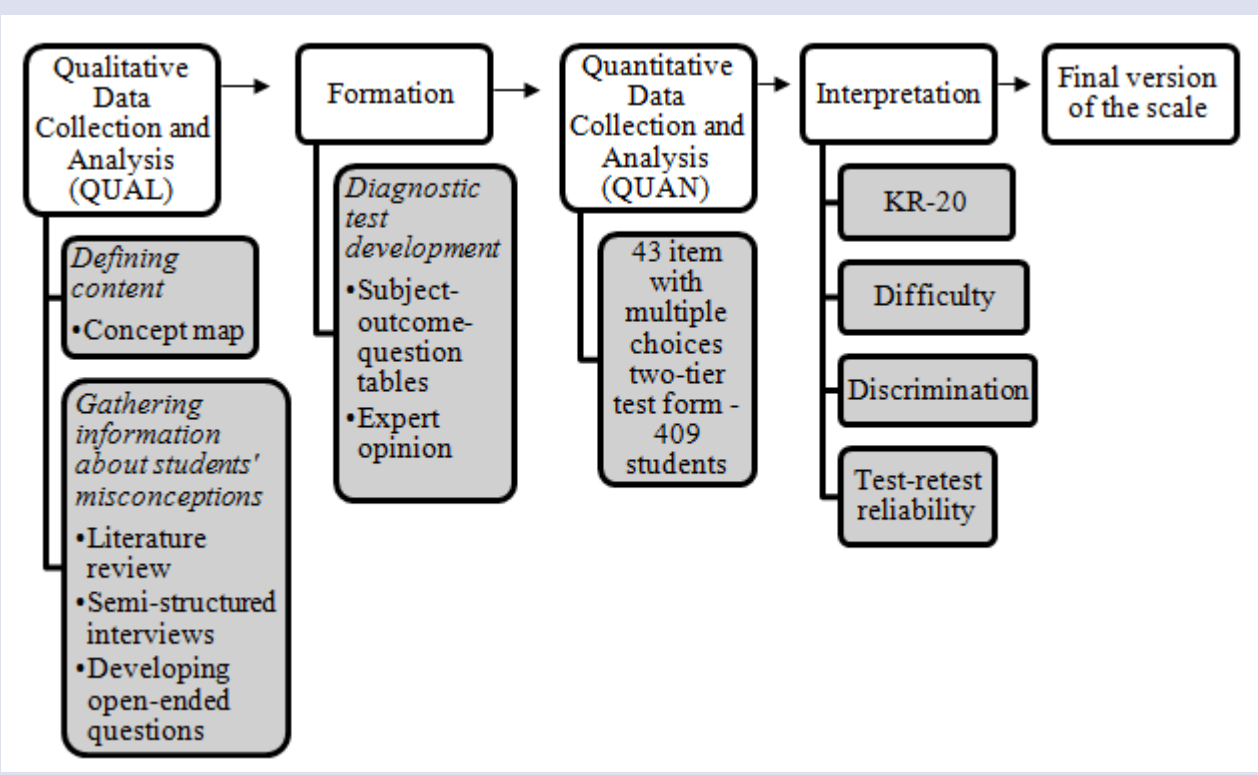


Figure 1. Exploratory sequential design of Socioscientific Issues Concept Test development

Gathering information about students' misconceptions: A literature review was conducted on the misconceptions of the determined subjects. For instance, in the study of Bodzin (2012), students evaluated natural gas and nuclear energy as renewable energy sources; Cebesoy and Karışan (2017) stated that students think that hydroelectric power plants will damage the ozone layer; Kapıcı and İlhan (2016) stated that students think nuclear power plant accidents are the biggest cause of global warming and glacial melting; Arsal (2010) found that the increase in acid rain causes the greenhouse effect, and they think that the greenhouse effect can be reduced by cleaning the beaches or by eating natural foods. Bahar & Aydın (2002), on the other hand, stated that students think that factories heat the

environment and increase human population as the cause of global warming. Bakırcı & Yıldırım (2017) determined that students think that there will be more earthquakes in the world as a result of the greenhouse effect. Apart from the literature review, semi-structured interviews were conducted with students from 8th grade to collect students' misconceptions. After content analyzing these interviews it was found that some students thought that GMO studies were carried out only on plants, that GMO was adding something harmful to a beneficial food, and that the vitamin of these products decreased. After all these procedures, open-ended questions were prepared.

Diagnostic test development: Subject-outcome-question tables were prepared and expert opinion was sought. Adjustments were made as a result of expert opinions. The questions which were prepared as open-ended, were applied to the students, turned into a two-tier multiple-choice test and analyzed.

Socioscientific Issues Concept Test is developed for science educators to use for determining students' misconceptions. This test includes questions about the topics in secondary school science curriculum. The achievements and related question numbers of the final version of the Socioscientific Issues Concept Test two-tier 30-question form are given in the table below (Table 4):

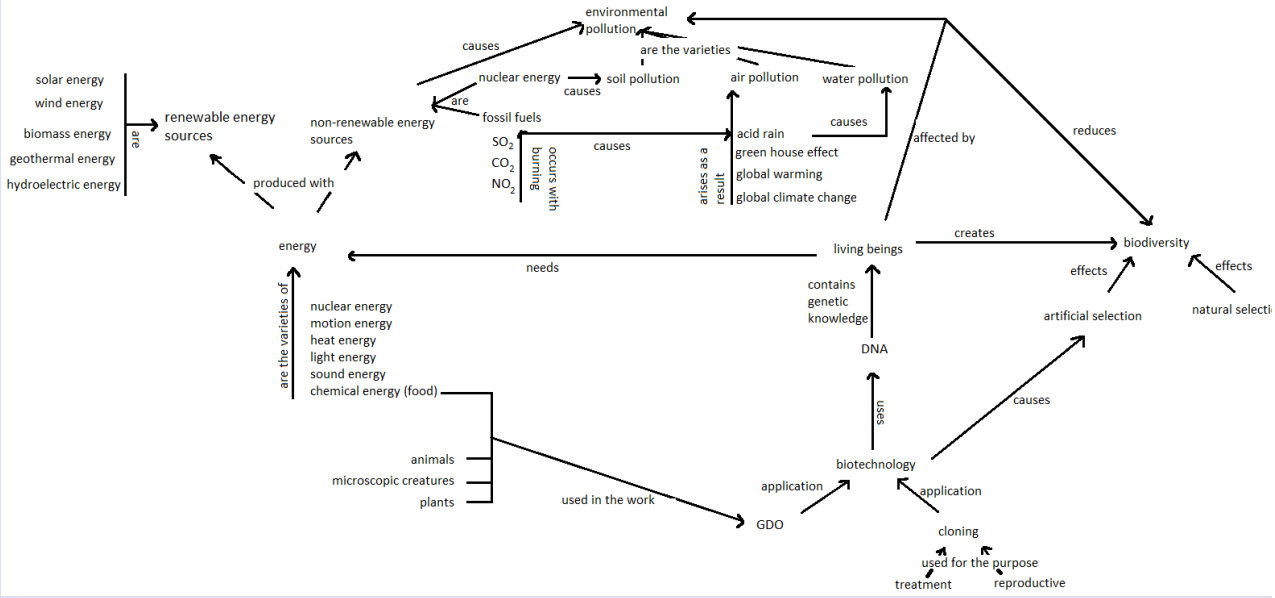


Figure 2. Socioscientific issues concept test concept map

Table 4. Socioscientific issues concept test outcome-question table (30 questions)

Outcome (MoNE, 2018a)	Question Number
F.5.6.1.1. Questions the importance of biodiversity for natural life. F.5.6.1.2. Discusses the factors that threaten biodiversity based on research data.	6, 13, 14, 15, 16, 17, 19, 25
F.5.6.2.1. Express the importance of interaction between human and environment.	6, 7, 8, 9, 13, 14, 16, 17
F.5.6.2.3. It makes inferences about environmental problems that may occur in the future as a result of human activities.	6, 7, 8, 9, 13, 14, 16, 17
F.5.6.2.4. Discusses the benefits and harm situations in human-environment interaction on examples.	6, 7, 8, 9, 13, 14, 15, 16, 17
F.6.4.4.1. Classifies fuels as solid, liquid and gaseous fuels and gives examples of commonly used fuels.	10, 18, 19
F.6.4.4.2. Discusses the effects of the use of different types of fuels for heating purposes on humans and the environment.	7, 8, 9, 10, 11, 12, 18, 19
F.8.2.5.1. Relates genetic engineering and biotechnology.	1, 2, 26, 28, 29, 30
F.8.2.5.2. Discusses the dilemmas created within the scope of biotechnological applications and the beneficial and harmful aspects of these applications for humanity.	1, 3, 27
F.8.4.4.7. Offers solutions for the prevention of acid rain.	13, 14, 15
F.8.6.3.3. Discuss the causes and possible consequences of global climate changes.	4, 5, 11, 12
F.8.7.3.3. Explains how electrical energy is produced in power plants.	20, 21, 24
F.8.7.3.4. Generates ideas about the advantages and disadvantages of power plants.	22, 23, 25

Table 5. Number of participants by gender

Gender	Open Ended	Percentage	First tier multiple choices	Percentage	Two-tier multiple choices	Percentage
Girl	25	58,13	24	58,53	232	56,72
Boy	17	39,53	17	41,46	171	41,80
Unspecified	1	2,32	0	0	6	1,46
Total	43	100	41	100	409	100

Table 6. Number of participants and grade levels by stage

Stage	Description	Number of Students	Grade Level
Forming first tier	Completely open ended	43	9
Forming second tier	First tier multiple choices, second tier open ended	41	9
Application of the two-tier form	Two-tier multiple choices	172 / 237	8 / 9
	Total	493	

Process

The Socioscientific Issues Concept Test includes questions on energy sources, global warming, genetic engineering and cloning. The literature on the mentioned subjects was searched, semi-structured interviews were held, and open-ended questions were prepared on socioscientific issues included in the MoNE Science course curriculum. Question number 7, 8, 11, 13, 14, 15, 17, 18, 20, 21, 23, 24, 25, 26, 27, 35, 36, 37, 38, 39, 40, 41, 42, 43 developed from interviews and 1, 2, 3, 4, 5, 6, 9, 10, 12, 16, 19, 22, 28, 29, 30, 31, 32, 33, 34 from literature (Çakırlar-Altuntaş, Yılmaz, & Turan, 2017; Demir,& Düzleyen, 2012; Babacan, 2017; Atabey, 2016; Namdar, 2018; Kutluca, 2012, Cansız, 2014; Saylan, 2014). Expert opinion was taken about these questions from three Science teachers, two of whom are doing PhD in Science Education, two academicians working in Science Education, and one Turkish teacher for clarity. After the arrangements made in line with the opinions received, the necessary permissions were obtained from the Istanbul Provincial Directorate of National Education and open-ended form with 43 questions was applied to the students. After multiple-choices formed for the first tier, new form was applied to the students of another class as the first-tier with multiple-choices and the second tier was open-ended. After both tiers were made multiple choices, final version of the test was applied to a large working group.

Participants

Eighth and ninth grade students participated in the development of the Socioscientific Issues Concept Test. As stated in the literature, regarding the number of samples required for the application of the scale should be at least 5 times of the number of items; It has been stated that 100 people are weak, 200 are moderate, 300 are good, 500 are very good, and 1000 are excellent (DeVellis, 2014; Şahin, & Boztunç-Öztürk, 2018; Yiğit, Bütüner, & Dertlioğlu, 2008). The participant information of the 43-item Socioscientific Issues Concept Test is given in Table 5.

Data Collection Tools

The first version of the Socioscientific Issues Concept Test developed by the researchers consisted of 43 open-ended questions. As stated in Table 6, the form consisting entirely of open-ended questions was applied to 43 ninth grade students. At the end of the application, the options of the first tier were created by using the answers from the students. The 43-question form, first tier with multiple-choices, was administered to 41 ninth grade students. Students were asked to explain the reason for the option they chose in the first tier. Based on the answers received, the second tier was also made multiple-choices. Socioscientific Issues Concept Test, both tiers of which became multiple choices, was applied to 409 students and analysis procedures were made.

Data Analysis

The Socioscientific Issues Concept Test consisting of 43 two-tier items (total of 86 questions) was administered to 409 students in total. Before the application, the students were informed about the procedure to be done. In the evaluation of the developed test, it was decided to give 1 point to those who answered both tiers correctly and 0 points to other markings (Arslan et al., 2012). Of the 409 students who answered the test, 24 students who left 10 or more of the 86 questions blank were not included in the analysis. The calculation of the KR-20 coefficient, item difficulties, item discrimination and Pearson Correlation values of the test was carried out using the data of 385 students and the SPSS 21.00 statistics program and the Excel program.

Results

Validity and Reliability

There are two explanations for the expression of reliability in studies: accuracy and consistency between the answers of individuals and the reliability coefficient shows that the scale is free from random errors (Büyüköztürk, 2014; Şeker, & Gençdoğan, 2014). Formulas such as KR-20 and Cronbach-alpha are used to examine the internal consistency between the scores obtained from the scale (Büyüköztürk, 2014). If items in a scale are expressed over two values, the Kuder-Richardson-20 formula is equivalent to the Cronbach alpha value (Büyüköztürk, 2014; DeVellis, 2014). KR-20 value varies between 0.00 and 1.00 (Şeker, & Gençdoğan, 2014).

Table 7. Socioscientific issues concept test item difficulty (pj) values (43 items)

Item	Pj	Item	Pj	Item	Pj	Item	Pj
1-1	0,605195	12-1	0,685714	23-1	0,636364	34-1	0,545455
1-2	0,555844	12-2	0,677922	23-2	0,644156	34-2	0,420779
1	0,488312	12	0,638961	23	0,58961	34	0,296104
2-1	0,6	13-1	0,428571	24-1	0,696104	35-1	0,457143
2-2	0,558442	13-2	0,831169	24-2	0,451948	35-2	0,402597
2	0,425974	13	0,374026	24	0,387013	35	0,337662
3-1	0,924675	14-1	0,657143	25-1	0,644156	36-1	0,345455
3-2	0,638961	14-2	0,672727	25-2	0,628571	36-2	0,353247
3	0,628571	14	0,631169	25	0,467532	36	0,155844
4-1	0,4	15-1	0,625974	26-1	0,441558	37-1	0,387013
4-2	0,488312	15-2	0,407792	26-2	0,464935	37-2	0,431169
4	0,374026	15	0,257143	26	0,32987	37	0,322078
5-1	0,283117	16-1	0,703896	27-1	0,490909	38-1	0,353247
5-2	0,909091	16-2	0,54026	27-2	0,384416	38-2	0,464935
5	0,25974	16	0,485714	27	0,192208	38	0,212987
6-1	0,371429	17-1	0,114286	28-1	0,535065	39-1	0,148052
6-2	0,451948	17-2	0,490909	28-2	0,542857	39-2	0,181818
6	0,27013	17	0,05974	28	0,431169	39	0,077922
7-1	0,576623	18-1	0,857143	29-1	0,680519	40-1	0,402597
7-2	0,592208	18-2	0,714286	29-2	0,553247	40-2	0,231169
7	0,537662	18	0,675325	29	0,490909	40	0,142857
8-1	0,732468	19-1	0,542857	30-1	0,581818	41-1	0,358442
8-2	0,348052	19-2	0,402597	30-2	0,542857	41-2	0,376623
8	0,244156	19	0,335065	30	0,451948	41	0,267532
9-1	0,425974	20-1	0,755844	31-1	0,623377	42-1	0,602597
9-2	0,418182	20-2	0,735065	31-2	0,314286	42-2	0,472727
9	0,392208	20	0,706494	31	0,223377	42	0,41039
10-1	0,774026	21-1	0,78961	32-1	0,61039	43-1	0,646753
10-2	0,748052	21-2	0,794805	32-2	0,768831	43-2	0,519481
10	0,724675	21	0,72987	32	0,574026	43	0,45974
11-1	0,862338	22-1	0,802597	33-1	0,246753		
11-2	0,818182	22-2	0,524675	33-2	0,283117		
11	0,787013	22	0,485714	33	0,064935	**Mean**	0,404591

Since the Developed Socioscientific Issues Concept Test was scored as 1-0, KR-20 was used in the reliability calculation. As a result of the analysis applied to the 43-item form, the KR-20 value was found to be 0.846. Test-retest method was also used in order to ensure the reliability of the scale. Pearson Correlation coefficient was calculated after final version with 30 items was applied for test-retest to find the relationship between scores taken from two applications.

In item analysis, validity outweighs all other features (Erkuş, 2014). Validity refers to how accurately the test measures the feature of the individual to be measured, and includes a number of classifications such as content validity, construct validity and face validity (Büyüköztürk, 2014). To ensure the validity of this study; expert opinion was taken, subject-outcome-question tables was created and analyzes were made. The fact that the developed test does not create rater bias, that the items are multiple choices and the number of them is sufficient increases the validity.

Item Difficulties and Item Discriminations

Item difficulty (43 items). In the scales, it is important how much of the participants know an item, and it is expressed as difficulty how much of the group succeeds in the relevant item (Erkuş, 2014). While calculating the item difficulty coefficient, the number of students who answered correctly for each tier and the net score of the item was divided by the total number of students (385). The difficulty values of the two-tier 43-question Socioscientific Subjects Concept Test according to the tiers are shown in Table 7.

Average difficulty (.40-.60) seems preferred for item difficulty (Büyüköztürk, Kılıç-Çakmak, Akgün, Karadeniz, & Demirel, 2014).

Item discrimination (43 items): While calculating the item discrimination value in the two-tier 43-question form, the upper and lower 27% of the sample (385) were calculated (103.95). The data of 104 students with the highest scores and 104 students with the lowest scores were used for discrimination calculation.

- When the scores were ranked from highest to lowest, it was seen that the 104th student had a score of 23. Again, individuals 105, 106 and 107, who received 23 points, were also included in the upper 27% group (top 27% 107 people)
- When we look at the data of the lower 27% group, which will start with the 282nd individual with 13 points, it is seen that the 281st individual has a score of 13. Therefore, the 281st individual was included in the lower 27% group (bottom 27 % 105 people)
- When calculating the discrimination, the correct answerers in the upper group are subtracted from those in the lower group and divided by half of the total number of people ([107+105]/2) for each question.

Table 8 shows the item discrimination values of the two-tier 43-question Socioscientific Issues Concept Test form.

KR-20, Item Difficulty and Discrimination (30 Items)

As a result of the analysis, items 3-5-6-8-10-17-27-33-36-38-39-40-41 were removed from the scale. For the final version, the KR-20 value was found to be 0.85 as shown in Table 9. The general acceptance for the KR-20 coefficient is at least 0.70 (Büyüköztürk, 2014; Şeker, & Gençdoğan, 2014). It is seen that the developed scale is reliable.

The difficulty and discrimination values of the remaining two-tier, 30-question Socioscientific Issues Concept Test form according to the tiers are shown in Table 10.

While calculating the item difficulty coefficient, the number of students who answered correctly for each tier and the net score of the item was divided by the total number of students (385). The average difficulty value of the test was found to be 0.46987. Reference values for Item Difficulty: average difficulty (.40-.60) is preferred (Büyüköztürk et al., 2014). It is seen that the Socioscientific Issues Concept Test is appropriate in terms of item difficulty.

The upper and lower 27% of this sample (385) were calculated (103.95) for the item discrimination calculation of the form consisting of 30 two-tier questions. It has been determined that the data of 104 students with the highest scores and 104 students with the lowest scores will be used for the discrimination calculation. Both tiers of the scale and net scores were evaluated separately.

Table 8. Socioscientific issues concept test item discrimination (Rjx) values (43 items)

Item	Rjx	Item	Rjx	Item	Rjx	Item	Rjx
1-1	0,349057	12-1	0,443396	23-1	0,556604	34-1	0,367925
1-2	0,518868	12-2	0,481132	23-2	0,5	34-2	0,358491
1	0,490566	12	0,528302	23	0,650943	34	0,396226
2-1	0,377358	13-1	0,301887	24-1	0,320755	35-1	0,339623
2-2	0,283019	13-2	0,396226	24-2	0,226415	35-2	0,396226
2	0,311321	13	0,377358	24	0,349057	35	0,5
3-1	0,160377	14-1	0,273585	25-1	0,603774	36-1	0,320755
3-2	0,188679	14-2	0,292453	25-2	0,575472	36-2	0,245283
3	0,198113	14	0,330189	25	0,716981	36	0,254717
4-1	0,349057	15-1	0,40566	26-1	0,320755	37-1	0,292453
4-2	0,358491	15-2	0,320755	26-2	0,481132	37-2	0,292453
4	0,396226	15	0,396226	26	0,462264	37	0,462264
5-1	0,254717	16-1	0,575472	27-1	0,301887	38-1	0,169811
5-2	0,150943	16-2	0,45283	27-2	0,311321	38-2	0,188679
5	0,245283	16	0,59434	27	0,235849	38	0,235849
6-1	0,235849	17-1	-0,12264	28-1	0,575472	39-1	-0,04717
6-2	0,292453	17-2	0,132075	28-2	0,45283	39-2	-0,08491
6	0,264151	17	0	28	0,669811	39	0,066038
7-1	0,481132	18-1	0,443396	29-1	0,566038	40-1	0,122642
7-2	0,396226	18-2	0,613208	29-2	0,45283	40-2	-0,10377
7	0,490566	18	0,735849	29	0,584906	40	0,056604
8-1	0,396226	19-1	0,283019	30-1	0,59434	41-1	0,103774
8-2	0,103774	19-2	0,273585	30-2	0,566038	41-2	0,188679
8	0,226415	19	0,320755	30	0,726415	41	0,216981
9-1	0,330189	20-1	0,424528	31-1	0,518868	42-1	0,396226
9-2	0,339623	20-2	0,490566	31-2	0,254717	42-2	0,509434
9	0,40566	20	0,575472	31	0,339623	42	0,59434
10-1	0,396226	21-1	0,537736	32-1	0,679245	43-1	0,528302
10-2	0,386792	21-2	0,528302	32-2	0,45283	43-2	0,433962
10	0,462264	21	0,679245	32	0,726415	43	0,566038
11-1	0,349057	22-1	0,45283	33-1	0,132075		
11-2	0,443396	22-2	0,433962	33-2	0,09434		
11	0,518868	22	0,509434	33	0,132075	**Mean**	0,418605

Table 9. Socioscientific issues concept test 30 item reliability analysis

	KR-20	Number of Item
1. tier	0,809	30
2. tier	0,780	30
Net score	0,850	30

Table 10. Socioscientific issues concept test item difficulty (pj) and discrimination (rjx) values (30 items)

Item	Pj	Rjx	Item	Pj	Rjx	Item	Pj	Rjx
1-1	0,605195	0,258621	16-1	0,703896	0,405172	28-1	0,535065	0,491379
1-2	0,555844	0,56621	16-2	0,54026	0,511416	28-2	0,542857	0,584475
1	0,488312	0,446512	16	0,485714	0,576744	28	0,431169	0,67907
2-1	0,6	0,258621	18-1	0,857143	0,25	29-1	0,680519	0,431034
2-2	0,558442	0,30137	18-2	0,714286	0,639269	29-2	0,553247	0,547945
2	0,425974	0,269767	18	0,675325	0,67907	29	0,490909	0,595349
4-1	0,4	0,293103	19-1	0,542857	0,224138	30-1	0,581818	0,517241
4-2	0,488312	0,465753	19-2	0,402597	0,383562	30-2	0,542857	0,60274
4	0,374026	0,4	19	0,335065	0,344186	30	0,451948	0,744186
7-1	0,576623	0,396552	20-1	0,755844	0,267241	31-1	0,623377	0,405172
7-2	0,592208	0,493151	20-2	0,735065	0,493151	31-2	0,314286	0,3379
7	0,537662	0,465116	20	0,706494	0,511628	31	0,223377	0,344186
9-1	0,425974	0,25	21-1	0,78961	0,344828	32-1	0,61039	0,517241
9-2	0,418182	0,392694	21-2	0,794805	0,557078	32-2	0,768831	0,484018
9	0,392208	0,4	21	0,72987	0,669767	32	0,574026	0,688372
11-1	0,862338	0,155172	22-1	0,802597	0,258621	34-1	0,545455	0,25
11-2	0,818182	0,538813	22-2	0,524675	0,511416	34-2	0,420779	0,465753
11	0,787013	0,474419	22	0,485714	0,511628	34	0,296104	0,427907
12-1	0,685714	0,267241	23-1	0,636364	0,387931	35-1	0,457143	0,267241
12-2	0,677922	0,547945	23-2	0,644156	0,520548	35-2	0,402597	0,465753
12	0,638961	0,530233	23	0,58961	0,613953	35	0,337662	0,502326
13-1	0,428571	0,198276	24-1	0,696104	0,206897	37-1	0,387013	0,267241
13-2	0,831169	0,474886	24-2	0,451948	0,319635	37-2	0,431169	0,374429
13	0,374026	0,316279	24	0,387013	0,344186	37	0,322078	0,493023
14-1	0,657143	0,155172	25-1	0,644156	0,474138	42-1	0,602597	0,336207
14-2	0,672727	0,392694	25-2	0,628571	0,584475	42-2	0,472727	0,593607
14	0,631169	0,27907	25	0,467532	0,660465	42	0,41039	0,595349
15-1	0,625974	0,284483	26-1	0,441558	0,293103	43-1	0,646753	0,405172
15-2	0,407792	0,374429	26-2	0,464935	0,520548	43-2	0,519481	0,502283
15	0,257143	0,362791	26	0,32987	0,437209	43	0,45974	0,55814
						First	0,613593	0,317241
						Second	0,56303	0,484932
						Net	0,46987	0,497364

Table 11. Test-retest Pearson Correlation Coefficient results

		Test Retest 1	Test Retest 2
Test Retest 1	Pearson Correlation	1	,776**
	Sig. (2-tailed)		,000
	N	55	55
Test Retest 2	Pearson Correlation	,776**	1
	Sig. (2-tailed)	,000	
	N	55	55

The scores obtained from the first tier of the questions in the scale are ranked from highest to lowest. When the scores were ranked from highest to lowest, it was seen that the 104th student had a score of 23. Again, individuals who scored 23 and ranked 105-106 were also included in the top 27% (top 27% 106 people). It was observed that the lowest 104 students had 16 points for the determination of the subgroup and 16 points were obtained up to the 126th place (lower 27% 126 people). While calculating the discrimination for each question, the number of correct students in the lower group was subtracted from the number of correct students in the upper group and divided by half of the total number of students in both groups ((106+126)/2). The mean values of the results obtained for the mean discrimination score of the first tiers were taken.

The scores obtained from the second tier of the questions in the scale are ranked from highest to lowest. The score of the 104th student in the upper group was 21. Since the scores obtained are 21 until the 115th ranked student, these students are also included in the top 27% group (top 27% 115 people). The score of the 104th student in the subgroup is 13 (bottom 27% 104 people). While calculating the discrimination for each question, the number of correct students in the lower group was subtracted from the number of correct students in the upper group and divided by half of the total number of students in both groups ((115+104)/2). The mean values of the results obtained for the mean discrimination score of the second tiers were taken.

For calculating the discrimination of the net scores obtained from the scale, the scores are ranked from highest to lowest. According to the 385 data included in the analysis, the score of the 104th student in the upper group was 19. Since the 105th and 106th students have a score of 19, these students are also included in the top 27% group (top 27 106 people). Since it was seen that the 104th student in the subgroup had a score of 10 and this score continued until the 109th student, these students were also included in the subgroup (Lower 27, 109 people). While calculating the discrimination for each question, the number of correct students in the lower group was subtracted from the number of correct students in the upper group and divided by half of the total number of students in both groups ((106+109)/2). The mean values of the results obtained for the mean discrimination score of the scale were taken.

Reference values for item discrimination are as follows (Büyüköztürk et al., 2014, p. 123):

- Very good if it is over 0.40,
- Good if it is between .30-.39,
- Medium between .20-.29,
- If it is below .20, it is bad

The mean discrimination of the Socioscientific Issues Concept Test was calculated as 0.497364 and this result seems to be appropriate. Another analysis performed to ensure the reliability of the scale is to calculate the Pearson Correlation coefficient. Test-retest was applied to 55 eighth grade students with 3-4 week intervals. The results of the analysis performed to calculate the Pearson Correlation coefficient for the obtained results are shown in Table 11.

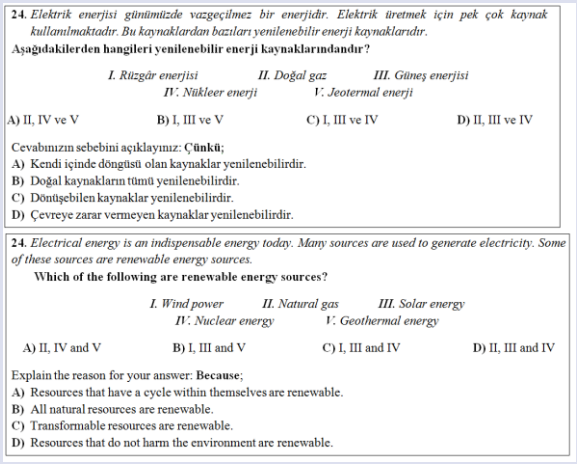
According to the results of the Pearson correlation analysis, the correlation coefficient between the test-retest results was found to be 0.776. Since this value found is greater than 0.70, it shows that the scores students get from test-retest applications are related (Tavşancıl, 2014). Büyüköztürk et al. (2014) stated that as the correlation coefficient gets closer to 1, the stability of the test increases. The results show that the two-tier Socioscientific Issues Concept Test is a stable and reliable scale.

Conclusions, Discussion and Suggestions

This study aims to develop a useful, valid and reliable two-tier socioscientific issues concept test. This study was conducted in an exploratory mixed design in which quantitative and qualitative methods were used together (Creswell, 2016). First, the topics were determined, and after the literature review, an open-ended item pool was created. After the expert opinion, it was applied to the students and the options of the tiers were created. There are 4 options in both tiers of the test. In the first tier, students answer the question and in the second tier they explain the reason for their answer. In the scoring of the test, if both tiers are answered correctly, 1 point is given and 0 points are given to other markings. The developed Socioscientific Issues Concept Test consists of 30 two-tier multiple-choice questions on energy sources, global warming, genetic engineering and cloning.

The KR-20 value of this test's final version is 0.85, Pearson correlation coefficient 0.776, the mean difficulty value is 0.46987, and the mean discrimination value is 0.497364. Şeker and Gençdoğan (2014) state that the general acceptance for the KR-20 coefficient and for the Pearson correlation coefficient, Tavşancıl (2014) states that the general acceptance is at least 0.70. is at least 0.70. It is seen that the KR-20 and Pearson correlation coefficient values of this test are appropriate. Büyüköztürk et al. (2014) stated that discrimination value is very good if it is over 0.40, good if it is between .30-.39, medium between .20-.29, if it is below .20, it is bad; and also stated that the average difficulty (.40-.60) should be preferred for item difficulty. It is seen that the developed test is suitable in terms of discrimination and difficulty. Some examples for the test items in both Turkish and English are shown in Figure 3 and 4.

In order to determine misconceptions, some researchers developed instruments with open ended questions (Ayvacı, & Şenel-Çoruhlu, 2009; Demirbaş, & Pektaş, 2009; Aydın, 2011; Artun, & Okur, 2015; Dawson, 2015). There are instruments with multiple choices, two or three tiers also (Arslan et al., 2012; Bodzin, 2012; Bakırcı, & Çalık, 2013; Bakırcı, & Yıldırım, 2017). We can also see that some researchers preferred scales with 3-5 point scales (Boyes, Stanisstreet, & Papantoniou, 1999; Bozkurt, & Cansüngü-Koray, 2002; Kılınç, Stanisstreet, & Boyes, 2008; Kılınç, Boyes, & Stainsstreet, 2013) and these instruments are about a specific topic. Arslan et al., (2012) stated that rather than trying to find a reason for the selected answer it is easier to answer just one multiple-choice test item correctly. Therefore, two-tier tests are decreasing luck success and giving that opportunity to evaluate the real knowledge about selected subject. For this reason, a two-tier instrument developed for determining misconceptions about socioscientific issues with this study. There are two ways of developing multiple-choice items as: a) asking students open-ended questions and generating choices from their answers b) asking test developers and advisors for a list of alternatives and with both tiers as multiple-choice, it provides the opportunity to apply to more people and increases reliability (Briggs et al., 2006). For future researches, to conduct studies with a larger number of participants in order to increase generalizability can be given as a suggestion. Researchers can make comparisons by applying test at different grade levels.



24. *Elektrik enerjisi günümüzde vazgeçilmez bir enerjidir. Elektrik üretmek için pek çok kaynak kullanılmaktadır. Bu kaynaklardan bazıları yenilenebilir enerji kaynaklarıdır.*
Aşağıdakilerden hangileri yenilenebilir enerji kaynaklarındandır?

I. Rüzgâr enerjisi *II. Doğal gaz* *III. Güneş enerjisi*
IV. Nükleer enerji *V. Jeotermal enerji*

A) II, IV ve V B) I, III ve V C) I, III ve IV D) II, III ve IV

Cevabınızın sebebini açıklayınız: **Çünkü;**
A) Kendi içinde döngüsü olan kaynaklar yenilenebilirdir.
B) Doğal kaynakların tümü yenilenebilirdir.
C) Dönüşebilen kaynaklar yenilenebilirdir.
D) Çevreye zarar vermeyen kaynaklar yenilenebilirdir.

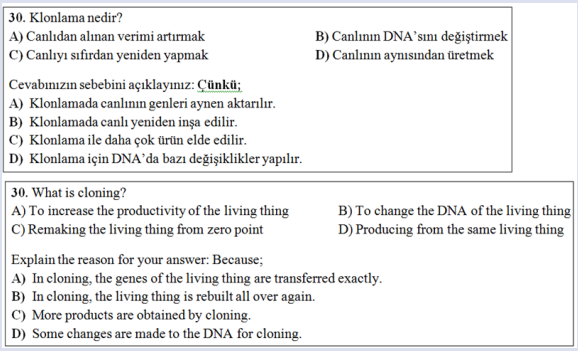
24. *Electrical energy is an indispensable energy today. Many sources are used to generate electricity. Some of these sources are renewable energy sources.*
Which of the following are renewable energy sources?

I. Wind power *II. Natural gas* *III. Solar energy*
IV. Nuclear energy *V. Geothermal energy*

A) II, IV and V B) I, III and V C) I, III and IV D) II, III and IV

Explain the reason for your answer: **Because;**
A) Resources that have a cycle within themselves are renewable.
B) All natural resources are renewable.
C) Transformable resources are renewable.
D) Resources that do not harm the environment are renewable.

Figure 3. Item 24 from final version of the test



30. Klonlama nedir?
A) Canlıdan alınan verimi artırmak B) Canlının DNA'sını değiştirmek
C) Canlıyı sıfırdan yeniden yapmak D) Canlının aynısından üretmek

Cevabınızın sebebini açıklayınız: **Çünkü;**
A) Klonlamada canlının genleri aynen aktarılır.
B) Klonlamada canlı yeniden inşa edilir.
C) Klonlama ile daha çok ürün elde edilir.
D) Klonlama için DNA'da bazı değişiklikler yapılır.

30. What is cloning?
A) To increase the productivity of the living thing B) To change the DNA of the living thing
C) Remaking the living thing from zero point D) Producing from the same living thing

Explain the reason for your answer: Because;
A) In cloning, the genes of the living thing are transferred exactly.
B) In cloning, the living thing is rebuilt all over again.
C) More products are obtained by cloning.
D) Some changes are made to the DNA for cloning.

Figure 4. Item 30 from final version of the test

Note

This study is adapted from first author's PhD thesis.

References

Anderson, D. L., Fisher, K. M. & Norman, G. J. (2002). Development and Evaluation of the Conceptual Inventory of Natural Selection. *Journal of Research in Science Teaching, 39*(10), 952-978.

Arsal, Z. (2010). İlköğretim öğretmen adaylarının sera etkisi ile ilgili kavram yanılgıları. *İlköğretim Online, 9*(1), 229-240.

Arslan, H. A., Çigdemoğlu, C., & Moseley, C. (2012). A three-tier diagnostic test to assess pre-service teachers' misconceptions about global warming, greenhouse effect, ozone layer depletion, and acid rain, *International Journal of Science Education, 34*(11), 1667-1686, DOI: 10.1080/09500693.2012.680618.

Artun, H., & Okur, M. (2015). Ortaokul öğrencilerinin çevre kavramına yönelik bilgi ve çevreyi anlama düzeylerinin belirlenmesi. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, 24, 277-293.

Atabey, N. (2016). *Sosyobilimsel konu temelli bir ünitenin geliştirilmesi: 7. sınıf öğrencilerinin konu alan bilgisi ve argümantasyon nitelikleri*. (Yayımlanmamış Doktora Tezi), Muğla Sıtkı Koçman Üniversitesi, Muğla.

Aubrecht, K.B. (2018). Teaching relevant climate change topics in undergraduate chemistry courses: Motivations, student misconceptions, and resources. *Current Opinion in Green and Sustainable Chemistry*, 13, 44-49.

Aydın, G. (2011). *Öğrencilerin "hücre bölünmesi ve kalıtım" konularındaki kavram yanılgılarının giderilmesinde ve zihinsel modelleri üzerinde yapılandırmacı yaklaşımın etkisi* (Yayımlanmamış Doktora Tezi). Dokuz Eylül Üniversitesi, İzmir.

Ayvacı, H. Ş. & Şenel-Çoruhlu, T. (2009). Öğrencilerin küresel çevre sorunlarına bakışları ve kavram yanılgılarının belirlenmesine yönelik gelişimsel bir araştırma. *Hasan Ali Yücel Eğitim Fakültesi Dergisi,* 12, 11-25.

Babacan, M. A. (2017). *Sosyobilimsel konulardaki etkinliklerin yedinci sınıf öğrencilerinin eleştirel düşünme becerilerine etkisi*. (Yayımlanmamış Yüksek Lisans Tezi), Ömer Halisdemir Üniversitesi, Niğde.

Bahar, M., & Aydın, F. (2002). Sınıf öğretmenliği öğrencilerinin sera gazları ve global ısınma ile ilgili anlama düzeyleri ve hatalı kavramlar. V. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi, Ankara.

Bakırcı, H., & Çalık, M. (2013). Adaptasyon ve doğal seçilim konusunda geliştirilen rehber materyallerin sekizinci sınıf öğrencilerinin alternatif kavramlarının giderilmesine etkisi. *Eğitim ve Bilim, 38*(168), 215-229.

Bakırcı, H., & Yıldırım, İ. (2017). Ortak bilgi yapılandırma modelinin sera etkisi konusunda öğrencilerin kavramsal anlamalarına ve bilginin kalıcılığına etkisi. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi (KEFAD)*, 18, 45-63.

Bayram, K. (2021). Kitap değerlendirmesi: Kuramdan uygulamaya sosyobilimsel konular, *Uluslararası Sosyal Bilimler Eğitimi Dergisi, 7*(1), 200-218. DOI: 10.47615/issej.915640

Bodzin, A. (2012). Investigating Urban Eighth-Grade Students' Knowledge of Energy Resources. *International Journal of Science Education, 34*(8), 1255-1275, DOI:10.1080/09500693.2012.661483.

Boyes, E., Stanisstreet, M. & V. S. Papantoniou (1999). The Ideas of Greek High School Students about the "Ozone Layer". *Learning*, 724-737.

Bozkurt, O., & Cansüngü-Koray, Ö. (2002). İlköğretim öğrencilerinin çevre eğitiminde sera etkisi ile ilgili kavram yanılgıları. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 23, 67-73.

Briggs, D. C., Alonzo, A. C., Schwab, C., & Wilson, M. (2006). Diagnostic Assessmentwith ordered multiple-choice items. *Educational Assessment, 11*(1), 33–63.

Büyüköztürk, Ş. (2014). *Sosyal Bilimler İçin Veri Analizi El Kitabı*. Ankara: Pegem Akademi Yayıncılık.

Büyüköztürk, Ş., Kılıç-Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2014). *Bilimsel Araştırma Yöntemleri*. Ankara: Pegem Akademi Yayıncılık.

Caleon, I. S., & Subramaniam, R. (2010). Do students know what they know and what they don't know? Using a four-tier diagnostic test to assess the nature of students' alternative conceptions. *Research in Science Education*, 40, 313-337. DOI 10.1007/s11165-009-9122-4.

Cansız, N. (2014). *Developing preservice science teachers' socioscientific reasoning through socioscientific issues-focused course*. (Yayımlanmamış Doktora Tezi), Orta Doğu Teknik Üniversitesi, Ankara.

Cebesoy, Ü. B., & Karışan, D. (2017). Fen bilgisi öğretmen adaylarının yenilenebilir enerji kaynaklarına yönelik bilgilerinin, tutumlarının ve bu kaynakların öğretimi konusundaki öz-yeterlik algılarının incelenmesi. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi, 14*(1), 1377-1415.

Chandrasegaran, A. L., Treagust, D. F.,& Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. *Chemistry Education Research and Practice, 8*(3), 293-307.

Creswell, J. W. (2016). *Araştırma Deseni Nitel, Nicel ve Karma Yöntem Yaklaşımları.* İkinci Baskı. Ankara: Eğiten Kitap.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research*. Thousand Oaks, CA: SAGE.

Çakırlar-Altuntaş, E., Yılmaz, M., & Turan, S. L. (2017). Biyoloji Öğretmen Adaylarının Sosyobilimsel Bir Konudaki Eleştirel Düşünmelerinin Empati Açısından İncelenmesi. *Bartın Üniversitesi Eğitim Fakültesi Dergisi, 6*(3), 915-931,

Dawson, V. (2015). Western Australian high school students' understandings about the socioscientific issue of climate change, *International Journal of Science Education, 37*(7), 1024-1043, DOI: 10.1080/09500693.2015.1015181.

Demir, B., & Düzleyen, E. (2012). *İlköğretim 8. sınıf öğrencilerinin GDO bilgi düzeylerinin incelenmesi*. X. Ulusal Fen Bilimleri ve Matemnatik Eğitimi Kongresi, 27-30 Haziran, Niğde.

Demirbaş, M., & Pektaş, H. M. (2009). İlköğretim öğrencilerinin çevre sorunu ile ilişkili temel kavramları gerçekleştirme düzeyleri. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi (EFMED), 3*(2), 195-211.

Demirci, N., Efe, S. (2007). İlköğretim öğrencilerinin ses konusundaki kavram yanılgılarının belirlenmesi. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi (EFMED), 1*(1), 23-56.

DeVellis, R. F. (2014). *Ölçek Geliştirme Kuram ve Uygulamalar*. Ankara: Nobel Akademik Yayıncılık.

Erkuş, A. (2014). *Psikolojide Ölçme ve Ölçek Geliştirme-1: Temel Kavramlar ve İşlemler*. Ankara: Pegem Akademi Yayıncılık.

Hodson, D. (1994). Seeking Directions for Change: the personalisation and politicisation of science education. *Curriculum Studies, 2*(1), 71-98, DOI:10.1080/0965975940020104

Jennings, B. D. (1982). *An Annotated Bibliography of Teaching Bioethics in the Public Secondary School,* Indiana University at South Bend, 1-50.

Kapıcı, H. Ö., & İlhan, G. O. (2016). Pre-service teachers' attitudes toward socio-scientific issues and their views about nuclear power plants. *Journal of Baltic Science Education, 15*(5), 642-652.

Kaplan, E.M. (2019). *Ortaokul öğrencilerinin nükleer enerji hakkındaki kavramsal yapıları*. (Yayımlanmamış Yüksek Lisans Tezi). Necmettin Erbakan Üniversitesi, Konya.

Karataş, F. Ö., Köse, S., & Coştu, B. (2003). Öğrenci yanılgılarını ve anlama düzeylerini belirlemede kullanılan iki aşamalı testler. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi, 1*(13), 54-69.

Kenan, O., & Özmen, H. (2014). Maddenin tanecikli yapısına yönelik iki aşamalı çoktan seçmeli bir testin geliştirilmesi ve uygulanması. *Eğitim ve Öğretim Araştırmaları Dergisi, 3*(3), 371-378.

Kılınç, A., Boyes, E., & Stainsstreet, M. (2013). Exploring students' ideas about risks and benefits of nuclear power using risk perception theories. *Journal of Science Education and Technology, 22*(3), 252-266.

Kılınç, A., Stanisstreet, M., & Boyes, E. (2008). Turkish Students' Ideas about Global Warming. *International Journal of Environmental & Science Education, 3*(2), 89-98.

Kutluca, A. Y. (2012). *Fen ve teknoloji öğretmen adaylarının klonlamaya ilişkin bilimsel ve sosyobilimsel argümantasyon kalitelerinin alan bilgisi yönünden incelenmesi.* (Yayımlanmamış Yüksek Lisans Tezi), Abant İzzet Baysal Üniversitesi, Bolu.

Lay, Y-F., Khoo, C-H., Treagust, D. F., & Chandrasegaran, A. L. (2013). Assessing secondary school students' understanding of the relevance of energy in their daily lives. *International Journal of Environmenal & Science Education, 8*(1), 199-

Ministry of National Education (Milli Eğitim Bakanlığı). (2018a). *Fen Bilimleri Dersi Öğretim Programı (İlkokul ve Ortaokul 3, 4, 5, 6, 7 ve 8. Sınıflar)*. Ankara.

Ministry of National Education (Milli Eğitim Bakanlığı). (2018b). *Bilim Uygulamaları Dersi Öğretim Programı (Ortaokul ve İmam Hatip Ortaokulu 5, 6, 7 ve 8. Sınıflar)*. Ankara.

Namdar, B. (2018). Teaching global climate change to pre-service middle school teachers through inquiry activities. *Research in Science & Technological Education*, DOI: 10.1080/02635143.2017.1420643.

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching, 41* (5), 513-536.

Saylan, A. (2014). *Relationships among pre-service science teachers' epistemological beliefs, knowledge level and trustworthiness on information sources: climate change, nuclear energy, and organ donation and transplantation*. (Yayımlanmamış Yüksek Lisans Tezi), Orta Doğu Teknik Üniversitesi, Ankara.

Sıbıç, O., Akçay, B., & Arık, M. (2020). Review of two-tier tests in the studies: creating a new pathway for development of two-tier tests. *International Journal of Contemporary Educational Research, 7*(2), 81-98. DOI: https://doi.org/10.33200/ijcer.747981

Şahin, M. G., & Boztunç-Öztürk, N. (2018). Eğitim alanında ölçek geliştirme süreci: Bir içerik analizi çalışması. *Kastamonu Üniversitesi Kastamonu Eğitim Dergisi, 26*(1), 191-199.

Şeker, H., & Gençdoğan, B. (2014). *Psikolojide ve Eğitimde Ölçme Aracı Geliştirme*. Ankara: Nobel Akademik Yayıncılık.

Tavşancıl, E. (2014). *Tutumların Ölçülmesi ve SPSS ile Veri Analizi*. Ankara: Nobel Akademik Yayıncılık.

Topaloğlu, M.Y., &. Kıyıcı, F. B. (2018). Okul dışı öğrenme ortamlarında yürütülen etkinliklerin öğrencilerin sosyobilimsel konulara ilişkin görüşlerine etkisi: Organ bağışı ve GDO. *E-Uluslararası Eğitim Araştırmaları Dergisi, 9*(1), 36-50

Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education, 10*(2), 159-169, DOI: 10.1080/0950069880100204.

Treagust, D. F., & Haslam, F. (1986). Evaluating secondary students' misconceptions of photosynthesis and respiration in plants using a two-tier diagnostic instrument. *59th Annual Meetingof the National Association for Research in Science Teaching*, 2-25.

Yaghmour, K. S., Obaidat, L. T., & Hamadneh, Q. M. (2016). The level of diagnostic tests' preparation skills among the teachers of the first three elementary grades' teachers at the directorate of education of bani kinana district. *Journal of Education and Practice, 7*(9), 155-164.

Yiğit, N., Bütüner, S. Ö., & Dertlioğlu, K. (2008). Öğretim amaçlı örütbağ sitesi değerlendirme ölçeği geliştirme *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi (EFMED), 2*(2), 38-51.

Zeidler, D. L. & Nichols, B. H. (2009). Socioscientific Issues: Theory and Practice. *Journal of Elementary Science Education, 21*(2), 49-58.

Araştırmanın Etik İzinleri

Yapılan bu çalışmada "Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi" kapsamında uyulması belirtilen tüm kurallara uyulmuştur. Yönergenin ikinci bölümü olan "Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler" başlığı altında belirtilen eylemlerden hiçbiri gerçekleştirilmemiştir.

Etik kurul izin bilgileri

Etik değerlendirmeyi yapan kurul adı = İstanbul Üniversitesi-Cerrahpaşa, Sosyal ve Beşeri Bilimler Araştırmaları Etik Kurulu Başkanlığı

Etik değerlendirme kararının tarihi= 07/01/2020

Etik değerlendirme belgesi sayı numarası= 4372