

Research Article/Araştırma Makalesi

Three Tier Concept Diagnostic Test Development and Validity Study: Let's Travel and Learn about the Living World Unit of Fifth Grade

Evşen AYMEN PEKER *¹  Erol TAŞ² 

¹ Tevfik İleri İmam Hatip Middle School, Samsun, Turkey, evsenaymen@hotmail.com

² Ordu University, Faculty of Education, Ordu, Turkey, eroltas@odu.edu.tr

* Corresponding Author: evsenaymen@hotmail.com

Article Info

Received: 28 August 2019

Accepted: 4 October 2019

Keywords: Misconceptions, three tier test, item analysis, reliability, validity

DOI: 10.18009/jcer.602223

Publication Language: Turkish

Abstract

The purpose of this study is to develop a valid and reliable three-tier concept diagnostic test which for the Let's Travel and Learn about the Living World unit of fifth grade Science lesson and which that provide the students with the level of success and their misconceptions related to the unit. The developed test was applied to 210 5th grade students attending at three secondary schools in the center of Samsun. A validity and reliability study were conducted for a three-tier concept diagnostic test consisting of a total of 23 items. The reliability of the concept diagnostic test was evaluated as 0.80 with the help of KR-20 reliability estimation method for the first stage questions. When the three items were evaluated together, the difficulty and discrimination values of the test were calculated as p_j : 0.47 and r_{jx} : 0.52. These data show that the developed test has a structure with reliable, medium difficulty and very good substances.



To cite this article: Aymen-Peker, E. & Taş, E. (2019). Investigation of in-school factors affecting distributed leadership practices. *Journal of Computer and Education Research*, 7(14), 515-539. DOI: 10.18009/jcer.602223

Üç Aşamalı Kavram Tanı Testi Geliştirme ve Geçerlik Çalışması: 5. Sınıf Canlılar Dünyasını Gezelim Tanıyalım Ünitesi

Makale Bilgisi

Geliş: 28 Ağustos 2019

Kabul: 4 Ekim 2019

Anahtar kelimeler: Kavram yanılgısı, üç aşamalı test, madde analizi, geçerlik, güvenilirlik

DOI: 10.18009/jcer.602223

Yayın Dili: Türkçe

Öz

Bu çalışmanın amacı, 5. sınıf fen bilimleri dersi "Canlılar Dünyasını Gezelim ve Tanıyalım" ünitesine yönelik; öğrencilerin ünite ile ilgili başarı düzeyleri ve sahip oldukları kavram yanılgılarının tespit edilmesini sağlayacak geçerli ve güvenilir bir üç aşamalı kavram tanı testi geliştirmektir. Geliştirilen test, Samsun il merkezinde bulunan üç ortaokulda, beşinci sınıfta öğrenim gören 210 öğrenciye uygulanmıştır. Toplam 23 maddeden oluşan üç aşamalı kavram tanı testi için geçerlilik güvenilirlik çalışması yapılmıştır. Kavram tanı testinin güvenilirliği birinci aşama sorularının analizinden KR-20 formülü ile hesaplanmış ve 0,80 olarak bulunmuştur. Üç madde birlikte değerlendirildiğinde testin güçlük ve ayırt edicilik değerleri sırasıyla p_j : 0.47 ve r_{jx} : 0.52 olarak hesaplanmıştır. Bu veriler geliştirilen testin güvenilir, orta güçlükte ve çok iyi maddelerden oluşan bir yapıya sahip olduğunu göstermektedir.

Summary

Three Tier Concept Diagnostic Test Development and Validity Study: Let's Travel and Learn about the Living World Unit of Fifth Grade

Introduction

Concept is the name given to the groups, when we group things, events, people and thoughts according to their similarities (Yağbasan & Gülçiçek, 2003). A child begins school life with many concepts that he has learned in the family environment. The child continues his journey of learning concept by encountering many new concepts during his school education. The researches reveal that preliminary knowledge and concepts created outside the school have a huge impact on learning (Novak, 1998).

One of the main objectives of the Science course, in which many concrete and abstract concepts are learned. (Ministry of Education, 2013; 2018), is to enable students to learn the concepts fully and correctly (Köse & Uşak, 2006). While student learn concepts, depending on both the preliminary knowledge and the experiences that student has in the learning process, he can misconfigure the knowledge scientifically. In such a case, it can be said that the student has misconceptions. However, it can be decided that the student has misconceptions, if the student explains his misconceptions and he is sure of his response (Karataş, Köse & Coştu, 2003).

When the literature on science education is examined, it is remarkable that students have misconceptions about different learning areas. It is observed that students have misconceptions in the researches about the classification of living things and environmental problems. Researchers such as Allen (2010), Çetinkaya (2010), Mutlu and Tokcan (2012), Özyılmaz-Akamca (2008), Saka, Ayas and Enginar (2002), Summers, Kruger, Childs and Mant (2000), Taş, Aymen Peker and Çetinkaya (2014), Tekkaya, Çapa and Yılmaz (2000), Yen, Yao and Mintzes (2007) have identified students' misconceptions about related topics.

In order to determine whether the student has misconceptions, tools such as analogy, concept map, concept caricature, prediction-observation-explanation, interview, drawing, word association, V diagram and misconception diagnostic tests can be used.

It was aimed to develop a valid and reliable measurement tool to determine the achievement levels and misconceptions of the 5th grade students about the "Let's Travel and Learn About The Living World" unit by taking into consideration the 2013 curriculum gains in this study.

Method

In the study, which was used the Survey method, validity and reliability studies were conducted for the development of "Concept Diagnosis Test". Three-stage process was followed in this study: (1) the development of the first-tier questions of the test, (2) the development of the second and third tier questions of the test and (3) the validity and reliability analysis of the three-tier concept diagnostic test.

Since the Let's Travel and Learn About The Living World unit was applied to 5th grade students for the first time in the 2013-2014 academic year and the Human and Environmental unit was taught at the 7th grade in the 2006 Science and Technology curriculum, test was applied to seventy-three students attending 7th and 8th grade. In the following period, the study was completed with 5th grade. Validity and reliability studies of the test were performed.

In the analysis of the data obtained during the development of the test, for each item, standard deviation, arithmetic mean, item distinctiveness and item difficulty. In the development of first tier test item, item analysis method based on the difference of 27% sub-upper group averages and double-series correlation calculation were used. After item analysis was completed, item distinctiveness and item difficulty were calculated. Taking these values into consideration, it was decided which test items should remain in the test.

After completing the elimination of the test's first tier item, the second and third tier questions of the test items were prepared. Three-tier concept diagnosis test was applied to 210 students. Students' responses to test items were first entered into MS Excel. In the 3-tier concept diagnostic test, considering the response to each tier of the question; Peşman and Eryılmaz (2010) and Arslan, Çiğdemoğlu and Moseley (2012) defined eight separate scores were calculated through the macro application developed in MS Excel program. After the

validity and reliability analysis of the obtained three-tier test, 23-item final three-tier test was obtained.

Results and Discussion

As a result of the study, a three-tier concept diagnosis test was developed in order to determine the level of grade 5 students' achievements in the Let's Travel and Learn About The Living World unit and what kind of misconceptions they have about this unit. At the same time, validity and reliability analyzes of the test were completed. Although the test was initially developed as 40 questions, the 23-item test was obtained after the analysis was completed. Microsoft Excel and SPSS 17 programs were used for data analysis.

Test reliability was calculated with KR-20 formula; KR-20 reliability level was found to be 0.80. When the three-tier were evaluated together, the difficulty and discrimination values of the test were calculated as p_j : 0.47 and r_{jx} : 0.52, respectively. These data show that the developed test is reliable, of medium difficulty and composed of very good test items.

Since the three-tier concept diagnosis test provides data on both academic achievement and conceptual diagnosis, it will enable researchers to obtain two separate data sets with a single scale. In the process of determining academic achievement, researchers may consider not only the first-tier questions of the scale, but also the three tiers. Researchers who use the scale to determine misconceptions should evaluate the three tiers together.

Giriş

Kavram, nesne ve düşüncelerin insan zihninde canlandırıldığı, soyut bir düşünce birimidir (Yağbasan & Gülçiçek, 2003). Öğrencilerin Fen Bilimleri derslerinde öğretilen kavramları öğrenme düzeylerini belirleyebilmek için başarı testlerinden yararlanılmaktadır. Açık uçlu testler (Öksüz & Güven Demir, 2019), kısa cevaplı testler, sınıflama gerektiren testler ile seçim gerektiren çoktan seçmeli testler (Çalık & Ayas, 2003) öğrencilerin öğretilen kavramları ne düzeyde öğrendiklerini belirlemede kullanılan testler arasında yer almaktadır.

Çoktan seçmeli testler nesnellik, uygulama-değerlendirme kolaylığı ve öğretilen üniteyle ilgili birçok alt kavramın öğrenilme düzeyini belirleme gibi özelliklere sahiptir (Bağcan-Büyükturan & Çıkrıkçı Demirtaşlı, 2012). Bu testlerde öğrencinin doğru seçeneği işaretlemesi, onun ilgili soruya yönelik bilgisinin tam ve doğru olması anlamına gelmediği gibi, bilgi eksikliği veya yaptığı bazı hatalardan dolayı çeldirici bulunan yanlış seçeneği işaretlemesi de kavram yanlışlığına sahip olduğu anlamına gelmemektedir. Bir öğrencinin kavram yanlışlığına sahip olduğuna karar verilebilmesi için, öğrencinin sahip olduğu kavram yanlışlığını açıklaması ve verdiği yanıtın emin olması gerekmektedir. (Karataş ve diğ., 2003). Çoktan seçmeli testlerin olumlu yönlerini taşıdığı gibi, olumsuzluklarını en aza indirdiği düşünüldüğü için kavram yanlışlıklarını belirlemek amacıyla 1980'li yıllarda iki aşamalı teşhis testleri geliştirilmeye başlanmıştır (Bernhisel, 1999; Chen, Lin & Lin, 2002; Odom & Barrow, 1995; Tan, Goh, Chia & Treagust, 2002). İki aşamalı kavram tanı testlerinin birinci aşamasında verilen yanıtın, ikinci aşamada nedenleri ile açıklanması istenmektedir (Karataş ve diğ., 2003).

Günümüzde ise öğrencilerin sahip oldukları kavram yanlışlıklarını belirlemek amacıyla üç aşamalı teşhis testleri kullanılmaya başlanmıştır. Bu testlerin ilk iki aşaması iki aşamalı teşhis testlerinin özelliklerine sahiptir. Üçüncü aşamada öğrencinin ilk iki aşamada verdiği yanıtın ne derecede emin olduğu sorgulanmaktadır. 3-aşamalı kavram tanı testlerinde öğrenci üç aşamada yer alan tüm sorularda kavram yanlışlığına götüren seçenekleri işaretlediğinde, öğrencinin kavram yanlışlığına sahip olduğu kabul edilmekte, ilk iki aşamada kavram yanlışlığına götüren seçenekler işaretlenmesine rağmen, öğrenci bu seçimleri emin olmayarak işaretlediğinde bu durum bilgi eksikliği olarak tanımlanmaktadır (Peşman & Eryılmaz, 2010).

Fen Bilimleri eğitiminde somut ve soyut birçok kavram öğretilmektedir. Literatür incelendiğinde öğrencilerin, farklı öğrenme alanları ile ilgili kavram yanlışlıklarının olduğu

dikkat çekmektedir. Canlıların sınıflandırılması ve çevre sorunları konuları ile ilgili yapılan araştırmalarda öğrencilerin kavram yanlışlarına sahip oldukları gözlenmektedir. Allen (2010), Çetinkaya (2010), Mutlu ve Tokcan (2012), Özyılmaz Akamca (2008), Saka, Ayas ve Enginar (2002), Summers, Kruger, Childs ve Mant (2000), Taş, Aymen Peker ve Çetinkaya (2014), Tekkaya, Çapa ve Yılmaz (2000) ile Yen, Yao ve Mintzes (2007) gibi araştırmacılar ilgili konulara yönelik öğrencilerin sahip oldukları kavram yanlışlarını belirlemiştir. Araştırmalarda daha çok canlıların sınıflandırılması konusu ile ilgili kavram yanlışlarının belirlendiği dikkat çekmektedir. 5. sınıf öğrencileri ile yürütülen çalışmaların ise 2013 öğretim programı öncesinde gerçekleştirildiği ve sadece canlıların sınıflandırılması konusu ile ilgili kavram yanlışlarının belirlendiği gözlenmektedir.

Bu çalışmanın amacı 2013 öğretim programı kazanımları göz önünde bulundurularak ortaokul 5. sınıf öğrencilerinin Canlılar Dünyasını Gezelim ve Tanıyalım ünitesi kazanımlarına ne düzeyde ulaştıklarını ve ünite ile ilgili ne tür kavram yanlışlarına sahip olduklarının tespit edilmesini sağlayacak geçerli ve güvenilir bir ölçme aracı geliştirmektir.

Yöntem

Araştırma Modeli

Tarama (Survey) yönteminin kullanıldığı bu çalışmada ortaokul 5.sınıf *Canlılar Dünyasını Gezelim Tanıyalım* ünitesine yönelik *Kavram Tanı Testi (KTT)* geliştirilmiştir. Test geliştirme sürecinde Şekil 1’de açıklanan işlem basamakları takip edilmiştir. Veri analizlerinde Microsoft Excel ve SPSS 17 programlarından yararlanılmıştır. Aritmetik ortalama ve standart sapma değerleri hesaplanan ölçek maddelerine %27’lik alt-üst grup ortalamaları farkına dayalı madde analizi uygulanmış, SPSS programı ile çift serili korelasyon değerleri hesaplanmıştır. Testin güvenilirliğini tespit etmede KR20 güvenilirlik katsayısından yararlanılmıştır. Test geliştirme sürecinde öncelikle birinci aşama soruları için geçerlik güvenilirlik çalışmaları yapılmış, birinci aşama soruları tespit edildikten sonra üç aşamalı test için analiz işlemleri tekrarlanmıştır.

Örneklem

Canlılar Dünyasını Gezelim ve Tanıyalım ünitesi, 2013-2014 eğitim öğretim yılında 5. sınıf öğrencilerine ilk kez uygulandığı ve 2006 yılı Fen ve Teknoloji dersi öğretim programında İnsan ve Çevre ünitesi yedinci sınıf seviyesinde yer aldığı için denemelik test 7.ve 8.sınıf öğrencilerine uygulanmıştır. Daha sonraki aşamalarda 5.sınıf öğrencileri ile

çalışma tamamlanarak, testin geçerlik ve güvenilirlik çalışmaları gerçekleştirilmiştir. Çalışma gruplarına ait bilgiler Tablo 1’de yer almaktadır.

Tablo 1. Test geliştirme aşaması katılımcı bilgileri

Test	Öğretim Yılı	Okul	Sınıf Seviyesi	Öğrenci Sayısı	Madde Sayısı
Tek Aşamalı test	2013-2014	A	7. ve 8.sınıf	73	40
İki Aşamalı Nihai Test	2013-2014	A	5.sınıf	62	30
Üç Aşamalı Test	2014-2015	A, B, C	5.sınıf	210	25

Bulgular

KTT Geliştirme Süreci

Çelik (2000) test geliştirilirken izlenmesi gereken adımları planlama, madde yazma, madde analizi ve madde seçimi şeklinde sıralamaktadır. Çoktan seçmeli başarı testi (Bakioğlu, Karamustafaoğlu & Karamustafaoğlu, 2014; Çalık & Ayas, 2003), iki aşamalı (Sesli & Kara, 2012) ve üç aşamalı kavram tanı testi (Peşman, 2005; Peşman & Eryılmaz, 2010) geliştirilme süreci ile ilgili araştırmalar incelenip, test geliştirme sürecinde izlenecek adımlar belirlenmiştir (Şekil 1).

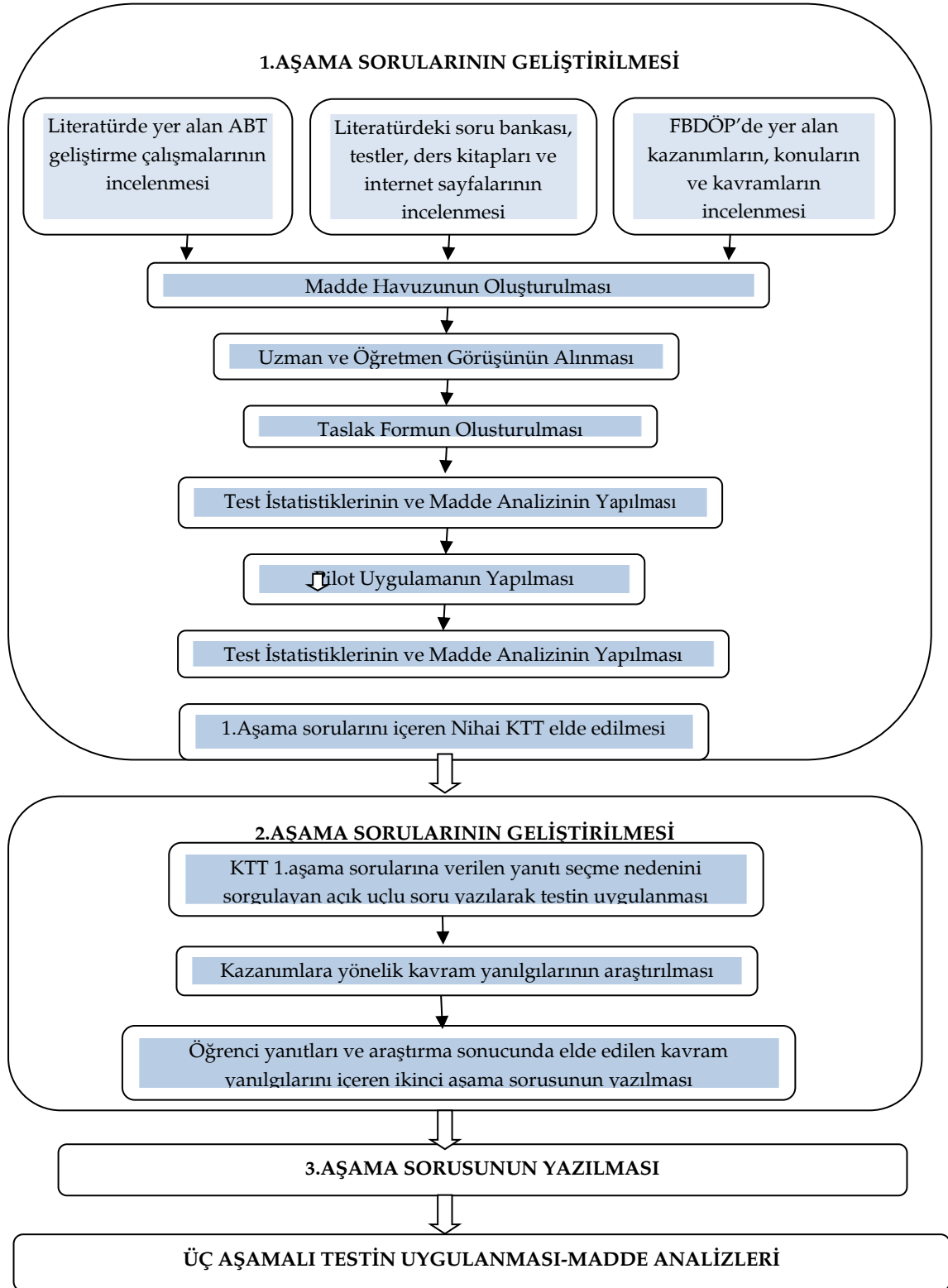
a. Kavram tanı testinin birinci aşama sorularının geliştirilmesi

KTT için ünitenin mevcut kazanımlarını kapsayacak şekilde 40 sorudan oluşan çoktan seçmeli *denemelik test* üç Fen Bilimleri öğretmeni ve üç uzman görüşü alınarak oluşturulmuştur.

Geliştirilen akademik başarı testi, test maddelerinin üniteye yer alan konulara ve kazanımlara göre dağılımı Tablo 2’de yer almaktadır.

Tablo 2. Taslak KTT’de yer alan soru maddelerinin konulara ve kazanımlara göre dağılımı

Konular	Kazanımlar	Madde Numarası	Madde Sayısı
Canlıları Tanıyalım	5.5.1.1. Canlılara örnekler vererek benzerlik ve farklılıklarına göre gruplandırır.	1,2,3,4,5,6,7,8,9,10, 11,12,13,14,15,16,17,18,23,	19
İnsan ve Çevre İlişkisi	5.5.2.1. İnsan faaliyetleri sonucunda oluşan çevre sorunlarını araştırır ve bu sorunların çözümüne ilişkin önerilerde bulunur.	20, 21, 22, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 40	16
	5.5.2.2. Yakın çevresindeki bir çevre sorununun çözümüne ilişkin proje tasarlar ve sunar.	19,24, 25, 38, 39	5



Şekil 1. Kavram tanı testi geliştirme süreci

Tablo 3’de taslak KTT ilk uygulama- test maddelerinin aritmetik ortalama ve standart sapma değerleri verilmiştir

Tablo 3. Taslak KTT ilk uygulama-soruların aritmetik ortalama-standart sapma değerleri

Madde	N	\bar{x}	SS	Madde	N	\bar{x}	SS	Madde	N	\bar{x}	SS	Madde	N	\bar{x}	SS
1	73	0.96	0.20	11	73	0.85	0.36	21	73	0.84	0.37	31	73	0.74	0.44
2	73	0.92	0.28	12	73	0.82	0.39	22	73	0.75	0.43	32	73	0.84	0.37
3	73	0.60	0.49	13	73	0.89	0.31	23	73	0.77	0.43	33	73	0.62	0.49
4	73	0.88	0.33	14	73	0.89	0.31	24	73	0.82	0.39	34	73	0.55	0.50
5	73	0.75	0.43	15	73	0.62	0.49	25	73	0.68	0.47	35	73	0.66	0.48
6	73	0.78	0.42	16	73	0.15	0.36	26	73	0.82	0.56	36	73	0.70	0.46
7	73	0.44	0.50	17	73	0.75	0.43	27	73	0.62	0.49	37	73	0.63	0.49
8	73	0.10	0.30	18	73	0.44	0.50	28	73	0.82	0.39	38	73	0.33	0.47
9	73	0.88	0.33	19	73	0.64	0.48	29	73	0.56	0.50	39	73	0.40	0.49
10	73	0.27	0.45	20	73	0.64	0.48	30	73	0.84	0.37	40	73	0.26	0.44

Taslak KTT ilk uygulamasında yer alan test maddelerinin analizinde, %27'lik alt-üst grup ortalamaları farkına dayalı madde analizi yöntemi kullanılmıştır.

Madde güçlüğü (p_i) 0-1 arasında değer almakta; 0'a yaklaştıkça madde zorlaşmakta, 1'e yaklaştıkça da kolaylaşmaktadır. Maddelerin güçlük düzeyleri ile ilgili olarak madde güçlük indeksi (p_j); 0.00-0.19 arasında ise madde çok zor, 0.20-0.34 arasında ise madde zor, 0.35-0.64 arasında ise madde orta güçlükte, 0.65-0.79 arasında ise madde kolay, 0.80-1.00 arasında ise madde çok kolay olarak kabul edilir (Sözbilir, 2010).

Bir maddenin bilenle bilmeyeni ne derece ayırdığını gösteren madde ayırt ediciliği (r_{jx}) ise, -1 ile 1 arasında değişmekte olup, maddenin ayırt ediciliğinin 0'a yaklaşması ayırt ediciliğinin düşmesi, 1'e yaklaşması ayırt ediciliğinin artması anlamına gelmektedir (Gönen, Kocakaya & Kocakaya, 2011). Ölçekten alınan toplam puanlara göre, en yüksek puandan en düşük puana doğru sıralandığında, üst grup ve alt-grup olan uç grupların, her bir maddeye verdikleri puan ortalamalarının karşılaştırılmasıdır (Tavşancıl, 2010).

Madde analizi sonuçlarında (Tablo 4), testte kalmasına karar verilecek maddelerin seçiminde madde ayırt edicilik indeksi (r_{jx}) olarak; $r_{jx} \leq 0.19$ ise madde kabul edilmez; 0.20–0.29 arasında ise madde düzeltilmeli; 0.30–0.39 arasında ise iyi bir maddedir/kabul edilir; $0.40 \leq r_{jx}$ ise çok iyi bir madde/ kabul edilir ölçütleri kullanılmıştır (Özçelik, 2010).

Tablo 4. Başarı testi birinci uygulama sonucu madde analizi tablosu

Madde No	Gruplar	A	B	C	D	Boş	Cevap Sayısı
1*	Üst Grup	-	-	20*	-	-	20
	Alt Grup	1	1	17*	1	-	20
2*	Üst Grup	-	-	-	20*	-	20
	Alt Grup	-	1	1	18*	-	20
3	Üst Grup	-	1	15*	4	-	20
	Alt Grup	1	4	8*	5	2	20
4	Üst Grup	20*	-	-	-	-	20
	Alt Grup	12*	2	5	-	1	20
5	Üst Grup	2	16*	-	-	2	20
	Alt Grup	2	10*	2	3	3	20

Madde No	Gruplar	A	B	C	D	Boş	Cevap Sayısı	
6	Üst Grup	18*	-	1	1	-	20	$P_f = 0.75$
	Alt Grup	12*	1	4	1	2	20	$r_{j_s} = 0.30$
7	Üst Grup	-	16*	4	-	-	20	$P_f = 0.53$
	Alt Grup	-	5*	14	1	-	20	$r_{j_s} = 0.55$
8	Üst Grup	-	5	3	11*	1	20	$P_f = 0.38$
	Alt Grup	4	5	4	4*	3	20	$r_{j_s} = 0.35$
9**	Üst Grup	-	20*	-	-	-	20	$P_f = 0.93$
	Alt Grup	-	17*	-	2	1	20	$r_{j_s} = 0.15$
10**	Üst Grup	10*	8	2	-	-	20	$P_f = 0.43$
	Alt Grup	7*	8	4	1	1	20	$r_{j_s} = 0.15$
11	Üst Grup	-	-	20*	-	-	20	$P_f = 0.83$
	Alt Grup	1	1	13*	3	2	20	$r_{j_s} = 0.35$
12	Üst Grup	-	-	20*	-	-	20	$P_f = 0.83$
	Alt Grup	-	2	13*	2	3	20	$r_{j_s} = 0.35$
13	Üst Grup	1	-	19*	-	-	20	$P_f = 0.78$
	Alt Grup	3	3	12*	1	1	20	$r_{j_s} = 0.35$
14	Üst Grup	-	-	20*	-	-	20	$P_f = 0.88$
	Alt Grup	2	3	13*	-	2	20	$r_{j_s} = 0.35$
15	Üst Grup	-	16*	1	2	1	20	$P_f = 0.58$
	Alt Grup	1	7*	2	9	1	20	$r_{j_s} = 0.45$
16**	Üst Grup	8	-	6*	3	3	20	$P_f = 0.18$
	Alt Grup	12	1	1*	2	4	20	$r_{j_s} = 0.25$
17	Üst Grup	-	-	-	18*	2	20	$P_f = 0.68$
	Alt Grup	3	1	4	9*	3	20	$r_{j_s} = 0.45$
18	Üst Grup	14*	3	-	2	1	20	$P_f = 0.53$
	Alt Grup	7*	4	5	2	2	20	$r_{j_s} = 0.35$
19	Üst Grup	-	2	17*	1	-	20	$P_f = 0.60$
	Alt Grup	4	2	7*	7	-	20	$r_{j_s} = 0.50$
20	Üst Grup	-	1	18*	-	1	20	$P_f = 0.63$
	Alt Grup	1	3	7*	5	4	20	$r_{j_s} = 0.55$
21	Üst Grup	-	-	1	19*	-	20	$P_f = 0.78$
	Alt Grup	2	3	2	12*	1	20	$r_{j_s} = 0.35$
22	Üst Grup	-	17*	-	1	2	20	$P_f = 0.68$
	Alt Grup	3	10*	2	3	2	20	$r_{j_s} = 0.35$
23	Üst Grup	-	20*	-	-	-	20	$P_f = 0.70$
	Alt Grup	4	8*	1	4	3	20	$r_{j_s} = 0.60$
24**	Üst Grup	-	1	-	19*	-	20	$P_f = 0.83$
	Alt Grup	-	3	2	14*	1	20	$r_{j_s} = 0.25$
25	Üst Grup	-	18*	-	2	-	20	$P_f = 0.53$
	Alt Grup	3	3*	3	7	4	20	$r_{j_s} = 0.75$
26	Üst Grup	-	-	-	20*	-	20	$P_f = 0.68$
	Alt Grup	4	5	2	7*	2	20	$r_{j_s} = 0.65$
27	Üst Grup	-	19*	-	1	-	20	$P_f = 0.58$
	Alt Grup	2	4*	7	4	3	20	$r_{j_s} = 0.75$
28	Üst Grup	-	20*	-	-	-	20	$P_f = 0.73$
	Alt Grup	3	9*	3	4	1	20	$r_{j_s} = 0.55$
29**	Üst Grup	-	-	7	12*	1	20	$P_f = 0.50$
	Alt Grup	2	3	4	8*	3	20	$r_{j_s} = 0.20$
30**	Üst Grup	1	17*	2	-	-	20	$P_f = 0.80$
	Alt Grup	2	15*	1	2	-	20	$r_{j_s} = 0.10$
31	Üst Grup	1	1	-	18*	-	20	$P_f = 0.68$
	Alt Grup	3	2	6	9*	-	20	$r_{j_s} = 0.45$
32	Üst Grup	-	-	18*	2	-	20	$P_f = 0.73$
	Alt Grup	5	-	11*	3	1	20	$r_{j_s} = 0.35$
33	Üst Grup	-	3	17*	-	-	20	$P_f = 0.60$
	Alt Grup	3	8	7*	-	2	20	$r_{j_s} = 0.50$
34**	Üst Grup	-	5	15*	-	-	20	$P_f = 0.65$
	Alt Grup	2	6	11*	1	-	20	$r_{j_s} = 0.20$
35	Üst Grup	-	-	19*	-	1	20	$P_f = 0.63$
	Alt Grup	4	5	6*	4	1	20	$r_{j_s} = 0.65$
36	Üst Grup	19*	-	-	-	1	20	$P_f = 0.68$
	Alt Grup	8*	3	3	4	2	20	$r_{j_s} = 0.55$
37	Üst Grup	-	-	19*	1	-	20	$P_f = 0.68$
	Alt Grup	2	7	8*	3	-	20	$r_{j_s} = 0.55$
38	Üst Grup	-	2	9	9*	-	20	$P_f = 0.23$
	Alt Grup	3	6	6	2*	3	20	$r_{j_s} = 0.35$
39	Üst Grup	16*	-	3	1	-	20	$P_f = 0.43$
	Alt Grup	1*	3	9	6	1	20	$r_{j_s} = 0.75$
40**	Üst Grup	12	-	-	8*	-	20	$P_f = 0.33$

Madde No	Gruplar	A	B	C	D	Boş	Cevap Sayısı
	Alt Grup	6	4	5	5'	-	20
							$r_{jx}=0.15$

*Test maddesinin doğru cevap seçeneğidir.

**Madde analizi sonucu testten çıkarılan maddeler

P_j = Maddenin güçlük indeksi: ((Maddeyi üst grupta doğru cevaplayanlar + Maddeyi alt grupta doğru cevaplayanlar) / Toplam cevaplayan)

r_{jx} = Maddenin ayırt edicilik gücü: ((Maddeyi üst grupta doğru cevaplayanlar-Maddeyi alt grupta doğru cevaplayanlar) / Üst ya da alt gruptan herhangi birisinin eleman sayısı)

Uygulamada veriler Microsoft Excel programına girilmiş, maddelerin güçlük dereceleri, standart sapmaları ve ayırıcılık dereceleri hesaplanarak, çok kolay ve çok zor olarak belirlenen sorular testten çıkarılmıştır. Ayırıcılık dereceleri 0.20 ve altında olan maddeler (1, 2, 9, 10, 29, 30, 34 ve 40) ile ayırıcılık derecesi 0.25 olmasına rağmen çok kolay ve çok zor madde olarak belirlenen 16. ve 24. sorular testten çıkarılmıştır (Tablo 5; Tablo 6).

Tablo 5. Üst ve alt grup öğrencilerinin doğru sayısına göre başarı test madde analizi

Madde	dü	da	pj	Güçlük düzeyi	rjx	Sonuç	Durum	Madde	dü	da	pj	Güçlük düzeyi	rjx	Sonuç	Durum
1	20	17	0.93	ÇK	0.15	Çık	-	21	19	12	0.78	K	0.35	İyi	√
2	20	18	0.95	ÇK	0.10	Çık	-	22	17	10	0.68	K	0.35	İyi	√
3	15	8	0.56	O	0.35	İ	√	23	20	8	0.70	K	0.60	Çİ	√
4	20	12	0.80	ÇK	0.40	Çİ	√	24	19	14	0.83	ÇK	0.25	Düz	-
5	16	10	0.65	K	0.30	İyi	√	25	18	3	0.53	O	0.75	Çİ	√
6	18	12	0.75	K	0.30	İyi	√	26	20	7	0.68	K	0.65	Çİ	√
7	16	5	0.53	O	0.55	Çİ	√	27	19	4	0.58	O	0.75	Çİ	√
8	11	4	0.38	O	0.35	İyi	√	28	20	9	0.73	K	0.55	Çİ	√
9	20	17	0.93	ÇK	0.15	Çık	-	29	12	9	0.53	O	0.15	Çık	-
10	10	7	0.43	O	0.15	Çık	-	30	17	15	0.80	ÇK	0.10	Çık	-
11	20	13	0.83	ÇK	0.35	İyi	√	31	18	9	0.68	K	0.45	Çİ	√
12	20	13	0.83	ÇK	0.35	İyi	√	32	18	11	0.73	K	0.35	İyi	√
13	19	12	0.78	K	0.35	İyi	√	33	17	7	0.60	O	0.50	Çİ	√
14	20	13	0.83	ÇK	0.35	İyi	√	34	15	11	0.65	K	0.20	Düz	-
15	16	7	0.58	O	0.45	Çİ	√	35	19	6	0.63	O	0.65	Çİ	√
16	6	1	0.18	ÇZ	0.25	Düz	-	36	19	8	0.68	K	0.55	Çİ	√
17	18	9	0.68	ÇK	0.45	Çİ	√	37	19	8	0.68	K	0.55	Çİ	√
18	14	7	0.53	O	0.35	İyi	√	38	9	2	0.23	Z	0.35	İyi	√
19	17	7	0.60	O	0.50	Çİ	√	39	16	1	0.43	O	0.75	Çİ	√
20	18	7	0.63	O	0.55	Çİ	√	40	8	5	0.33	Z	0.15	Çık	-

ÇK: Çok Kolay, K: Kolay, O: Orta, Z: Zor, ÇZ: Çok zor, Çİ: Çok iyi, İ: İyi, Çık: Çıkarılmalı, Düz: Düzeltilmeli

Veriler SPSS programına aktarılmış, başarı testinin toplamından alınan puan (sürekli değişken) ile testin her maddesinden alınan puan arasındaki ilişki çift serili korelasyon şeklinde hesaplanmıştır. Çift serili korelasyon katsayısı, sürekli bir değişken ile gerçekte sürekli, ancak yapay olarak iki kategorili süresiz bir duruma getirilen bir değişken arasındaki ilişki miktarı hesaplanırken kullanılmaktadır (Büyüköztürk, Çokluk & Köklü, 2010).

Cevaplayıcıların ölçme aracından aldığı toplam puan ile her bir maddeden aldığı puan arasındaki ilişkiyi madde toplam korelasyonu açıklar. Korelasyonun pozitif ve yüksek olması, testin iç tutarlılığının yüksek olduğunu ve ölçek maddelerinin benzer davranışları

örneklediğini gösterir (Büyüköztürk, 2010). Herhangi bir maddenin, toplam puan ile korelasyonu düşük ise, o maddenin diğer maddelerden farklı bir özelliği ölçtüğü şeklinde ilişki yorumlanır. Madde toplam korelasyonunun negatif olmaması ve en az 0.20 olması gerekmektedir. Negatif değer olması ölçeğin toplanabilirlik özelliğini bozmaktadır.

Tablo 6. Birinci uygulamada yer alan maddelerin çift serili korelasyon katsayısı değerleri

Madde No	N	p	q	\bar{Y}_p	\bar{Y}_q	SS	pq/y	r_c
1	73	0.96	0.04	26.82	18.00	0.20	0.45	0.42
2*	73	0.92	0.08	26.76	23.17	0.28	0.50	0.19
3	73	0.60	0.40	28.11	23.97	0.49	0.62	0.30
4	73	0.88	0.12	27.86	16.56	0.33	0.53	0.64
5	73	0.75	0.25	27.71	22.67	0.43	0.59	0.32
6	73	0.78	0.22	27.67	22.19	0.42	0.58	0.34
7*	73	0.44	0.56	28.75	24.68	0.50	0.63	0.27
8*	73	0.10	0.90	31.00	25.98	0.30	0.51	0.27
9*	73	0.88	0.12	26.88	23.56	0.33	0.53	0.19
10*	73	0.27	0.73	28.20	25.81	0.45	0.60	0.15
11*	73	0.85	0.15	27.61	20.00	0.36	0.55	0.44
12	73	0.82	0.18	27.72	20.69	0.39	0.57	0.43
13	73	0.89	0.11	27.43	18.63	0.31	0.52	0.49
14	73	0.89	0.11	27.18	20.63	0.31	0.52	0.36
15	73	0.62	0.38	28.22	23.64	0.49	0.62	0.30
16*	73	0.15	0.85	29.82	25.87	0.36	0.55	0.23
17	73	0.75	0.25	31.20	24.00	0.43	0.59	0.45
18*	73	0.44	0.56	27.84	25.39	0.50	0.63	0.16
19	73	0.64	0.36	28.45	22.88	0.48	0.62	0.37
20	73	0.64	0.36	28.64	22.54	0.48	0.62	0.40
21	73	0.84	0.16	27.84	19.50	0.37	0.55	0.49
22	73	0.75	0.25	28.11	21.44	0.43	0.59	0.42
23	73	0.77	0.23	28.79	18.82	0.43	0.58	0.61
24*	73	0.82	0.18	27.20	23.08	0.39	0.56	0.25
25	73	0.69	0.31	29.38	20.13	0.47	0.61	0.60
26	73	0.78	0.22	28.74	18.38	0.56	0.58	0.64
27	73	0.62	0.38	29.58	21.46	0.49	0.62	0.54
28	73	0.82	0.18	28.05	19.15	0.39	0.56	0.53
29*	73	0.56	0.44	28.07	24.41	0.50	0.63	0.25
30*	73	0.84	0.16	26.98	23.83	0.37	0.55	0.18
31	73	0.74	0.26	28.24	21.42	0.44	0.59	0.43
32	73	0.84	0.16	27.38	21.83	0.37	0.55	0.32
33	73	0.62	0.38	28.51	23.18	0.49	0.62	0.35
34*	73	0.55	0.45	28.28	24.27	0.50	0.63	0.27
35	73	0.66	0.34	29.15	21.32	0.48	0.61	0.51
36	73	0.70	0.30	28.29	21.55	0.46	0.60	0.43
37	73	0.63	0.37	29.07	22.04	0.49	0.62	0.46
38	73	0.33	0.67	29.42	25.02	0.47	0.61	0.30
39	73	0.40	0.60	31.24	23.32	0.49	0.62	0.52
40*	73	0.26	0.74	27.53	26.09	0.44	0.59	0.09

*Madde toplam puan korelasyonu 0.30'un altında olan maddeler

Taslak başarı testinde yer alan her bir maddenin çift serili korelasyon katsayısı değeri hesaplandığında 2, 7, 8, 9, 10, 16, 18, 24, 29, 30, 34 ve 40. maddelerin korelasyon katsayısı değerleri 0.30'un altındadır (Tablo 6). Bu soru maddelerinden 2, 9, 10, 16, 24, 29, 30, 34 ve 40 daha önce yapılan madde güçlük ve ayırt edicilik hesaplamasında düşük çıkarak elenen maddelerdir. 2, 8 ve 18. maddelerin korelasyon değeri 0.20 üzerinde olduğu için taslak test

içerisinde kalmasına karar verilmiştir. Maddeler arasındaki korelasyonun yüksek olması, maddelerin homojen olduğunu ve dolayısıyla güvenilirliğinde yüksek olduğunu göstermektedir (Tavşancıl, 2010).

Madde analizi sonucunda 30 maddeye düşen testin KR-20 güvenilirlik testi sonucu Formül-3 ile hesaplanıp, 0.86 olarak belirlenmiştir. Bir ölçme aracının güvenilir kabul edilebilmesi için KR-20 güvenilirlik katsayısının 0.70 ve üzerinde olması beklenmektedir (Özçelik, 2010). Bu sebeple madde analizi gerçekleştirilen testin güvenilir olduğu kabul edilmiştir.

$$\text{Formül 2} \\ KR20 = \frac{K}{K-1} \times \left(1 - \frac{\sum pq}{S_x^2}\right)$$

K= Testteki Soru Sayısı
p=Test Maddesinin Güçlük Derecesi (Doğru Cevap Sayısı /Toplam Öğrenci Sayısı)
q=Test Maddesinin Zorluk Derecesi [q=1-p]
S_x²= Testin varyansı

Geliştirilen birinci aşama başarı testi madde elemesi tamamlandıktan sonra, test maddelerinin üniteye yer alan konu ve kazanımlara göre dağılımı Tablo 7’de yer almaktadır.

Tablo 7. Nihai birinci aşama başarı testinde yer alan soru maddelerinin konulara ve kazanımlara göre dağılımı

Konular	Kazanımlar	Madde Numarası	Madde Sayısı
<i>Canlıları Tanıyalım</i>	5.5.1.1.	1,2,3,4,5,6,7,9,11,18,27,28,29,30	14
<i>İnsan ve Çevre İlişkisi</i>	5.5.2.1.	8,12,13,15,16, 19,20,21,23,24,25,26	12
	5.5.2.2.	10,14,17,22	4

30 sorudan oluşan nihai test pilot uygulamada 62 beşinci sınıf öğrencisine uygulanmıştır. Başarı testinin, ikinci uygulamasındaki test maddelerinin aritmetik ortalama ve standart sapma değerleri Tablo 8’de verilmiştir.

Tablo 8. Taslak başarı testi ikinci uygulamasındaki soruların aritmetik ortalama ve standart sapma değerleri

Madde No	N	\bar{x}	SS	Madde No	N	\bar{x}	SS	Madde No	N	\bar{x}	SS
1	62	0.90	0.30	11	62	0.73	0.45	21	62	0.37	0.49
2	62	0.79	0.41	12	62	0.56	0.50	22	62	0.45	0.50
3	62	0.69	0.46	13	62	0.68	0.47	23	62	0.44	0.50
4	62	0.73	0.45	14	62	0.58	0.50	24	62	0.47	0.50
5	62	0.65	0.48	15	62	0.66	0.48	25	62	0.39	0.49
6	62	0.68	0.47	16	62	0.60	0.49	26	62	0.37	0.49
7	62	0.74	0.44	17	62	0.65	0.48	27	62	0.24	0.43
8	62	0.48	0.50	18	62	0.56	0.50	28	62	0.66	0.48
9	62	0.53	0.50	19	62	0.50	0.50	29	62	0.32	0.47
10	62	0.55	0.50	20	62	0.52	0.50	30	62	0.44	0.50

Birinci uygulamada olduğu gibi elde edilen test puanlarına %27’lik alt-üst grup ortalamaları farkına dayalı madde analizi uygulanmıştır. (Tablo 9; Tablo 10).

Tablo 9. Başarı testi ikinci uygulama sonucu madde analizi tablosu

Madde No	Gruplar	A	B	C	D	Boş	Cevap Sayısı	
1	Üst Grup	17*	-	-	-	-	17	P _j = 0.85
	Alt Grup	12*	-	4	-	1	17	r _{jx} =0.29
2	Üst Grup	-	1	16*	-	-	17	P _j = 0.79
	Alt Grup	2	1	11*	2	1	17	r _{jx} =0.29
3	Üst Grup	-	16*	1	-	-	17	P _j = 0.65
	Alt Grup	-	6*	5	5	1	17	r _{jx} =0.59
4	Üst Grup	17*	-	-	-	-	17	P _j = 0.68
	Alt Grup	6*	5	3	2	1	17	r _{jx} =0.65
5	Üst Grup	1	1	15*	-	-	17	P _j = 0.68
	Alt Grup	-	3	8*	5	1	17	r _{jx} =0.41
6	Üst Grup	3	-	14*	-	-	17	P _j = 0.62
	Alt Grup	5	2	7*	2	1	17	r _{jx} =0.41
7	Üst Grup	-	1	16*	-	-	17	P _j = 0.62
	Alt Grup	1	6	5*	4	1	17	r _{jx} =0.65
8	Üst Grup	-	-	16*	1	-	17	P _j = 0.53
	Alt Grup	5	4	2*	5	1	17	r _{jx} =0.82
9	Üst Grup	-	17*	-	-	-	17	P _j = 0.59
	Alt Grup	2	3*	2	9	1	17	r _{jx} =0.82
10	Üst Grup	-	1	14*	2	-	17	P _j = 0.56
	Alt Grup	7	-	5*	4	1	17	r _{jx} =0.53
11	Üst Grup	3	2	5	7*	-	17	P _j = 0.38
	Alt Grup	4	2	4	6*	1	17	r _{jx} =0.06
12	Üst Grup	-	1	-	15*	1	17	P _j = 0.59
	Alt Grup	4	3	4	5*	1	17	r _{jx} =0.59
13	Üst Grup	-	16*	1	-	-	17	P _j = 0.76
	Alt Grup	4	10*	-	2	1	17	r _{jx} =0.35
14	Üst Grup	-	1	16*	-	-	17	P _j = 0.59
	Alt Grup	6	4	4*	2	1	17	r _{jx} =0.71
15	Üst Grup	-	-	1	16*	-	17	P _j = 0.62
	Alt Grup	4	1	6	5*	1	17	r _{jx} =0.65
16	Üst Grup	-	16*	-	-	1	17	P _j = 0.71
	Alt Grup	2	8*	3	3	1	17	r _{jx} =0.47
17	Üst Grup	-	16*	1	-	-	17	P _j = 0.56
	Alt Grup	5	3*	6	2	1	17	r _{jx} =0.76
18	Üst Grup	-	16*	-	-	1	17	P _j = 0.53
	Alt Grup	8	2*	-	6	1	17	r _{jx} =0.82
19	Üst Grup	1	16*	-	-	-	17	P _j = 0.50
	Alt Grup	8	1*	1	6	1	17	r _{jx} =0.88
20	Üst Grup	-	-	15*	1	-	17	P _j = 0.56
	Alt Grup	2	3	4*	7	1	17	r _{jx} =0.65
21	Üst Grup	-	-	17*	-	-	17	P _j = 0.74
	Alt Grup	-	3	8*	5	1	17	r _{jx} =0.53
22	Üst Grup	-	2	-	15*	-	17	P _j = 0.47
	Alt Grup	8	3	4	1*	1	17	r _{jx} =0.82
23**	Üst Grup	9*	3	3	2	-	17	P _j = 0.47
	Alt Grup	7*	4	2	3	1	17	r _{jx} =0.14
24	Üst Grup	-	2	13*	2	-	17	P _j = 0.47
	Alt Grup	2	1	3*	10	1	17	r _{jx} =0.59
25**	Üst Grup	9*	4	4	-	-	17	P _j = 0.44
	Alt Grup	6*	4	5	1	1	17	r _{jx} =0.18
26**	Üst Grup	1	-	3	13*	-	17	P _j = 0.41
	Alt Grup	4	3	2	7*	1	17	r _{jx} =0.12
27	Üst Grup	1	1	2	10*	3	17	P _j = 0.38
	Alt Grup	1	7	5	3*	1	17	r _{jx} =0.41
28	Üst Grup	-	16*	-	1	-	17	P _j = 0.65
	Alt Grup	4	6*	2	4	1	17	r _{jx} =0.59
29**	Üst Grup	1	2	12*	2	-	17	P _j = 0.65
	Alt Grup	3	2	10	2	-	17	r _{jx} =0.12
30**	Üst Grup	8*	6	3	-	-	17	P _j = 0.41
	Alt Grup	6*	6	4	1	-	17	r _{jx} =0.12

*Test maddesinin doğru cevap seçeneğidir. **Madde analizi sonucu testten çıkarılan maddeler
P_j= Maddenin güçlük indeksi ((Maddeyi üst grupta doğru cevaplayanlar + Maddeyi alt grupta doğru cevaplayanlar) / Toplam cevaplayan)

$r_{jx} = \text{Maddenin ayırt edicilik gücü} ((\text{Maddeyi üst grupta doğru cevaplayanlar} - \text{Maddeyi alt grupta doğru cevaplayanlar}) / \text{Üst ya da alt gruptan herhangi birisinin eleman sayısı})$

Birinci uygulama test geliştirme sürecinde olduğu gibi bu uygulamada da elde edilen veriler Microsoft Excel programına girilerek, test maddelerinin güçlük dereceleri, standart sapmaları ve ayıricılık dereceleri hesaplanmıştır. Ayıricılık dereceleri 0.20 ve altında olan maddeler (11, 23, 25, 29 ve 30) testten çıkarılmıştır (Turgut, 1992) (Tablo 10;Tablo 11).

Tablo 10. Üst ve alt gruptaki öğrencilerin doğru sayısına göre ikinci uygulama başarı testinin madde analizi

Madde	dü	da	pj	Güçlük düzeyi	r _{jx}	Sonuç	Duru	Madde	dü	da	pj	Güçlük düzeyi	r _{jx}	Sonuç	Duru
1	17	12	0.85	ÇK	0.29	Düz	-	16	16	8	0.71	K	0.47	Çİ	√
2	16	11	0.79	K	0.29	Düz	-	17	16	3	0.56	O	0.76	Çİ	√
3	16	6	0.65	K	0.59	Çİ	√	18	16	2	0.53	O	0.82	Çİ	√
4	17	6	0.68	K	0.65	Çİ	√	19	16	1	0.50	O	0.88	Çİ	√
5	15	8	0.68	K	0.41	Çİ	√	20	15	4	0.56	O	0.65	Çİ	√
6	14	7	0.62	O	0.41	Çİ	√	21	17	8	0.74	K	0.53	Çİ	√
7	16	5	0.62	O	0.65	Çİ	√	22	15	1	0.47	O	0.82	Çİ	√
8	16	2	0.53	O	0.82	Çİ	√	23	9	7	0.47	O	0.14	Çık	-
9	17	3	0.59	O	0.82	Çİ	√	24	13	3	0.47	O	0.59	Çİ	√
10	14	5	0.56	O	0.53	Çİ	√	25	9	6	0.44	O	0.18	Çık	-
11	7	6	0.38	O	0.06	Çık	-	26	13	7	0.59	O	0.35	İ	√
12	15	5	0.59	O	0.59	Çİ	√	27	10	3	0.38	O	0.41	Çİ	√
13	16	10	0.76	K	0.35	İ	√	28	16	6	0.65	K	0.59	Çİ	√
14	16	4	0.59	O	0.71	Çİ	√	29	12	10	0.65	K	0.12	Çık	-
15	16	5	0.62	O	0.65	Çİ	√	30	8	6	0.41	O	0.12	Çık	-

ÇK: Çok Kolay, K: Kolay, O: Orta, Z: Zor, ÇZ: Çok zor, Çİ: Çok iyi, İ: İyi, Çık: Çıkarılmalı, Düz: Düzeltmeli

Veriler için, birinci uygulamada olduğu gibi; çift serili korelasyon hesaplanmıştır.

Tablo 11. Taslak başarı testi ikinci uygulamasında yer alan maddelerin çift serili korelasyon katsayısı değerleri

Madde No	N	p	q	Ȳp	Ȳq	SS	pq/y	r _c
1	62	0.90	0.10	17.40	10.33	0.30	.513	0.56
2	62	0.79	0.21	17.86	13.38	0.41	.576	0.40
3	62	0.69	0.31	18.77	12.74	0.46	.353	0.33
4	62	0.73	0.27	19.02	11.35	0.45	.331	0.39
5	62	0.73	0.27	18.91	11.65	0.45	.331	0.37
6*	62	0.68	0.32	18.55	13.50	0.47	.358	0.28
7	62	0.74	0.26	18.76	11.63	0.44	.324	0.36
8	62	0.48	0.52	21.20	12.91	0.50	.396	0.51
9	62	0.53	0.47	20.70	12.62	0.50	.398	0.49
10	62	0.55	0.45	19.47	13.82	0.50	.396	0.34
11*	62	0.37	0.63	15.48	17.05	0.49	.378	0.09
12	62	0.56	0.44	19.77	13.22	0.50	.394	0.40
13	62	0.44	0.56	19.74	14.74	0.50	.396	0.30
14	62	0.58	0.42	20.22	12.35	0.50	.391	0.47
15	62	0.66	0.34	19.44	12.00	0.48	.366	0.42
16	62	0.60	0.40	19.65	12.88	0.49	.386	0.40
17	62	0.65	0.35	19.80	11.68	0.48	.370	0.46
18	62	0.56	0.44	20.23	12.63	0.50	.394	0.46
19	62	0.50	0.50	21.61	12.23	0.50	.399	0.58

Madde No	N	p	q	Ȳp	Ȳq	SS	pq/y	r _c
20	62	0.52	0.48	20.19	13.43	0.50	.398	0.41
21	62	0.44	0.56	20.93	13.83	0.50	.394	0.43
22	62	0.45	0.55	21.14	13.44	0.50	.396	0.47
23*	62	0.37	0.63	18.35	16.08	0.49	.378	0.13
24	62	0.47	0.53	19.76	14.42	0.50	.398	0.33
25*	62	0.39	0.61	18.63	15.84	0.49	.384	0.16
26	62	0.32	0.68	21.80	14.60	0.47	.358	0.40
27	62	0.24	0.76	22.20	15.23	0.43	.311	0.33
28	62	0.66	0.34	18.98	12.90	0.48	.366	0.34
29*	62	0.68	0.32	18.38	13.85	0.47	.358	0.25
30*	62	0.65	0.35	18.33	14.36	0.48	.370	0.23

*Madde toplam puan korelasyonu 0.30'un altında olan maddeler

Başarı testinde yer alan her bir maddenin çift serili korelasyon katsayı değeri hesaplandığında 6, 11, 23, 25, 29 ve 30. maddelerin korelasyon katsayı değerleri 0.30'un altındadır (Tablo 11). Bu soru maddelerinden 11, 23, 25, 29 ve 30. maddeler daha önce yapılan madde güçlük ve ayırt edicilik hesaplamasında düşük çıkarak elenen maddelerdir. 6. madde değeri 0.28 olarak belirlendiği için testte yer almasına karar verilmiştir.

Bu uygulama sonunda test madde sayısı 25'e düşürülmüştür. Formül-3 kullanılarak hesaplanan KR-20 güvenilirlik katsayısı 0.87 olarak tespit edilmiştir.

Tablo 12. Madde analizi sonucunda bulunan KTT istatistikleri

Kavram Tanı Testi (KTT) 1.Aşama Soruları	Madde Sayısı	N	\bar{x}	Varyans (S ²)	Standart Sapma (S)	Ort. Güçlük (P)	KR-20
1.Uygulama	40	73	26.47	43.09	6.56	0.65	0.85
2.Uygulama	30	62	16.92	42.27	6.50	0.59	0.86
Nihai Ölçek	25	62	14.62	30.73	5.54	0.61	0.87

KTT aritmetik ortalama, standart sapma, varyans, ortalama güçlük ve güvenilirlik hesaplamaları Tablo 12'de görülmektedir. Öğrenci başarıları ölçümünde kullanılacak testlerde testin ortalama güçlük indisinin 0.50 veya ona yakın bir değer olması gerekir (Çepni, Baki, Ayas, Demircioğlu & Akyıldız, 2009). 25 soru şeklinde uygulanan test bu açıdan ölçülen özelliğe hizmet etmektedir. Aynı zamanda belirlenen testin güvenilirlik katsayısı değeri de iç tutarlılığın yeterli düzeyde olduğunu gösteren 0.70 değerinden yüksektir. Bu durum da testin iç tutarlılığının beklenen seviyede olduğunu göstermektedir.

b. Kavram tanı testinin ikinci ve üçüncü aşama sorularının geliştirilmesi

KTT 2. aşama soruları öğrencilerin birinci soruya verdikleri cevabın nedenini keşfetmeye yöneliktir. 2. aşama sorularının elde edilebilmesi için madde analizi sonrası 25 soru olarak düzenlenen test kullanılmıştır. Her bir test sorusundan sonra, öğrencilerin birinci aşama sorusuna verdikleri yanıtı seçme nedenini sorgulayan açık uçlu soru yazılmıştır.

Öğrencilerin açık uçlu sorulara verdiği yanıtlar ve literatür taraması sonrasında üniteye ait belirlenen kavram yanlışları dikkate alınarak ikinci aşama sorusunun seçenekleri belirlenmiştir. En sık tekrarlanan ifadeden başlamak üzere dört seçenek oluşturulmuş, beşinci olarak diğer seçeneği eklenmiştir. Test öğrencilere uygulanırken, öğrencinin birinci aşama sorusuna verdiği yanıtın nedeni, ona göre verilen diğer seçeneklerde yoksa öğrencinin bu şıkkı işaretlemesi ve kendi nedenini yazması istenmiştir (Şekil 2).

<p>Aşağıdakilerden hangisi mikroskopik canlıların özellikleri ile ilgili <u>değildir</u>?</p> <p>A) Sıcak ve soğuk ortamlarda yaşayabilir.</p> <p>B) Canlı vücudunda yaşayamaz.</p> <p>C) Toprakta, tatlı ve tuzlu sularda yaşayabilir.</p> <p>D) Uygun nem, sıcaklık ve besin bulunduğu hızla çoğalır.</p>	<p>Cevabımşıkkı.</p> <p>Bu seçeneği neden seçtiniz?</p> <p>Açıklayınız.....</p> <p>.....</p>
<ul style="list-style-type: none"> • Mikroskopik canlılar canlı vücudunda yaşayabilir. • Toprak ve tatlı sularda yaşayamazlar. • Mikroskopik canlılar sıcak ve soğuk ortamlarda yaşayamazlar. • Mikroskopik canlılar sadece sıcak ortamda yaşar. • Birçok mikroskopik canlı vücudumuzda yaşar. • İnsan vücudunda bulunur ve bizi hasta eder. • İnsan vücudu dahil birçok canlı vücudunda yaşar. • Virüs ve bakteriler yüzünden hasta oluruz. • Canlı vücudunda yaşayıp, hastalıklara yol açar. 	<div style="border: 1px solid black; padding: 5px; text-align: center;"> <p>Öğrencilerin açık uçlu sorulara verdiği yanıtlar</p> </div>

Şekil 2. İkinci aşama sorusu geliştirilmesi

Testin 3. aşama soruları ise öğrencinin 2. aşama sorusuna verdiği yanıtın ne kadar emin olduğunu tespit etmeye yöneliktir. Testteki bir test maddesi Şekil 3 Şekil 3’de görülmektedir.

Soru 3.1.

Aşağıdakilerden hangisi mikroskopik canlıların özellikleri ile ilgili değildir?

A) Sıcak ve soğuk ortamlarda yaşayabilir.

B) Canlı vücudunda yaşayamaz.

C) Toprakta, tatlı ve tuzlu sularda yaşayabilir.

D) Uygun nem, sıcaklık ve besin bulunduğu hızla çoğalır.

Soru 3.2.

Bir önceki soruda seçtiğiniz cevabın nedeni aşağıdakilerden hangisidir?

A. Mikroskopik canlılar canlı vücudunda yaşayabilir.

B. Toprak ve tatlı sularda yaşayamazlar.

C. Mikroskopik canlılar sıcak ve soğuk ortamlarda yaşayamazlar.

D. Mikroskopik canlılar sadece sıcak ortamda yaşar.

E. Diğer

Soru 3.3.

Bir önceki soruda verdiğiniz cevaptan ne kadar eminsiniz?

A) Eminim B) Emin Değilim C) Tahmin Ettim

Şekil 3. KTT’de yer alan üç aşamalı soru örneği

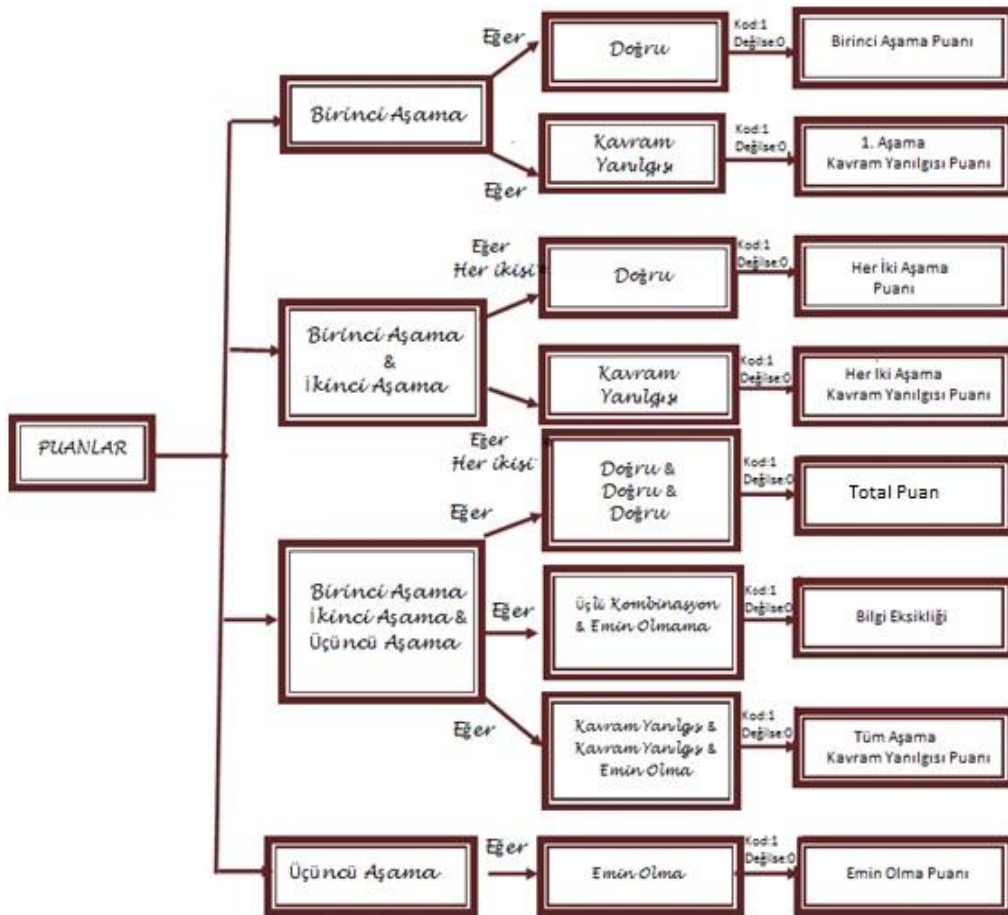
Öğrencinin üç aşamalı testte yer alan her üç aşama sorusuna verdiği yanıt dikkate

alınarak, öğrencilerin test maddesinde ölçülmeye çalışılan kazanım ile ilgili bir kavram yanılıgına sahip olup olmadığı belirlenmelidir.

3 Aşamalı KTT'nin Geçerlik ve Güvenirlik Analizi

KTT toplam 210 öğrenciye uygulanmıştır. Test maddelerine öğrencilerin verdikleri yanıtlar MS Excel programına girilip, bu programda geliştirilen makro uygulamayla ilgili puanlar hesaplanmıştır. 3 aşamalı KTT'de sorunun her bir aşamasına verilen yanıt dikkate alınarak, Peşman ve Eryılmaz (2010) ile Arslan, Çiğdemoglu ve Moseley'in (2012) tanımladığı gibi sekiz ayrı puan hesaplaması yapılmıştır. Bunlar; 1. aşama puanı, her iki aşama puanı, toplam puan, emin olma puanı, bilgi eksikliği puanı, 1.aşama; her iki aşama ve tüm aşama kavram yanılıgı puanları şeklindedir.

Bu puanların tespit edilmesinde Şekil 4'de yer alan kodlamalardan yararlanılmıştır.



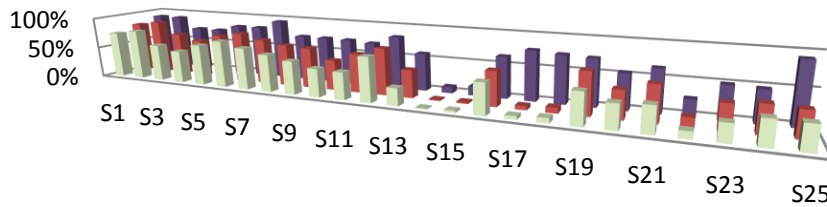
Şekil 4. Puanlama ve Kodlama prosedürü. Üçlü kombinasyon 'yanlış-doğru', 'doğru-yanlış' ve 'yanlış-yanlış' (Arslan ve diğerlerinden (2012) uyarlanmıştır).

Şekil 4’de gösterilen ve Tablo 13’de belirtildiği gibi gerçekleştirilen kodlamalar sonucunda, üç aşamalı test ile bilimsel bilgi, kavram yanlılığı, bilgi eksikliği ve şanslı cevap olmak üzere dört kategori için veriler elde edilmektedir (Tablo 13)

Tablo 13. Üç aşamalı testte mümkün olan olası yanıtlar ve bunların kategorileri

Birinci Aşama	İkinci Aşama	Üçüncü Aşama	Kategoriler
Doğru	Doğru	Emin	Bilimsel Bilgiye Sahip
Doğru	Yanlış	Emin değil	Bilgi Eksikliği
Yanlış	Doğru	Emin değil	Bilgi Eksikliği
Yanlış	Yanlış	Emin değil	Bilgi Eksikliği
Doğru	Doğru	Emin değil	Tahmini Şanslı cevap
Doğru	Yanlış	Emin	Kavram Yanlılığı (Pozitif yönde Yanlış)
Yanlış	Doğru	Emin	Kavram Yanlılığı (Negatif yönde Yanlış)
Yanlış	Yanlış	Emin	Kavram Yanlılığı

Tek aşama, iki aşama ve üç aşama durumlarına göre öğrencilerin madde başarı oranları Şekil 5’de görülmektedir.



	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20	S21	S22	S23	S24	S25
Üç Aşama Birlikte	74	80	59	51	65	74	66	59	52	44	43	70	27	1%	3%	48	5%	8%	49	37	40	11	27	37	36
Birinci ve İkinci Aşama	82	89	70	61	72	79	71	66	64	48	61	73	43	2%	3%	54	6%	9%	65	43	55	14	38	42	39
Sadece Birinci Aşama	91	91	73	73	82	82	95	72	72	72	70	83	59	10	14	65	78	74	72	55	65	26	50	48	92

Şekil 5. KTT madde doğru cevaplandırma oranları

Şekil 5 incelendiğinde soru tek aşama olarak değerlendirildiğinde doğru cevaplandırma oranı en yüksek düzeydedir. 14, 15, 17, 18 ve 22. sorular iki aşama ve tüm aşamalar birlikte değerlendirildiğinde başarı oranı düşük olan sorular olduğu görülmektedir. KTT ile ilgili genel tanımlayıcı istatistikler SPSS paket programına veriler aktarıldıktan sonra elde edilmiştir. Testte üç aşama birlikte değerlendirildiğinde başarı ortalamasının tek aşamalı duruma göre düştüğü, KR-20 güvenilirlik düzeyinin arttığı görülmektedir (Tablo 14). Güvenirlik katsayısı 1’e yaklaştıkça, güvenilirlik düzeyi artar.

Tablo 14. KTT soruların güçlük ve ayırt edicilik değerleri

Madde No:	1.Aşama Puanı		Sadece 2. Aşama Puanı		Her iki Aşama Puanı		Total Puan	
	pj	rjx	Pj	rjx	pj	rjx	pj	rjx
1	0.89	0.23	0.75	0.47	0.76	0.44	0.68	0.58
2	0.91	0.14	0.85	0.23	0.85	0.23	0.79	0.25
3	0.72	0.49	0.69	0.47	0.69	0.47	0.59	0.61
4	0.71	0.51	0.59	0.58	0.59	0.58	0.49	0.63

Madde No:	1.Aşama Puanı		Sadece 2. Aşama Puanı		Her iki Aşama Puanı		Total Puan	
	pj	rjx	Pj	rjx	pj	rjx	pj	rjx
5	0.82	0.30	0.70	0.46	0.70	0.46	0.61	0.56
6	0.82	0.32	0.75	0.32	0.75	0.32	0.72	0.35
7	0.94	0.12	0.69	0.54	0.69	0.54	0.66	0.54
8	0.72	0.46	0.64	0.65	0.64	0.65	0.56	0.74
9	0.73	0.47	0.60	0.60	0.60	0.60	0.53	0.67
10	0.67	0.39	0.47	0.35	0.48	0.33	0.47	0.32
11	0.70	0.56	0.59	0.79	0.60	0.77	0.43	0.72
12	0.76	0.40	0.67	0.60	0.66	0.61	0.63	0.63
13	0.58	0.67	0.46	0.77	0.46	0.77	0.34	0.65
14	0.10	0.05	0.02	0.04	0.02	0.04	0.01	0.02
15	0.11	0.05	0.04	0.04	0.04	0.04	0.03	0.02
16	0.61	0.74	0.53	0.77	0.53	0.77	0.52	0.86
17	0.74	0.49	0.04	0.09	0.04	0.07	0.02	0.04
18	0.68	0.56	0.05	0.11	0.04	0.09	0.04	0.07
19	0.66	0.68	0.56	0.84	0.56	0.84	0.50	0.89
20	0.57	0.40	0.43	0.61	0.44	0.60	0.38	0.61
21	0.60	0.63	0.52	0.68	0.52	0.68	0.42	0.74
22	0.30	0.28	0.25	0.21	0.25	0.21	0.22	0.20
23	0.52	0.58	0.40	0.49	0.41	0.47	0.32	0.47
24	0.47	0.46	0.41	0.51	0.42	0.49	0.34	0.51
25	0.90	0.19	0.47	0.46	0.47	0.46	0.41	0.44
ORTALAMA	0.65	0.40	0.49	0.46	0.49	0.47	0.43	0.48

Verilerin analizi sonucunda her maddenin güçlük ve ayırt edicilik değerleri I. aşama, sadece II. aşama, her iki aşama puanı ve total puanlar için ayrı ayrı hesaplanmış Tablo 15'de sunulmuştur.

Tablo 15. Genel tanımlayıcı istatistikler

	Tek Aşama Puanı	İki Aşama Puanı	Üç Aşama Puanı
Öğrenci Sayısı	210	210	210
Madde Sayısı	25	25	25
Ortalama	16.64	12.49	10.66
Standart Sapma	4.002	4.539	4.788
Varyans	16.022	20.605	22.933
Çarpıklık	-.575	-.398	-.046
Basıklık	-.328	-.701	-1.050
Minimum	4.00	2.00	2.00
Maksimum	23.00	21.00	21.00
Güvenilirlik (KR-20)	0.78	0.81	0.83

Tablo incelendiğinde 14, 15, 17 ve 18. soruların rjx değerlerinin 0.20'nin altında olduğu, pj değerinin de çok düşük olduğu gözlenmektedir. Madde güçlük ve ayırt edicilik değerleri, dört ayrı şekilde hesaplandığında, yüzde değerleri açısından farklılıklar olmakla birlikte, 14 ve 15. maddelerin ilgili puanlarının dört puanlamada da düşük olduğu dikkat çekmektedir.

Veriler SPSS programına aktarıldıktan sonra, testin 1. aşama puanları dikkate alınarak testin toplamından alınan puan (sürekli değişken) ile testin her maddesinden alınan puan arasındaki ilişki çift serili korelasyon şeklinde hesaplandığında 14 ve 15. maddelerin korelasyon katsayı değerleri 0.30'un altındadır (Tablo 16). Bu maddeler, güçlük ve ayırt

edicilik hesaplamasında da düşük çıkararak elenen maddeler (Tablo 15) olduğu için, testten çıkarılmış, test madde sayısı 23'e düşürülmüştür.

Tablo 16. KTT 1.aşama puanları-maddelerin çift serili korelasyon katsayı değerleri (n=210)

Madde	p	Q	Yp	Yq	SS	pq/y	r _c	Madde	p	Q	Yp	Yq	SS	pq/y	r _c
1	0.91	0.09	16.99	13.11	0.29	0.504	0.49	14	0.10	0.90	15.40	16.77	0.29	0.513	-0.18
2	0.91	0.09	16.94	13.63	0.29	0.504	0.42	15	0.14	0.86	16.13	16.72	0.35	0.541	-0.08
3	0.73	0.27	17.69	13.82	0.45	0.596	0.58	16	0.65	0.35	18.50	13.22	0.48	0.614	0.81
4	0.73	0.27	17.72	13.74	0.45	0.596	0.59	17	0.78	0.22	17.71	12.91	0.42	0.58	0.70
5	0.82	0.18	17.26	13.84	0.39	0.563	0.48	18	0.74	0.26	17.85	13.13	0.44	0.593	0.70
6	0.82	0.18	17.24	13.84	0.38	0.563	0.48	19	0.72	0.28	18.14	12.71	0.45	0.599	0.81
7	0.95	0.05	16.85	12.40	0.21	0.461	0.51	20	0.55	0.45	17.81	15.22	0.50	0.625	0.40
8	0.72	0.28	17.72	13.81	0.45	0.599	0.58	21	0.65	0.35	18.14	13.82	0.48	0.614	0.66
9	0.72	0.28	17.82	13.55	0.45	0.599	0.64	22	0.30	0.60	17.70	16.27	0.44	0.593	0.21
10	0.72	0.28	17.56	14.29	0.45	0.599	0.49	23	0.50	0.50	18.32	14.95	0.50	0.627	0.53
11	0.70	0.30	17.92	13.70	0.46	0.604	0.64	24	0.48	0.52	18.14	15.25	0.50	0.626	0.45
12	0.83	0.17	17.56	12.17	0.38	0.568	0.77	25	0.92	0.08	16.89	13.76	0.27	0.495	0.39
13	0.59	0.41	18.47	14.00	0.49	0.622	0.69								

Son durumda hesaplanan KR-20 güvenilirlik katsayısı 0.80'dir. Üç madde birlikte değerlendirildiğinde testin güçlük ve ayırt edicilik değerleri ise sırasıyla pj: 0.47 ve rjx: 0.52'dir.

Tablo 17. Nihai KTT'de yer alan soru maddelerinin konulara ve kazanımlara göre dağılımı

Konular	Kazanımlar	Madde Numarası	Madde Sayısı
Canlıları Tanıyalım	5.5.1.1.	1,2,3,4,5,6,7,9, 18, 23, 25	11
İnsan ve Çevre İlişkisi	5.5.2.1.	8, 10, 11,12,13, 17, 19,20,21, 22, 24	11
	5.5.2.2.	16	1

Sonuç

Bu çalışma sonucunda 5.sınıf öğrencilerinin *Canlılar Dünyasını Gezelim ve Tanıyalım* ünitesi kazanımlarına ne düzeyde ulaştıklarını ve ünite ile ilgili ne tür kavram yanlışlarına sahip olduklarını tespit edebilmek amacıyla kullanılacak üç aşamalı başarı testi geliştirilip, ölçeğin geçerlik ve güvenilirlik çalışmaları tamamlanmıştır. Başlangıçta 40 madde olarak geliştirilen teste, geçerlik ve güvenilirlik çalışmaları sonucunda 23 madde içerecek şekilde son şekli verilmiştir.

Test güvenilirliği KR-20 formülü ile hesaplanmış; KR-20 güvenilirlik düzeyinin ,80 olduğu tespit edilmiştir. Üç madde birlikte değerlendirildiğinde testin güçlük ve ayırt edicilik değerleri sırasıyla pj: 0.47 ve rjx: 0.52 olarak hesaplanmıştır. Bu veriler geliştirilen

testin güvenilir, orta güçlükte ve çok iyi maddelerden oluşan bir yapıya sahip olduğunu göstermektedir.

Testte test maddeleri tek aşama, iki aşama ve üç aşama birlikte değerlendirilerek ayrı ayrı başarı ortalaması hesaplanmıştır. Başarı ortalaması üç aşama birlikte değerlendirildiğinde tek aşamalı duruma göre düşmesine rağmen, KR-20 güvenilirlik düzeyinin arttığı tespit edilmiştir (Şekil 5). Bu sonuç elde edilen testin güvenilirliğini bir kez daha ortaya koymaktadır.

Tek aşamalı başarı testleri araştırmacı ya da öğretmenlere, kavramın öğrenilme düzeyini belirleyebilmeleri konusunda bilgi vermektedir. Ancak araştırma sonucu geliştirilen üç aşamalı test, araştırmacı ya da öğretmenlere kavramın öğrenilme düzeyini tespit etmenin yanı sıra, bilimsel bilgi, kavram yanılığı, bilgi eksikliğine sahip olma ya da sorunun şans eseri cevaplanıp cevaplanmadığını tespit etme fırsatı sunmaktadır. Bu veriler de ilgili kişiler için daha fazla yol gösterici olacaktır.

Öneriler

- Yapılan analiz işlemleri sonucunda geçerliği ve güvenilirliği ortaya konulan kavram tanı testi, 5.sınıf öğrencilerinin “Canlılar Dünyasını Gezelim ve Tanıyalım” ünitesi kazanımlarına ulaşip ulaşmadıklarını ve hangi kavram yanılıklarına sahip olduğunu belirlemek için kullanılabilir özelliktedir.
- Kavram tanı testi araştırmacıların tek bir ölçekle iki ayrı veri seti (akademik başarı ve kavram bilgisi) elde etme kolaylığı sağlayacaktır.
- Araştırmacılar, akademik başarı belirleme sürecinde ölçeğin sadece birinci aşama sorularını dikkate alabilecekleri gibi, üç aşamayı da dikkate alarak değerlendirme yapabilir. Kavram yanılığı belirlemek amacıyla ölçekten yararlanacak araştırmacılar ise üç aşamayı birlikte değerlendirmelidir.
- Çalışma, üç aşamalı başarı testi/kavram tanı testi geliştirecek araştırmacılar için yol göstericidir.

Bilgilendirme

Bu çalışma Ondokuz Mayıs Üniversitesi Bilimsel Araştırma Projeleri kapsamında PYO.EGF.1904.14.006 no’lu proje olarak desteklenen ve birinci yazarın doktora tezinin bir parçasından oluşmaktadır.

Kaynaklar

- Allen, M. (2010). *Misconceptions in primary science*. Berkshire: Open University Press.
- Arslan, H. O., Çiğdemoğlu, 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
- Bağcan-Büyükturan, E. & Çıkrıkçı Demirtaşlı, S. (2012). Çoktan seçmeli testler ile yapılandırılmış gridlerin psikometrik özellikleri bakımından karşılaştırılması. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 45 (1), 395-415.
- Bakioğlu, B., Karamustafaoğlu, S. & Karamustafaoğlu, O. (2014, Mayıs). 5. sınıf "Vücudumuzun bilmecesini çözelim" ünitesi başarı testi: geçerlik ve güvenilirlik. Konya: International Conference On Education in Mathematics, Science & Technology- (ICEMST 2014) bildiri kitabı, 16-18 Mayıs, Konya.
- Bernhisel, S. M. (1999). *Measuring preservice and inservice biology teachers' understanding of selected biological concepts*. Unpublished Dissertation. Utah: Utah State University
- Büyüköztürk, Ş. (2010). *Sosyal Bilimler için veri analizi el kitabı* (12. bs.) Ankara: Pegem Akademi.
- Büyüköztürk, Ş., Çokluk, O. & Köklü, N. (2010). *Sosyal bilimler için istatistik* (6.bs.). Ankara: Pegem Akademi.
- Chen, C.C., Lin, H.S., & Lin, M.L. (2002). *Developing a two-tier diagnostic instrument to assess high school students' understanding-the formation of images by a plane mirror*. Proceedings of the National Science Council, 12 (3), 106-121.
- Çalık, M. & Ayas, A. (2003). Çözümlerde kavram başarı testi hazırlama ve uygulama. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 14 (2), 1-17.
- Çelik, D. (2000). *Okullarda ölçme ve değerlendirme nasıl olmalı?* İstanbul: Milli Eğitim Basımevi.
- Çepni, S., Baki, A., Ayas, A., Demircioğlu, G., & Akyıldız, S. (2009). *Ölçme ve değerlendirme*. Trabzon: Celepler Matbaacılık.
- Çetinkaya, M. (2010). *Canlıların sınıflandırılması konusu için web destekli kavram haritaları ve anlam çözümleme tablolarının öğrenme üzerindeki etkisinin araştırılması* (Yayımlanmamış yüksek lisans tezi), Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü, Samsun.
- Gönen, S., Kocakaya, S. & Kocakaya, F. (2011). Dinamik konusunda geçerliliği ve güvenilirliği sağlanmış bir başarı testi geliştirme çalışması. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 8 (1), 40-57.
- Karataş, F. Ö., Köse S. & Coştu, B., (2003). Öğrenci yanılgılarını ve anlama düzeylerini belirlemede kullanılan iki aşamalı testler. *Pamukkale Eğitim Fakültesi Dergisi*, 1 (13), 54-69.
- Mutlu, M. & Tokcan, H. (2012). İlköğretim 7. sınıf öğrencilerinin toprak kirliliği hakkındaki düşünceleri. *International Journal of Social Science Research*, 1 (1), 65-75.
- Novak, J. D. (1998). Metacognitive strategies to help students learning how to learn. *Research Matters-to the Science Teacher*, No.9802, Mart National Association for Research in

- Science Teaching. Erişim adresi:
<https://www.narst.org/publications/research/Metacogn.cfm>
- Odom, A.L., & Barrow, L.H. (1995). Development and application of a two tier diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction, *Journal of Research in Science Teaching*, 32 (1), 45-61.
- Öksüz, Y., & Güven Demir, E. (2019). Açık uçlu ve çoktan seçmeli başarı testlerinin psikometrik özellikleri ve öğrenci performansı açısından karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 34(1), 259-282. doi: 10.16986/HUJE.2018040550
- Özçelik, D. A. (2010). *Test hazırlama kılavuzu* (4. baskı). Ankara: PegemA Yayıncılık.
- Özyılmaz-Akamca, G. (2008). *İlköğretimde analogiler, kavram karikatürleri ve tahmin-gözlem-açıklama teknikleriyle desteklenmiş fen ve teknoloji eğitiminin öğrenme ürünlerine etkisi* (Yayımlanmamış doktora tezi). Dokuz Eylül Üniversitesi Eğitim Bilimleri Enstitüsü, İzmir.
- Peşman, H. (2005). *Development of a three-tier test to assess ninth grade students' misconceptions about simple electric circuits* (Yayımlanmamış yüksek lisans tezi). Orta Doğu Teknik Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
- Peşman, H. & Eryılmaz, A. (2010). Development of a three-tier test to assess misconceptions about simple electric circuits. *Journal of Educational Research*, 103 (3), 208-222.
- Saka, A., Ayas, A. & Enginar, İ. (2002, Eylül). *Öğrencilerin omurgalı-omurgasız canlılar ile ilgili görüşlerinin yaşlara göre değişimi*. 5. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi. 16-18 Eylül, Ankara: ODTÜ.
- Sesli, E. & Kara, Y. (2012). Development and application of a two-tier multiple-choice diagnostic test for high school students' understanding of cell division and reproduction, *Journal of Biological Education*, 46 (4), 214-225.
- Sözbilir, M. (2010). *Madde analizi ve test geliştirme*. Erişim adresi:
<https://olcmevedegerlendirme.files.wordpress.com/2010/09/7-madde-analizi-ve-test-gelistirme.pdf>
- Summers, M., Kruger, C., Childs, A. & Mant, J. (2000). Primary school teachers' understanding of environmental issues: An interview study. *Environmental Education Research*, 6 (4), 293-312.
- Tan, K-C. D., Goh, N-K., Chia, L. S. & Treagust, D. F. (2002). Development and application of a two-tier multiple-choice diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis, *Journal of Research in Science Teaching*, 39 (4), 283-301.
- Taş, E., Aymen Peker, E. & Çetinkaya, M. (2014). Determining of misconceptions by means of tree-tier test about the unit human and environment. *Proceedings of 2nd International Conference on Economics and Social Sciences*, 85-90.
- Tavşancıl, E. (2010). *Tutumların ölçülmesi ve SPSS ile veri analizi*. Ankara: Nobel Yayın.
- Tekkaya, C., Çapa, Y. & Yılmaz, Ö. (2000). Biyoloji öğretmen adaylarının genel biyoloji konularındaki kavram yanlışları. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 18, 140-147.

- Yağbasan R. & Gülçiçek Ç. (2003). Fen öğretiminde kavram yanlışlarının karakteristiklerinin tanımlanması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 13, 110 –128.
- Yen, C-F., Yao, T-W. & Mintzes, J-J. (2007). Taiwanese students' alternative conceptions of animal biodiversity. *International Journal of Science Education*, 29 (4), 535–553.