



Two-Stage Test Development Study Supported by Concept Cartoon on Oral and Dental Health¹

Sena KURTOĞLU GÜLDALI², Fethiye KARSLI BAYDERE³

Research Article

Received: 08 December 2022

Accepted: 26 June 2023

ABSTRACT

The aim of this research is to develop a concept test supported by concept cartoons for the subject of "oral and dental health" in the 9th grade health information lesson and to conduct validity-reliability studies of this test. The two-stage test was applied to a total of 268 9th grade students studying at two different high schools in Giresun city center. There are a total of 23 items in the two-stage test. Reliability, validity and item analysis studies were conducted for the test. KR-20 and reliability coefficient (Cronbach Alpha) values were calculated for the reliability analysis of the two-stage test. As a result of the analyzes, the KR-20 value was calculated as 0.807 and the Cronbach Alpha value as 0.808. As a result of the item analysis, it was seen that the item difficulty (p) was 0.40, and the discrimination index (d) was 0.42. It has been concluded that the two-stage test developed in the light of these values is a valid and reliable concept test that can be used as an alternative to traditional tests and to determine the alternative concepts of 9th grade students about oral and dental health.

Keywords: Oral and dental health, Two-stage concept test, Test development

Ethical Committee Date / Number : Giresun University Ethical Committee , 01 September 2021 , No:14/14

EXTENDED ABSTRACT

Purpose and Significance

Oral and dental health is one of the basic subjects of science and health science courses. Especially school-age children have many health problems such as oral and dental diseases (Ceylan & Turan, 2003). These problems may not be seen as important by both school-age children and parents (Ceylan & Turan, 2003). In addition, it is a matter that is accepted by everyone how much weight should be given to the issue of health information during the pandemic process we live in. In order to prevent health problems from continuing to increase, the need to increase the number of researches to investigate and prevent the conditions that cause diseases in children at school age is increasing day by day. Detailed information on oral and dental health is included in the 9th grade textbook (MEB, 2018). As a matter of fact, it is thought that detailing the subject of oral and dental health at such a late learning level will lead to incomplete learning in individuals. It can also be a prerequisite for understanding many science subjects such as the mouth and teeth and their health, the digestive system. In addition, when the relevant literature is examined, no attempt has been made to develop a validated and reliable test that teachers can easily access on this subject. In addition, no research has been found in which misconceptions about oral and dental health are defined comprehensively. In this context, it is thought that there is a need for a measurement tool that can determine students'

¹ This research was produced from the master thesis of Sena KURTOĞLU GÜLDALI under the supervision of Fethiye KARSLI BAYDERE.

² Master student, Giresun University, Graduate School of Natural and Applied Sciences, sena.krtgl@gmail.com  0000-0002-9803-6591

³ Assoc. Prof. Dr., Giresun University, Graduate School of Natural and Applied Sciences, fethiyekarsli28@gmail.com  0000-0003-0994-0974

conceptual understanding of oral and dental health and their alternative concepts, if any. In this context, it is believed that this research will fill this gap in the literature.

The aim of this research is to develop a concept test supported by concept cartoons for the subject of "oral and dental health" in the 9th grade health information lesson and to conduct validity-reliability studies of this test.

Methods

This research is a study that includes the development of a test for oral and dental health and the validity, reliability and item analysis of the developed test. When we look at test development researches, processes such as preparation, application and reporting are completed (Çalıřkan & Kaplan, 2009). In this study, in a general order; preparation of test questions, pilot implementation, validity-reliability analysis and finally test development steps in which item analysis was carried out (Burns Okey & Wise, 1985; Karalı & Ayas 2013; Temiz, Tařar & Tan, 2006; Feyziođlu, Demirdađ, Akyıldız & Altun, 2012; Karalı-Baydere & Yigit, 2020). In addition, during the test development and validation process, certain procedures were used to create the two-stage multiple choice format (Treagust, 1985, 1988, 1995). In the 2021-2022 academic year, the two-stage test developed on "Oral and dental health" was piloted to 9th grade students consisting of 60 students after the test items were written. The reliability analysis of the test was made based on the data of 208 9th grade high school students, independent of the pilot application group of 60 people. The application of the measurement tool was carried out during the normal course hours during the academic year. These students are students from different high schools with different achievement levels and socioeconomic backgrounds. In these schools, all students take biology and health science lessons in the 9th grade. Test development and validation procedures generally include determining the target of the scale, defining the content boundaries of the targeted concepts, identifying original alternative concepts related to the targeted concepts, writing the test items and pilot testing (Treagust, 1985, 1988, 1995). In this section, the application of the steps to the test developed within the scope of the research is presented respectively.

Results, Discussion and Conclusions

The two-stage test, which was developed in 24 items, was presented to the opinion of experts consisting of two science educators, a biology teacher and a dentist. Necessary adjustments were made in the test in the light of the feedback received from the experts. In the light of the dentist's opinion, item 24 was removed from the test. Before the pilot application, the developed test was taught to two students other than the sample, and it was desired to determine whether there were any places that were difficult to understand. In the light of the feedback received, it was determined that there was no problem in the intelligibility of the test. In the light of all these regulations, the test has been made ready for application. The validity analysis of the developed test was carried out with a total of 60 students studying in the 9th grade. As a result of the feedback received from the application group, the defects detected were corrected and the test was re-examined by field and field education experts. In the light of the feedback received, the experts stated that the test items were appropriate for the age and grade levels of the students, the oral and dental health subject was sufficient to measure the achievements, and there were no scientific errors in the question roots and distractors. The mean item difficulty of the developed test was calculated as 0.40. This means that the item difficulty index of the test is of medium difficulty. The mean discrimination index of the developed test was calculated as 0.42. When the discrimination index results of the two-stage test are examined, it is seen that the discrimination power of the test is very good.

When the results of the reliability study of the research were examined, KR-20 and Cronbach alpha values were calculated for the internal consistency reliability of the two-stage test. As a result of the data analysis, the reliability coefficient (Cronbach Alpha) of the test was calculated as 0.808. This value calculated in educational research is interpreted as the test is reliable (Bykztrk, 2013). In addition, the KR-20 value of the test was calculated as 0.807. Tests with a reliability coefficient above 0.70 are described as tests with high reliability (Bykztrk, 2007; Bykztrk, Kılıç Çakmak, Akgn, Karadeniz, & Demirel, 2018; Bykztrk, 2013; Erkuř, 2006; zelik, 2010). Based on this, it was concluded that the test developed within the scope of the research was reliable based on the reliability coefficient and KR-20 value.

Ağız ve Diş Sağlığı Konusuna Yönelik Kavram Karikatürü Destekli İki Aşamalı Test Geliştirme Çalışması¹

Sena KURTOĞLU GÜLDALI², Fethiye KARSLI BAYDERE³

Araştırma Makalesi

Başvuru Tarihi: 08 Aralık 2022

Kabul Tarihi: 26 Haziran 2023

ÖZET

Bu çalışmada amaç, 9. Sınıf sağlık bilgisi dersi "ağız ve diş sağlığı" konusuna yönelik kavram karikatürü destekli bir kavram testi geliştirmek ve bu testin geçerlik-güvenirlik çalışmalarını yapmaktır. İki aşamalı test Giresun il merkezinde iki farklı lisede öğrenim gören toplam 268 9. Sınıf öğrencisine uygulanmıştır. İki aşamalı testte toplam 23 madde bulunmaktadır. Test için güvenirlik, geçerlik ve madde analizi çalışması yapılmıştır. İki aşamalı testin güvenirlik analizi için KR-20 ve güvenirlik katsayısı (Cronbach Alpha) değerleri hesaplanmıştır. Analizler sonucunda KR-20 değeri 0,807 ve Cronbach Alpha değeri 0,808 olarak hesaplanmıştır. Yapılan madde analizi sonucunda madde gücünün (p) 0.40, ayırtedicilik indeksinin (d) ise 0,42 olduğu görülmüştür. Bu değerler ışığında geliştirilen iki aşamalı testin 9. Sınıf öğrencilerinin ağız ve diş sağlığı konusu ile ilgili alternatif kavramlarını belirlemede ve geleneksel testlere alternatif olarak kullanılabilecek geçerliğe ve güvenirliğe sahip bir kavram testi olduğu sonucuna ulaşılmıştır.

Anahtar Kelimeler: : Ağız ve diş sağlığı, İki aşamalı test, Test geliştirme

Etik Kurul İzni Tarih / Sayı: Giresun Üniversitesi Etik Komisyonu, 01Eylül 2021, No: 14/14


1. Giriş


Bilim ve Teknolojinin gelişmesinin, yaşamın her noktası üzerinde olumlu etkiler bıraktığı bir gerçektir. Bilim ve teknolojinin bu gelişiminde ise, fen bilimleri eğitimi önemli bir rol oynamaktadır. Fen bilimleri eğitiminin temelinde fen okuryazarı bireyler yetiştirmek amaçlanmıştır (Milli Eğitim Bakanlığı [MEB], 2018). Öğrencilerin fen okuryazarı olarak yetiştirilmesinin altında yatan hususlar biliş, duyuş ve psikomotor bakımından öğrencilerin ön öğrenmelerinin sağlanmasıdır. Aynı zamanda kazandırılması hedeflenen bu becerilerin bireylerde ne derecede anlam kazandığının tespit edilmesi, eksik ya da yanlış öğrenmelerinin ortadan kaldırılması ve bireye geri dönütler sunulması oldukça önemlidir (Açıkgöz ve Karlı 2015; Karaca, 2008; Erdoğan ve Kurt, 2012). Bahsedilen tüm bu durumlar eğitimde ölçme ve değerlendirmenin kullanılmasıyla çözüme ulaşır (İlhan ve Hoşgören, 2017). Nitekim ölçme ve değerlendirme işlemi eğitimin vazgeçilmez bir basamağıdır (MEB, 2018; Erden, 1998; Burke, 2005; Hlebowitsh, 2005; Oliva, 2005; Sönmez, 2005; Açıkgöz ve Karlı 2015).

MEB'e göre ölçme ve değerlendirme işlemi öğretim programının kazanımlarına ve kısıtlamalara uyularak gerçekleştirilmelidir (MEB, 2018). Buradan yola çıkılarak geliştirilen testler, MEB'in belirlediği kazanımları içeren, bilişsel becerilerin ölçüldüğü aynı zamanda alternatif yöntem ve tekniklerin tercih edildiği bir ölçme ve değerlendirme araçlarıdır. Eğitimde ölçme ve değerlendirmede kullanılan alternatif tekniklere tanılayıcı dallanmış ağaç, yapılandırılmış grid, anlam çözümleme tablosu, kavram haritası ve kavram karikatürü gibi teknikler örnek verilebilir (Karlı, 2015). Testlerde bu alternatif tekniklerin kullanılmasındaki amaç; her öğrencinin ön bilgileri ve kişisel özellikleri aynı değildir, işte bu farklılığı (MEB, 2018; Özmen, 2004) bahsi geçen alternatif tekniklerin kullanılması ile gidermeye çalışmaktır. Eğitimde kullanılan kavram testleri, fen eğitimi alanında da oldukça fazla tercih edilen bir ölçme ve değerlendirme aracıdır.

Her öğrencinin okul dışı yaşamda soyut fen kavramları hakkında edindikleri yanlış düşünce ve inançlara sahip oldukları ve formal eğitime bu düşünce ve inanışlar ile geldikleri yapılan çalışmalarda tespit edilmiştir (Amir ve Tamir, 1994). Bununla birlikte daha önceden yapılan çalışmalar (örneğin, Wandersee,

¹ Bu araştırma Sena KURTOĞLU GÜLDALI'nın Fethiye KARSLI BAYDERE danışmanlığında yaptığı yüksek lisans tezinden üretilmiştir

² Yüksek Lisans Öğrencisi, Giresun Üniversitesi, Fen Bilimleri Enstitüsü, sena.krtgl@gmail.com  0000-0002-9803-6591

³ Doç.Dr., Giresun Üniversitesi, Fen Bilimleri Enstitüsü, fethiyekarsli28@gmail.com  0000-0003-0994-0974

Mintzes ve Novak, 1994), đrencilerin fen alanında formal đrenme ortamlarına girdiklerinde fen kavramları ve olaylar hakkında ok eřitli alternatif kavramları barındırdıklarına iřaret etmiřtir. Ayrıca, bu kavramların kkeni, đrencilerin gzlem, algı, kltr ve dil farklılıklarından oluřabilmekte, bununla birlikte nceki đretmenlerin aıklamaları ve nceki đretim materyallerini ieren eřitli kiřisel deneyimleri de bunun sebepleri olarak grlebilmektedir. Literatrde, kiřinin nceki kavram yanılıđlarını gidermenin veya bu yanılıđların giderilerek yeniden yapılandırmanın kolay olmadıđını gsteren nemli kanıtlar da vardır. đrenciler, eđer ki gerekli ve uygun mdahale yapılmaz ise bu alternatif kavramlara inatla tutunmaktadır. Tm bu yanılıř n bilgiler, formal đretimde sunulan her Őeyle etkileřime girebilir ve đrenciler tarafından eřitli istenmeyen durumlarla sonulanabilir.

Ausubel (1968), anlamlı bir Őekilde đrenmek iin bireylerin yeni bilgileri halihazırda bildikleri ilgili kavram ve nermelerle iliřkilendirmeyi Őemeleri gerektiđine iřaret etmiřtir. Anlamlı đrenmede bu sreci kolaylařtırmanın yolu đretmenlerin, đrencilerinin kavramlarının bařlangı noktasını bilmeleri, bunu tespit etmeleri ve buna gre yeni Őeyler đretmeleri veya đrenmeleridir. đrencilerin kavramları hakkında derinlemesine bilgi sahibi olmanın yolu onlarla bireysel mlakatlar yapılmasıdır (Wandersee ve diđerleri, 1994). Ancak bireysel mlakat yapmak, bunları transkript etmek ve yorumlamak ok fazla zaman aldıđı iin, her ne kadar bireysel mlakatlar deđerli olarak grlse de đretmenler iin bunları yrtmek neredeyse imkansızdır. Bu nedenle đretmenlerin đrencilerin n bilgilerini belirlemede daha kullanılıř ve verimli aralar tercih etmesi zorunluluk olarak deđerlendirilebilir. Bununla ilgili olarak đrencilerin bilimsel kavramları anlamalarını arařtırmak iin farklı deđerlendirme prosedrleri savunulurken, đrenciler tarafından dođru kabul edilen hatalı bilgilerle ilgili bilgilerin eldirici ieren testlere dahil edilmesi eřitli arařtırmalarda tavsiye edilmiřtir (rneđin, Lin, 2004). Bu tavsiyeler bizi đrencilerin kavramlarına dayalı olarak eldiricileri olan oktan Őemeli test maddelerinin geliřtirilmesine dođru itmektedir. rneđin Treagust (1988), Haslam ve Treagust (1987), Odom ve Barrow (1995), (Lin, 2004) ve daha birok arařtırmacı, bilinen alternatif kavramlar da dahil olmak zere đrenci muhakemesine dayalı sorulara alternatif oktan Őemeli maddeler kullanmıřlardır. Bu arařtırmacılar đrencilerin kavramlarını belirlemek iin alternatif oktan Őemeli araların geliřtirilmesinin, alternatif deđerlendirme alanına nemli bir katkı yapma potansiyeline sahip olduđunu belirtmiřlerdir. rneđin, eđitmenlerin đrencilerinin kavramsal durumlarına iliřkin farkındalıđını artırabilir; đeler grup tartıřmaları iin kullanılabilir; mfredat revizyonu iin faydalı bilgiler sađlayabilir; araların geliřtirilmesi, sınıfta kolayca kullanılabilir arařtırma bulgularını ierir.

đrencilerin yanılıř dřnce ve inanların tespiti noktasında kullanılması uygun olan tekniklerden biri de kavram karikatrleridir (Erdođan ve zsevge, 2012). Kavram karikatrleri daha ok alternatif kavramların tespit edilmesinde kullanılan bir alternatif tekniktir (Canpolat ve Pınarbařı, 2011; Karatař, Kse ve Cořtu, 2003; Treagust, 1988). Naylor ve McMurdo (1990), kavram karikatrlerini,  veya daha fazla izgi karakterin tartıřmalarının resmedilmesi Őeklinde ilk kez tanımlamıřtır. Fen eđitimi boyutundaki kavram karikatr ise, gnlk hayattaki olaylara karřı farklı dřncelerin savunulduđu ve ilgi ekici karikatrlerin kullanıldıđı resim veya izimler olarak tanımlanabilir (Keogh ve Naylor, 1999a; Martinez, 2004; İnel, Balım ve Evrekli, 2009). Kullanılan karakterler, bilimsel olarak kabul gren ve bilimsel olarak kabul grmeyen fikirleri (Kabapınar, 2007) ve aynı zamanda her biri farklı fikirleri savunacak Őekilde oluřturulur. Eđitimde, karikatrler đrencilere sorgulatma becerisinin kazandırılmasında ve var olan alternatif kavramların belirlenmesi iin kullanılır (Keogh ve Naylor, 1999b). Kavram karikatrlerinde đrenciler karakterlerin savunduđu fikirlerden bir veya birden fazlasına katılım gsterir. Eđer bu fikir bilimsel dođrulara aykırı ise bu durum đrencide var olan alternatif kavramına iřaret eder. Bylece kavram karikatr tekniđi yardımıyla đrencilerin alternatif kavramlarının tespit edilmesi sađlanır (Keogh ve Naylor, 1997; 1999; Martinez, 2004; İnel, Balım ve Evrekli, 2009; Kabapınar, 2005). Bu bađlamda literatrdeki ilgili alıřmalar incelendiđinde, kavram karikatrlerinin ođunlukla alternatif kavramlarının ve kararsızlıkların tespitinde kullanıldıđı grlmektedir (İnel, Balım ve Evrekli, 2009; Demir, Uzođlu ve Bykkařap, 2012; Ekici, Ekici ve Aydın, 2007; Atasoy ve Akdeniz, 2009; Ekim, 2007). Aynı zamanda kavram karikatr tekniđinin đrencilerin derse karřı dikkatlerini artırdıđı ve motivasyonlarını ykselttiđi ynnde arařtırmalar da bulunmaktadır (rneđin, Naylor ve Keogh, 2013). Fen bilimlerinde yaygın olarak alternatif kavramların tespit edilmesinde nerilen bir diđer teknik ise iki ařamalı testlerdir (Aykutlu ve Ően, 2012; etinkaya ve Tař, 2016; Can Polat ve Pınarbařı, 2011; Treagust, 1998, Tan, Goh, Chia ve Treagust, 2002; Voska ve Heikkinen, 2000; Tyson, Treagust ve Bucat, 1999; Mann ve Treagust, 1998; Odom ve Barrow, 1995; Garnett ve Treagust, 1992; Peterson, Treagust ve Garnett, 1989; Haslam ve Treagust, 1987). İki ařamalı testlerin birinci ařaması genellikle oktan Őemeli yapıda olup, eldiricilerin ve bilimsel

dođrunun olduđu seeneklerden oluřmaktadıř. İkinci ařama ise ođrencilere ilk ařamada verdikleri cevaplara aıklama yapma imkanı sunar. Bu ařama ođrencilerin verdikleri dođru cevabı řans eseri verip vermediklerinin ve yaptıkları aıklama ile varsa alternatif kavramının tespit edildiđi bir ařamadır. İki ařamalı testlerin farklı trleri bulunur bu trler ve ierikleri ařađıda Tablo1 de sunulmuřtur.

Tablo 1

İki ařamalı test trleri

Trler	1. Ařama	2. Ařama
İki ařaması da oktan semeli testler	oktan semeli	oktan semeli (+Aık ulu bir seenek olabilir.)
Aık ulu iki ařamalı testler	oktan semeli	Aık ulu
Sınıflamaya dayalı iki ařamalı testler	Dođru ve yanlış seenekli	oktan semeli (+Aık ulu bir seenek olabilir.)

Ađız ve diř sađlıđı konusu fen bilimleri dersi ve sađlık bilgisi derslerinin temel konularından birisidir. Özellikle okul ađı ocuklarının ađız ve diř hastalıkları gibi birok sađlık sorunları bulunmaktadır (Ceylan ve Turan, 2003). Bu sorunlar gerek okul ađı ocukları gerekse ebeveynler tarafından gerektiđi kadar önemli grlmeyebilmektedir (Ceylan ve Turan, 2003). Ayrıca yařadığımız pandemi srecinde sađlık bilgisi konusuna ne kadar ađırlık verilmesi gerektiđi herkese kabul gren bir konudur. Sađlık sorunlarının daha da artarak devam etmemesi iin ocuklara okul ađlarında hastalıklara sebep olan durumların arařtırılmasına ve engellenmesine ynelik arařtırmaların sayısının artmasına olan ihtiya gn getike artmaktadır. Eđitim sisteminde bu konuya ne kadar yer verildiđi konusunda okul kitapları incelendiđinde ađız ve diř sađlıđı ile ilgili; 5. sınıfta sadece diřlerin yapısının ođretildiđi (MEB, 2018), ađız ve diř sađlıđı zerinde durulmadığı, 7. sınıf sindirim sistemi konusunda da sindirim sisteminin bařlangıcı ađız olmasına rađmen ders kitaplarında bu konuya iliřkin gerekli nemin verilmediđi dikkat ekmektedir (MEB, 2018). Ađız ve diř sađlıđı konusuna ynelik detaylı bilgi 9. Sınıf ders kitabında yer almaktadır (MEB, 2018). Nitekim ađız ve diř sađlıđı konusunun bu kadar ge bir ođrenim seviyesinde detaylandırılmasının bireylerde eksik ođrenmelere yol aacađı dřnlmektedir. Ađız ile diřler ve bunların sađlıđı, sindirim sistemi gibi birok fen konularını anlamının da n kořulu olabilir. Ayrıca ilgili literatr incelendiđinde ođretmenlerin kolayca eriřebileceđi geerliliđi ve gvenirliđi sađlanmış test geliřtirmeye ynelik herhangi bir giriřimde bulunulmamıřtır. Buna ek olarak, ađız ve diř sađlıđı konusunda kapsamlı bir Őekilde alternatif kavramlarının tanımlandığı arařtırmaya da rastlanılmamıřtır. Bu bađlamda ođrencilerin ađız ve diř sađlıđı konusunda kavramsal anlama dzeylerini ve varsa alternatif kavramlarını belirleyebilecek bir lme aracına ihtiya duyulduđu dřnlmektedir. Bu kapsamda arařtırmanın literatrdeki bu bořluđu dolduracađına inanılmaktadır.

Bu alıřma kapsamında geliřtirilen ađız ve diř sađlıđı kavram testi (ADSKT) iki ařamalı yapıda hazırlanmıřtır. Buradaki ama iki ařamalı test tekniđi sayesinde ilk ařamada řans faktrne karřı ikinci ařamadaki aık ulu gereke kısmında ođrencinin verdiđi cevabın bilimsel olarak ne lde dođru olduđunun ve var olan alternatif kavramlarının belirlenmesini sađlamaktır. Bununla birlikte hazırlanan kavram testinde madde kkleri kavram karikatr Őeklinde oluřturulmuřtur. Soru kklerinin kavram karikatr Őeklinde oluřturulması sayesinde ođrencilerin test maddelerine karřı dikkatinin ekilerek, motivasyonlarının arttırılmasının sađlanması hedeflenmektedir. Bir lme aracının geliřtirilmesi, ođrencilere kazandırılması hedeflenen kazanımların ne lde kazandırılabilirdiđinin tespit edilmesi aısından olduka önemlidir. Bu da yapılan arařtırmanın nedenini oluřturmaktadır (İlhan ve Hořgren, 2017). Ayrıca geliřtirilen testin ađız ve diř sađlıđı konusu ile iliřkili yapılacak alıřmalarda katkı sađlayacađı dřnlmektedir.

1.1. Arařtırmanın amacı

Bu arařtırmada ama, 9. Sınıf “ađız ve diř sađlıđı” konusuna ynelik kavram karikatr destekli bir kavram testi geliřtirmek ve bu testin geerlik-gvenirlik alıřmalarını yapmaktır.

2. Yöntem

Araştırmanın yöntemi tarama yöntemidir. Bu kapsamda araştırmada, ağız ve diş sağlığı konusuna yönelik bir test geliştirme ve geliştirilen testin geçerlik, güvenirlik ve madde analizini içeren bir çalışmadır.

2.1. Etik Kurul İzni

Giresun Üniversitesi Rektörlüğü

Sosyal Bilimler Fen ve Mühendislik Bilimleri Araştırmaları Etik Kurulu

Sayı :E-50288587-050.01.04-42559 13.09.2021

Konu :01 Eylül 2021 tarih ve 14/14 sayılı Etik Kurul Kararı

2.2. Araştırmanın Modeli

Test geliştirme araştırmalarına bakıldığında hazırlık, uygulama ve rapor etme şeklinde süreçler tamamlanmaktadır (Çalışkan ve Kaplan, 2009). Bu çalışmada ise genel bir sıralamayla; test sorularının hazırlanması, pilot uygulama, geçerlik-güvenirlik analizi ve son olarak madde analizinin gerçekleştirildiği test geliştirme adımları esas alınmıştır (Burns Okey ve Wise, 1985; Karlı ve Ayas 2013; Temiz, Taşar ve Tan, 2006; Feyzioğlu, Demirdağ, Akyıldız ve Altun, 2012; Karlı-Baydere ve Yiğit, 2020). Ayrıca test geliştirme ve doğrulama sürecinde, iki aşamalı çoktan seçmeli formatı oluşturmak için belirli prosedürlerden de faydalanılmıştır (Treagust, 1985, 1988, 1995).

2.3. Çalışma Grubu

2021-2022 eğitim öğretim yılında "Ağız ve diş sağlığı" konusunda geliştirilen iki aşamalı test, test maddelerinin yazılmasının ardından 60 kişiden oluşan 9. sınıf öğrencilerine pilot olarak uygulanmıştır. Testin güvenirlik analizi 60 kişilik pilot uygulama grubundan bağımsız olarak 208 lise 9. sınıf öğrencisinin verilerine dayalı olarak yapılmıştır. Ölçme aracının uygulaması öğretim yılı içerisinde bir ders saatinde gerçekleştirilmiştir. Uygulamaya katılan tüm öğrenciler 9. sınıfta biyoloji dersi ve sağlık bilgisi derslerini almaktadırlar.

2.4. İki Aşamalı ADSKT'nin Geliştirilmesi

Test geliştirme ve doğrulama prosedürleri genel olarak, ölçeğin hedefinin belirlenmesi, hedeflenen kavramların içerik sınırlarının tanımlanmasını, hedeflenen kavramlarla ilgili özgün alternatif kavramların belirlenmesini, testin maddelerinin yazılmasını ve pilot testini içermektedir (Treagust, 1985, 1988, 1995). Bu kısımda adımların araştırma kapsamında geliştirilen teste uygulanışları sırasıyla sunulmuştur.

2.4.1. Ölçeğin amacının belirlenmesi:

Öncelikle testin kullanım amacı belirlenmiştir. Ölçeğin geçerliği bakımından, ölçmede kullanılacak testin amaca uygun olması oldukça önemlidir (Fraenkel ve Wallen, 2009). Literatür taramaları sonucunda 'ağız ve diş sağlığı' konusuna yönelik hazırlanmış bir kavram testine rastlanılamamıştır. Bu bağlamda ağız ve diş sağlığı konusuna yönelik öğrencilerde var olan alternatif kavramlarının belirlenmesi amacıyla bir kavram testi geliştirilmiştir.

2.4.2. Ünite/konunun belirlenmesi:

Bu araştırmada daha önce üzerinde herhangi bir test geliştirme çalışmasının yürütülmediği 'ağız ve diş sağlığı' konusu seçilmiştir. Geliştirilen testin maddeleri ortaöğretim sağlık bilimleri dersi 'ağız ve diş sağlığı' konusunun kazanımlarına ve bu kazanımlara ek olarak yazılan kazanımlara uygun olarak hazırlanmıştır. Ağız ve Diş sağlığı konusunu ele alan MEB ortaokul ve lise ders kitapları incelenmiş ve bu bölümde yer alan metinler ve resimler gözden geçirilmiştir. Ağız ve Diş sağlığı konusunun içerik sınırları, bir önermesel bilgi ifadeleri listesi ile tanımlanmıştır. Önermelerin gerçek öğretim bağlamlarını doğrulanmasına uzmanlık alanları fen eğitimi olan iki akademisyen katılmıştır. İncelemelere göre bazı önermeler değiştirilmiş ya da silinmiştir. Kavram testlerinin amacına ulaşmasında hem geçerli bir içerik hem de çeldiricilerde kullanılmak üzere uygun cevaplar gerekmektedir. Araştırmacılar, öğrencilerin ağız ve diş sağlığı konusuyla ilgili kavramlarını keşfetmek için bir grup öğrenci ile görüşmeler yapmıştır. Görüşme yapılan öğrencilerin

yanlıř anlama ya da anlayamama kategorisine giren alternatif kavramları ve arařtırmacıların kendi deneyimlerinden yararlanarak yazdıkları, çeldiricileri oluřturmuřtur.

2.4.3. Test maddelerinin yazılması ve belirtke tablosunun oluřturulması:

Arařtırmada kullanılan konunun kazanımlarını ieren soru tipleri belirlenmiřtir. Testin ilk ařaması oktan semeli olarak hazırlanmıřtır. Testte oktan semeli sorular beř seenekten oluřmaktadır. Maddeler kavram karikatr řeklinde 'StoryboardThat' adlı bilgisayar programı kullanılarak oluřturulmuřtur. Testin ikinci ařamasında ise birinci ařamada seilen seeneđin neden seildiđinin yazılması istenmiřtir. Testte toplam 24 madde bulunmaktadır. oktan semeli seeneklerin dođru cevap seeneklerinin tm teste eřit dađılmasına dikkat edilmiřtir. Ađız ve diř sađlıđına ynelik kazanımlar ve her kazanımla iliřkili olarak hazırlanan soru sayısının yer aldıđı belirtke tablosu ařađıda sunulmuřtur.

Tablo 2

İki ařamalı ADSKT'te llecek kazanımlara gre madde sayısı ve numaraları

Kazanım No	Kazanımlar	Madde Sayısı	Madde Numarası
1.1.7.	Ađız ve diř sađlıđı iin yapılması gerekenleri aıklar.	3	9., 11., 12.
1.1.7.1	Ađız florasında yayılıř gsteren mikroorganizmalar ve sebep olduđu hastalıklar hakkında bilgi verir.	3	1., 3., 20.
1.1.7.2	Diři oluřturan katmanları řekil zerinde gsterir.	3	14., 19., 22.
1.1.7.3	Diřler ve diřlerin grevleri hakkında bilgi verir.	3	2., 17., 18.
1.1.7.4	Diř sađlıđının bozulmasında etkili olan sebepleri rneklerle aıklar.	4	4., 5., 6., 7.,
1.1.7.5	Diř fıralama zamanı, usul, diř fırası seimi ve diř macununun zellikleri ve ne kadar kullanılması gerektiđi hakkında bilgi verir.	8	9., 10., 11., 12., 13., 15., 16., 23.
1.1.7.6	Vcudumuzdaki hastalıkların ađız sađlıđına etkilerinin neler olduđunu aıklar.	2	8., 21.

Tablo 2.' de grldđ gibi her kazanıma uygun en az  sorunun hazırlanmasına dikkat edilmiřtir (Akbulut ve epni, 2013). Bu sayede gvenirlik analizi alıřmaları sonrasında ıkarılmasının uygun grldđ maddelerin testten atılması sonucunda her kazanıma ynelik en az bir sorunun testte kalması sađlanmış olacaktır.

2.4.4. Uzman grřnn alınması:

Kavram testi, son hali verildikten sonra soruların tm kazanımları kapsayıp kapsamadıđının belirlenmesi (Webb, 1997), bilimsel olarak dođruluđu, yazım ve noktalama hatalarının belirlenmesi, soruların đrenci seviyesine uygunluđu, çeldiricilerin uygun řekilde hazırlanıp hazırlanmadıđı, cevap anahtarının dođruluđu aılarından iki fen eđitimcisi ve bir diř hekiminin grřlerine bařvurulmuřtur. Geerliliđi belirlemek iin iki ana soru ele alınmıřtır: (a) madde, đretim programında tanımlanan ieriđi ve ele alınan kazanımları deđerlendiriyor mu? ve (b) soru lise đrencileri iin uygun dzeyde mi? Bu kriterler karřılanmaması durumunda, madde testten ıkarılmıřtır. Literatrde bařka ađız ve diř sađlıđı konulu testlerinin mevcut olmaması nedeniyle, tahmin ve korelasyon geerliliđi yapılamamıřtır.

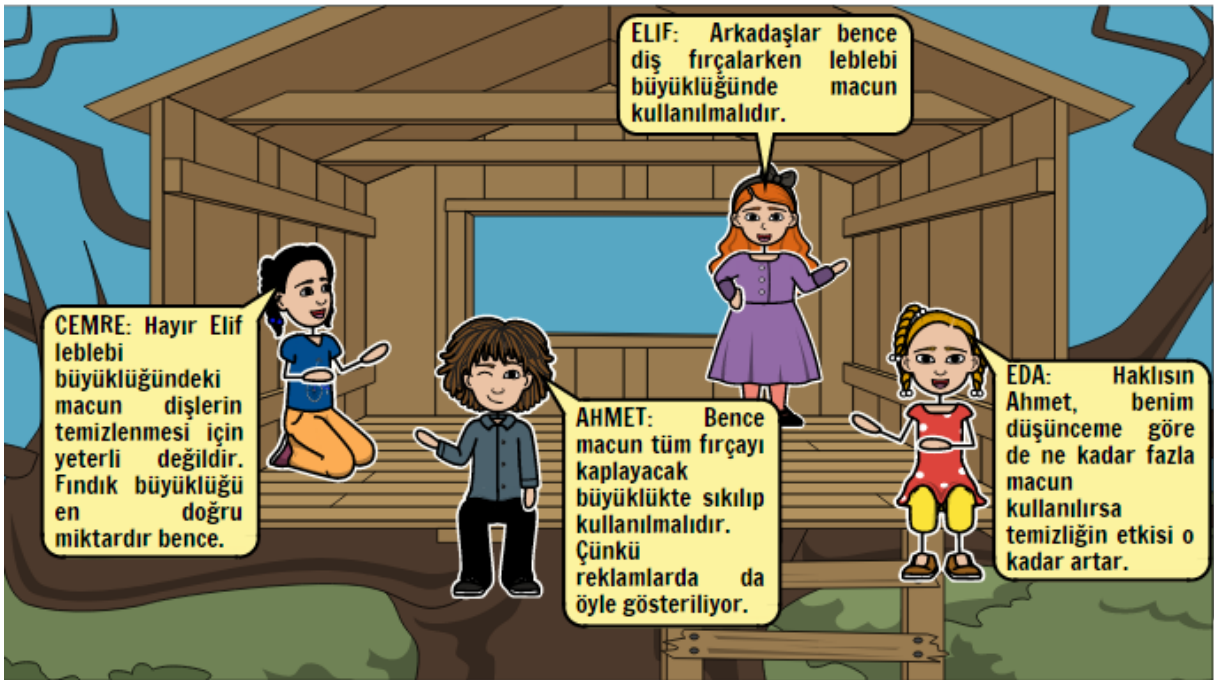
2.4.5. leđin yeniden dzenlenmesi:

Uzmanlardan alınan geri dntler iřıđında test tekrar gzden geirilir ve gerekli dzenlemeler yapılır (zgven, 1998). Bu bađlamda ADSKT uzman dntlerine gre tekrar dzenlenmiřtir. Testteki 24. soru diř hekiminden alınan dntler iřıđında bilimsel olarak dođru olmadıđı řeklindeki aıklaması dikkate alınarak testten ıkarılmıřtır. Gerekli tm dzeltmeler yapıldıktan sonra, maddelerin tekrar incelenmesi, test maddelerinin anlaşılabilirliđi, anlaşılması g yerlerin testten ıkarılması veya dzeltilmesi ve testin

cevaplandırılması için geçen sürenin tespit edilmesi için 23 maddelik test toplam 60 9. sınıf öğrencisine pilot olarak uygulanmıştır. Öğrencilerin uygulama esnasında maddelerle ilgili soruları üzerine testteki 10. maddede bulunan diş macunu boyutunun belirtilmek istendiği “nohut büyüklüğü” ifadesi yerine “leblebi büyüklüğü” ifadesi yazılmıştır. 15. maddenin a ve c şıklarının aynı olduğu yine aynı seçenekler içinde doğru cevabın hiçbir şıkta var olmadığı fark edilmiştir. Testteki bu maddenin aynı seçenekleri değiştirilmiş, e seçeneğine doğru cevap eklenmiştir. 22. maddenin b seçeneğinde ve 23. maddenin a seçeneğindeki yazım yanlışları düzeltilmiştir. Gerekli düzenlemeler yapıldıktan sonra teste son hali verilmiştir. Uygulama sonunda testin yaklaşık 40 dakika sürede cevaplandırıldığı sonucuna ulaşılmıştır. Son olarak testin cevap anahtarları kontrol edilmiş ve tüm testte eşit dağılan cevaplar olmasına dikkat edilmiştir. Testte bulunan örnek bir sorunun ilk hali ve uzman görüşü sonrasındaki hali aşağıda sunulmuştur.

Ölçeğin uzman görüşüne sunulmadan önceki hali;

SORU-10



Yukarıdaki dört arkadaş diş fırçalarken kullanılması gereken doğru diş macunu miktarı hakkında konuşmaktadırlar. Sizce hangisi/ hangilerinin söyledikleri kesinlikle doğrudur?

a-Elif

b- Eda

e- Ahmet ve Eda

c-Cemre

d- Ahmet

Çünkü;

.....
.....

Testteki 10. maddede bulunan Cemre'nin ifadesindeki “fındık büyüklüğü” ifadesinin “mercimek tanesi büyüklüğü” şeklinde değiştirilmesi şeklinde dönüt alınmıştır. 10. maddenin uzman görüşüne sunulduktan sonra gerekli düzenlemenin yapıldığı hali;

SORU-10



Yukarıdaki dört arkadaş diş fırçalarken kullanılması gereken doğru diş macunu miktarı hakkında konuşmaktadırlar. Sizce hangisi/ hangilerinin söyledikleri kesinlikle doğrudur?

- a-Elif
c-Cemre
Çünkü;
- b- Eda
d- Ahmet
- e- Ahmet ve Eda

ADSKT'deki onuncu soru incelendiğinde uzmanlardan birisinin Cemre adlı karakterin ifadesindeki, fındık ve leblebi büyüklüğünün aynı boyutlarda olduğunu bu yüzden öğrenciler tarafından karıştırılabilecek bir seçenek oluşturduğunu bu yüzden düzeltilmesi gerektiği şeklindeki önerisi doğrultusunda araştırmacılar Cemre karakterinin ifadesindeki fındık büyüklüğünü mercimek tanesi olarak değiştirmişlerdir.

2.4.6. Pilot Uygulama Sonrası Yapılan İşlemler:

60 kişilik uygulama sonunda testteki sorularda yazım ve imla yanlışları olup olmadığının, anlaşılabilirliğin belirlenmesinin, kapsam ve görünüş geçerliğinin tespit edilmesi amacıyla test iki fen eğitimcisi, bir Biyoloji öğretmeni ve bir Türkçe öğretmeninden oluşan uzman görüşüne tekrar sunulmuştur.

Uzmanlardan alınan geri dönütler ışığında 1. maddedeki "protista alemi" ifadesi "monera alemi" olarak değiştirilmiştir. 2. maddedeki " yiyeceklerin mekanik sindirimi dişler tarafında gerçekleştirilir" ifadesi "dişler yiyecekleri mekanik sindirime uğratar" şeklinde düzeltilmiştir. 13. madde için " arkadaşlar diş parlatma tozları, gerekenden fazla kullanıldığında diş hekimine başvurulmalıdır." çeldirisinin çok kolay olduğu yönünde bildirilen görüş ışığında maddenin seçenekleri yeniden düzenlenmiştir. 17. maddenin a ve c şıkkının aynı olduğu belirtilmiştir. Bu yanlışlığı düzeltmek için c şıkkı değiştirilmiştir. Yine başka bir uzman görüşü doğrultusunda karikatürlere diş ve temizleme ürünlerini içeren görseller eklenmiştir. 22. maddede yer alan "bence sıcak ve soğğun hissedildiği tabaka pulpadır." ifadesi "bence sıcak ve soğğun hissedildiği tabaka pulpa ve dentindir." şeklinde revize edilmiştir. Madde kökü ile ilgili alınan dönütlere göre 1., 8., 9., 10., 13. ve 21. madde kökünde yapılan yazım yanlışlarına ilişkin gerekli değişiklikler ve düzeltmeler yapılmıştır. Ardından son halı verilen test geçerlik, güvenilirlik ve madde analizleri yapılması için 208 kişilik uygulama grubuna uygulanmıştır. Madde analizi sonucunda hiçbir madde testten çıkarılmamıştır. Son olarak test maddeleri ve cevap anahtarı gözden geçirilmiştir.

3. Bulgular

Bu çalışmanın katılımcılarına güvenilirlik ve geçerlik çalışmaları için Ağız ve Diş sağlığı kavram testinin son hali uygulanmıştır. 208 kişilik uygulama grubundan elde edilen verilere göre geçerlik, güvenilirlik ve madde analizi çalışmaları aşağıda sıralanmıştır.

3.1. İki Aşamalı ADSKT'nin Geçerliğine İlişkin Bulgular

Araştırmalarda kullanılan test kapsam ve görüş geçerliğinin sağlanması amacıyla uzmanlara sunulmuştur. Uzmanlara sunulmadan önce test araştırmacılar tarafından detaylıca incelenmiş ve maddelerin konu kazanımlarına uygunluğu tespit edilmiştir. Testte toplam 24 madde bulunmaktadır. Geliştirilen test iki fen eğitimcisi, bir biyoloji öğretmeni ve bir diş hekiminden oluşan uzmanların görüşüne sunulmuştur. Testi inceleyen uzmanların test ile ilgili görüş birliği/ayrılığı oranı hesaplanmış ve %85 oranında görüş birliğine varıldığı görülmüştür. Uzmanlardan alınan geri dönütler ışığında testte gerekli düzenlemeler yapılmıştır. Diş hekiminin görüşü ışığında 24. madde testten çıkarılmıştır. Geliştirilen test pilot uygulama öncesinde örneklem dışında iki öğrenciye okutulmuş ve anlaşılması güç yerlerin olup olmadığı tespit edilmek istenmiştir. Alınan geri dönütler ışığında testin anlaşılabilirliğinde herhangi bir sorunun olmadığı saptanmıştır. Tüm bu düzenlemeler ışığında test uygulamaya hazır hale getirilmiştir. Geliştirilen testin geçerlik analizi 9. sınıfta öğrenim gören toplam 60 öğrenci ile gerçekleştirilmiştir. Uygulama grubundan alınan geri dönütler sonunda tespit edilen aksaklıklar düzeltilmiş test, alan ve alan eğitimi uzmanları tarafından tekrar incelenmiştir. Alınan geri dönütler ışığında 23 maddelik test hakkında uzmanlar, test maddelerinin öğrencilerin yaş ve sınıf düzeylerine uygun olduğu, ağız ve diş sağlığı konusunun kazanımları ölçmede yeterli olduğu, soru kökleri ve çeldiricilerde herhangi bir bilimsel hata olmadığı yönünde görüş bildirmişlerdir.

3.2. İki Aşamalı ADSKT'nin Madde Analizine İlişkin Bulgular

İki aşamalı ADSKT'nin madde analizi yapılırken 208 kişilik uygulama grubu test maddelerinden alınan öğrenci puanları hesaplanmıştır. Bu puanlar büyükten küçüğe doğru sıralanmıştır. Öğrencilerin %27'si üst grubu %27'si alt grubu oluşturacak şekilde sıralanmıştır. Belirlenen alt ve üst grupların ardından madde güçlüğü ($p = (Dü + Da) / 2N$) ve madde ayırt edicilik indeksi ($d = (Dü - Da) / N$) formülleri kullanılarak hesaplamalar yapılmıştır. (p: madde güçlüğü, d: madde ayırt edicilik indeksi Dü: maddeyi doğru cevaplayan üst grup öğrenci sayısı, Da: maddeyi doğru cevaplayan alt grup öğrenci sayısı, N: toplam öğrenci sayısı). İki aşamalı ADSKT'nin üst ve alt grupta bulunan öğrencilerin seçeneklere göre dağılımları, güçlük ve ayırt edicilik indeksleri hesaplanmıştır. Asıl uygulama verileri kullanılarak oluşturulan madde analizi sonuçları Tablo 3.'te sunulmuştur.

Tablo 3

İki aşamalı ADSKT'ye ilişkin madde analizi bulguları

M	Grup	A	B	C	D	E	Boş	p	d	M	Grup	A	B	C	D	E	Boş	p	d
1	üst	1	1	1	4	49	0	0,72	0,3	13	üst	34	1	1	3	17	0	0,49	0,23
	alt	2	10	0	12	32	0				alt	21	9	12	8	6	0		
2	üst	33	7	5	9	2	0	0,41	0,35	14	üst	3	37	7	6	3	0	0,37	0,57
	alt	13	22	9	12	0	0				alt	7	5	10	20	14	0		
3	üst	4	40	1	5	6	0	0,51	0,39	15	üst	9	2	2	10	33	0	0,41	0,35
	alt	13	18	15	5	5	0				alt	6	0	7	30	13	0		
4	üst	8	5	34	4	5	0	0,33	0,55	16	üst	3	31	3	6	13	0	0,42	0,25
	alt	17	9	3	24	3	0				alt	13	17	13	5	8	0		
5	üst	19	4	2	26	5	0	0,29	0,33	17	üst	5	3	10	31	7	0	0,34	0,41
	alt	20	5	16	7	8	0				alt	10	9	9	8	20	0		
6	üst	34	7	10	1	4	0	0,34	0,51	18	üst	1	7	2	7	39	0	0,41	0,57

	alt	5	9	18	19	5	0		alt	10	13	8	18	7	0	
	üst	9	10	32	3	2	0,5		üst	5	6	6	36	3	0,35 0,57	
7	alt	9	16	4	14	13	0	0,32	19	alt	13	12	8	4	19	0
	üst	4	3	8	2	39	0,53		20	üst	0	7	41	7	1	0,48 0,5
8	alt	15	12	8	12	9	0	0,42		alt	10	11	13	13	9	0
	üst	23	13	2	2	16	0,25		21	üst	10	36	3	2	5	0,34 0,58
9	alt	9	16	7	9	15	0	0,28		alt	15	3	11	9	18	0
	üst	42	6	3	2	3	0,25		22	üst	1	38	4	1	2	0,40 0,55
10	alt	19	9	11	9	8	0	0,54		alt	7	7	7	8	27	0
	üst	2	11	11	25	7	0	0,41	23	üst	1	5	13	1	36	0 0,47 0,33
11	alt	8	11	4	2	31	0	0,24		alt	5	16	9	9	17	0
	üst	6	36	9	1	4	0	0,32								
12	alt	10	18	4	6	18	0	0,48								

M: Madde no; p: Madde güçlüğü; d: Ayırt edicilik

Tablo 3.'teki verilere göre iki aşamalı ADSKT'nin ortalama madde güçlüğü 0,40 olarak hesaplanmıştır. Bu da testin madde güçlük indeksinin orta güçlükte olduğu anlamına gelmektedir.

Tablo 3 incelendiğinde iki aşamalı ADSKT'nin madde analizi sonuçlarına göre 23 maddenin; ortalama ayırt ediciliği indeksi ise 0,42 olarak hesaplanmıştır. İki aşamalı ADSKT'nin ayırt edicilik indeks sonuçlarına bakıldığında geliştirilen testin ayırt edicilik gücünün ise çok iyi olduğu görülmektedir.

Yapılan tüm bu hesaplamaların ışığında geliştirilen iki aşamalı ADSKT'nin 9. Sınıf öğrencilerinin ağız ve diş sağlığı konusu ile ilgili kavramsal anlamalarını ölçebilecek güçlükte ve ayırt edicilik değerinde bir kavram testi olduğu görülmektedir.

3.3. İki Aşamalı ADSKT'nin Güvenirlik Analizinin Yapılması

Araştırmalarda önemli olan bir diğer unsur güvenilirliktir. Bir ölçme aracının güvenilir olduğunu gösteren unsurlar bulunmaktadır. Güvenilir testler farklı zaman aralıklarında elde edilen cevaplar arasında tutarlılık oluşturur (Büyüköztürk, 2007). Ölçme aracında güvenilirliği sağlamak için kullanılan yöntemler iç tutarlılık güvenilirliği, test tekrar test güvenilirliği, paralel testler güvenilirliği ve gözlemciler arası güvenilirlik şeklindedir (Şencan, 2005).

İki aşamalı ADSKT'nin güvenilirlik analizi için iç tutarlılık güvenilirliği esas alınmıştır. Test puanları arasındaki iç tutarlılık hesaplamalarında, Kuder R'chardson-20, 21 (KR-20, 21) ve Cronbach alfa katsayıları kullanılmaktadır (Büyüköztürk, 2007). KR-20 ve KR-21 çoktan seçmeli test ve maddelerde kullanılmaktadır (Şencan, 2005). İki aşamalı ADSKT'nin birinci aşamasının iç tutarlılık güvenilirliği için KR-20 hesaplanmıştır. Testin iki aşaması birlikte güvenilirlik değeri için Cronbach alfa katsayısı hesaplanmıştır.

Testteki çoktan seçmeli sorulara verilen doğru cevaplar 1 yanlış cevaplar 0 şeklinde kodlanmıştır. Bu bağlamda öğrencilerin çoktan seçmeli sorulardan alabilecekleri maksimum puan 23'tür. Yapılan veri analizi sonucunda testin güvenilirlik katsayısı (Cronbach Alpha) 0,808 olarak hesaplanmıştır. Ayrıca testin KR-20 değeri 0,807 olarak hesaplanmıştır.

23 maddeden oluşan ADSKT'nin iç tutarlılığına ilişkin hesaplanan güvenilirlik katsayısı (Cronbach Alpha) ve KR-20 değeri Tablo 4.'te verilmiştir.

Tablo 4
İki aşamalı ADSKT'in iç tutarlılık güvenilirlik analizine ilişkin veriler

Güvenirlik analizi türü	N	Madde sayısı	Hesaplanan değer
Cronbach Alpha	208	23	0,808
KR-20	208	23	0,807

Tablo 4. incelendiđinde Cronbach Alpha gvenirlik katsayısı ve KR-20 deđerinin 0,70 ve st olduđu gzlenmektedir. Nitekim gvenirlik katsayısı 0,70'in stnde olan testler gvenirliđi yksek testler olarak nitelendirilmektedir (Bykztrk, 2007; Bykztrk, Kılıç Çakmak, Akgn, Karadeniz ve Demirel, 2018; Bykztrk, 2013; Erkuş, 2006; zelik, 2010). Buradan yola çıkılarak arařtırma kapsamında geliřtirilen iki ařamalı ADSKT'nin Cronbach Alpha ve KR-20 gvenirlik katsayısı deđerleri baz alınarak testin gvenilir bir test olduđu sonucuna varılmıřtır.

4. Sonular ve Tartıřma

Bu arařtırmada "ađız ve diř sađlıđı" konusuna ynelik geliřtirilen iki ařamalı kavram testi yardımıyla lise 9. sınıf đrencilerinin konuya iliřkin ne tr kavramsal anlamalar edindiklerini tespit etmeye ynelik geliřtirilen testin geerlik, madde analizi ve gvenirlik alıřmaları yapılmıřtır.

Madde gçlđ 0 ve 1 arasında bir sayısal deđer almaktadır. Madde gçlđ 0'a yaklařtıđıca madde zorlařmakta, 1'e yaklařtıđıca ise madde kolaylařmaktadır (Çepni, Ayař, Ekiz ve Akyıldız, 2008). İki ařamalı kavram testinin madde analiz sonularına incelendiđinde, ortalama madde gçlđ 0,40 olarak hesaplanmaktadır. Bu da testin madde gçlk indeksinin orta gçlkte olduđu anlamına gelmektedir. Madde ayırt ediciliđi ise 1'e yaklařtıđında ayırt edicilik indeksi artmakta, 0'a yaklařtıđında ise ayırt edicilik indeksi dřmektedir. Eđer maddenin ayırt edicilik indeksi negatif (-) bir deđer alırsa bu madde testin gvenirliđini dřrmektedir (Gnen, Kocakaya ve Kocakaya, 2011). zelik (2010) yaptıđı alıřmada madde ayırt edicilik indeksini (d) 1 ile 0 arasında aldıđı bazı deđerlere gre sınıflamıřtır. Bunlar; 0,19 ve daha kkse, madde kabul edilmez; 0,20-0,29 deđer aralıđında arasında ise, madde dzeltilmeli; 0,30-0,39 deđer aralıđında arasında ise, iyi bir madde ve kabul edilir; 0,40 ve daha bykse, ok iyi bir maddedir. 23 maddelik testin; ortalama ayırt ediciliđi indeksi ise 0,42 olarak hesaplanmıřtır. İki ařamalı ADSKT'nin ayırt edicilik indeks sonularına bakıldıđında geliřtirilen testin ayırt edicilik gcnn ise ok iyi olduđu sonucuna varılmaktadır.

Arařtırmanın gvenirlik alıřması sonuları incelendiđinde, iki ařamalı ADSKT'nin i tutarlılık gvenilirliđi iin KR-20 ve Cronbach alfa deđerleri hesaplanmıřtır. Yapılan veri analizi sonucunda testin gvenirlik katsayısı (Cronbach Alpha) 0,808 olarak hesaplanmıřtır. Eđitim arařtırmalarında hesaplanan bu deđer testin gvenilir olduđu řeklinde yorumlanmaktadır (Bykztrk, 2013). Ayrıca testin KR-20 deđeri 0,807 olarak hesaplanmıřtır. Gvenirlik katsayısı 0,70'in stnde olan testler gvenirliđi yksek testler olarak nitelendirilmektedir (Bykztrk, 2007; Bykztrk, Kılıç Çakmak, Akgn, Karadeniz ve Demirel, 2018; Bykztrk, 2013; Erkuş, 2006; zelik, 2010). Buradan yola çıkılarak arařtırma kapsamında geliřtirilen iki ařamalı ADSKT'nin ADSKT'nin Cronbach alfa gvenirlik katsayısı ve KR-20 deđeri baz alınarak testin gvenilir olduđu sonucuna varılmıřtır.

Kaynaklar

- Aıkgz, M. ve Karalı, F. (2015). Alternatif lme-deđerlendirme yaklařımları kullanılarak iř ve enerji konusunda geliřtirilen bařarı testinin geerlilik ve gvenirlik analizi. *Amasya niversitesi Eđitim Fakltesi Dergisi*, 4(1), 1-25.
- Akbulut, H. İ. ve Çepni, S. (2013). Bir niteye ynelik bařarı testi nasıl geliřtirilir? :İlkđretim 7. sınıf kuvvet ve hareket nitesine ynelik bir alıřma. *Amasya niversitesi Eđitim Fakltesi Dergisi*, 2(1), 18-44.
- Amir, R. ve Tamir, P. (1994), "In-Depth Analysis of Misconceptions as a Basis For Developing Research-based Remedial Instruction: *The Case of Photosynthesis*", *The American Biology Teacher*, 56(2), 94-100.
- Atasoy, ř. ve Akdeniz A.R. (2009). *Kavram karikatrlerinin etki-tepki kuvvetleri ile ilgili yanılıđları gidermeye etkisi*. 3. Uluslararası Bilgisayar ve đretim Teknolojileri Sempozyumu, Karadeniz Teknik niversitesi, Trabzon.
- Ausubel, D.P. (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart and Winston.
- Aykutlu, I., ve řen, A. İ. (2012).  ařamalı test, kavram haritası ve analogi kullanılarak lise đrencilerinin elektrik akımı konusundaki kavram yanılıđlarının belirlenmesi. *Eđitim ve Bilim*, 37(166), 275-288.
- Burke, K. (2005). *How to authentics learning (4th Ed)*. Thousand Oaks, Calif.: Corwin Press.
- Burns, J. C., Okey, J. R. and Wise, K. C. (1985). Development of an integrated process skill test: TIPS II. *Journal of Research in Science Teaching*, 22(2), 169-177.
- Can řen, H. ve Eryılmaz, A. (2011). Bir bařarı testi geliřtirme alıřması: basit elektrik devreleri bařarı testi geerlik ve gvenirlik arařtırması. *Yznc Yıl niversitesi Eđitim Fakltesi Dergisi*, 8 (1), 1-39.

- Canpolat, N. ve Pınarbaşı, T. (2011). Bazı kimya kavramlarına yönelik iki kademeli çoktan seçmeli bir testin geliştirilmesi ve uygulanması. *Erzincan Eğitim Fakültesi Dergisi*, 13(1), 55-80.
- Çalık, M. ve Ayas, A. (2003). Çözeltilerde kavram başarı testi hazırlama ve uygulama. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 2(14), 1-17.
- Çalışkan, İ. Ö. ve Kaptan, F. (2009). Constructing science process skills test. *Çağdaş Eğitim Dergisi*, 34, 369, 27-34.
- Çetinkaya, M. ve Erol, T. A. Ş. (2016). "Vücudumuzda sistemler" ünitesine yönelik üç aşamalı kavram tanı testi geliştirilmesi. *ODÜ Sosyal Bilimler Araştırmaları Dergisi (ODÜSOBİAD)*, 6(15), 317-330.
- Demir, İ. (2019). *Yaşam Temelli Öğretimin Ortaokul 7.Sınıf Öğrencilerinin Ağız Ve Diş Hijyeni Konusunda Kavram Öğrenmelerine, Fen Bilimlerine Karşı Tutumlarına Ve Motivasyonlarına Etkisi*. Yüksek Lisans Tezi, Marmara Üniversitesi, İstanbul.
- Demir, Y., Uzoğlu, M. ve Büyükkasap, E. (2012). Fen bilgisi öğretmen adaylarının kuvvet ve hareket ile ilgili sahip olduğu kavram yanlışlarının belirlenmesinde kullanılan karikatürlerin ve çoktan seçmeli soruların etkililiğinin karşılaştırılması. *Eğitim ve Öğretim Araştırmaları Dergisi*, 1,1, 88-102.
- Ekici, F., Ekici, E., and Aydın, F. (2007). Utility of Concept Cartoons in Diagnosing and Overcoming Misconceptions Related to Photosynthesis. *International of Journal of Environmental and Science Education*, 2(4), 111-124.
- Ekim, K.F. (2007). *İlköğretim fen öğretiminde kavramsal karikatürlerin öğrencilerin kavram yanlışlarını gidermedeki etkisi*. Yayımlanmamış Yüksek Lisans Tezi, Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Erden, M. (1998). *Eğitimde program değerlendirme (3. Baskı)*. Ankara: Anı Yayıncılık.
- Erdoğan, A. ve Cerrah Özsevgeç, L. (2012). Kavram karikatürlerinin öğrencilerin kavram yanlışlarının giderilmesi üzerindeki etkisi: sera etkisi ve küresel ısınma örneği. *Turkish Journal of Education*, 1(2), 38-50.
- Erdoğan, M. Y. and Kurt, F. (2012). Öğretmenlerin ölçme ve değerlendirme yeterlik algılarının bazı değişkenler açısından incelenmesi. *Eğitim Bilimleri Dergisi*, 1(2), 23-36.
- Evrekli, E., İnel, D., Balım, A. G. ve Kesercioğlu, T. (2009). Fen öğretmen adaylarına yönelik yapılandırıcı yaklaşım tutum ölçeği: geçerlik ve güvenilirlik çalışması. *Türk Fen Eğitimi Dergisi*, 6(2),134-148.
- Feyzioğlu, B., Demirdağ, B., Akyıldız, M. ve Altun, E. (2012). *Ortaöğretim öğrencilerine yönelik bilimsel süreç becerileri testi geliştirilmesi: geçerlik ve güvenilirlik çalışması*. Kuram ve Uygulamada Eğitim Bilimleri, 12(13), 1887-1906.
- Fraenkel, J. R. and Wallen, N. E. (2009). *How to design and evaluate research in education*. New York: McGraw-Hill Companies.
- Garnett, P.J. and Treagust, D.F. (1992). *Conceptual Difficulties Experienced by Senior High School Students of Chemistry: Electrochemical (Galvanic) and Electrolytic Cells*, Journal of Research in Science Teaching, 29, 10, 1079-1099.
- Haslam, F. and Treagust, D. F. (1987). *Diagnosing Secondary Students' Misconceptions of Photosynthesis and Respiration in Plants Using a Two-Tier Multiple Choice Instrument*, Journal of Biological Education, 21, 3, 203-211.
- Hlebowitsh, P. S. (2005). *Desinging the school curriculum*. USA: Pearson Education.
- İlhan, N., ve Hoşgören, G. (2017). Fen bilimleri dersine yönelik yaşam temelli başarı testi geliştirilmesi: Asit baz konusu. *Fen Bilimleri Öğretimi Dergisi*, 5(2), 87-110.
- Kabapınar, F. (2005). Effectiveness of teaching via concept cartoons from the point of view of constructivist approach. *Kuram ve Uygulamada Eğitim Bilimleri Dergisi*, 5(1), 135-146.
- Kabapınar, F. (2007). *Fen öğretiminde kavram karikatürleri: oluşturmacı bir öğretim yöntemi, ilköğretim çağına genel bir bakış*. İstanbul: Morpa Kültür Yayınları, 243-264.
- Karaca, E. (2008). An investigation of primary and high school teachers' perception levels of efficacy of measurement and evaluation in education in Turkey. *Social Behavior and Personality*, 36(8), 1111-1122.
- Karataş, F. Ö., Köse, S. Ve Coştu, B. (2003). Öğrenci yanlışlarını ve anlama düzeylerini belirlemede kullanılan iki aşamalı testler. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*. 13(1), 54-69.
- Karlı Baydere, F. ve Yiğit, M. (2020). Hidrokarbonlar Konusuna Yönelik Bir Kavram Tanı Testi Geliştirilmesi. *Karaelmas Eğitim Bilimleri Dergisi*, 8(2), 366-379.
- Keogh, B., and Naylor, S. (1997). *Thinking about Science Posters: Developing Ideas Through Concept Cartoon Posters*. Millgate House Publishers.
- Keogh, B., and Naylor, S. (1999a). Science Goes Underground. *Adults Learning*, 10(5), 6-8.
- Keogh, B., and Naylor, S. (1999b). Concept Cartoons, Teaching and Learning in Science: An Evaluation. *International Journal of Science Education*, 21(4), 431-446.
- Lin, S. W. (2004). Development and application of a two-tier diagnostic test for high school students' understanding of flowering plant growth and development. *International Journal of Science and Mathematics Education*, 2(2), 175-199.
- Mann, M. and Treagust, D. F. (1998). A Pencil and Paper Instrument to Diagnose Students' Conception of Breathing, Gas Exchange and Respiration, *Australian Science Teachers Journal*, 44, 2, 55-59.
- Martinez, Y. M. (2004). *Does The K-W-L Reading Strategy Enhance Student Understanding in Honors High School Science Classroom?*. (Unpublished masters thesis). Fullerton: California State University.
- Millî Eğitim Bakanlığı (2018). *Fen bilimleri dersi öğretim programı (ilkokul ve ortaokul 3, 4, 5, 6, 7 ve 8. sınıflar) öğretim programı*. Ankara: Millî Eğitim Bakanlığı
- Naylor, S., and Keogh, B. (2013). Concept Cartoons: what have we learnt?. *Journal of Turkish Science Education*, 10(1), 3-11.
- Naylor, S., and McMurdo, A. (1990). Supporting science in schools. *Timperley. Nussbaum*.

- Odom, A. L., Barrow, H. L. (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.
- Oliva, P. F. (2005). *Developing the curriculum (6. Edition)*. Newyork: Pearson Education.
- zgven, İ. E. (1998). Psikolojik Testler. Ankara: PDREM Yayınları.
- zmen, H. (2004). Fen đretiminde đrenme teorileri ve teknoloji destekli yapılandırmacı (Constructivist) đrenme. *The Turkish Online Journal Of Educational Technology*, 3, 14, 100-109.
- Peterson, R.F., Treagust, D.F., and Garnett, P.J. (1989). *Development and Application of a Diagnostic Instrument to Evaluate Grade-11 and -12 Students' Concepts of Covalent Bonding and Structure Following a course of Instruction*, Journal of Research in Science Teaching, 26, 4, 301-314.
- Snmez, V. (2005). *Program geliřtirme đretmen el kitabı (13. Baskı)*. Ankara: Anı Yayıncılık.
- Tan, K. C. D., Goh, K. N., Chia, S. L. and 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.
- Temiz, B. K., Tařar, M. F. ve Tan, M. (2006). Development and Validation of a Multiple Format Test of Science Process Skills. *International Education Journal*, 7, 7, 1007-1027.
- 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.
- Turgut, M. F. ve Baykul, Y. (2012). *Eđitimde lme ve deđerlendirme(4. Baskı)*. Ankara: Pegem Akademi.
- Tyson, L., Treagust, D. F. and Bucat, R. B. (1999). The complexity teaching and learning chemical equilibrium, *Journal of Chemical Education*, 76(4), 554-558.
- Voska, K. W. and Heikkinen, H. W. (2000). *Identification and analysis of student conception used to solve chemical equilibrium problems*. Journal of Research in Science Teaching, 37(2), 160-176.
- Wandersee, J.H., Mintzes, J.J. and Novak, J.D. (1994). *Research on alternative conceptions in science*. In D.L. Gabel (Ed.), Handbook of research on science teaching and learning (pp. 177-210). New York: Macmillan.
- Webb, N. L. (1997). *Determining alignment of expectations and assessments in mathematics and science education. NISE brief 1(2)*. Madison, WI: University of Wisconsin-Madison, National Institute for Science Education.

Arařtırmacıların Katkı Oranı Beyanı:

Yazarlar arařtırmaya eřit oranda katkı sađlamıřlardır.