



## Maddenin Yapısı ve Özellikleri Konusunda Başarı Testi Geliştirme: Geçerlik ve Güvenirlik Çalışması

### Developing an Achievement Test on the Structure and Properties of the Substance: A Validity and Reliability Study

Merve Nur KÖROĞLU<sup>1</sup>, Melek KARACA<sup>2</sup>, Oktay BEKTAŞ<sup>3</sup>

**Makale Türü:** Araştırma Makalesi

**Başvuru Tarihi:** 20.09.2021

**Kabul Tarihi:** 21.11.2022

**Atf İçin:** Köroğlu, M. N., Karaca, M. ve Bektaş, O. (2023). Maddenin yapısı ve özellikleri konusunda başarı testi geliştirme: Geçerlik ve güvenilirlik çalışması. *Anadolu Üniversitesi Eğitim Fakültesi Dergisi (AUJEF)*, 7(1), 1-32.

**ÖZ:** Bu çalışmanın temel amacı yedinci sınıf “*Maddenin Yapısı ve Özellikleri*” konusunda geçerli ve güvenilir bir başarı testi geliştirmektir. Bu amaç doğrultusunda konu ile ilgili kazanımlara uygun, alanyazına dayalı 23 çoktan seçmeli sorudan oluşan başarı testi hazırlanmıştır. Bu sorular, ortaöğretim sınavlarındaki çıkmış sorularından ve geçerli, güvenilir başarı testi çalışmalarından yola çıkarak hazırlanmıştır. Hazırlanan sorulardan elde edilen taslak başarı testi 2020-2021 eğitim-öğretim bahar yarıyılında Kayseri ili Melikgazi ilçesinde yedinci sınıfa devam etmekte olan 268 öğrenciye uygulanmıştır. Mevcut çalışma, nicel araştırma yönteminin tarama desenine dayalı yürütülmüştür. Elde edilen veriler SPSS 25 ve “*factor 11.04.02*” programı kullanılarak analiz edilmiştir. Klasik test teorisi rehber alınarak oluşturulan başarı testinin kapsam geçerliğinde Bloom’un bilişsel alan taksonomisi basamakları dikkate alınmış, sorular alanında uzman kişilerin kontrolüne sunulmuştur. Testin normal dağılım varsayımlarına bakılıp, yapı geçerliği için, açımlayıcı ve doğrulayıcı faktör analizleri yürütülmüştür. Yapılan analizler sonucunda testten elde edilen puanların KR-20 güvenilirlik katsayısı 0,86 olarak hesaplanmıştır. Sonuç olarak “*Maddenin Yapısı ve Özellikleri*” konusunda geçerli ve güvenilir bir başarı testi elde edilmiştir.

**Anahtar sözcükler:** Fen eğitimi, madde ve yapısı, başarı testi, test geliştirme, geçerlik ve güvenilirlik

**ABSTRACT:** This study aims to develop a valid and reliable achievement test for the seventh-grade “*Structure and Properties of Substance*”. For this purpose, an achievement test consisting of 23 multiple-choice questions

<sup>1</sup> Yüksek Lisans Öğrencisi, Erciyes Üniversitesi, Eğitim Fakültesi, Matematik ve Fen Bilimleri Eğitimi Bölümü, [4050530084@erciyes.edu.tr](mailto:4050530084@erciyes.edu.tr), [mn.koroglu.mn@gmail.com](mailto:mn.koroglu.mn@gmail.com), ORCID: 0000-0001-9517-2435

<sup>2</sup> Dr., Bağımsız Araştırmacı, [melekcaraca38@gmail.com](mailto:melekcaraca38@gmail.com), ORCID: 0000 0002 6957 5932

<sup>3</sup> Prof. Dr., Erciyes Üniversitesi, Eğitim Fakültesi, Matematik ve Fen Bilimleri Eğitimi Bölümü, [obektas@erciyes.edu.tr](mailto:obektas@erciyes.edu.tr), ORCID: 0000 0002 2562 2864

<sup>4</sup> Etik Kurulu onayı alınmıştır.

based on the literature and appropriate for the objectives related to the subject was prepared. We have prepared these questions based on questions from secondary education exams and valid and reliable achievement test studies. The draft achievement test was administered to 268 students attending the seventh grade in the Melikgazi district of Kayseri province in the 2020-2021 academic spring semester. The present study was conducted according to the survey, which is one of the quantitative research designs. The data were analyzed using the SPSS 25, and “*factor 11.04.02*” program. Bloom's taxonomy of cognitive steps was taken into account in the content validity of the achievement test, which was created by taking the classical test theory as a guide, and we have presented it to the control of experts in the field. Considering the normal distribution assumptions of the test, exploratory and confirmatory factor analyses were run for the construct validity. As a result of the analysis, the KR-20 reliability coefficient of the scores obtained from the test was calculated as 0.86. As a result, a valid and reliable achievement test was obtained on the subject of “*Structure and Properties of Substance*”.

**Keywords:** Science education, substance and its structure, achievement test development, validity, and reliability

## 1. GİRİŞ

Bir eğitim sisteminin hedef felsefesine yönelik çalışmaların sağlam temellere dayanması, o eğitim sisteminin güvenilir değerlendirme ölçütleri geliştirmesi ve kullanmasıyla mümkündür (Gardner, 1989). Öğretimin kalitesi, kullanılan ölçme değerlendirme yöntemleri ve araçlarıyla bağlantılıdır (Tünkler ve Güven, 2019). Bir eğitim sisteminin başarısı da ölçme ve değerlendirmenin kalitesiyle doğru orantılıdır (Balcı ve Tekkaya, 2000).

Türkiye öğretim programı; öğretim hedefleri, eğitim durumları ve değerlendirme olmak üzere üç ana unsuru hedef almaktadır (MEB, 2018). Bu bağlamda, öğrencilerin belirlenmiş programın uygulanmasının sonunda öğrenme seviyelerini belirlemek amacıyla yürütülen çalışmalar ölçme ve değerlendirme olarak isimlendirilir. Ölçme ve değerlendirme, tahmin edilen sonuçlar ile gerçek sonuçların mukayese edilmesini sağlayan eğitimin bütünleyici bir modülüdür. Ölçme ve değerlendirmenin gerçekleştirilmemesi, öğretme ve öğrenmenin yapılandırılmasını ve geliştirilmesini engeller. Ölçme ve değerlendirmenin temel amacı, eğitim ve öğretimin niteliğini arttırmaktır. Bu nedenle geçerli ve güvenilir ölçme yöntemlerinin kullanılması önem arz etmektedir (Balcı ve Tekkaya, 2000). Öğrenme çıktılarını bakılarak kazanımlara ne düzeyde ulaşıldığını görmek için ölçme değerlendirme sürecinin özenli olarak yürütülmesi gerekmektedir (Karaca vd., 2016).

Tercih edilen ölçme araçları incelendiğinde, her birinin avantaj ve dezavantajları olduğu söylenebilir. Ölçme araçlarından; kısa yanıtlı sorular, açık uçlu sorular, boşluk doldurma soruları, doğru yanlış testleri, yazılı sınavlar öğrencilerin başarılarını ölçüp değerlendirmek adına kullanılanlardan birkaçıdır (Balcı ve Tekkaya, 2000). Çoktan seçmeli testler, öğrencilerin belirli konu veya kazanımdaki öğrenme düzeylerini tespit etmek amacıyla uygulanan psikometrik araçlardır. Öğrencilerin belirlenmiş bir konu ya da kavram ile ilgili bilgilerinin ortaya konması amacıyla en sık tercih edilen araç olduğu söylenebilir (Marzo ve Andrés, 2014). Çoktan seçmeli testler öğrencilerin bilgilerini ölçmesinin yanı sıra karar verme ve kıyaslama becerisi gibi farklı becerilerini de ölçebilmektedir. Her eğitim düzeyinde kullanılabilen ve çok sayıda soru bulundurması sayesinde de puanlama güvenirliliği yüksek ölçme araçlarıdır. Ayrıca çoktan seçmeli testlerdeki soruların doğru ve yanlış cevaplarının tek ve kesin olması puanlamanın daha nesnel olmasını sağlamaktadır (Temizkan ve Sallabaş, 2011). Bu gibi avantajlarla beraber aktarılması beklenen konu veya kazanım hakkında bilgisi olmayan öğrencinin soruyu doğru cevaplama olasılığı gibi bir dezavantajı da vardır. Bu nedenle konu hakkında bilgi düzeyleri farklı olan öğrencileri ayırt etmek güçleşmektedir (Haladyna, 1992). Bu dezavantajlar göz önünde bulundurulduğunda ölçme değerlendirme uzmanları iyi hazırlanmış bir çoktan seçmeli sınavın öğrenci başarısını ölçmede kullanılmasının uygun olduğu görüşünü savunmaktadırlar (Balcı ve Tekkaya, 2000).

Klasik test teorisinde başarı ölçümü yapılırken öğrencinin her bir sorudan aldığı puanlar toplanır. Puanlama yapılırken hazırlanmış her bir soruya eşit puan dağılımı verilir (Traub, 1997). Mevcut çalışma Klasik test teorisine göre geliştirilmiştir.

Her branşta olduğu gibi fen eğitiminde de başarı testleri, öğrencilerin işlenen konuları ne seviyede öğrendiklerinin saptanması yönünden objektif dönüt sağlayan ölçme değerlendirme araçlarıdır. Bu nedenle fen eğitiminde de başarı testlerinin öğrenme hedefine yönelik hazırlanması büyük önem taşımaktadır (Balcı ve Tekkaya, 2000).

Öğretmenler başarı testini hazırlarken tasarımı, soru hazırlama ve soruları son olarak kontrolden geçirmeye yeterli özeni gösterdiği taktirde öğrencilerin öğrenme seviyeleri hakkında doğru bilgiye ulaşacaklardır (Childs, 2020). Çoktan seçmeli testlerin yanı sıra yukarıda da ifade edildiği gibi çok farklı

ölçme araçları vardır. Örneğin açık uçlu testlerin öğrencilerin kavram yanılgılarını belirlemek, farklı alternatif anlatım tarzlarını görebilmek gibi avantajları bulunmaktadır (Karataş vd., 2003). Bunun yanı sıra uygulanma ve puanlanmasının zaman alması, kapsam geçerliğinin sağlanmasının zor olması, objektif değerlendirme yapmanın zor olması gibi dezavantajları barındırmaktadır (Romagnano, 2001). Çoktan seçmeli testler ise diğer ölçme araçlarına göre daha fazla kişiye uygulama şansı, objektif değerlendirme sağlanması, kapsam geçerliğinin yüksek olması gibi avantajlara sahiptir (Doğruöz, 2022). Farklı disiplinlere ve hemen hemen her biliş seviyesine rahatlıkla uyarlanabilmesi yönünden ölçme ve değerlendirme sürecinde çoktan seçmeli sorulara sıklıkla yer verilir (Marzo ve Andrés, 2014). Bu nedenlerle çoktan seçmeli testin avantajları göz önünde bulundurularak, mevcut çalışmada çoktan seçmeli bir başarı testi geliştirilmesi tercih edilmiştir.

Başarı testleri geliştirilirken testin hedefi, kapsamı, geçerlik ve güvenirlik çalışmaları göz önünde bulundurulmalıdır (Balcı ve Tekkaya, 2000; Haladyna, 1997). Hazırlanan testin öğrenme amaçlarına uygunluğu doğru bir ölçme ve değerlendirme yapmak dikkat edilmesi gereken önemli noktalardan biridir (Balcı ve Tekkaya, 2000). Bloom'un bilişsel alan taksonomisine göre zihinsel gelişim düzeyi en basitten en karmaşığa göre; bilgi, kavrama, uygulama, analiz, sentez ve değerlendirme basamaklarından oluşmaktadır (Bloom vd., 1956). Bu çalışmada kazanıma ilişkin soru yazılırken Bloom'un bilişsel alan taksonomisi basamakları dikkate alınmıştır. Geliştirilen test ile ortaokul yedinci sınıf düzeyindeki "*Maddenin Yapısı ve Özellikleri*" konusunda öğrencilerin üst düzey düşünme becerilerini ölçerek başarılarını değerlendirmek amaçlanmıştır. Aynı zamanda bu alanda çalışacak olan araştırmacılara katkı sağlaması umulmaktadır.

Fen bilimleri dersi öğrenmekte güçlük çekilen konuları barındırmaktadır (Noh ve Scharmann, 1997). "*Maddenin Yapısı ve Özellikleri*" de öğrencilerin öğrenmekte zorlandığı fen konularından biridir (Şimşek vd., 2016). Maddenin yapısını anlamak fen bilimlerindeki diğer konuların anlaşılması açısından önem taşımaktadır (Bouwma-Gearhart vd., 2009). Ortaokul müfredatından başlayıp lisede de devam eden "*Maddenin Yapısı ve Özellikleri*" konusu öğrencilerin zorlandığı bir konu olarak görülmektedir. Buna rağmen ortaokul müfredatından başlayıp lisede devam eden bu konuda, öğrencilerin ortaokul yıllarında konuyu karmaşık bulması sebebiyle ilerleyen yıllarda halen zorlandıkları görülmüştür (Sözbilir, 2002). Alanyazın incelendiğinde yedinci sınıf düzeyinde "*Maddenin Yapısı ve Özellikleri*" ile ilgili geliştirilmiş başarı testleri sayısının sınırlı olduğu görülmektedir (Nacaroğlu vd., 2019; Saraç, 2018; Üçüncü ve Sakız, 2020).

Başarı testlerinde iki kategorili bir sınıflama söz konusu olduğundan tetrakorik korelasyon kullanılması önerilmektedir (Hambleton ve Swaminathan, 1985; Olsson, 1979; Uebersax, 2015). Bunu yapmak için de SPSS yerine "*factor analysis*" gibi programlar kullanılmalıdır. Fen bilimleri yazınında maddenin yapısını konu alan başarı testi çalışmalarında tetrakorik korelasyona dayalı analiz yapılmadığı gözlenmiştir (Eren vd., 2020; Nacaroğlu vd., 2019; Saraç, 2018; Üçüncü ve Sakız, 2020). Mevcut çalışmada ise tetrakorik korelasyon matrisi kullanılarak başarı testi geliştirilmiştir. Bunun yanı sıra belirlenen kazanımlar Bloom'un bilişsel alan taksonomisi basamaklarına göre sınıflandırılmıştır. Bu konu ile ilgili başarı testi geliştirme çalışmalarının çoğunda da Bloom'un bilişsel alan taksonomisinin tercih edildiği gözlenmiştir (Avcı, 2019; Coşkun ve Sarıkaya, 2020; Demir vd., 2018; Kaplan vd., 2022; Kızılkapan ve Bektaş, 2018; Nacaroğlu ve Bektaş, 2019; Saraç, 2018; Üçüncü ve Sakız, 2020). Maddenin yapısını konu alan başarı testi geliştirme çalışmalarının çoğunda çalışmada kullanılan test kuramından bahsedilmemiştir (Eren vd., 2020; Saraç, 2018; Üçüncü ve Sakız, 2020). Bu çalışmada ise klasik test teorisi kullanılmış, bu teorelin tercih nedeni de açıklanmıştır.

Geliştirilen başarı testiyle MEB 2018 kazanımlarına uygun daha güncel bir çalışma ortaya konacağı düşünülmektedir. Bu bağlamda, mevcut araştırmada yedinci sınıf “*Maddenin Yapısı ve Özellikleri*” konusunda geçerli ve güvenilir bir başarı testi geliştirmek amaçlanmıştır. Bu amaç doğrultusunda aşağıdaki sorulara cevap aramıştır:

1. Yedinci sınıf öğrencilerinin “*Maddenin Yapısı ve Özellikleri*” konusunda başarı seviyelerinin belirlenmesinde geliştirilen başarı testi geçerli midir?

2. Yedinci sınıf öğrencilerinin “*Maddenin Yapısı ve Özellikleri*” konusunda başarı seviyelerinin belirlenmesinde hazırlanan başarı testi güvenilir midir?

## 2. YÖNTEM

Mevcut çalışma, 324 nolu başvuru ile Erciyes Üniversitesi Sosyal ve Beşerî Bilimler Etik Kurulu tarafından 29/06/2021 tarihinde verilen etik kurul onayı kapsamında yürütülmüştür.

### 2.1. Araştırma Deseni

Ortaokul yedinci sınıf öğrencilerinin “*Maddenin Yapısı ve Özellikleri*” konusunu öğrenme düzeylerini belirlemek amacı doğrultusunda nicel araştırma desenlerinden biri olan tarama deseni kullanılmıştır. Diğer desenlere kıyasla tarama deseni, bir konu veya duruma ilişkin hedef kitlenin ilgi, yetenek, görüş gibi kavramların ortaya çıkarılmasında etkili olan bir desendir (Creswell, 2009). Bu çalışmada ortaokul yedinci sınıf öğrencilerinin “*Maddenin Yapısı ve Özellikleri*” konusunda geçerli, güvenilir bir teste tabi tutulmalarını sağlamak amacı göz önünde bulundurularak tarama deseni kullanılmıştır. Ayrıca, test geliştirme çalışmaları incelendiğinde çalışmaların geliştirdikleri testlerin sonuçlarını bir evrene genelleme kaygısı taşıdıkları ve bu sebepten dolayı nicel araştırma yönteminin bir deseni olan tarama desenini kullandıkları belirlenmiştir (Gönültaş vd., 2020; Vurgun ve Bektaş, 2019). Bu çalışmada da yazarlar alanyazında yer alan tarama deseni kullanan çalışmaları doğrulamak ve evrene genelleme yapmak amacıyla tarama desenini tercih etmişlerdir.

### 2.2. Çalışma Grubu

Çalışmanın ulaşılabilir evreni, 2020-2021 eğitim ve öğretim döneminde Kayseri ili Melikgazi ilçesindeki yedinci sınıf öğrencilerinden oluşmaktadır. Küme örnekleme çalışma grubundaki bireylerin bulunduğu konumdan rastgele bir seçim yapılarak uygulanan veri toplama aracından heterojen sonuç almayı sağlar (Özdamar, 2001). Mevcut çalışma küme örnekleme kullanılarak Kayseri ili Melikgazi ilçesi yedinci eğitim bölgesinden seçilmiş üç ortaokulda bulunan ve yedinci sınıfta öğrenim gören 268 öğrenciden oluşmaktadır. Çalışma grubunda 116 erkek, 152 kız öğrenci yer almıştır. Genel bir yargıya göre örneklem büyüklüğü, soru sayısının en az beş katını karşılamalıdır (Büyüköztürk, 2002). Mevcut çalışma en az beş katı kuralını karşılamakta, yaklaşık on katı kuralına da uymaktadır. Dolayısıyla yeterli örneklem büyüklüğüne sahip olduğu söylenebilir. Çalışma kapsamında geliştirilen başarı testi 31 sorudan oluşturulmuş olup ardından yapılan analizler sonucunda 23 soruya indirgenip başarı testinin son hali oluşturulmuştur.

### 2.3. Veri Toplama Süreci

Mevcut araştırmada 2020-2021 eğitim öğretim dönemi bahar yarısında “*Maddenin Yapısı ve Özellikleri*” konusunun yedinci sınıflarda öğretilmesinin ardından Akademik Başarı Testi (ABT) uygulanmıştır. Başarı testi 2020-2021 eğitim öğretim yılında yüz yüze yapılmasının yanı sıra pandemi

sebebiyle eğitime getirilen kısıtlamalardan dolayı çevrimiçi olarak Google Form aracılığı ile de uygulanmıştır.

## 2.4. Ölçek Geliştirme Süreci ve Verilerin Analizi

Başarı testinin geliştirilme aşamasında geçerlik, güvenilirlik çalışmalarını yürütmek amacıyla *factor 11.04.02* ve SPSS 25 programları kullanılmıştır. Veriler SPSS 25 programına girilmiş ve normal dağılım, geçerlik, güvenilirlik, alt ve üst gruplar için t testine bakılmıştır. AFA ve DFA ise *factor 11.04.02* programıyla gerçekleştirilmiştir. Başarı testinin hedeflenen örnekleme uygulanmasının ardından katılımcıların cevap kağıtları okunmuş, aynı zamanda çevrimiçi katılan kişilerin yanıtları Google Forms üzerinden Excel programına aktarılıp kontrol edilmiştir. Bu süreçte doğru yanıtlar için “1”, yanlış ve boş yanıtlar için ise “0” puan değerleri verilmiştir. Böylece her soruyu doğru yapan öğrencinin “23” puan alması beklenirken hiçbir soruya doğru yanıt vermeyen öğrencinin “0” puan alması beklenmiştir.

Geçerlik ve güvenilirlik analizleri yapılmadan önce normallığe bakılmıştır. Faktör analizinin varsayımlarından biri de normal dağılımdır (Tabachnick ve Fidell, 2013). Normal dağılım varsayımını karşılamak için en çok kullanılan tekniklerden biri de basıklık-çarpıklık katsayılarının kontrolüdür (Pallant, 2007, s. 95-99). Bu bağlamda, betimsel istatistik değerleri, basıklık ve çarpıklık değerleri, histogram grafikleri incelenmiştir. Bulgularda ise basıklık ve çarpıklık değerlerine yer verilmiştir. Basıklık ve çarpıklık değerleri -2 ve +2 değer aralığında değerlendirilmiştir. Basıklık ve çarpıklık katsayı değerinin -1 ile +1 aralığında olması beklenir fakat -2 ile +2 değer aralığı da kabul edilebilir düzeydedir (George ve Mallery, 2010, s. 19). Mevcut testteki her bir soru Lorenzo-Seva ve Ferrando (2021) tarafından geliştirilen “*factor 11.04.02*” programı ile analiz edildiğinde basıklık ve çarpıklık değerlerinin -2 ile +2 arasında olduğu görülmüş; dolayısıyla bu testin normal dağılım varsayımını karşıladığı belirlenmiştir. Testin geçerliğini kontrol etmek için öncelikle madde katsayıları hesaplanmıştır. O nedenle yapı geçerliği doğrultusunda her bir soru için madde güçlük katsayıları, madde ayırt edicilik katsayıları hesaplanmış; açımlayıcı faktör analizi (AFA) ve doğrulayıcı faktör analizi (DFA) “*factor 11.04.02*” programı ile yapılmıştır (Lorenzo-Seva ve Ferrando, 2021). AFA yaparken amaç; daima en doğru sayıda faktör ile en yüksek açıklayıcılık yüzdesine sahip boyutlara ulaşmak ve soruları faktörler altında bir araya getirmektir. AFA sonucunda elde edilen alt boyutların geçerliğini göstermek üzere DFA uygulanmaktadır. DFA sonucunda elde edilen uyum katsayıları baz alınarak, söz konusu faktörlerin geçerli bir yapıya sahip olup olmadıkları değerlendirilmektedir (Pallant, 2016). Üst ve alt gruplar belirlenerek t testi ile soruların ayırt edicilikleri kontrol edilmiştir. Ardından KR-20 güvenilirlik katsayısı hesaplanmıştır.

## 3. BULGULAR

### 3.1. Geçerliğe İlişkin Bulgular

Belirli bir hedefe yönelik testin o amaca uygun olması ve zihinsel becerileri ölçme yeteneği geçerlik olarak adlandırılır (Pallant, 2007, s. 81-95). Nicel çalışmalarda geçerlik o çalışmanın doğruluğunu ve kalitesini gösterir (Heale ve Twycross, 2015). Bu çalışmada da kapsam ve yapı geçerliğine yönelik bulgulara yer verilmiştir.

#### 3.1.1. Kapsam Geçerliğine İlişkin Bulgular

Başarı testindeki çoktan seçmeli soruların hazırlanması sürecinde konuyla ilgili alanyazın taraması yapılmış olup araştırmalardan elde edilen sonuçlar doğrultusunda 46 soru içeren soru havuzu



oluşturulmuştur. Sorular Bloom'un bilişsel alan taksonomisi basamakları dikkate alınarak oluşturulmuştur. Taslak test, geçerliği ve güvenirligi kontrol edilmiş çalışmalardan (Ateş, 2018; Coştu vd., 2003; Güler, 2020; Karagöz ve Korkmaz 2015; Kızılcapan ve Bektaş, 2018; Minaslı, 2009; Özelma, 2019) ve MEB liselere geçiş sınavlarındaki çıkmış sorulardan (2010 SBS, 2011 SBS, 2013 PYBS, 2014 PYBS, 2015 PYBS) yararlanılarak oluşturulmuştur. Bu soru havuzunda 46 soru için bilgi basamağında altı, kavrama basamağında 10, analiz basamağında 12, uygulama basamağında 11 ve değerlendirme basamağında yedi soru bulunmaktadır.

Kapsam geçerliği, belirlenen konuda uzman kişilerin görüşlerinden yararlanılmasıdır (Heale ve Twycross, 2015). Kazanımlar Bloom'un bilişsel alan taksonomisi basamaklarına göre sınıflandırılıp MEB (2018) kazanımlarındaki karşılıkları belirtilmiştir. Ardından sorular bir ölçme değerlendirme uzmanı, bir fen bilimleri uzmanı, iki fen bilimleri öğretmeni ve bir Türkçe öğretmenin kontrolüne sunulmuştur. Türkçe öğretmeni soruları; anlatım bozukluklarının tespiti, yazım ve imla kuralları, dil ve anlatım yönüyle incelemiştir. Fen bilimleri öğretmenleri; soruların kazanımlarla ilişkisini, sınıf seviyesine göre soruların zorluğunu, aynı zamanda Bloom'un bilişsel alan taksonomisi basamaklarına göre incelemiştir. Ölçme değerlendirme uzmanı, oluşturulan soru havuzundan itibaren soruların kazanım ve Bloom'un bilişsel alan taksonomisi ilişkisini inceleyip düzenli aralıklarla dönütler vermiş ve genel bir değerlendirme yapmıştır. Örneğin, Karagöz ve Korkmaz'ın (2015) çalışmasındaki sorunun şekli, netleştirilip kullanılmıştır. Ateş'in (2018) çalışmasındaki 6. ve 21. sorularda okuma yönünden problemler olduğu görülmüş ve düzeltilmiştir. Sonuç olarak, 46 soru içeren soru havuzundan 15 tanesi uzman önerileri doğrultusunda testten çıkarılmıştır. Kalan sorular, ilgili MEB kazanımının işaret ettiği basamaklar göz önünde bulundurularak, her bir bilişsel basamakta soru bulunacak şekilde seçilip ölçme ve değerlendirme uzmanının kontrolüne sunularak başarı testi uygulamaya hazır hale getirilmiştir.

Bu çalışmada kazanıma uygun olarak hazırlanan sorular bilgi, kavrama, uygulama, analiz ve değerlendirme basamaklarında hazırlanmıştır. Bloom'un bilişsel alan taksonomi basamakları basitten karmaşığa doğru ilerleyerek bilginin ve hedef becerilerin gelişmesini ele alır (Bloom vd., 1956). Bloom'un bilişsel alan taksonomisi basamaklarından sentez basamağında soru eklenememiştir. Sentez basamağı, fikirlerin belirli kurallar çerçevesinde yeni bir ürün oluşturulması yeteneğini ifade eder (Birgin, 2016). Sentez basamağı yeni fikir, icat ve ürün oluşturulması olduğundan çoktan seçmeli testlerde yer alması, bu basamağın ve çoktan seçmeli testin doğasına uygun değildir.

Başarı testi soruları ve cevap anahtarı Ek.1 ile çalışmaya eklenmiştir. MEB (2018) kazanımları ise Tablo 1'de sunulmuştur. Taslak testin kazanımları, soruların bulunduğu taksonomik basamaklar ve sorulara karşılık gelen MEB kazanımları Tablo 2'de verilmiştir. Başarı testi sorularının kazanımları, bulunduğu taksonomi basamağı Tablo 2'de verilmiştir.

**Tablo 1: MEB (2018) Kazanımları Tablosu**

Kazanım Numarası	MEB (2018) Kazanımları
MB1	Atomun yapısını ve yapısındaki temel parçacıklarını söyler.
MB2	Geçmişten günümüze atom kavramı ile ilgili düşüncelerin nasıl değiştiğini sorgular.
MB3	Bilimsel bilginin zamanla değişebileceğine vurgu yapılıır.

<b>MB4</b>	Aynı veya farklı atomların bir araya gelerek molekül oluşturacağını ifade eder.
<b>MB5</b>	Saf maddeleri, element ve bileşik olarak sınıflandırarak örnekler verir.
<b>MB6</b>	Periyodik sistemdeki ilk 18 elementin ve yaygın elementlerin (altın, gümüş, bakır, çinko, kurşun, cıva, platin, demir ve iyot) isimlerini, sembollerini ve bazı kullanım alanlarını ifade eder.
<b>MB7</b>	Yaygın bileşiklerin formüllerini, isimlerini ve bazı kullanım alanlarını ifade eder.
<b>MB8</b>	Karışımları, homojen ve heterojen olarak sınıflandırarak örnekler verir.
<b>MB9</b>	Homojen karışımların çözelti olarak da ifade edilebileceği vurgulanır.
<b>MB10</b>	Günlük yaşamda karşılaştığı çözücü ve çözünenleri kullanarak çözelti hazırlar.
<b>MB11</b>	Çözünme hızına etki eden faktörleri deney yaparak belirler.
<b>MB12</b>	Temas yüzeyi, karıştırma ve sıcaklık faktörlerine değinilir.
<b>MB13</b>	Bağımlı, bağımsız ve kontrol edilen değişken kavram gruplarına vurgu yapılır.

**Tablo 2: Başarı Testine Ait Belirtke Tablosu**

Soru Numarası	Kazanımlar						MEB Kazanımları
		Bilgi	Kavrama	Uygulama	Analiz	Değerlendirme	
1	Geçmişten günümüze atom kavramı ile ilgili atom teorilerini bilir.	X					MEB2
2	Molekül kavramının ne olduğunu ve özelliklerini bilir.	X					MEB4
3	Bilimsel bilgilerin zamanla değişebileceği çıkarımını yapar.					X	MEB3
4	Çözünme hızına etki eden faktörleri yorumlar.					X	MEB11
5	Atomun yapısında bulunan tanecikleri kavrar.		X				MEB1
6	Element sembollünü bilir.	X					MEB5
7	Çözelti, çözünen, çözen kavramlarını bilir.	X					MEB9
8	Elementler, sembolleri ve günlük yaşamdaki kullanımlarını kavrar.		X				MEB10
9	Atom ve molekülü kavrar.		X				MEB4
10	Yaygın elementlerin sembollerini, günlük hayattaki kullanım yerlerini ifade eder.		X				MEB7
11	Maddenin tanecikli yapıda olduğuna örnekler verir.		X				MEB10
12	Saf maddeleri element ve bileşik olarak sınıflandırarak yorumlar.		X				MEB5
13	Atomların elektron alma verme eğilimlerini kavrar.		X				MEB6
14	Molekülleri modellerine göre sınıflandırır.			X			MEB5
15	Element, bileşik, karışımların özelliklerini model üzerinde ayırt eder.			X			MEB5
16	Periyodik sistemdeki ilk 18 elementin ve yaygın elementlerin (altın, gümüş, bakır, çinko, kurşun, cıva, platin, demir ve iyot) isimlerini, sembollerini problem durumlarına göre uyarlar.			X			MEB5
17	Maddelerin taneciklerinin hareketinde sıcaklığın etkisini çözümler.			X			MEB11



18	Günlük yaşamda karşılaştığı maddelerin formüllerini çözümler.	X					MEB7
19	Atomun boyutunu günlük hayattan örneklerle karşılaştırır.	X					MEB10
20	Günlük yaşamda karşılaştığı çözücü ve çözünenlere örnekler verir.		X				MEB10
21	Elementlerin sembollerle gösterilmesinin sebeplerini yorumlar.			X			MEB7
22	Atomun yapısını, proton elektron nötron özelliklerini model üzerinde çözümler.	X					MEB1
23	Elementlerin günlük hayattaki kullanımını, formülünü, özelliklerini analiz eder.		X				MEB4
24	Saf maddeleri, element ve bileşik olarak sınıflandırarak model üzerinde gösterir.	X					MEB12
25	Sıvı ve gazların sıkıştırılabilme özellikleri olduğunu kanıtlar.			X			MEB12
26	Elementlerin özelliklerinin hangisine ait olduğunu ayırt eder.		X				MEB4
27	Bileşik, çözücü, çözen, çözünen kavramlarını analiz eder.		X				MEB8
28	Verilen karışımların çözücü olup olmadığına karar verir.		X				MEB10
29	Çözücü, çözünen, çözücü gibi maddeleri verilen örneklerle analiz eder.		X				MEB10
30	Çözünme hızına etki eden faktörleri analiz eder.		X				MEB12
31	Çözünme hızına etki eden faktörlerin hangilerinin etkili olduğuna karar verir.		X				MEB13
<b>Toplam</b>			<b>4</b>	<b>7</b>	<b>8</b>	<b>8</b>	

Yapı geçerliğine ilişkin analizlerden sonra 1., 9., 18., 19., 22., 26., 28. ve 30. sorular çıkartılmış ve teste 23 soru kalmıştır. Dolayısıyla geliştirilen başarı testinin Bloom'un bilişsel alan taksonomisi ve MEB kazanımları açısından kapsam geçerliğini koruduğu söylenebilir.

### **3.1.2 Yapı Geçerliğe İlişkin Bulgular**

Yapı geçerliği kapsamında uzman görüşüne başvurulmuştur. Fen eğitiminde uzman olan öğretim elemanları mevcut başarı testinin maddelerinin kazanıma uygun soru olup olmaması yönüyle değerlendirmeler sunmuşlardır. Örneğin, ölçme değerlendirme uzmanı yedinci sınıf 2013 PYBS sorusunu Türkçe dersine yönelik kazanımı barındırdığı düşüncesiyle soru havuzundan çıkarmayı doğru bulmuştur. Fen bilimleri öğretmenleri ise soruların öğrencilerin seviyesine uygunluğu bakımından ve doğru kazanımı ölçüp ölçmemesi yönüyle değerlendirmeler yapmışlardır. Türkçe öğretmeni Türkçe dil bilgisi kurallarını dikkate alarak soruları yazım ve imla kurallarına göre incelemiş, anlaşılabilirliğe katkı sağlamıştır. Örneğin, soruların daha anlaşılır bir şekilde oluşturulmasını sağlamıştır. Soru köklerindeki olumsuz ifadelerin koyu renkle vurgulanmasını önermiştir. Çoktan seçmeli testin cevap anahtarı oluşturulurken doğru cevaplar için her seçenek sayısının birbirine yakın olması sağlanmıştır. Örneğin sorulardan doğru cevabı "A" olan beş tane, cevabı "B" olan sekiz tane, cevabı "C" olan dokuz tane ve cevabı "D" olan dokuz soru belirlenmiştir. Ölçme değerlendirme uzmanı taslak soruları incelemiş ve başarı testinin son halinin uygulanmaya hazır olduğunu belirtmiştir.

### 3.1.3. Madde Analizleri

Testin geçerliğini kontrol etmek için güçlük katsayıları ve ayırt edicilik katsayılarına bakılmıştır. Bu hedef doğrultusunda, öğrencilerin başarı testi sonucunda elde ettikleri puanlar en yüksek puan alandan, en düşük puan alana doğru sıralama yapıldıktan sonra katılımcı sayısından %27'lik üst grup ile %27'lik alt grup belirlenmiştir (Beuchert ve Mendoza, 1979). Alt ve üst grupta 72 kişi vardır. Madde güçlük katsayısı, alt ve üst gruptan soruya doğru verilen yanıt sayısının alt ve üst gruptaki toplam kişi sayısına bölünmesi sonucu elde edilmiştir. Madde ayırt edicilik katsayısı ise üst gruptan soruya doğru cevap verenlerin sayısından alt gruptan soruya doğru cevap verenlerin sayısı arasındaki farkın alt grup sayısına bölümüyle elde edilerek bulunur.

Her soru için üst gruptan doğru cevap verenlerin sayısından alt gruptakilerde soruya doğru cevap verenlerin sayısı çıkarılıp, sonuç alt ya da üst gruplardaki kişi sayılarından birine bölünmüştür. Bu çözüm yöntemi her soru için Excel programı kullanılarak uygulanmıştır. Soruların güçlük katsayısı ve madde ayırt edicilik katsayısı Tablo 3'te gösterilmiştir:

**Tablo 3: Soruların Güçlük ve Madde Ayırt Edicilik Katsayısı**

Soru	Güçlük Katsayısı (p)	Madde Ayırt Edicilik Katsayısı (r)	Soru	Güçlük Katsayısı (p)	Madde Ayırt Edicilik Katsayısı (r)	Soru	Güçlük Katsayısı (p)	Madde Ayırt Edicilik Katsayısı (r)	Soru	Güçlük Katsayısı (p)	Madde Ayırt Edicilik Katsayısı (r)
1	0,10	0,13	9	0,11	0,09	17	0,30	0,72	25	0,30	0,65
2	0,26	0,58	10	0,30	0,77	18	0,16	0,09	26	0,11	0,22
3	0,25	0,65	11	0,24	0,38	19	0,12	0,08	27	0,24	0,56
4	0,40	0,43	12	0,28	0,68	20	0,28	0,66	28	0,12	0,18
5	0,29	0,66	13	0,20	0,38	21	0,31	0,58	29	0,23	0,61
6	0,31	0,59	14	0,31	0,65	22	0,10	0,11	30	0,16	0,29
7	0,29	0,72	15	0,22	0,61	23	0,23	0,50	31	0,30	0,62
8	0,20	0,52	16	0,25	0,61	24	0,23	0,50			

Buradan yola çıkarak dördüncü soru hariç diğer tüm soruların zor olduğu, dördüncü sorunun ise orta güçlükte olduğu söylenebilir. Her bir sorunun güçlük katsayısı değerlerinin toplamının madde sayısına oranının 0,50 civarlarında olması beklenir. Bu değer o testin ortalama zorluk seviyesinde olduğunu gösterir (Pallant, 2013). Testin geneline bakıldığında bu işlem sonucunda 0,23 ile testin zor bir test olduğu söylenebilir. Yalnızca madde güçlük katsayısına bakılarak soru çıkarmak doğru değildir (Nacaroğlu vd., 2020). Bu nedenle testteki soruların ayırt edicilik katsayıları da hesaplanmıştır.

Madde ayırt edicilik katsayısı, değerler kontrol edilip yorumlanırken; her soru bazında ulaşılan değer 0,40 ve üzeri ise maddenin ayırt edicilik değeri yüksek, 0,30 ve 0,39 arasında ise orta düzeyde ayırt edici soru olduğu söylenmektedir. Bunun yanı sıra 0,20 ve 0,29 arasında ise sorunun düzeltilmesi gerektiğini, 0,19 ve altında ise o sorunun ayırt ediciliğinin oldukça düşük olduğunu ve testten atılması gerektiğini göstermektedir (Pallant, 2007, s. 179-199).

Elde edilen verilere göre, 11. ve 13. soruların ayırt ediciliklerinin orta düzeyde, 1, 9, 18, 19, 22 ve 28. soruların ayırt ediciliklerinin düşük düzeyde ve bu sorular dışındaki diğer soruların ayırt ediciliklerinin yüksek düzeyde oldukları bulunmuştur. Testteki 26 ve 30. sorular ise düzeltilip teste eklenmiştir.

Mevcut başarı testi soru dağılımları belirtke tablosuna göre Bloom'un bilişsel alan taksonomisine göre oluşturulmuş olup orta güçlükte bir test hedeflenmiştir. Testin 0,50 güçlük katsayısı civarlarında bir

değere sahip olması beklenirken, örneklemdaki öğrencilerin genel başarı seviyelerinden ötürü bu değer daha düşük çıkmıştır. Bu durum testin zor denebilecek düzeyde bir değer elde etmesine sebep olmuştur.

### 3.1.4. Alt ve Üst Gruplar için t Testi

Soruların ayırt ediciliği ve geçerliği için kullanılacak bir diğer yöntem de t testidir (Saidi ve Siew, 2019). Belirlenen iki grup arasında bir fark olup olmadığını test etmek amacı ile t testi uygulanır (Kim, 2015; Pallant, 2007, s. 232). %27'lik dilimler ile alt ve üst gruplar belirlendikten sonra bu iki grup arasında anlamlı bir fark olup olmadığına bakılmış ve Tablo 4 oluşturulmuştur.

**Tablo 4:** Alt ve Üst Gruplar için Bağımsız Örneklem T Testi

Sorular	Gruplar	N	Ortalama	t	Standart Hata	Standart Hata Ortalaması	p
S1	üst grup	72	,26	2,13	,44	,05	0,010
	alt grup	72	,13		,33	,04	
S2	üst grup	72	,78	8,56	,42	,05	0,026
	alt grup	72	,19		,40	,05	
S3	üst grup	72	,81	10,28	,40	,05	0,025
	alt grup	72	,15		,40	,04	
S4	üst grup	72	,97	6,91	,16	,02	0,040
	alt grup	72	,54		,50	,06	
S5	üst grup	72	,88	10,73	,33	,04	0,029
	alt grup	72	,21		,41	,05	
S6	üst grup	72	,89	9,11	,32	,04	0,031
	alt grup	72	,29		,46	,05	
S7	üst grup	72	,92	12,61	,28	,03	0,029
	alt grup	72	,19		,40	,05	
S8	üst grup	72	,65	7,67	,48	,06	0,020
	alt grup	72	,13		,33	,04	
S9	üst grup	72	,26	1,42	,44	,05	0,011
	alt grup	72	,17		,38	,04	
S10	üst grup	72	,96	15,12	,20	,02	0,030
	alt grup	72	,18		,39	,05	
S11	üst grup	72	,65	5,05	,48	,06	0,024
	alt grup	72	,26		,44	,05	
S12	üst grup	72	,88	11,11	,33	,04	0,028
	alt grup	72	,19		,40	,05	
S13	üst grup	72	,57	5,23	,50	,06	0,020
	alt grup	72	,18		,39	,05	
S14	üst grup	72	,92	10,57	,28	,03	0,031
	alt grup	72	,26		,44	,05	
S15	üst grup	72	,72	9,41	,45	,05	0,022
	alt grup	72	,11		,32	,04	
S16	üst grup	72	,78	9,22	,42	,05	0,025
	alt grup	72	,17		,37	,04	
S17	üst grup	72	,93	12,70	,26	,03	0,030
	alt grup	72	,21		,41	,05	
S18	üst grup	72	,36	1,26	,48	,06	0,016
	alt grup	72	,26		,44	,05	
S19	üst grup	72	,28	1,17	,45	,05	0,012
	alt grup	72	,19		,40	,05	
S20	üst grup	72	,86	10,69	,35	,04	0,028
	alt grup	72	,19		,40	,05	
S21	üst grup	72	,88	8,74	,33	,04	0,031
	alt grup	72	,29		,46	,05	
S22	üst grup	72	,25	1,69	,44	,05	0,010
	alt grup	72	,14		,35	,04	
S23	üst grup	72	,69	6,94	,46	,05	0,023

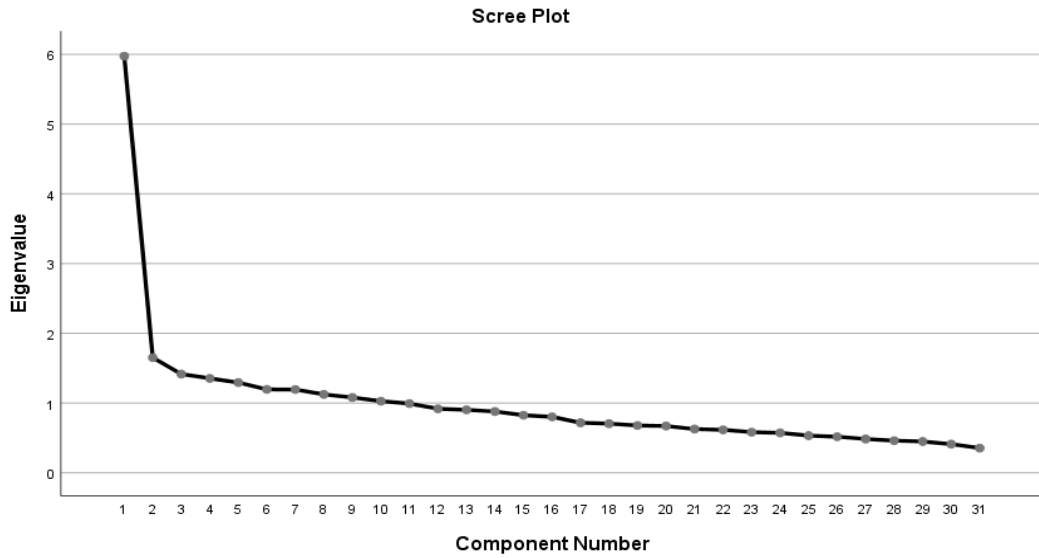
	alt grup	72	,19		,40	,05	
S24	üst grup	72	,68	6,97	,47	,05	0,023
	alt grup	72	,18		,39	,05	
S25	üst grup	72	,89	10,41	,32	,04	0,030
	alt grup	72	,24		,43	,05	
S26	üst grup	72	,33	3,30	,47	,06	0,011
	alt grup	72	,11		,32	,04	
S27	üst grup	72	,74	8,31	,44	,05	0,024
	alt grup	72	,17		,37	,04	
S28	üst grup	72	,32	2,62	,47	,05	0,012
	alt grup	72	,14		,35	,04	
S29	üst grup	72	,75	9,29	,44	,05	0,023
	alt grup	72	,14		,35	,04	
S30	üst grup	72	,46	3,95	,50	,06	0,016
	alt grup	72	,17		,37	,04	
S31	üst grup	72	,89	9,73	,32	,04	0,030
	alt grup	72	,26		,44	,05	

Bağımsız örneklem t testi dikkate alındığında testin bütün sorularının üst grup lehine olduğu görülmektedir. Bu da testin ayırt edicilik katsayısının yeterli düzeyde olduğunu göstermektedir.

### 3.1.5. Açımlayıcı Faktör Analizi (AFA)

Yapı geçerliği, faktörlerin uyumu ve ayırt ediciliği hakkında yapılan çalışmaları kapsar (Harvey vd., 2004). Testte bulunan soruların yapı geçerliğine bakmak için AFA (Açımlayıcı faktör analizi) yürütülmüştür. AFA bir ölçek veya başarı testinde, altta yatan konuyu açığa çıkarıp daha iyi bir şekilde tanımlanmasını sağlayan istatistiksel bir uygulamadır (Marsh vd., 2009). AFA'nın amacı, değişkenleri boyutlandırmak veya faktör analizi sonucunda bir veya birden fazla faktör elde etmektir (Akyüz, 2018).

Mevcut testin AFA'sı "*factor 11.04.02*" programı kullanılarak yapılmıştır (Lorenzo-Seva ve Ferrando, 2021). Normal dağılım varsayımını sağlayan değişkenlerin iki kategori şeklinde gruplandırılması durumunda faktör yapısı tetrakorik korelasyona dayalı olarak incelenir (Hambleton ve Swaminathan, 1985; Olsson, 1979; Uebersax, 2015). Tetrakorik korelasyon matrisine dayalı faktör analizi verileri ".dat" dosyasına dönüştürülerek "*factor 11.04.02*" programında yapılmıştır (Lorenzo-Seva ve Ferrando, 2021). Faktör analizinin yapılması için Bartlett's test değerinin anlamlı olması ve Kaiser-Meyer-Olkin (KMO) değerinin 0,6'dan yüksek çıkması beklenmektedir (Pallant 2007, s. 179-199). Bu testin KMO değerinin 0,79 olduğu görülmüştür. Mevcut başarı testinin yeterli örneklem varsayımına ulaştığı tespit edilmiştir. Ayrıca Bartlett's test değerinin anlamlı olduğu belirlenmiştir (Tablo 5) (Pallant, 2007, s.179-199; Tabachnick ve Fidell, 2013, s.657). Tüm bunlardan hareketle, yazarlar faktör analizinin yürütülebileceğine karar verip tetrakorik korelasyon yöntemini kullanmışlardır.



Şekil 1: Yamaç Birikinti Grafiği

Tablo 5: Açıklanan Toplam Varyans

Bileşen	Toplam Açıklanan Varyans						
	Özdeğer			Yüklenen Faktörlerin Karelerin Dağılımı			Yüklenen Faktörlerin Döndürme Dağılımı
	Toplam	Yüzde Varyans	Birikinti Yüzdesi	Toplam	Yüzde Varyans	Birikinti Yüzdesi	Toplam
1	5,97	19,27	19,27	5,97	19,27	19,27	4,08
2	1,65	5,33	24,60	1,65	5,33	24,60	1,50
3	1,41	4,56	29,16	1,41	4,56	29,16	1,42
4	1,35	4,37	33,53	1,35	4,37	33,53	1,56
5	1,29	4,17	37,70	1,29	4,17	37,70	2,64
6	1,19	3,86	41,56	1,19	3,86	41,56	1,58
7	1,19	3,85	45,40	1,19	3,85	45,40	3,11
8	1,12	3,62	49,03	1,12	3,62	49,03	2,48
9	1,08	3,48	52,51	1,08	3,48	52,51	1,18
10	1,03	3,31	55,82	1,03	3,31	55,82	2,12

AFA'yı yürütmek için eğik döndürme işlemi yapılmıştır. Yazarlar faktör yapılarından ve faktörlere dahil olacak sorulardan emin olamadıkları için dik döndürme yerine eğik döndürme (direct oblimum) tekniği kullanmışlardır. Ayrıca, yazarlar faktörlerin birbirleri arasındaki korelasyonlarının 0,30'dan büyük çıkma ihtimallerini göz önünde bulundurarak bu döndürme tekniğini tercih etmişlerdir. Yapılan ilk analizde bazı faktörler arasındaki korelasyonların 0,30'dan büyük olması bu ihtimali doğrulamıştır (Pallant, 2007; Seçer, 2015). Yapılan ilk analiz sonucunda 10 faktörlü bir yapı elde edilmiş ve bu yapının varyansın %55,82'sini açıkladığı görülmüştür. Açıklanan varyansa ve faktörlere ait özdeğerlere ilişkin değerler Tablo 5'te sunulmuştur.

Faktörlerin içerisindeki maddeleri çıkarmak için binişikliklerine bakıldı. Maddelerin arasında binişiklik olmadığından yamaç birikinti grafiği esas alınmış ve yukarıdaki yamaç birikinti grafiğinde de görüldüğü gibi tek faktörlü bir yapı olacağına karar verilmiştir. Buradan hareketle tek faktörlü yapı üzerinden tetrakorik korelasyon faktör analizi yürütülmüştür.

Şekil 1'deki yamaç birikinti grafiğine göre eğim birinci faktörden hemen sonra düzleşmeye başlamaktadır. Bu değerden sonra faktörlerin varyansa olan katkısının yetersiz olacağı görülmektedir (Pallant, 2007, s.182-183). Yamaç birikinti grafiği bu testin tek faktörlü olduğunu göstermektedir (Tablo 7). Buradan hareketle, 31 sorudan oluşan tek faktörlü yapı için tetrakorik korelasyon faktör analizi gerçekleştirilmiştir. Analiz bulgularına göre 1, 9, 18, 19, 22, 26, 28 ve 30. soruların faktör yüklerinin 0,30'un altında olduğu belirlenmiş, bu sorular testten çıkartılmıştır. Kalan 23 soru için fen eğitiminde uzman iki kişi testin kapsam geçerliğini kontrol etmiş ve uygun bulmuştur. Tekrarlanan faktör analizi sonucunda elde edilen tek faktörlü yapıya ait faktör yükleri Tablo 6'da verilmiştir.

**Tablo 6:** Başarı Testine Ait Faktör Yük Değerleri

Soru Numarası	Faktör 1
2	0,56
3	0,61
4	0,47
5	0,65
6	0,61
7	0,68
8	0,48
10	0,73
11	0,35
12	0,63
13	0,39
14	0,65
15	0,64
16	0,64
17	0,62
20	0,65
21	0,62
23	0,52
24	0,48
25	0,61
27	0,55
29	0,59
31	0,58

Yapılan AFA analizi sonucunda varyansın %37'sini açıklayan, 23 sorudan oluşan tek faktörlü bir yapı olduğu bulgusuna ulaşılmıştır. Tek faktörlü ölçeklerde açıklanan varyansın %30 ve daha fazla olması yeterli görülmektedir (Büyüköztürk, 2002; Çokluk vd., 2010).



**Tablo 7: Özdeğer ve Varyans Oranı**

Bileşen	Özdeğer	Varyans Oranı	Kümülatif Varyans Oranı
1	8,549	0,371	0,371
2	1,489	0,064	
3	1,271	0,055	
4	1,229	0,053	
5	1,126	0,048	
6	1,052	0,045	
7	0,964	0,041	
8	0,883	0,038	
9	0,794	0,034	
10	0,670	0,029	
11	0,629	0,027	
12	0,608	0,026	
13	0,554	0,024	
14	0,529	0,023	
15	0,488	0,021	
16	0,395	0,017	
17	0,367	0,015	
18	0,359	0,015	
19	0,338	0,014	
20	0,227	0,009	
21	0,216	0,009	
22	0,157	0,006	
23	0,097	0,004	

### 3.1.6. Doğrulayıcı Faktör Analizi (DFA)

Bu çalışmada geliştirilen testin DFA'sı, "factor 11.04.02" programı ile gerçekleştirilmiştir (Lorenzo-Seva ve Ferrando, 2021). DFA neticesinde ulaşılan Ki-kare (411.152) serbestlik derecesine (230) bölündüğünde sonucun beşin altında olması beklenmektedir (Hu ve Bentler, 1999). Bu oran seçilmiş örneklem grubundan alınmış olup, değerün üçün altında olması mükemmel uyumu göstermektedir (Jöreskog ve Sörbom, 1993, s. 10-11). Mevcut çalışmada bu sonucun 1,78 olduğu belirlenmiş, verinin mükemmel uyuma sahip olduğu bulgusuna ulaşılmıştır (Tablo 8).

DFA'da kullanılan RMSEA değerinin 0,00-0,05 arasında olması kabul edilebilir uyum, 0,08-0,10 vasat uyum, 0,10'dan büyük değerler kabul edilemez olarak belirtilmiştir (Schermelleh-Engel, Moosbrugger ve Muller, 2003). Bu testten elde edilen RMSEA değerinin 0,04 olması, bu testin uyumunun mükemmel olduğunu göstermektedir.

**Tablo 8: Doğrulayıcı Faktör Analizi Uyum İyiliği İndeks Tablosu**

Uyum Ölçüsü	Mükemmel Uyum	Kabul Edilebilir Uyum	Gözlenen Değer	Uyum Durumu
$\chi^2 / sd$	$0 \leq \chi^2 / sd \leq 2$	$2 \leq \chi^2 / sd \leq 3$	1,78	Mükemmel
p değeri	$0,05 \leq p \leq 1,00$	$0,01 \leq p \leq 0,05$	1,00	Mükemmel
RMSEA	$0 \leq RMSEA \leq 0,05$	$0,05 \leq RMSEA \leq 0,08$	0,04	Mükemmel
NNFI	$0,97 \leq NNFI \leq 1,00$	$0,95 \leq NNFI \leq 0,97$	0,98	Mükemmel
CFI	$0,97 \leq CFI \leq 1,00$	$0,95 \leq CFI \leq 0,97$	0,99	Mükemmel
GFI	$0,95 \leq GFI \leq 1,00$	$0,90 \leq GFI \leq 0,95$	0,96	Mükemmel
AGFI	$0,90 \leq AGFI \leq 1,00$	$0,80 \leq AGFI \leq 0,90$	0,95	Mükemmel

DFA analizi sonucunda, AFA ile belirlenen tek faktörlü yapı doğrulanmıştır.

### 3.2.1. Güvenirlğe İlişkin Bulgular

Güvenirlk bir testin birden fazla defa uygulanmasından sonraki ölçme sonuçlarının tutarlı olmasıdır. İç tutarlılık, bir testteki puanların güvenirlğini kontrol etmeyi amaçlar (Cronbach, 1951). Bir testin iç tutarlılık kat sayısını hesaplamının yollarından biri de Kuder Richardson (KR-20) kat sayısının hesaplanmasıdır (Heale ve Twycross, 2015). KR-20 başarı testlerinin iç tutarlılığını ölçmek için kullanılırken Cronbach alfa ise birden beşe kadar puanlanabilen ölçeklerde kullanılmaktadır (McMillan, 1992). KR-20, Cronbach alfa ile aynı alfa testini temsil etmektedir (Bardhoshi ve Erford, 2017). Genel kamı başarı testlerinde KR-20, ölçeklerde ise Croanbach alfa testinin yapılması yönündedir (Cronbach, 1951).

**Tablo 9: KR-20 Güvenirlk Katsayısı Değeri**

KR-20	KR 20 Standartlaştırılmış Maddelere Dayalı	N
0,86	0,86	23

Tablo 9 incelendiğinde KR-20 değerinin 0,86 olduğu görülmektedir. KR-20 değerinin 0,70 ve üzerinde olması o testin güvenilir bir test olduğunu göstermektedir (Heale ve Twycross, 2015). Elde edilen veriler ışığında, mevcut testin güvenilir bir test olduğu bulgusuna ulaşılmıştır.

**Tablo 10: İlgili Soru Çıkartıldığında Testin Aldığı KR-20 Değerleri**

Soru	KR-20 soru çıkartılırsa
S2	0,858
S3	0,857
S4	0,861
S5	0,856
S6	0,857
S7	0,855
S8	0,860
S10	0,854
S11	0,864
S12	0,856
S13	0,863
S14	0,856
S15	0,857
S16	0,856
S17	0,857
S20	0,856
S21	0,857
S23	0,859
S24	0,860
S25	0,857
S27	0,859
S29	0,858
S31	0,858

Tablo 10'da her bir soru testten çıkarıldığında testin KR-20 değerindeki değişiklik gösterilmiştir. KR-20 değerinin 0,86 olduğundan yola çıkılarak her bir sorunun güvenirlk katsayısı kontrol edilmiştir. Bu değer toplam güvenirlği etkilemediği için güvenirlk katsayılarına ilişkin soru çıkarılmamıştır.

AFA'dan önce ve sonra güvenirlige bakılmıştır. Herhangi bir maddenin engel teşkil etmediği düşünülerek AFA yürütülmüştür. Bunun ardından AFA'da son hâli verilen testin tekrar puanlarının güvenirligine bakılıp sonuçlar Tablo 10'da sunulmuştur.

#### **4. TARTIŞMA ve SONUÇ**

Bu çalışma yedinci sınıf fen bilimleri dersi “*Maddenin Yapısı ve Özellikleri*” konusunda güvenirligi ve geçerligi olan bir başarı testi geliştirmek hedefi doğrultusunda gerçekleştirilmiştir. Mevcut başarı testi geliştirilirken testin amacı, ünite ve konu kazanımlarının belirlenmesi, belirtke tablosu ve başarı testi soru havuzunun oluşturulması, soruların uzman kontrolünden geçirilmesi ve ardından soruların uzmanların dönütleri sonucu düzenlenmesi ile kapsam geçerligi kontrolleri yapılmıştır. Ulaşılabilir evren ve örneklem belirlenerek, uzaktan eğitim süreci sebebiyle başarı testi hem çevrimiçi ortamda hem de yüz yüze uygulanmıştır.

SPSS 25 programı kullanılarak elde edilen verilerin normal dağılımı kontrol edilmiş, testin güvenirlilik analizleri yapılmıştır. Faktör programı kullanılarak testin yapı geçerligi kontrol edilmiştir.

Alanyazında fen eğitimi çalışmaları kapsamında geliştirilen başarı testlerinin de benzer hedefleri dikkate aldıkları görülmüştür (Demir vd., 2016; Nacaroglu vd., 2019). Fakat yapılan bazı çalışmalarda hangi test teorisinin rehber alındığı belirtilmemiştir (Ayvaci ve Durmuş, 2016; Çardak ve Selvi, 2018; Saraç, 2017; Üçüncü ve Sakız, 2020; Yıldız vd., 2019). Mevcut çalışmada ise öğrencilerin her soruya eşit önemi vermeleri hedef olarak belirlenip bu hedef doğrultusunda klasik test teorisi rehber alınmıştır. Başarı testi soruları Bloom'un bilişsel alan taksonomisinin bilgi, kavrama, uygulama, analiz ve değerlendirme basamaklarından hazırlanmıştır. Bu çalışmanın yeni araştırmacılara bu yönüyle rehber olması umulmaktadır. Geliştirilen testin kapsam geçerligini sağlamak için konuyla ilgili alanyazın taraması yapılmış ve araştırma sonuçları doğrultusunda 46 soru içeren soru havuzu oluşturulmuştur. Kazanımlar Bloom'un bilişsel alan taksonomisi basamaklarına göre belirlendikten sonra uzman kontrolüne sunulmuştur.

Alanyazındaki başarı testi geliştirme çalışmaları incelendiğinde; kazanımların belirlenmesi, kazanımlara uygun soruların hazırlanması, uzman görüşünün alınması basamaklarının uygulandığı görülmüştür (Altay ve Kışla, 2019; Avcı, 2019; Aygan ve Zengin, 2017; Ayvaci, Durmuş, 2016; Çardak ve Selvi 2019; Karlı, vd., 2019; Özkan ve Yadigaroglu, 2020; Saraç 2018; Üçüncü ve Sakız, 2020; Yıldız, vd., 2019). Öte yandan çalışmaların bazılarında uzman kontrolüne yer verilmediği belirlenmiştir (Narlı ve Baser, 2008; Varış ve Cesur, 2012). Mevcut çalışmada ise sorular bir ölçme değerlendirme uzmanı, bir fen bilimleri uzmanı, iki fen bilimleri öğretmeni ve bir Türkçe öğretmeni olmak üzere toplam beş adet uzmanın değerlendirmesine başvurulmuştur. Bu yönüyle çalışmanın farklı alanlardaki uzmanların görüşlerine başvurularak kapsam geçerligi için daha derin bir bakış açısı sunduğu söylenebilir.

Bir başarı testinin güçlülüğünün 0,50 dolaylarında olması kabul görmektedir (Haladyna, 1997). Geliştirilen testin güçlük katsayısı 0,23 olup genel anlamda testin güçlülüğünün yüksek olduğu görülmüştür. Bu durumun sebeplerinden biri olarak, yazarlar teste cevap veren örneklemin başarı düzeyinin düşük olmasını göstermişlerdir. Örneklemdaki öğrencilerin Kayseri'de orta düzeyde bir başarı gösteren okullarda okumuş olmaları da bu duruma bir sebep olabilir. Diğer bir sebep olarak, yazarlar testin cevaplanma biçimini göstermişlerdir. Örneklemdaki öğrencilerin bir kısmı çevrimiçi olarak, diğer

bir kısmı ise yüz yüze testi cevaplamışlardır. Çevrimiçi testi cevaplayan öğrencilerin sayısı (55) düşük olmasına rağmen bu öğrencilerin puanlarının daha düşük olduğu tespit edilmiştir. Öte yandan, özellikle çevrimiçi testi cevaplayan öğrencilerin motivasyonlarının düşüklüğünün de testin zorluğuna bir etkisinin olabileceği yazarlar tarafından düşünülmüştür (Bayındır, 2021; Yüksel Şahin, 2021). AFA ve DFA birçok başarı testinde yapılmamakta iken (Akbulut ve Çepni, 2013; Demir vd., 2016; Şen ve Eryılmaz, 2011) mevcut çalışmada her iki faktör analizine de yer verilerek testin geçerliğine yönelik tüm analizler yapılmıştır. SPSS’te kullanılan DFA eşit aralıklı ölçekler için kullanılmaktadır. Eşit aralıklı ölçekler için DFA’yı LISREL ya da AMOS programı kullanarak yapmayı tercih eden çalışmalar vardır (Akyüz, 2018; Aytaç ve Öngen, 2012; Erdoğan vd., 2007; Ergül ve Yılmaz, 2020). Bu çalışmada ise ordinal ölçekler için tercih edilen tetrakorik korelasyon analizi yapılmıştır (Ekström, 2009; Hutchinson, 1993). Öte yandan alanyazında geliştirilen birçok başarı testi çalışmalarında tetrakorik korelasyon analizi yapılmadığı görülmektedir (Akbulut ve Çepni, 2013; Demir vd., 2016; Şen ve Eryılmaz, 2011). Dolayısıyla testin geçerliği hakkında güvenilir sonuçlara ulaşılamamaktadır. Oysaki bu çalışmada tetrakorik korelasyon uygulanmıştır. Dolayısıyla mevcut çalışmanın özellikle başarı testi geliştirme çalışması için acemi olan araştırmacılara doğru analiz seçimi konusunda yol gösterici olacağı düşünülmektedir. Analizler sonucunda, 23 sorudan oluşan varyansın %37’sini açıklayan tek faktörlü bir yapı elde edilmiştir. Alanyazında tek faktörden oluşan başarı testlerinin kabul edilebilir açıklanan varyans değerinin minimum %30 olması gerektiği belirtilmiştir (Büyüköztürk, 2002; Çokluk vd., 2010). Bu bağlamda, araştırma kapsamında geliştirilen testin yapı geçerliğinin sağlandığı söylenebilir.

Yapılan geçerlik analizlerinden sonra başarı testinin KR-20 güvenilirlik katsayısı 0,86 olarak hesaplanmıştır. Bu bağlamda, geliştirilen testin yüksek güvenilirliğe sahip olduğu sonucuna ulaşılmıştır. (Altay ve Kışla 2019; Avcı 2019; Kara ve Çelikler, 2015; Karlı vd., 2019; Özaşkın Arslan ve Karamustafaoğlu, 2019; Özkan ve Yadigaroğlu, 2020; Saraç 2018; Varış ve Cesur, 2012). Mevcut çalışmanın bu yönüyle alanyazın ile paralellik gösterdiği söylenebilir.

“Maddenin Yapısı ve Özellikleri” konusundaki başarı testi ortaokul yedinci sınıf düzeyinde olup, çıkmış sorular içerisinden, aynı zamanda geçerliği ve güvenilirliği sağlanmış çalışmalardan alınıp uzman görüşleriyle seçilerek hazırlanmıştır. Geçerliği ve güvenilirliği kanıtlanmış olan bu testi fen bilimleri öğretmenleri ve başarı testi geliştirecek araştırmacılar kullanabilirler.

## 5. ÖNERİLER

### 5.1. Bulgulara yönelik öneriler

- İleride yapılacak çalışmalar için ayırt ediciliği yüksek olan bu test ölçüt test olarak kabul edilip benzer kazanımlara yönelik sorular hazırlanabilir.
- Geliştirilen başarı testi yedinci sınıf seviyesindeki öğrencilere hitap etmektedir. İleride yapılacak çalışmalardaki başarı testlerinin farklı sınıf seviyelerindeki öğrenciler için de geliştirilebileceği önerilmektedir.
- Mevcut başarı testi “Maddenin Yapısı ve Özellikleri” konusunda geliştirilmiş olup yapılacak olan başka çalışmaların farklı konularda gerçekleştirilebileceği düşünülmektedir.

## 5.2. Sınırlıklara yönelik öneriler

• Geliştirilen başarı testi Kayseri ili Melikgazi ilçesi yedinci eğitim bölgesinde üç ortaokul ile sınırlandırılmış olup bir sonraki araştırmaların evren ve örnekleminin daha geniş tutulabileceği ve böylece daha doğru sonuçlar elde edileceği öngörülmektedir.

## KAYNAKLAR

- Akbulut, H. İ. ve Çepni, S. (2013). Bir üniteye yönelik başarı testi nasıl geliştirilir? İlköğretim 7. sınıf kuvvet ve hareket ünitesi. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, 2(1), 18-44. Retrieved from <https://dergipark.org.tr/tr/download/article-file/19600>
- Akyüz, H. E. (2018). Yapı geçerliliği için doğrulayıcı faktör analizi: Uygulamalı bir çalışma. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 7(2), 186-198. <https://doi.org/10.17798/bitlisfen.414490>
- Altay, G. ve Kışla, T. (2019). Programlama temelleri dersine yönelik başarı testi geliştirme çalışması. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 13(2), 680-703. <https://doi.org/10.17522/balikesirnef.523595>
- Ateş, A. (2018). *7.sınıf fen ve teknoloji dersi "maddenin tanecikli yapısı ve saf maddeler" hakkında artırılmış gerçeklik teknolojileri kullanılarak oluşturulan öğrenme materyalinin akademik başarıya etkisi*. (Yayımlanmış Yüksek Lisans Tezi). Niğde Ömer Halisdemir Üniversitesi, Eğitim Bilimleri Enstitüsü, Niğde.
- Avcı, F. (2019). Yenilenmiş bloom taksonomisine göre madde ve ısı başarı testi: geçerlik güvenirlik çalışması. *Iğdır Üniversitesi Sosyal Bilimler Dergisi*, 21, 263-292. Retrieved from <https://dergipark.org.tr/tr/pub/igdirsosbilder/issue/66828/1045181>
- Aygan, M. B. ve Zengin, F. K. (2017). Yenilenebilir enerji konusunda geçerliği ve güvenirligi sağlanmış bir başarı testi geliştirme çalışması. *International Congress Of Eurasian Social Sciences (ICOESS), Özel Sayısı*, 8(28), CXXXIX-CLV.
- Ayvacı, H. Ş. ve Durmuş, A. (2016). Bir başarı testi geliştirme çalışması: ısı ve sıcaklık başarı testi geçerlik ve güvenirlik araştırması. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi* 35(1), 87-103. doi: 10.7822/omuefd.35.1.8 Retrieved from <https://dergipark.org.tr/tr/pub/omuefd/issue/26353/277708>
- Balcı, E. ve Tekkaya, C. (2000). Ölçme ve değerlendirme tekniklerine yönelik bir ölçeğin geliştirilmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 18, 42-50.
- Bardhoshi, G. ve Erford, B. T. (2017). Processes and procedures for estimating score reliability and precision. *Measurement and Evaluation in Counseling and Development*, 50(4), 256-263. <https://doi.org/10.1080/07481756.2017.1388680>
- Bayındır, N. (2021). Çevrimiçi öğretim sürecinde motivasyon faktörü. *Gaziantep Üniversitesi Eğitim Bilimleri Dergisi*, 5(2), 291-303.
- Beuchert, A. K. ve Mendoza, J. L. (1979). A monte carlo comparison of ten item discrimination indices. *Journal of Education Measurement*, 16(2), 109-118. <https://doi.org/10.1111/j.1745-3984.1979.tb00092.x>
- Birgin, O. (2016). *Matematik öğrenmeye ilişkin inanç ölçeğinin geliştirilmesi: geçerlik ve güvenirlik çalışması*. 1. International Academic Research Congress, Antalya.
- Bloom, B.S. vd. (1956) A Taxonomy of Educational Objectives: Handbook I The Cognitive Domain. Longman, Green Co., New York.
- Bouwma-Gearhart, J., Stewart, J. ve Brown, K. (2009). Student misapplication of a gas-like model to explain particle movement in heated solids: implications for curriculum and instruction toward students' creation and revision of accurate explanatory models. *International Journal of Science Education*, 31(9), 1157-1174. <https://doi.org/10.1080/09500690902736325>
- Büyüköztürk, S. (2002). Hand book of data analysis. Ankara: Pegem A Publishing.

- Childs, K. R. (2020). Write away: writing across the curriculum and beyond. *Texas Association for Literacy Education Yearbook*, 7, 44-48.
- Çoşkun, I., ve Sarıkaya, R. (2020). Dördüncü sınıf maddenin özellikleri akademik başarı testi geliştirme: bir geçerlik ve güvenilirlik çalışması. *Turkish Studies-Educational Sciences*, 15(5), 3203-3228.
- Coştu, B., Karataş, F. Ö. ve Ayas, A. (2003). Kavram öğretiminde çalışma yapraklarının kullanılması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi* 2(4), 33-48. Retrieved from <https://dergipark.org.tr/tr/pub/pauefd/issue/11129/133090>
- Creswell, J. W. (2009). *Research Design*. Third Edition, SAGE Publications, Inc.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Çardak, S. Z. ve Selvi, K. (2018). Öğretim ilke ve yöntemleri dersi için bir başarı testi geliştirme. *Akdeniz Eğitim Araştırmaları Dergisi*, 12(26), 379-406. <https://doi.org/10.29329/mjer.2018.172.19>
- Çokluk, Ö., Şekercioğlu, G. ve Büyüköztürk, Ş. (2010). *Sosyal Bilimler İçin Çok Değişkenli SPSS ve LISREL Uygulamaları*. Pegem A Yayıncılık, Ankara.
- Demir, N., Kızılay, E. ve Bektaş, O. (2016). 7.sınıf çözümler konusunda başarı testi geliştirme: geçerlik ve güvenilirlik çalışması. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 10(1), 209-237. <https://doi.org/10.17522/nefemed.52947>
- Doğruöz, E. (2022). Çoktan seçmeli testler. İçinde M.K. Yöntem (Ed.) *Uygulama örnekleri ile etkili üniversite eğitimi* (s.167-170). İstanbul: Eğitim Yayınevi.
- Ekström, J. (2009). Contributions to the theory of measures of association for ordinal variables. *Doctoral dissertation, ACTA Universitatis Upsaliensis*, (44)2, 169-172.
- Erdoğan, Y., Bayram, S., ve Deniz, L. (2007). Web tabanlı öğretim tutum ölçeği: açıklayıcı ve doğrulayıcı faktör analizi çalışması. *Uluslararası İnsan Bilimleri Dergisi*, 4(2), 1-14.
- Eren, A. A., Önal, Ö. Ü. N. T., ve Büyük, U. (2020). Elementler ve bileşikler konusu için geçerli ve güvenilir bir başarı testi geliştirme çalışması. *Pearson Journal Of Social Sciences & Humanities*, 6(6), 152-167. <http://dx.doi.org/10.46872/pj.52>
- Ergül, B., ve Yılmaz, V. (2020). COVID-19 salgını süresince aile içi ilişkilerin doğrulayıcı faktör analizi ile incelenmesi. *IBAD Sosyal Bilimler Dergisi*, 38-51. <https://doi.org/10.21733/ibad.733909>
- Gardner, E. (1989). *Five common issues of tests*, ERIC clearinghouse on test measurement and evaluation. Washington DC., American Institutes for Research.
- George, D. ve Mallery, P. (2010). *SPSS for Window Step by Step Answers to Selected Exercises*.
- Gönen, S., Kocakaya, S. ve Kocakaya, F. (2011). Dinamik konusunda geçerliliği ve güvenilirliği sağlanmış bir başarı testi geliştirme çalışması. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 8(1),40-57. Retrieved from <https://dergipark.org.tr/tr/pub/yyuefd/issue/13707/165951>
- Gönültaş, M., Ezberci Çevik, E., ve Güneri, E. (2020). Fen bilimleri öğretmen adaylarının kök hücreye yönelik algılarının belirlenmesi. *Asya Öğretim Dergisi*, 8(2), 52-63. <https://doi.org/10.47215/aji.732735>
- Güler, T. (2020). *Artırılmış gerçeklik destekli argümantasyon yönteminin ilköğretim 7.sınıf öğrencilerinin maddenin tanecikli yapısı ve saf maddeler konusundaki akademik başarılarına etkisi*. (Yayımlanmış Yüksek Lisans Tezi). Erciyes Üniversitesi, Eğitim Bilimleri Enstitüsü, Kayseri.
- Hambleton, R. K., ve Swaminathan, H. (1985). Item response theory: principles and applications. Boston, MA: Kluwer-Nijhoff. <https://doi.org/10.1177%2F014662168500900315>
- Haladyna, T.M. (1992). The effectiveness of several multiple-choice formats. *Applied Measurement in Education*, 5(1), 73-88. [https://doi.org/10.1207/s15324818ame0501\\_6](https://doi.org/10.1207/s15324818ame0501_6)
- Haladyna, T. M. (1997). *Writing test items to evaluate higher-order thinking*. London: Allyn & Bacon.
- Harvey, B., Pallant, J. ve Harvey, D. (2004). An evaluation of the factor structure of the frost multidimensional perfectionism scale. *Educational and Psychological Measurement*, 64(6), 1007-1018. DOI: 10.1177/0013164404264842
- Heale, R. ve Twycross, A. (2015). Validity and reliability in quantitative studies. *Evid Based Nurs* 18(3), 66-67. <http://dx.doi.org/10.1136/eb-2015-102129>



- Hu, L., ve Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55. doi:10.1080/10705519909540118.
- Hutchinson, T. P. (1993). Kappa muddles together two sources of disagreement: tetrachoric correlation is preferable. *Research in Nursing & Health*, 16(4), 313-316.
- Jöreskog, K. ve Sörbom, D. (1993). *Lisrel 8: Structural Equation Modeling with the Simplis Command Language*. Scientific Software International. Inc.
- Kaplan, E., Bektaş, O., ve Karaca, M. (2022). Madde ve ısı ünitesi başarı testi geliştirme çalışması. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, (63), 78-116.
- Kara, F. ve Çelikler, D. (2019). 5.sınıf “maddenin değişimi” ünitesinde kullanılan bağlam temelli öğrenmenin öğrencilerin başarısına etkisi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 15(1), 216-245. Doi: 10.17860/mersinefd.428001
- Karaca, M., Bektaş, O. ve Saraçoğlu, S. (2016). Kimyasal tepkimeler konusunda açık uçlu ve çoktan seçmeli test geliştirme: geçerlik ve güvenilirlik çalışması. *Tarih Okulu Dergisi*, 9(25), 1117-1154. <https://doi.org/10.14225/Joh840>
- Karagöz, F. ve Korkmaz, S. D. (2015). Fen ve teknoloji dersinde web destekli öğretim yönteminin 7.sınıf öğrencilerinin akademik başarılarına ve öğrendikleri bilgilerin kalıcılığına etkisi. *International Periodical for the Languages, Literature and History of Turkish or Turkic*, 10(11), 927-948. <https://doi.org/10.7827/TurkishStudies.8686>
- Karataş, F. Ö., Köse, A. G. S., ve Coştu, A. G. B. (2003). Öğrenci yanılgılarını ve anlama düzeylerini belirlemede kullanılan iki aşamalı testler. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 13(13), 54-69.
- Karlı, G., Karamustafaoğlu, S. ve Kurt, M. (2019). Yenilenen fen bilimleri dersi öğretim programına yönelik 7. sınıf “hücre ve bölünmeler” ünitesi başarı testi: geçerlik ve güvenilirlik. *Fen Bilimleri Öğretimi Dergisi*, 7(1), 68-98.
- Kızılcapan, O. ve Bektaş, O. (2018). Yedinci sınıf maddenin yapısı ve özellikleri ünitesi başarı testi geliştirme: geçerlik ve güvenilirlik çalışması. *The Journal of International Lingual, Social and Educational Sciences* 4(2), 186-202. <https://doi.org/10.34137/jilses.431971>
- Kim, T. K. (2015). T test as a parametric statistic. *Korean J Anesthesiol*, 68(6), 540-546. <https://doi.org/10.4097/kjae.2015.68.6.540>
- Lorenzo-Seva, U. ve Ferrando, P. J. (2021). “Factor” program. <https://psico.fcep.urv.cat/utilitats/factor/Description.html>. 18.05.2022 tarihinde erişilmiştir.
- Marsh, H.W., Muthén, B., Asparouhov, T., Lüdtke, T., Robitzsch, A., Morin, A.J.S., ve Trautwein, U. (2009) Exploratory Structural Equation Modeling, Integrating CFA and EFA: Application to Students' Evaluations of University Teaching, *Structural Equation Modeling. A Multidisciplinary Journal*, 16(3), 439-476, DOI:10.1080/10705510903008220 <https://doi.org/10.1080/10705510903008220>
- Marzo, P. F. ve Andrés, A. M. (2014). Multiple-choice tests: inferences based on estimators of maximum likelihood. *Open Journal of Statistics*, 4, 466-483. <http://dx.doi.org/10.4236/ojs.2014.46045>
- McMillan, J. H. (1992). Educational research: fundamentals for the consumers. New York: Harper Collins.
- Milli Eğitim Bakanlığı [MEB], (2018). <http://mufredat.meb.gov.tr/Dosyalar/201812312311937-FEN%20B%C4%B0L%C4%B0MLER%C4%B0%20C3%96%C4%9ERET%C4%B0M%20PROGRA MI2018.pdf> sayfasından 18.06.2021 tarihinde 01.00 saatinde erişildi.
- Minaslı, E. (2009). *Fen ve teknoloji dersi maddenin yapısı ve özellikleri ünitesinin öğretilmesinde simülasyon ve model kullanılmasının başarıya, kavram öğrenmeye ve hatırlamaya etkisi*. (Yayımlanmamış Yüksek Lisans Tezi). Marmara Üniversitesi, Eğitim Bilimleri Enstitüsü, İstanbul.
- Nacaroğlu, O., Bektaş, O. ve Kızılcapan, O. (2020). Madde döngüleri ve çevre sorunları konusunda başarı testi geliştirme: geçerlik ve güvenilirlik çalışması. *Kastamonu Education Journal*, 28(1), 36-51. doi:10.24106/kefdergi.3374

- Nacaroğlu, O. ve Bektaş, O. (2019). Fen bilimleri dersindeki madde ve değişim ünitesine yönelik geçerli ve güvenilir başarı testi geliştirme: BİLSEM örneği. *Maarif Mektepleri Uluslararası Eğitim Bilimleri Dergisi*, 3(2), 67-88. <https://doi.org/10.46762/mamulebd.633540>
- Narlı, S., ve Baser, N. E. (2008). " Küme, bağıntı, fonksiyon" konularında bir başarı testi geliştirme ve bu test ile üniversite matematik bölümü 1. sınıf öğrencilerinin bu konulardaki hazırbulunuşluklarını betimleme üzerine nicel bir araştırma. *Dokuz Eylül Üniversitesi Buca Eğitim Fakültesi Dergisi*, (24).
- Noh, T. ve Scharmann, L. C. (1997). Instructional influence of a molecular-level pictorial presentation of matter on students' conceptions and problem-solving ability. *Journal of Research in Science Teaching*, 34(10), 199-217. [https://psycnet.apa.org/doi/10.1002/\(SICI\)1098-2736\(199702\)34:2%3C199::AID-TEA6%3E3.0.CO;2-O](https://psycnet.apa.org/doi/10.1002/(SICI)1098-2736(199702)34:2%3C199::AID-TEA6%3E3.0.CO;2-O)
- Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 44(4), 443-460. <https://doi.org/10.1007/BF02296207>
- Özaşkın Arslan, A. G. ve Karamustafaoğlu, S. (2019). 2018 Fen bilimleri öğretim programı kapsamındaki 7. sınıf güneş sistemi ve ötesi ünitesine yönelik bir başarı testi geliştirme. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi*, 38(2), 172-205. DOI: 10.7822/omuefd.528571.
- Özdamar K. Örneklem Yöntemleri. SPSS ile Biyoistatistik. 4ncü Baskı, 2001, 261-65.
- Özelma, E. İ. (2019). "Maddenin tanecikli yapısı" konusunun öğretiminde argümantasyon tabanlı öğretim yönteminin fen başarısına ve tartışma istekliliğine etkisi. (Yayımlanmamış Yüksek Lisans Tezi). Sivas Cumhuriyet Üniversitesi, Eğitim Bilimleri Enstitüsü, Sivas.
- Özkan, S. ve Yadigaroglu, M. (2020). Başarı testi geliştirme: asit-baz başarı testi geçerlik ve güvenilirlik araştırması. *Turkish Studies- Education*, 15(2), 1141-1163. <http://dx.doi.org/10.29228/TurkishStudies.41734>
- Pallant, J. (2007). *SPSS survival manual a step-by-step guide to data analysis using SPSS for Windows*. Third Edition, Open University Press.
- Pallant, J. (2016). *SPSS Survival Manual A Step By Step Guide to Data Analysis Using SPSS Program* (6th ed.). London, UK McGraw-Hill Education.
- Romagnano, L. (2001). Implementing the assessment standards: The myth of objectivity in mathematics assessment. *The Mathematics Teacher*, 94(1), 31-37.
- Saraç, H. (2018). Fen bilimleri dersi 'maddenin değişimi' ünitesi ile ilgili başarı testi geliştirme: geçerlik ve güvenilirlik çalışması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18(1), 416-445. <https://doi.org/10.17240/aibuefd.2018.-388815>
- Saidi, S. S. ve Siew, N. M. (2019). Investigating the validity and reliability of survey attitude towards statistics instrument among rural secondary school students. *International Journal of Educational Methodology* 5(4), 651-661. <https://doi.org/10.12973/ijem.5.4.651>
- Seçer, İ. (2015). *SPSS ve LISREL ile Pratik Veri Analizi. Analiz ve Raporlaştırma*. Anı Yayıncılık.
- Şen, H. C. ve Eryılmaz, A. (2011). Bir başarı testi geliştirme çalışması: basit elektrik devreleri başarı testi geçerlik ve güvenilirlik araştırması. *Yüzüncü Yıl Üniversitesi, Eğitim Fakültesi Dergisi*, 8(1), 1-39. Retrieved from <https://dergipark.org.tr/pub/yyuefd/issue/13707/165950>
- Schermelleh Engel, K., Moosbrugger, H. ve Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23-74.
- Sözbilir, M. (2002). Öğrencilerin maddenin tanecikli yapısı hakkındaki bilgileri ve yanlış anladıkları noktalar: seçilmiş araştırmaların bir özeti. *Kastamonu Eğitim Dergisi* 10(2), 345-350. <https://doi.org/10.17860/mersinefd.306003>
- Şimşek, Ü., Doğan, A. ve Kılıç, M. A. (2016). Jigsaw yönteminin maddenin tanecikli yapısı ünitesinin öğretiminde öğrenci başarısına etkisi. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, 18(2), 1357-1379. DOI: 10.17556/jef.31750
- Tabachnick, B.G. ve Fidell L.S. (2013). *Using Multivariate Statistics* (6th ed.). Boston, MA: Pearson.
- Temizkan, M., ve Sallabaş, M. E. (2011). Okuduğunu anlama becerisinin değerlendirilmesinde çoktan seçmeli testlerle açık uçlu yazılı yoklamaların karşılaştırılması. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, (30), 207-220.

- Traub, R. E. (1997). Classical test theory in a historical perspective. *Educational Measurement: Issues and Practice*, 16, 8-14.
- Tünkler, V. ve Güven, C. (2019). Mikroöğretim uygulamasının öğretmen adaylarının tamamlayıcı ölçme-değerlendirme tekniklerine yönelik okuryazarlık düzeylerine etkisi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 34(2), 541-564. <https://doi.org/10.16986/HUJE.2018043466>
- Uebersax, J. S. (2015). The tetrachoric and polychoric correlation coefficients. Statistical Methods for Rater Agreement. <http://john-uebersax.com/stat/tetra.htm> adresinden 18.05.2022 tarihinde erişilmiştir.
- Üçüncü, G. ve Sakız, G. (2020). Başarı testi geliştirme süreci: ilkokul dördüncü sınıf maddeyi tanıyalım ünitesi örneği. *Kastamonu Education Journal*, 28(1), 82-94. doi:10.24106/kefdergi.3440
- Variş, Y. A. ve Cesur, D. (2012). Ortaöğretim müzik dersine yönelik başarı testinin geliştirilmesi. *International Periodical For The Languages, Literature and History of Turkish or Turkic* 7(4), 3189-3198. <http://dx.doi.org/10.7827/TurkishStudies.4144>
- Vurgun, F. ve Bektaş, O. (2019). Altıncı sınıf öğrencilerinin fen'e yönelik girişimciliklerinin belirlenmesi. *Fen, Matematik, Girişimcilik ve Teknoloji Eğitimi Dergisi*, 2(2), 60-78. Retrieved from <https://dergipark.org.tr/tr/pub/fmgted/issue/50283/583832>
- Yıldız, S., Keçeci, G. ve Kırbağ Zengin, F. (2019). Dengeli beslenme akademik başarı testi: geçerlik ve güvenilirlik araştırması. *YYÜ Eğitim Fakültesi Dergisi*, 16(1) 848-868. <http://efdergi.yyu.edu.tr> <http://dx.doi.org/10.23891/efdyyu.2019.144>
- Yüksel-Şahin, F. (2021). Psikolojik danışmanların covid-19'a, yüz yüze ve çevrimiçi psikolojik danışma yapmaya, yüz yüze ve çevrimiçi eğitim almaya ilişkin görüşlerinin incelenmesi. *IBAD Sosyal Bilimler Dergisi*, (9), 496-522.

**Ek.1. MADDEİNİN YAPISI VE ÖZELLİKLERİ BAŞARI TESTİ**

1) Moleküller ile ilgili aşağıdakilerden hangisi doğrudur?

- A) En az iki çeşit atom bulunmalıdır.  
 B) Bir anyon bir kation bulunmalıdır.  
 C) Aynı ya da farklı atomların birleşmesi ile oluşurlar.  
 D) Bir hidrojen molekülü bir hidrojen atomundan daha küçüktür.



Zeynep

Atom modeli önceleri üzümlü keke benzetilirken, zamanla elektron bulutu olduğu yönünde görüşler gelişmiştir.

2) Yukarıda Zeynep'in atom modelleri ile ilgili konuşması yer almaktadır. Zeynep'in konuşmasına bakarak aşağıdaki hangi çıkarımı yapmak **en doğru** olacaktır?

- A) Bilimsel bilgi zamanla değişebilir.  
 B) Atom modeli üzümlü keke benzetilmekteydi.  
 C) Atom modeli ile ilgili birçok görüş ortaya atılmıştır.  
 D) Atom modeli günümüzde elektron bulutuna benzetilmektedir.

3) Aşağıdaki bardakların hangisinde tuzun çözünme hızı **en fazladır**?

A) 10g tuz  
50g su  
25°C



B) 10g tuz  
50g su  
20°C



C) 10g tuz  
50g su  
45°C



D) 10g tuz  
50g su  
30°C



4) Atomun yapısında bulunduğu bilinen taneciklere ait bazı bilgiler aşağıda numaralar yoluyla verilmiştir. Bu numaraların ve taneciklerin eşleşmesi aşağıdaki seçeneklerin hangisinde doğru olarak verilmiştir?

- 1- Negatif yüklüdür, katmanlarda bulunur  
 2- Yüksüzdür, atomun merkezinde bulunur  
 3- Pozitif yüklüdür, atomun merkezinde bulunur

- |             | <u>1</u> | <u>2</u> | <u>3</u> |
|-------------|----------|----------|----------|
| A) Nötron   | Elektron | Proton   |          |
| B) Elektron | Nötron   | Proton   |          |
| C) Elektron | Proton   | Nötron   |          |
| D) Proton   | Nötron   | Elektron |          |

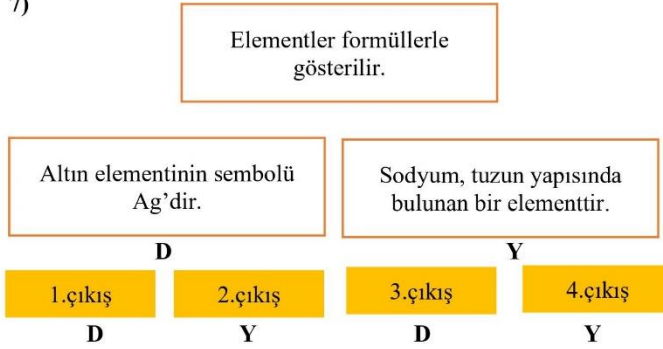
5) Aşağıdakilerden hangisi element sembolü değildir?

- A) Cl B) S C) Au D) OH<sup>-</sup>

6) Tuzlu su ile ilgili aşağıdakilerden hangisi **söylenemez**?

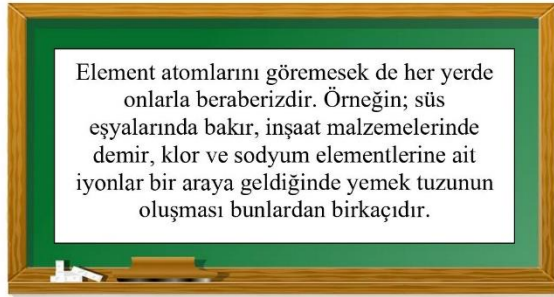
- A) Tuz çözünendir. B) Tuz çözücü maddedir.  
 C) Tuz iyonlarına ayrılmıştır. D) Katı-sıvı çözeltiye örnektir.

7)



Yukarıdaki ifadelerden doğru olanlar için D, yanlış olanlar için Y yolu izlenirse kaçınıcı çıkışa ulaşılır?

- A) 1.çıkış  
B) 2.çıkış  
C) 3.çıkış  
D) 4.çıkış



8) Tahtaya yazılı metinde geçen elementler ve bu elementlere ait sembollerin tamamı aşağıdakilerden hangisinde doğru verilmiştir?

- A) 

Bakır	Ba
Demir	D
Tuz	T
Sodyum	Na

B) 

Bakır	Ba
Demir	D
Klor	Cl
Tuz	T
- C) 

Bakır	Cu
Demir	Fe
Klor	Cl
Sodyum	Na

D) 

Bakır	Cu
Demir	D
Tuz	Kl
Sodyum	S

9) Aşağıda günlük hayattaki olaylardan bazı örnekler verilmiştir.

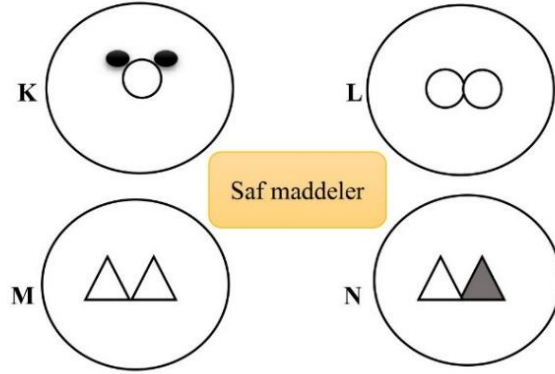
I. Odada bir kişi parfüm sıktığında parfüm kokusunun odaya yayılması.

II. Süte kakao ve şeker atılıp karıştırıldığında tadının her yerinde aynı olması.

III. Mutfakta pişen pastanın kokusunun eve yayılması. Buna göre, yukarıdaki olaylardan hangileri maddelerin tanecikli yapıda olduğunu kanıtlar?

- A) I ve II B) I ve III  
C) II ve III D) I, II ve III

10) Merve Öğretmen; K, L, M ve N saf maddelerinin atom modellerini aşağıdaki gibi göstererek Bade' den element modellerini, Deniz' den ise bileşik modellerini seçmelerini istemiştir.



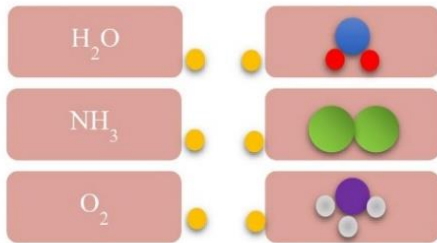
Buna göre Bade ve Deniz, hangi modelleri seçmelidir?

- | Bade      | Deniz  |
|-----------|--------|
| A) K ve L | M ve N |
| B) K ve M | L ve N |
| C) L ve M | K ve N |
| D) M ve N | K ve L |

Atom	İyon
K	$K^{+2}$
L	$L^{-1}$
M	$M^{+1}$
N	$N^{-2}$

11) Çizelgeye göre hangi atomlar **elektron vermiştir**?

- A) K ve M  
B) L ve N  
C) K ve L  
D) M ve N



12) Yukarıdaki formüllerle modellerin doğru eşleştirilmesi aşağıdaki seçeneklerin hangisinde doğru verilmiştir?

- A) B) C) D)



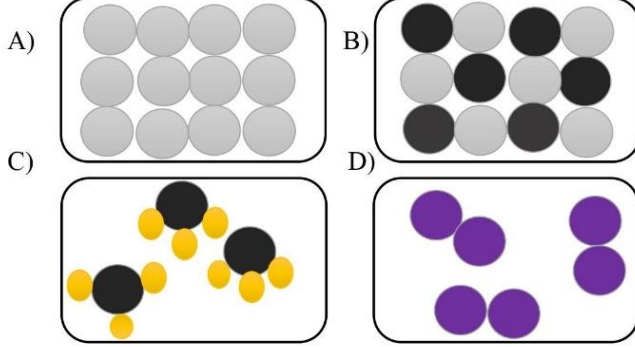
13) Ne bileşim

Ne de karışım

Atomik yapıdayım

Özdeşir atomlarım

Dörtlükte kendini tanıtan maddenin tanecik modeli aşağıdakilerden hangisidir?



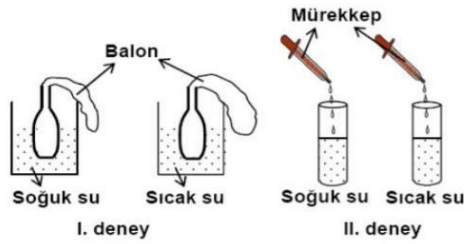
<ul style="list-style-type: none"><li>• Neon</li><li>• Silisyum</li><li>• Berilyum</li></ul> <p><u>İsim</u> Nesibe</p>	<ul style="list-style-type: none"><li>• Helyum</li><li>• Lityum</li><li>• Azot</li></ul> <p><u>İsim</u> ?</p>
----------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------

1.kart

2.kart

14) Karttaki isim Nesibe olarak belirlendiğine olduğuna göre, öğrencinin 2.kartta türettiği isim aşağıdakilerden hangisidir?

- A) Halime  
B) Nihal  
C) Halit  
D) Helin



15) Yukarıdaki deneyleri yapan bir öğrenci;

I. deneyde balonun sıcak suda şiştiğini,

II. deneyde mürekkebin sıcak suda daha çabuk dağıldığını gözlemliyor.

Öğrenci bu deneylerin sonucunda aşağıdakilerden hangisini söyleyebilir?

- A) Katı, sıvı ve gaz taneciklerin aynı şekilde hareket ettiğini  
B) Maddeleri oluşturan taneciklerin hareketli olup olmadığını  
C) Sıcaklığın, maddelerin taneciklerinin hareketini nasıl etkilediğini  
D) Madde, sıvıdan gaz hâline geçtiğinde taneciklerin hareketinin nasıl değiştiğini

16) Aşağıda verilen ifadeler hangi tür çözeltilere örnektir?

- I) Tuzun suda çözünmesi  
 II) Balıkların suda yaşamasını sağlayan göldeki oksijen  
 III) Alkol ve suyun birleşerek kolonyayı oluşturması.

	Sıvı-Katı	Sıvı-Sıvı	Sıvı-Gaz
A) I	III	II	
B) II	I	III	
C) I	II	III	
D) III	II	I	

17)

Dil	Elementin İsmi	Sembolü
Türkçe	Berilyum	Be
İngilizce	Beryllium	Be
İspanyolca	Berilio	Be

Yukarıdaki tablo incelendiğinde yapılacak **en doğru** yorum aşağıdakilerden hangisidir?

- A) Elementler tek tür atoma sahiptir.  
 B) Elementlerin kendine özgü sembolü vardır.  
 C) Elementler her ülkede farklı isimlerle ifade edilse de sembolleri aynıdır.  
 D) Elementlerin sembollerini kullanmak bilimsel çalışmalarda iletişimi kolaylaştırır.

18) Aşağıda özellikleri verilen X elementi ve sembolü hangi seçenekte doğru olarak verilmiştir?

- Canlıların yapısında bulunur.
- Kurşun kalemlerin ucunda bulunur.
- Fosil yakıtların (Kömür, Petrol...) yapısında bulunur.

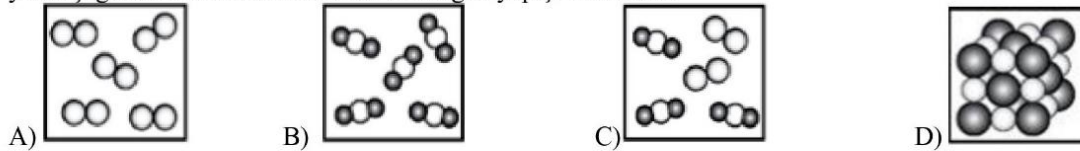
Element	Sembolü
A) Azot	At
B) Karbon	K
C) Azot	N
D) Karbon	C

**Özellikleri:**

- Saf bir maddedir.
- Molekül yapılıdır.
- İki farklı element atomundan oluşur.

?

19) Şekildeki gibi hazırlanan karta bir maddenin özellikleri yazılmıştır. Buna göre «?» ile gösterilen yere aşağıdaki tanecik modellerinden hangisi yapılandırılmalıdır?



20) Aşağıda verilen maddelerden hangisinin sıkıştırılma özelliği **en fazladır**?

- A) Oksijen  
 B) Su  
 C) Kola  
 D) Limonata

21) Ali öğretmen derse bir bardak su ve biraz şeker getirerek sınıfta bunları karıştırmıştır. Daha sonra öğrencilerine dönerek suyun içerisinde çözünen şekeri ayırıp ayıramayacaklarını sormuştur. Ali öğretmenin sorusuna bazı öğrencilerin verdikleri cevaplar şunlardır;

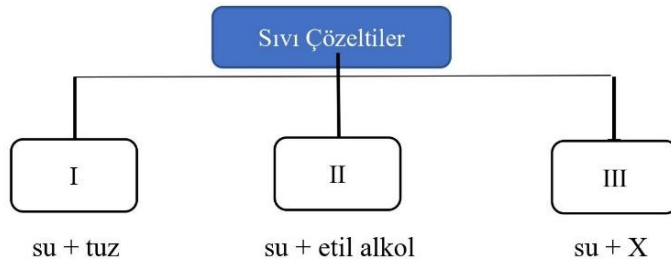
**Kaan:** Yeni bir bileşik oluşmuştur, bu yüzden şeker ve su artık fiziksel yöntemlerle ayıramaz

**Özgür:** Bir çözelti oluşmuştur, çözeltideki şeker sudan buharlaştırma yöntemi ile ayrılabilir.

**Can:** Bir heterojen karışım oluşmuştur, şeker sudan yoğunluk farkından yararlanılarak ayrılabilir.

Buna göre hangi öğrencilerin yorumu doğrudur?

- A) Yalnız Kaan
- B) Yalnız Özgür
- C) Kaan ve Can
- D) Özgür ve Can



22) Buna göre, III. Sınıfa ait örnekteki X maddesi ile ilgili aşağıdakilerden hangisi doğrudur?

- A) Suda çözünen bir katıdır.
- B) Suda çözünen bir gazdır.
- C) Suda çözünmeyen bir gazdır.
- D) Suda çözünmeyen bir sıvıdır.

23) Çağdaş öğretmen derste şeker çözme yarışması düzenlemiştir. Yarışmada eşit miktardaki şekerin bir bardak su içerisinde **en hızlı çözünmesini** sağlayan yarışmacı birinci olacaktır. Yarışmacıların hazırladıkları çözeltiler şu şekildedir.

Yarışmacı	Sıcaklık	Tanecik Boyutu	Karıştırma
Pelin	20 °C	Küp şeker	Var
Yağmur	30 °C	Toz şeker	Yok
Egemen	20 °C	Küp şeker	Yok
Yağız	30 °	Toz şeker	Var

Buna göre bu yarışmayı yarışmacılardan hangisi kazanır?

- A) Pelin
- B) Yağmur
- C) Egemen
- D) Yağız

<i>Başarı Testi Cevap Anahtarı</i>	
<b>Soru Numarası</b>	<b>Cevap</b>
<b>1</b>	C
<b>2</b>	A
<b>3</b>	C
<b>4</b>	B
<b>5</b>	D
<b>6</b>	B
<b>7</b>	C
<b>8</b>	C
<b>9</b>	D
<b>10</b>	C
<b>11</b>	B
<b>12</b>	A
<b>13</b>	A
<b>14</b>	D
<b>15</b>	C
<b>16</b>	A
<b>17</b>	C
<b>18</b>	D
<b>19</b>	B
<b>20</b>	A
<b>21</b>	B
<b>22</b>	B
<b>23</b>	D

### **EXTENDED ABSTRACT**

Measurement and evaluation are important parts of the education system. It is a method used to confirm to what extent the student has learned a subject. Multiple choice questions are one of the methods frequently used in measurement and evaluation. It is known that these exams, which are generally prepared by teachers, are used to measure students' knowledge levels (Balcı & Tekkaya, 2000).

With classical test theory, by giving equal points distribution to each question, students are provided to give equal importance to each question (Traub, 1997). While preparing for the achievement test, it was aimed that the students give equal importance to each question, so that the classical test theory was taken as a guide while preparing for the current achievement test.

If the teachers' pay enough attention to the question preparation part while preparing an achievement test, they will have made an accurate assessment and evaluation (Kamshia & Childs, 2020). The questions for this study were prepared according to Bloom's taxonomy of cognitive steps. Questions were formed at the steps of Bloom's taxonomy of cognitive, except for the synthesis step. Because in the synthesis step, students are asked to come up with a product within the framework of a certain subject or achievement. This is not possible for multiple-choice tests (Birgin, 2016).

Achievement test questions were obtained from articles including valid and reliable tests on the seventh-grade secondary school " *Structure and Properties of the Substance*" (Ateş, 2018; Coştu et al., 2003; Güler, 2020; Karagöz & Korkmaz 2015; Kızılkapan & Bektaş, 2018; Minaslı, 2009; Özelma, 2019) and the questions in the MEB high school entrance exams (2010 SBS, 2011 SBS, 2013 PYBS, 2014 PYBS, 2015 PYBS). Content validity was ensured by taking the opinions of five experts on the prepared questions.

Science course includes subjects that are difficult to learn (Noh & Scharmann, 1997). The structure and properties of the substance are one of the science subjects that students have difficulty in learning

(Şimşek, Doğan, & Kılıç, 2016). Understanding the structure of the substance is important not only for science but also for many different fields (Bouwma-Gearhart, Stewart, & Brown, 2009). Despite this, students find the subject complex in the secondary school years and still have difficulties in the following years in this subject, which starts in the secondary school curriculum and continues in high school (Sözbilir, 2002). When the literature is examined, it is seen that the number of achievement tests developed on "Structure and Properties of the Substance" at the seventh-grade level is limited (Nacaroğlu et al., 2019; Saraç, 2018; Third & Sakız, 2020). In this context, the present study aims to develop a valid and reliable achievement test for the seventh-grade "Structure and Properties of the Substance". For this purpose, we have sought answers to the following questions:

1. Is the achievement test developed to determine the achievement levels of seventh-grade students in the subject of "*Structure and Properties of the Substance*" valid?

2. Is the achievement test developed to determine the achievement levels of seventh-grade students on the "*Structure and Properties of the Substance*" reliable?

To determine the learning levels of secondary school seventh-grade students on the subject of "*Structure and Properties of the Substance*" and to develop an achievement test with validity and reliability, a survey design, which is one of the quantitative research designs, was used. The universe of the present study consists of seventh-grade students in the Melikgazi district of Kayseri province. The number of questions in the achievement test applied to a total of 268 students was 23, and the obtained data were analyzed using SPSS 25 and the «*factor 11.04.02*» program (Lorenzo-Seva & Ferrando, 2021). As a result of Kaiser-Meyer-Olkin (KMO), it was observed that it met the sufficient sample assumption (Pallant 2007, pp. 179-199).

In the construct validity part of the current study; analyses based on EFA, and CFA were run. As a result of the analysis, the KR-20 reliability coefficient of the scores obtained from the test was calculated as 0.86. A KR-20 value of 0.70 and above indicates that that test is reliable (Heale & Twycross, 2015). Item discrimination indices and item difficulty indexes were calculated for each question in the achievement test, and it was observed that the questions of this achievement test had the desired level of discrimination. It is accepted that the difficulty of an achievement test is around 0.50 (Haladyna, 1997). The difficulty index of the developed test was 0.23, and it was seen that the difficulty of the test was high in general. With EFA, KMO and Barlett's statistics were calculated and factor loadings of each question were found. The 1<sup>st</sup>, 9<sup>th</sup>, 18<sup>th</sup>, 19<sup>th</sup>, 22<sup>nd</sup>, 26<sup>th</sup>, 28<sup>th</sup>, and 30<sup>th</sup> questions which had factor loadings lower than .30 were removed from the current achievement test, and a single-factor structure was obtained that explained 37% of the variance at the 23-item test. In the literature, it has been stated that the acceptable explained variance value of scales consisting of a single factor should be at least 30% (Buyukozturk, 2002; Cokluk et al., 2010). In this context, it can be said that the construct validity of the test developed within the scope of the research was ensured.

The developed achievement test is limited to three secondary schools in the region of Melikgazi district of Kayseri province, and it is predicted that the population and sample of the next studies can be kept wider and thus more accurate results will be obtained. It is thought that achievement test studies to be carried out in the future will contribute to the literature by multiple choices of more science-related subjects for each grade level, and science teachers will accurately measure students' achievements.

The present study aims to develop a valid and reliable achievement test to confirm that the subject of "*Structure and Properties of the Substance*" in the secondary school science course is important and

therefore whether the students have learned the subject sufficiently or not, and this goal has been achieved. The 23-question multiple-choice achievement test prepared is a test that can be used by researchers who want to work on the "*Structure and Properties of the Substance*", as well as by science teachers to accurately measure and evaluate students' achievements.