

Saraç, H. (2018). Fen Bilimleri Dersi 'Maddenin Değişimi' Ünitesi İle İlgili Başarı Testi Geliştirme: Geçerlik ve Güvenirlik Çalışması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18 (1), 416-445.

Geliş Tarihi: 17/07/2017

Kabul Tarihi: 02/02/2018

## FEN BİLİMLERİ DERSİ 'MADDENİN DEĞİŞİMİ' ÜNİTESİ İLE İLGİLİ BAŞARI TESTİ GELİŞTİRME: GEÇERLİK VE GÜVENİRLİK ÇALIŞMASI \*

Hakan SARAÇ\*\*

### ÖZET

Araştırmanın amacı, beşinci sınıf Fen Bilimleri dersi 'Maddenin Değişimi' ünitesine yönelik bir başarı testi geliştirmektir. Bu amaçla ilk olarak kazanımlar belirlenmiş ve bu kazanımlara uygun 55 sorudan oluşan madde havuzu hazırlanmıştır. Fen eğitimi alanında uzman görüşlerine başvurularak soru sayısı 30'a düşürülmüştür. Hazırlanan 30 soruluk test 2014-2015 eğitim öğretim yılında 230 tane beşinci sınıf öğrencisine uygulanmıştır. Başarı testinin geçerliği ve güvenirligi incelenmiştir. Başarı testinin kapsam geçerliği için belirtke tablosu hazırlanmıştır. Öğrencilerin testte yer alan sorulara verdikleri cevaplar doğrultusunda madde analizi yapılarak her bir maddenin ayırt edicilik ve güçlük indeksleri ayrı ayrı hesaplanmıştır. Soruların madde ayırt edicilik ve güçlük indeksleri arasında bağımsız gruplar t- testi sonuçlarına da bakılmıştır. Madde analizi sonucunda 2 madde testten çıkarılmış ve 28 sorudan oluşan 'Maddenin Değişimi' ünitesi başarı testi oluşturulmuştur. Testin ortalama güçlüğü 0.501, ortalama ayırt ediciliği 0.488 olarak hesaplanmıştır. Başarı testinin güvenirlilik analizinde, Kuder Richardson-20 tekniği kullanılmıştır. Hesaplanan güvenirlilik katsayısı (r) 0.714 olarak tespit edilmiştir. Araştırmadan elde edilen bulgular doğrultusunda önerilerde bulunulmuştur.

**Anahtar Kelimeler:** Başarı testi, fen bilimleri, maddenin değişimi, geçerlik, güvenirlilik

## DEVELOPMENT OF ACHIEVEMENT TEST ABOUT SCIENCE LESSON 'CHANGE OF MATTER' UNIT: VALIDITY AND RELIABILITY STUDY

### ABSTRACT

The aim of the research is to develop an achievement test for the 'Change of Matter' unit of the fifth-grade science course. For this purpose, firstly the gains were determined, and an item pool consisting of 55 questions suitable for these gains was prepared. The number of questions has been reduced to 30 by applying for expert opinions in science education. The 30-question test was applied to 230 fifth-year students in the 2014-2015 academic year. The validity and reliability of the achievement test have been examined. A statement table is prepared for the validity of the achievement test. By analyzing the items according to the answers given by the students to the questions in the test, the indivisibility and difficulty indices of each item are calculated separately. Independent groups t-test results were also examined among item indifference and difficulty indices of the questions. As a result of the item analysis, 2 items were removed from the test and the 'Change of Matter' unit of 28 questions was tested for achievement. The mean strength of the test was 0.501 and the mean discrimination was 0.488. In the reliability analysis of the achievement test, Kuder Richardson-20 technique was used. The calculated reliability coefficient (r) was determined as 0.714. Suggestions were made in the direction of the findings obtained from the research.

**Key Words:** Achievement test, science, change of matter, validity, reliability

\* Bu makale yazarın Doktora çalışmasından yararlanılarak hazırlanmıştır.

\*\* Fizik Öğretmeni, MEB, hknsrcmv@gmail.com

## 1.GİRİŞ

İnsan davranışları üzerinde değişiklikler ortaya çıkarma süreci olarak ifade edilen eğitimin, süreç içerisinde başarılı olup olmadığı ya da davranışlar üzerinde ne düzeyde değişimler ortaya çıkardığının belirlenmesi önemli bir gereksinimdir (Karaca, 2008; Erdoğan ve Kurt, 2012). Eğitim sürecindeki bu gereksinim ölçme ve değerlendirme ile karşılanmaktadır. Bu nedenle de ölçme ve değerlendirme eğitim sürecinin vazgeçilmez bir ögesi olarak karşımıza çıkmaktadır (Yeşilyurt, 2012).

Ölçme ve değerlendirme birbiriyle ilgisi olan farklı kavramlardır (Ayaydın, 2010). Ölçme, herhangi bir niteliği gözlemek ve gözlem sonucunu sayı veya sıfatlarla ifade etmektir (Turgut ve Baykul, 2012). Değerlendirme ise ölçmede elde edilen sonuçların bir ölçütle karşılaştırılması sonucunda bir karara ulaşılması olarak ifade edilmektedir (Bahar, Nartgün, Durmuş ve Bıçak, 2012).

Fen eğitiminde belirlenen hedef ve davranışların gerçekleşme düzeylerini saptamak ve öğrenci başarılarını belirlemek amacıyla ilgili tüm öğretim kazanımlarını hedefleyen geçerliği ve güvenilirliği yüksek ölçme araçlarına ihtiyaç vardır (Gönen, Kocakaya ve Kocakaya, 2011). Öğrenci başarısını ölçmek ve değerlendirmek amacı ile anketler, sözlü sınavlar, doğru yanlış testleri, çoktan seçmeli testler, eşleştirmeli testler, boşluk doldurmalı testler, ölçekler, kısa yanıtı testler, yazılı yoklamalar, açık uçlu sorular, iki aşamalı testler eğitimin tüm alanlarında kullanılmaktadır (Şimşek, 2007; Ogan Bekiroğlu, 2004). Öğretilen bilgilerin hepsini yoklamaya olanak sağlayan, hem basit hem de karmaşık kavramları ölçme imkânı tanıyan çoktan seçmeli testler öğrenci başarısını ölçmek için günümüzde sıklıkla kullanılan ölçme araçlarından biridir. Ogan Bekiroğlu (2004), öğrencilerin belirli konu ya da kavram hakkındaki anlama düzeylerini ortaya çıkarmada mülakatlardan sonra en yaygın kullanılan ölçme aracının çoktan seçmeli testler olduğunu ifade etmektedir. Özçelik (1998) ise, çoktan seçmeli testleri bugüne kadar bulunabilmiş en üstün ölçme aracı olarak tanımlamaktadır. Çoktan seçmeli soruları içeren testler, öğrencilerin eleştirel düşünme becerilerini ve yaratıcılıklarını belirlemede sınırlı olsa da özellikle öğrencilerin kavram yanlışlarını ve başarılarını belirlemek için tercih edilen ölçme araçlarıdır (Küçükahmet, 2002).

Öğrencilerin eksik ve yanlış öğrenmelerinin ortaya çıkarılmasında kullanılan çoktan seçmeli testler, daha kullanışlı olmakla birlikte sınıf ortamında da kolayca uygulanabilmektedir (Kan, 2014). Uygulanması ve puanlaması diğer ölçme araçlarına göre daha kolay olan bu testler iyi hazırlanmış çeldiriciler ile öğrencilerin eksik ve yanlış öğrenmelerinin belirlenmesini ve sonuçlarının genellenebilmesini sağlamaktadır (Demirci ve Efe, 2007; Kan, 2014). Ancak çoktan seçmeli testlerde sınırlı sayıda seçenek olduğundan öğrenciler fikirlerini belirli kalıplar dışında ifade etmede yetersiz kalmaktadırlar (Baykul ve Turgut, 2014; Karataş, Köse ve Coştu, 2003). Ayrıca öğrencilerin sorunun doğru cevabını bilmeden tesadüfi olarak doğru seçeneği işaretleyebilmesi çoktan seçmeli testlerin dezavantajlarından biridir (Baykul ve Turgut, 2014; Mintzes, Wandersee ve Novak, 2000).

Fen eğitim araştırmalarında çoktan seçmeli soruları içeren başarı testlerinin geliştirilmesine yönelik birçok çalışma bulunmaktadır. Ulusal alan yazında alanda fen eğitimi araştırmaları kapsamında, Vücutumuzun Bilmecesini Çözelim (Şener ve Taş, 2017), Çözümler (Demir, Kızılay ve Bektaş, 2016), Isı ve Sıcaklık (Ayvacı ve Durmuş, 2016), İş ve Enerji (Açıkgöz ve Karşlı, 2015), Maddenin Tanecikli Yapısı (Kenan ve

Özmen, 2014), Modern Fizik Konuları (Demir ve Akarsu, 2014), Kuvvet ve Hareket (Akbulut ve Çepni, 2013), Dinamik (Gönen, Kocakaya ve Kocakaya, 2011), Basit Elektrik Devreleri (Şen ve Eryılmaz, 2011), Genetik (Çakır ve Aldemir, 2011), Çözeltiler ve Fiziksel Özellikler (Tosun ve Taşkesenligil, 2011) konu başlıkları hakkında geliştirilen başarı testleri mevcuttur. Bu çalışmalardan elde edilen geçerlik ve güvenilirlik sonuçları genel olarak Tablo 1’de gösterilmiştir.

**Tablo 1.**

*Fen Eğitim Araştırmaları Alanında Yapılan Başarı Testi Çalışmaları*

Araştırmacılar	Araştırma İçeriği	Örnek-lem Sayısı (N)	Madde Sayısı (n)	Ortalama Madde Güçlüğü İndeksi (p <sub>jx</sub> )	Ortalama Madde Ayırt edicilik İndeksi (r <sub>jx</sub> )	Güvenirlik Katsayısı (r/a)
Şener ve Taş (2017)	Vücudumuzun Bilmecesini Çözelim	178	46	.520	.440	.620
Demir, Kızılay ve Bektaş (2016)	Çözeltiler	100	20	.584	.485	.738
Ayvacı ve Durmuş (2016)	Isı ve Sıcaklık	47	19	.490	.340	.760
Açıkgöz ve Karşı (2015)	İş ve Enerji	70	18	.670	.440	.689
Kenan ve Özmen (2014)	Maddenin Tanecikli Yapısı	80	20	-	-	.887
Demir ve Akarsu (2014)	Modern Fizik Konuları	150	30	.400	.330	.603
Akbulut ve Çepni (2013)	Kuvvet ve Hareket	74	33	.560	.420	.970
Gönen, Kocakaya ve Kocakaya (2011)	Dinamik	131	25	.360	.430	.807
Şen ve Eryılmaz (2011)	Basit Elektrik Devreleri	307	30	.554	.447	.896
Çakır ve Aldemir (2011)	Genetik	120	16	.480	.420	.730
Tosun ve Taşkesenligil (2011)	Çözeltiler ve Fiziksel Özellikler	160	31	.410	.400	.770

### 1.1 Araştırmanın Amacı

Araştırmanın amacı, beşinci sınıf Fen Bilimleri dersi ‘Maddenin Değişimi’ ünitesine yönelik başarı testi geliştirmektir. Bunun için, ‘Maddenin Değişimi’ ünitesine yönelik geliştirilen çoktan seçmeli soruların geçerlik ve güvenilirlik analizi yapılmıştır.

## 1.2. Araştırmanın Önemi

Fen eğitimi alanında geliştirilen başarı testlerinin çoğunlukla geliştirildiği konu alanlarındaki kavram yanlışlarını belirlemeye odaklandığı görülmektedir. Örneğin; kütle-ağırlık, ısı-sıcaklık, bileşik-karışım, erime-donma, madde ve değişim gibi kavramlar öğrencilerin en çok kavram yanlışısına sahip olduğu alanlardır (Açıkgöz ve Karşlı, 2015; Ayvacı ve Durmuş, 2016; Demir, Kızılay ve Bektaş, 2016; Kenan ve Özmen, 2014; Tosun ve Taşkesenligil, 2011). Dolayısıyla, 'Maddenin Değişimi' ünitesine ait bölüm, konu ve kavramlar hakkında başarı testlerinin geliştirilmesinin Fen öğretimine katkı sağlayacağına inanılmaktadır. Fen Bilimleri Dersi Öğretim Programında (MEB, 2013), öğrencilerin süreç içerisinde izlenmesi, yönlendirilmesi, öğrenme güçlüklerinin belirlenerek giderilmesi, anlamlı ve kalıcı öğrenmenin desteklenmesi amacıyla sürekli geri bildirim sağlanmasına yönelik bir ölçme-değerlendirme anlayışı benimsenmiştir. Eğitim sürecinde elde edilen sayısal değerlerin anlam kazanabilmesi, öğrencinin gelişiminin izlenmesi ve bu gelişime bağlı olarak öğrencinin yönlendirilmesi, çoktan seçmeli testler ile mümkün olabileceği düşünülmektedir. Ayrıca alan yazında Fen araştırmaları kapsamında beşinci sınıf 'Maddenin Değişimi' ünitesine yönelik başarı testi geliştirme çalışmasına rastlanmamıştır. Bu yüzden çalışmanın alan yazına katkı sağlayacağı düşünülmektedir. Geliştirilecek olan başarı testinin MEB (2013) tarafından güncellenen Fen bilimleri öğretim programına uygun olması önemlidir. Çünkü geliştirilen bu başarı testi, ortaokulda çalışan öğretmenler için öğrencilerin 'Maddenin Değişimi' ünitesine yönelik başarılarını belirleme aşamasında kullanılabilir. Ayrıca, araştırmacılar için bu konuda öğrencilerin başarılarını araştırma sürecinde de kullanılabilir.

## 2.YÖNTEM

Bu araştırma bir başarı testi geliştirme çalışmasıdır. Eğitim araştırmalarında yapılan başarı testi geliştirme çalışmaları incelendiğinde test geliştirme süreci hazırlık, uygulama ve raporlaştırma aşamalarından oluşmakla birlikte bu aşamalarda çeşitli adımlar takip edilmektedir (Gömlüksiz ve Erkan, 2010).

### 2.1. Araştırma Modeli

Bu çalışma, nicel araştırma yöntemlerinden tarama deseni kullanılarak yapılmıştır. Tarama deseni, bir evren içinden seçilen bir örneklem üzerinde yapılan çalışmalar yoluyla evren genelindeki eğilim, tutum veya görüşlerin nicel olarak betimlemesini sağlar (Creswell ve Plano Clark, 2011). Tarama deseninde bilimin gözleme kaydetme, olaylar arasındaki ilişkileri tespit etme, kontrol edilen değişmez ilişkiler üzerinde genellemelere varma durumları incelenir (Karaşar, 2009).

### 2.2. Örneklem

Bu çalışmanın örneklemini 2014-2015 eğitim-öğretim yılında İstanbul ili Ümraniye ilçesine bağlı bir ortaokulda altıncı sınıfta öğrenim gören 230 öğrenci oluşturmaktadır. Araştırmalarda çalışma grubunun büyüklüğü arttıkça, gerçek puanlara daha fazla yaklaşılabileceği, daha doğru tahminler yapılabileceği ifade edilmektedir. Çalışma Grubu büyüklüğünün belirlenmesinde faktör analizi, madde analizi gibi işlemler dikkate alınarak madde sayısının en az iki kat, hatta tercihen 10 kat olması tavsiye edilmektedir (Büyükoztürk, Kılıç Çakmak, Akgün, Karadeniz ve Demirel, 2012). Bu araştırma kapsamında 30 madde halinde hazırlanan başarı testi, yaklaşık olarak madde sayısının 8

katı sayıda örnekleme uygulanmıştır. Bu bağlamda geliştirilen testin geçerlilik ve güvenilirlik analizlerinde örneklem sayısı olarak 230 öğrencinin yeterli olduğu söylenebilir. Araştırmaya katılan öğrencilerin ortaokul altıncı sınıf öğrencilerinden seçilmesinde ortaokul beşinci sınıf öğrencilerinin başarı testi uygulanmadan önce bu üniteyi işlememiş olmaları veya yeni işlemiş olmaları nedeniyle soruları boş bırakma eğiliminde olacakları düşünülerek araştırmada bu sorunu en aza indirme ve tüm sorularının cevaplanmasını sağlamak amacı ile örneklem ortaokul altıncı sınıf öğrencilerinden oluşturulmuştur.

### **2.3. Başarı Testi Geliştirme Aşamaları**

#### **2.3.1. Testin amacının belirlenmesi**

Yapılan alan yazın taramasında beşinci sınıf Fen Bilimleri dersi ‘Maddenin Değişimi’ ünitesine ait başarı testi çalışmalarına rastlanmamıştır. Bu nedenle araştırmanın özgün olacağı, alan yazına katkı sağlayacağı ve bu alanda çalışma yapacak olan araştırmacılara yararlı olacağı düşünülerek beşinci sınıf Fen Bilimleri dersi ‘Maddenin Değişimi’ ünitesine ait bir başarı testi geliştirmek amaç edinilmiştir.

#### **2.3.2. Konunun belirlenmesi**

Günlük hayatta sıkça karşılaşılan ve daha önceki çalışmalarda geliştirilen testlerin kapsamında yer almayan ‘Maddenin Değişimi’ ünitesinin alt başlıklarına ait konu ve kavramlar bu çalışmada geliştirilen başarı testinin kapsamına alınmıştır. Bu ünite öğrencilerin; maddenin hâl değiştirmesi sürecinde oluşan erime, donma, kaynama, yoğunlaşma, buharlaşma, süblimleşme ve kırılgılaşma olaylarını ısı alınıp verilmesi temelinde açıklamaları ve erime, donma ve kaynama noktalarını kullanarak saf maddeleri ayırt etmeleri amaçlanmaktadır (MEB, 2013).

#### **2.3.3. Madde havuzu**

Beşinci sınıf Fen Bilimleri dersi ‘Maddenin Değişimi’ ünitesiyle ilgili öğrenci kazanımları ve bilimsel süreç becerileri dikkate alınarak test soruları Fen Bilimleri ders ve test kitaplarından ve sanal ortamdan araştırılarak araştırmacı tarafından araştırmanın amacına uygun olacak şekilde geliştirilmiştir. Başlangıçta 55 maddeden oluşan bir soru havuzu oluşturulmuştur.

#### **2.3.4. Yazım denetimi ve uzman görüşü alma**

Hazırlanan 55 madde yazım kuralları ve bilimsel açıdan doğruluklarının değerlendirilmesi için, fen eğitimi alanında uzman üç öğretim üyesi ve beş fen bilimleri dersi öğretmenine soruların doğruluğu, belirlenen hedef kazanımlara ve öğrencilerin seviyelerine uygunluğu ile konu bütünlüğünü kapsayıp kapsamadığı incelenmiş ve görüşler doğrultusunda soru sayısı 30’a indirilmiştir.

#### **2.3.5. Testin kapsam geçerliğinin incelenmesi**

Başarı testinin kapsam geçerliği için; uygulamadan önce Fen Bilimleri öğretim programında (MEB, 2013) yer alan ‘Maddenin Değişimi’ ünitesine ait bölümler ve kazanımlara ilişkin belirtke tablosu hazırlanmıştır.

### 2.3.6. Testin uygulanması

Uygulama için hazır hale getirilen başarı testi, 2014-2015 eğitim öğretim yılı güz döneminde İstanbul ili Ümraniye ilçesinde bulunan bir ortaokulda altıncı sınıfta öğrenim görmekte olan toplam 230 öğrenciye uygulanmıştır. Testi doğru cevaplayan öğrencilerin cevapları SPSS.22 istatistik programına '1', yanlış ve boş cevaplar ise '0' olarak girilmiştir. Dolayısıyla bir katılımcının testten alabileceği en yüksek puan 30, en düşük puan ise sıfır olarak belirlenmiştir.

### 2.3.7. Madde güçlük ve ayırt edicilik analizi

Madde güçlüğü ( $P_{jx}$ ); sıfır'a yaklaştıkça soru zorlaşır, bir'e yaklaştıkça ise soru kolaylaşır. Soruların madde güçlüklerinin .50 civarında olması testin geçerlilik ve güvenilirliğini artırır. Madde ayırt ediciliği ( $R_{jx}$ ); soruların ölçülen özelliklerle ilgili bilen ve bilmeyen bireyleri ne derece ayırt ettiğini gösterir. Madde ayırt edicilik indeks değeri, .30 'un altında olan sorular ya ölçekten çıkartılmalı ya da sorular düzeltilerek kullanılmalıdır (Büyüköztürk, 2012). Madde analizi kapsamında öğrencilerin puanları yüksekten düşüğe doğru sıralanmıştır. Puan sıralamasına göre öğrencilerden % 27'lik alt ve % 27'lik üst grup olmak üzere iki grup belirlenmiştir. Her bir soru için alt ve üst gruplara göre soruların madde güçlük ve ayırt edicilik indeksleri gerekli formüller kullanılarak hesaplanmıştır;

$$P_{jx} = \frac{D_{\bar{U}} + D_A}{N_{\bar{U}} + N_A} \quad R_{jx} = \frac{D_{\bar{U}} - D_A}{N_{\bar{U}} \text{ veya } N_A}$$

$P_{jx}$ : Madde güçlük indeksi,

$R_{jx}$ : Madde ayırt edicilik indeksi,

$D_{\bar{U}}$ : Maddeyi %27'lik üst grupta doğru cevaplayan öğrencilerin sayısı

$D_A$ : Maddeyi %27'lik alt grupta doğru cevaplayan öğrencilerin sayısı

$N_{\bar{U}}$ : Üst gruptaki %27'lik öğrenci sayısı

$N_A$ : Alt gruptaki %27'lik öğrenci sayısı

Soruların alt ve üst gruplardaki ayırt edicilik düzeylerini ölçmek için bir başka analiz olarak, verilere bağımsız gruplar t-testi yapılmıştır.

### 2.3.8. Güvenirlilik analizi

Bir testten öğrencilerin aldığı puanların güvenilirliğine hizmet edebilecek farklı istatistik teknikler vardır. KR-20 ve KR-21 istatistik teknikleri, doğru yanıtı bir, yanlış yanıtı sıfır puan verilen testlerde kullanılmak üzere geliştirilmiştir. Bunlardan KR-20 testteki her bir maddenin güçlük indekslerinin bilinmesi durumunda, KR-21 ise bilinmemesi durumunda kullanılır. Bir testten öğrencilerin aldığı puanların güvenilirliğine hizmet edebilecek bir başka istatistik tekniği ise Cronbach Alfa yöntemidir. Bu yöntem KR-20 ile aynı mantık üzerine kurulu olduğundan, doğru yanıtı bir, yanlış yanıtı sıfır puan verilen testlerde kullanılabilir (Atılğan, 2013). Maddenin Değişimi ünitesi başarı testi çalışmasında güvenilirlik katsayısı ( $r$ ), Kuder Richardson-20 (KR-20) yöntemi ile hesaplanmıştır.

### 2.3.9. Maddelerin sınıflanması

Testte bulunan her bir maddenin bilimsel süreç boyutu ve bilgi birikimi boyutuna göre sınıflama tablosu revize edilmiş Bloom Taksonomisi tablosu hazırlanmıştır.

### 3.BULGULAR

Çalışmanın bu kısmında, beşinci sınıf Fen Bilimleri dersi, ‘Maddenin Değişimi’ ünitesi için geliştirilen başarı testinin geçerlik ve güvenilirlik ile ilgili analiz sonuçlarına yer verilmiştir.

#### 3.1. Testin Geçerlik Çalışması

##### 3.1.1. Kapsam geçerliği

Kapsam geçerliği, başarı testlerini geliştirmede başlangıç noktasını oluşturur. Testler için belirtke tablosu hazırlamak testin kapsam geçerliğini arttırmada kullanılan bir yoldur (Büyükoztürk ve ark., 2012). Bu amaçla çalışmada uygulanacak testin kapsam geçerliğini sağlamak için ilk olarak, uygulamadan önce soruların kapsadığı bölüm ve konularla ilgili belirtke tablosu hazırlanmıştır. ‘Maddenin Değişimi’ ünitesine ait bölümler, kazanımlar ve konular (MEB, 2013) ile ilgili belirtke tablosu Tablo 2’de verilmiştir.

**Tablo 2.**

*Başarı Testinin Belirtke Tablosu*

Bölüm	Kazanımlar*	Konular/Kavramlar	Soru
<b>Maddenin Hal Değişimi</b>	Maddelerin ısı etkisiyle hâl değiştirebileceğine yönelik deneyler yapar, elde ettiği verilere dayalı çıkarımlarda bulunur.	Erime ve Donma	8, 14
		Kaynama, Buharlaştırma ve Yoğuşma	1, 4, 10, 13, 16, 18, 19
<b>Maddenin Ayırt Edici Özellikleri</b>	Saf maddelerin ayırt edici özelliklerinden erime, donma ve kaynama noktalarını, yaptığı deneyler sonucunda belirler.	Süblimleşme ve Kırağılaşma	26, 28
		Erime-Donma Sıcaklığı	20, 24, 25
<b>Isı ve Sıcaklık</b>	Isı ve sıcaklık arasındaki temel farkları açıklar. Sıcaklığı farklı olan sıvıların karıştırılması sonucu ısı alışverişi olduğuna yönelik deneyler yapar ve sonuçlarını yorumlar.	Kaynama Sıcaklığı	21, 23, 30
		Isı ve Sıcaklık	5, 11, 15
<b>Isı Maddeleri Nasıl Etkiler</b>	Günlük yaşamdan örneklerle genleşme ve büzülme olayları arasındaki ilişkiyi fark eder.	Katılarda Genleşme	9, 12, 17
		Sıvılarda Genleşme	3, 22
		Gazlarda Genleşme	2, 29

\*İlköğretim Kurumları (İlkokullar ve Ortaokullar) Fen Bilimleri Dersi Öğretim Programı (MEB, 2013)

Belirtke tablosunun hazırlanması sürecinde, alanında uzman üç fen eğitimcisi ve beş fen bilgisi öğretmeni ile çalışılmıştır. Böylece test sorularının kapsam geçerliği sağlanmıştır.

**3.1.2. Madde analizi**

Madde analizinde öğrencilerin puanları yüksekte doğru sıralanarak % 27'lik alt ve % 27'lik üst grup olmak üzere iki grup belirlenmiştir. Üst grubun belirlenmesi aşamasında 230 öğrenciden puanı en yüksek olan 62 öğrencinin puanlarına bakılmıştır. Bu öğrencilerden 62. sıradaki öğrencinin puanının 23 olduğu tespit edilmiştir. Aynı zamanda 23 puan almış olan dört öğrenci daha olduğu görülmüştür. Bu sebeple üst grubun öğrenci sayısı 66 olarak alınmıştır. Alt grubun belirlenmesi aşamasında da aynı duruma dikkat edilmiştir. Bunun sonucunda alt gruptaki öğrenci sayısı ise 64 olarak alınmıştır. Alt ve üst gruplara göre soruların madde güçlük indeksleri (P<sub>jx</sub>) ve madde ayırt edicilik indeksleri (R<sub>jx</sub>) hesaplanmıştır. Ayrıca bağımsız gruplar t testi yapılmıştır. Sorulara ilişkin P<sub>jx</sub>, R<sub>jx</sub> değerleri ve p anlamlılık değeri Tablo 3'te verilmiştir.

**Tablo 3.**  
*Başarı Testinin Madde Güçlük ve Ayırt Edicilik İndeksleri*

Sorular	Ortalama Madde Güçlük İndeksi (P <sub>jx</sub> )	Ortalama Madde Ayırt Edicilik İndeksi (R <sub>jx</sub> )	Anlamlılık Değeri (p)	Sonuç
1	.78	.41	0.003	Başarı testinde kullanıldı
2**	.27	.01	0.076	Başarı testinden çıkarıldı
3**	.24	.24	0.057	Başarı testinden çıkarıldı
4	.60	.63	0.001	Başarı testinde kullanıldı
5	.57	.58	0.002	Başarı testinde kullanıldı
6	.46	.64	0.001	Başarı testinde kullanıldı
7	.57	.46	0.008	Başarı testinde kullanıldı
8	.74	.53	0.003	Başarı testinde kullanıldı
9*	.24	.32	0.042	Düzeltilme Yapılarak Başarı testinde kullanıldı
10	.81	.53	0.000	Başarı testinde kullanıldı
11	.45	.63	0.004	Başarı testinde kullanıldı
12	.51	.70	0.001	Başarı testinde kullanıldı
13	.63	.61	0.007	Başarı testinde kullanıldı
14	.48	.43	0.001	Başarı testinde kullanıldı
15	.51	.42	0.006	Başarı testinde kullanıldı
16*	.35	.30	0.035	Düzeltilme Yapılarak Başarı testinde kullanıldı
17*	.28	.30	0.040	Düzeltilme Yapılarak Başarı testinde kullanıldı
18	.35	.45	0.001	Başarı testinde kullanıldı
19*	.49	.35	0.020	Düzeltilme Yapılarak Başarı testinde kullanıldı
20	.38	.46	0.002	Başarı testinde kullanıldı
21	.53	.60	0.003	Başarı testinde kullanıldı
22	.59	.53	0.000	Başarı testinde kullanıldı
23	.39	.43	0.010	Başarı testinde kullanıldı
24	.38	.43	0.003	Başarı testinde kullanıldı



**Tablo 3. Devamı**

Sorular	Ortalama Madde Güçlük İndeksi (P <sub>jx</sub> )	Ortalama Madde Ayırt Edicilik İndeksi (R <sub>jx</sub> )	Anlamlılık Değeri (p)	Sonuç
25	.58	.45	0.002	Başarı testinde kullanıldı
26	.42	.53	0.000	Başarı testinde kullanıldı
27	.53	.50	0.001	Başarı testinde kullanıldı
28	.48	.54	0.003	Başarı testinde kullanıldı
29	.49	.51	0.014	Başarı testinde kullanıldı
30	.37	.42	0.002	Başarı testinde kullanıldı

\* Düzeltme Yapılarak Başarı testinde kullanıldı

\*\* Başarı testinden çıkarıldı

Madde güçlük indeksinin 1'e yaklaşması o maddenin kolay olduğu, 0'a yaklaşması zor olduğu, 0.50 olması ise sorunun orta güçlükte olduğu anlamına gelir. Ayrıca eğer bir testin konuyu bilenlerle bilmeyenler arasındaki değişkenliği doğru bir şekilde tanımlaması isteniyorsa orta güçlükte maddelerden oluşması gerekir (Gömleksiz ve Erkan, 2010). Bu değerler ve açıklamalar göz önüne alındığında Tablo 3'e göre soruların kolay, orta ve zor sorulardan oluşmasına rağmen çoğunun orta güçlükte (.58, .51, .53 vb.) olduğu görülmektedir. Buna göre testin sorularının güçlük düzeyi istenen durumu karşılamaktadır.

Bir testin ortalama güçlük indeksi de teste ilişkin bazı bilgileri verir. Testin ortalama güçlüğü'nün 0.50 den küçük olması testin öğrencilere zor geldiğini, 0.50 den büyük olması testin öğrencilere kolay geldiğini göstermektedir. Buna bağlı olarak bir testin ortalama güçlük indeksinin 0.50 civarında (orta güçlükte) olması gerekmektedir (Tekin, 2010). Uygulanan testin ortalama güçlük indeksine bakıldığında ( $P_{jx} = 0.501$ ) testin orta güçlükte olduğu söylenebilir.

Madde ayırt ediciliği; soruların ölçülen özellikle ilgili bilen ve bilmeyen bireyleri ne derece ayırt ettiğini gösterir. Madde ayırt edicilik indeksinin bir madde için; 0.40 ve üstünde olması o maddenin ayırt etme gücünün yüksek olduğunu, 0.30 ile 0.39 arasında olması maddenin ayırt etme gücünün orta düzeyde olduğu ve maddelerde düzeltme yapılıp geliştirilmesi gerektiği, madde ayırt edicilik indeks değeri, 0.30'un altında olan maddelerin ise ölçekten çıkartılması gerektiği ifade edilmektedir (Büyüköztürk, 2012). Uygulanan testin madde ayırt edicilik indeksine Tablo 5'de görüldüğü üzere ( $R_{jx} = 0.488$ ) testin ayırt etme gücünün yüksek olduğu söylenebilir. Tablo 3 incelendiğinde madde ayırt etme gücü 0.30'un altında kalan 2 soru başarı testinden çıkarılmış, .30 ile .40 arasında kalan 4 soru ise küçük düzeltmeler yapılarak testte kullanılmıştır. Böylece 'Maddenin Değişimi' ünitesi başarı testinin son hali toplam 28 sorudan oluşmuştur.

Test sorularının alt ve üst gruplar için ayırt edici olup olmadığının belirlenmesinde kullanılacak bir başka analiz ise bağımsız gruplar t testidir (Büyüköztürk, 2012; Kan, 2014). Alt ve üst gruplar arasında 28 tane soru açısından anlamlı bir fark olması ( $p < .05$ ) bu soruların ayırt edici olduğunu göstermektedir.

Yapılan madde analizleri sonucunda Tablo 5'e göre başarı testinin ortalama güçlüğü .501 ve ortalama ayırt etme gücü ise .488 olarak hesaplanmıştır. Bu sonuçlara göre başarı testinin güçlük derecesi .50 civarında olduğu için zorluk derecesi bakımından ideal

olduğu ifade edilebilir. Ayırt edicilik bakımından ise ortalama değer 0.40'dan büyük olduğu için ayırt edicilik özelliği çok iyi olarak ifade edilebilir (Büyüköztürk, 2012).

### 3.2. Testin Güvenirlik Çalışması

Güvenirlik, testin ölçmek istediği özelliği ne derece doğru ölçtüğüyle ilgili bir kavramdır. Güvenirlik katsayısı sıfır ile bir arasında değişen bir sayı ile ifade edilir. Bu değer bir yaklaşması oranında testten öğrencilerin aldığı puanların güvenilirliği artar (Gömlüksiz ve Erkan, 2010). Cronbach Alfa veya güvenilirlik katsayısının 0.40'tan düşük olması testten öğrencilerin aldığı puanların güvenilir olmadığını, 0.40-0.60 arasında olması testten öğrencilerin aldığı puanların güvenilirliğinin düşük olduğunu, 0.60-0.90 arasında olması testten öğrencilerin aldığı puanların oldukça güvenilir olduğunu, 0.90 nın üstünde olması ise testten öğrencilerin aldığı puanların yüksek derecede güvenilir olduğunu gösterir (Can, 2014). 'Maddenin Değişimi' ünitesi başarı testi çalışmasında güvenilirlik katsayısı (r), Kuder Richardson-20 (KR-20) yöntemi ile hesaplanmış ve testin güvenilirlik değeri 0.714 olarak bulunmuştur. Bu değere bakılarak öğrencilerin testten aldığı puanların güvenilir olduğu söylenebilir.

Testte bulunan her bir maddenin bilimsel süreç boyutu ve bilgi birikimi boyutuna göre sınıflama tablosu revize edilmiş Bloom Taksonomisi tablosu hazırlanmıştır. Hazırlanan revize edilmiş 'Bloom Taksonomi Tablosu' Tablo 4'te gösterilmiştir.

**Tablo 4.**

*Başarı Testi Bloom Taksonomisi Tablosu*

Düzeyleler	BİLGİ BİRİKİMİ BOYUTU			
	Olgusal Bilgi	Kavramsal Bilgi	İşlemsel Bilgi	Biliş Ötesi Bilgi
<b>Hatırlama</b>	1, 14, 26, 28			
<b>BİLİŞSEL SÜREC BOYUTU</b>	<b>Anlama</b>	4, 10, 13, 15, 19, 20, 24	5, 8, 16	2, 3, 7, 12, 25
	<b>Uygulama</b>	17, 18		22, 27
	<b>Analiz Etme</b>	23	6	
	<b>Değerlendirme</b>		21	9, 11

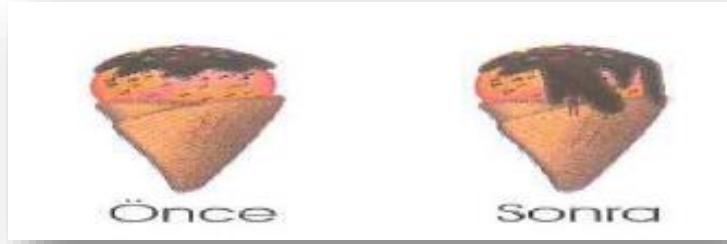
➤ *Hatırlama - Olgusal Bilgi düzeyi örnek soru (1)*

Çaydanlıkta ısıtılan su ( I ).....olayı ile su buharına dönüşür. Su buharı soğutulduğunda ise tekrar su hâline geçer. Bu olaya ( II ).....denir.

Yukarıda verilen paragraftaki I ve II numaralı boşluklara aşağıdakilerden hangisi gelmelidir?

I	II
A) Erime	Yoğuşma
B) Buharlaştırma	Donma
C) Buharlaştırma	Yoğuşma
D) Donma	Erime

➤ *Anlama - Kavramsa Bilgi düzeyi örnek soru (8)*



Yukarıdaki olay ile ilgili,

- I. Dondurma ısı vererek hal değiştirmiştir.
- II. Dondurma katı halden sıvı hale geçmiştir.
- III. Dondurmanın yukarıdaki değişimi erime olayına örnektir.

Yargılarından hangisi veya hangileri doğrudur?

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

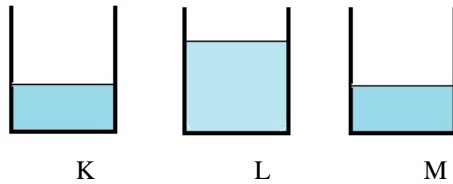
➤ *Uygulama – Bilgi Ötesi Bilgi düzeyi örnek soru (22)*

- 1. Termometre                      2. Yangın alarm sistemleri
- 3. Konserve                        4. Sıcak hava - seyahat balonları

Yukarıda verilen madde ve olaylardan hangisi ya da hangileri maddelerin genişleme özelliğinden yararlanılarak yapılmıştır?

- A) 1 ve 2                              B) 2 ve 3  
C) 1, 2 ve 3                        D) 1, 2 ve 4

➤ *Analiz Etme – İşlemsel Bilgi düzeyi örnek soru (6)*

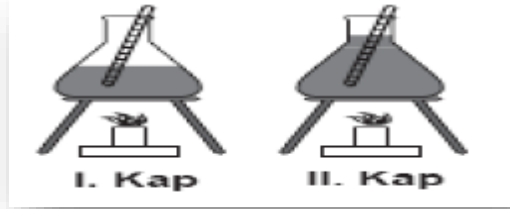


Yukarıdaki kaplarda aynı sıcaklıkta saf su bulunmaktadır. Kaplar özdeş ısıtıcılar kullanılarak 10 dakika ısıtılıyor. Buna göre termometre yardımıyla sıvıların sıcaklıkları ölçüldüğünde aşağıdakilerden hangisi yanlış olur?

- A) Sıcaklığı en fazla olan L'dir.
- B) K ve M'nin sıcaklığı eşittir.
- C) K, L ve M eşit ısı almıştır.
- D) K'nın sıcaklığı L'den fazladır.

➤ *Değerlendirme – Biliş Ötesi Bilgi düzeyi örnek soru (11)*

Neşe öğretmen, aşağıda şekilde özdeş kaplarda bulunan aynı sıcaklıktaki suları özdeş ısıtıcılarla eşit süre ısıtıyor. Daha sonra öğrencilerden suyun sıcaklıkları hakkında yorum yapmalarını istiyor.



Öğrencilerin yaptığı yorumlardan hangisi doğrudur?

- A) I. Kaptaki suyun sıcaklığı daha düşüktür.
- B) II. Kaptaki suyun sıcaklığı daha düşüktür.
- C) Kaplardaki maddeler aynı olduğundan sıcaklıkları eşittir.
- D) Kaplardaki suyun miktarları farklı olduğundan bir şey söylenemez.

Araştırmada 'Maddenin Değişimi' ünitesine yönelik geliştirilen başarı testine ait istatistiksel sonuçlar genel olarak Tablo 5'te verilmiştir.

**Tablo 5.**

*Başarı Testine Ait İstatistiksel Sonuçlar*

<b>MDÜ-BT Soru Sayısı</b>	28
<b>Uygulanan kişi sayısı (N)</b>	230
<b>Kuder Richardson-20 güvenirlik katsayısı (r)</b>	0.714
<b>Ortalama madde güçlüğü (P<sub>jx</sub>)</b>	0.501
<b>Ortalama madde ayırt ediciliği (R<sub>jx</sub>)</b>	0.488

Tablo 5 incelendiğinde, geliştirilen başarı testinin ortalama güçlük indeksinin 0.501 ve ayırt edicilik indeksinin ise 0.488 olarak hesaplandığı görülmektedir. MDÜ-BT'nin analizi sonucunda KR-20 güvenirlik katsayısı 0.714 olarak hesaplanmıştır. Yapılan analizler sonucunda, araştırmada beşinci sınıf Fen Bilimleri dersi 'Maddenin Değişimi' ünitesine ait geçerli ve puanların güvenilir olduğu bir başarı testi geliştirildiği söylenebilir. Testin son hali Ek-1'de verilmiştir.

#### 4.TARTIŞMA ve SONUÇ

Bu çalışmanın amacı beşinci sınıf Fen Bilimleri dersi 'Maddenin Değişimi' ünitesiyle ilgili geçerliği ve güvenirliği kanıtlanmış çoktan seçmeli sorulardan oluşan bir başarı testi geliştirmektir. Bu amaçla geliştirilen test oluşturulurken bazı basamaklar izlenmiştir. Bu basamaklar; testin amacı ve konusunun belirlenmesi, maddelerin yazılması ve soru havuzunun oluşturulması, yazım denetiminin yapılması ve uzman görüşünün alınması, testin kapsam geçerliğinin incelenmesi ve uygulamanın yapılması, uygulama sonrası madde güçlük ve ayırt edicilik analizlerinin yapılması, daha sonra testin güvenirlik analizinin yapılması ve son olarak testte bulunan her bir maddenin bilimsel süreç boyutu

ve bilgi birikimi boyutuna göre sınıflama tablosu olan revise edilmiş Bloom Taksonomisi tablosu hazırlanması şeklindedir. Alan yazında Fen eğitimi araştırmaları kapsamında benzer basamaklar dikkate alınarak geliştirilen birçok başarı testi geliştirme çalışması bulunmaktadır (Açıkgöz ve Karşlı, 2015; Akbulut ve Çepni, 2013; Ayvacı ve Durmuş, 2016; Çakır ve Aldemir, 2011; Demir ve Akarsu, 2014; Demir, Kızılay ve Bektaş, 2016; Gönen, Kocakaya ve Kocakaya, 2011; Kenan ve Özmen, 2014; Şen ve Eryılmaz, 2011; Şener ve Taş, 2017; Tosun ve Taşkesenligil, 2011). Alan yazındaki bu çalışmalara bakıldığında mevcut çalışmaya benzer test geliştirme basamaklarının kullanıldığı görülmektedir.

Belirtilen çalışmalardan ve bu çalışmadan farklı olarak Kenan ve Özmen'in (2014), Demir ve Akarsu'nun (2014) ve Gönen, Kocakaya ve Kocakaya'nın (2011) çalışmalarında sorular sadece uzman kontrolüne sunulurken kapsam geçerliği sağlanmaya çalışılmıştır.

Madde analizi kısmında, uygulanan başarı testi için her bir maddenin ayrı ayrı madde güçlük ve ayırt edicilik indeksleri, ayrıca ortalama madde güçlük indeksleri ve madde ayırt edicilik indeksleri hesaplanmıştır. Analiz sonucunda, başarı testinin ortalama madde güçlük indeksi değeri 0.501 ve ortalama madde ayırt edicilik indeksi 0.488 olarak bulunmuştur. Testin ortalama madde güçlük indeksi değeri 0.50'den küçükse test öğrencilere zor, 0.50'den büyük ise testin öğrencilere kolay geldiğini söylenebilir. Bu nedenle madde ortalama güçlük indeksinin 0.50 düzeyinde olması gerekmektedir (Gömleksiz ve Erkan, 2010). Bu durumda araştırmada öğrencilere uygulanan başarı testinin ortalama madde güçlük indeksinin 0.50'ye yakın olması, testin öğrencilerin düzeyine uygun olduğunu göstermektedir. İlgili alan yazın incelendiğinde Fen eğitimi araştırmaları kapsamında geliştirilen başarı testlerinin madde güçlük endeksinin .50 civarında olan benzer çalışmalara da (Akbulut ve Çepni, 2013; Demir, Kızılay ve Bektaş, 2016; Şen ve Eryılmaz, 2011; Şener ve Taş, 2017) rastlanmaktadır. Ayrıca alan yazında Fen eğitimi araştırmaları kapsamında geliştirilen başarı testlerinin madde güçlük endekslerinin .50'nin çok üzerinde olan çalışmalar (Açıkgöz ve Karşlı, 2015) olduğu gibi, 0.50'nin altında olan çalışmalarda (Ayvacı ve Durmuş, 2016; Çakır ve Aldemir, 2011; Demir ve Akarsu, 2014; Gönen, Kocakaya ve Kocakaya, 2011; Tosun ve Taşkesenligil, 2011) mevcuttur.

Ayrıca geliştirilen başarı testinin madde ayırt edicilik indeks değeri 0.40 ve üzerinde ise maddenin ayırt etme gücünün yüksek olduğu sonucuna ulaşılır (Tekin, 2010). Araştırmada geliştirilen başarı testinin ortalama madde ayırt edicilik indeksinin 0.488 olduğu tespit edilmiştir. Bu değer, hazırlanan başarı testinin madde ayırt edicilik düzeyinin yüksek olduğunu göstermektedir. İlgili alan yazın incelendiğinde Fen eğitimi araştırmaları kapsamında geliştirilen başarı testlerinin madde ayırt edicilik indekslerinin .40'ın üzerinde olan benzer çalışmalara da (Açıkgöz ve Karşlı, 2015; Akbulut ve Çepni, 2013; Çakır ve Aldemir, 2011; Demir, Kızılay ve Bektaş, 2016; Gönen, Kocakaya ve Kocakaya, 2011; Şen ve Eryılmaz, 2011; Şener ve Taş, 2017; Tosun ve Taşkesenligil, 2011) rastlanmaktadır. Ayrıca alan yazında Fen eğitimi araştırmaları kapsamında geliştirilen başarı testlerinin madde güçlük indekslerinin .40'ın altında olan (Ayvacı ve Durmuş, 2016; Demir ve Akarsu, 2014) çalışmalarda mevcuttur. Alan yazındaki bazı çalışmalarda (Kenan ve Özmen, 2014) ise geliştirilen başarı testinin madde güçlük ve madde ayırt edicilik indekslerinin hesaplanmadığı tespit edilmiştir.

Bir başarı testi geliştirilirken dikkate alınması gereken bir diğer önemli nokta testten elde

edilen puanların güvenilirliğidir. Güvenirlik analizinde en çok tercih edilen yöntemler KR-20, KR-21 ve Cronbach Alpha güvenirlilik katsayılarının hesaplanmasıdır. KR-20 yöntemi, Cronbach Alfa yöntemi ile aynı mantık üzerine kurulu olduğundan, doğru yanıtla bir, yanlış yanıtla sıfır puan verilen testlerde kullanılabilir (Atılğan, 2013). Yapılan araştırmada beşinci sınıf öğrencilerine uygulanan 'Maddenin Değişimi' ünitesi başarı testi sonuçları için güvenirlilik analizleri yapılmış ve başarı testindeki her bir maddenin güçlük indeks değeri bilindiği için KR-20 güvenirlilik katsayısı (r) hesaplanmıştır ve elde edilen değer 0.714'tür. Yapılan araştırmada bulunan değer 0.60-0.90 arasında olduğu için başarı testinden öğrencilerin aldığı puanların güvenilir olduğu söylenebilir. Alan yazında Fen eğitimi araştırmaları kapsamında geliştirilen başarı testi çalışmaları (Açıkgöz ve Karşlı, 2015; Akbulut ve Çepni, 2013; Ayvaci ve Durmuş, 2016; Çakır ve Aldemir, 2011; Demir ve Akarsu, 2014; Demir, Kızılay ve Bektaş, 2016; Gönen, Kocakaya ve Kocakaya, 2011; Kenan ve Özmen, 2014; Şen ve Eryılmaz, 2011; Şener ve Taş, 2017; Tosun ve Taşkesenligil, 2011) güvenirlilik katsayısı dikkate alınarak incelendiğinde ulaşılan sonuçlar ile araştırmanın sonucu örtüşmektedir.

Sonuç olarak beşinci sınıf düzeyinde Fen Bilimleri dersi Maddenin Değişimi ünitesine ait alt bölümler, kazanımlar ve kavramlar çerçevesinde alan yazına katkı sağlayabilecek geçerli ve güvenilir bir başarı testinin geliştirildiği söylenebilir.

#### 4.1. Öneriler

Benzer çalışmaya katılacak öğrenciler aynı seviyede fakat farklı eğitim-öğretim kurumlarından seçilebilir ve geçerlik güvenirlilik çalışmaları yapılırken örneklem sayısı artırılabilir. Öğrencilerin 'Maddenin Değişimi' ünitesine yönelik kazanımlarını ölçebilmek amacıyla alternatif ölçme değerlendirme teknikleri kullanarak iki aşamalı, üç aşamalı çoktan seçmeli testler veya açık uçlu testler geliştirilebilir. Öğrencilerin 'Maddenin Değişimi' ünitesine yönelik tutumlarını belirlemek amacıyla likert tipi ölçekler geliştirilebilir.

### KAYNAKÇA

- Açıkgöz, M. ve Karşlı, F. (2015). Alternatif Ölçme-Değerlendirme Yaklaşımları Kullanılarak İş ve Enerji Konusunda Geliştirilen Başarı Testinin Geçerlilik ve Güvenirlik Analizi. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, 4(1), 1-25.
- Akbulut, H. İ. ve Çepni, S. (2013). Bir Üniteye Yönelik Başarı Testi Nasıl Geliştirilir? : İlköğretim 7. Sınıf Kuvvet ve Hareket Ünitesine Yönelik Bir Çalışma. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, 2(1), 18-44.
- Ayaydın, A. (2010). Desen Eğitiminde Ölçme ve Değerlendirme Üzerine Bir Araştırma. *Ahi Evran Üniversitesi Eğitim Fakültesi Dergisi*, 11(2), 159-172.
- Ayvacı, H. Ş. ve Durmuş, A. (2016). Bir Başarı Testi Geliştirme Çalışması: Isı ve Sıcaklık Başarı Testi Geçerlik ve Güvenirlik Araştırması. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi*, 35(1), 87-102.
- Bahar, M., Nartgün, Z., Durmuş, S. ve Bıçak, B. (2012). *Geleneksel-tamamlayıcı ölçme değerlendirme teknikleri* (5. Baskı). Ankara: Pegem Akademi.
- Baykul, Y. ve Turgut, M. F. (2014). *Eğitimde Ölçme ve Değerlendirme*, Ankara: PegemA.
- Büyüköztürk, Ş. (2012). *Sosyal bilimler için veri analizi el kitabı*. Ankara. Pegem A Yayıncılık.
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. ve Demirel, F. (2012). *Bilimsel araştırma yöntemleri* (13. Baskı). Ankara: Pegem Akademi
- Can, A. (2014). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi* (2. Baskı). Ankara: Pegem A Yayıncılık.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research*. Los Angeles: Sage.
- Çakır, M. ve Aldemir, B. (2011). İki Aşamalı Genetik Kavramlar Tanı Testi Geliştirme ve Geçerlik Çalışması. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 8(16).
- Demir, B. ve Akarsu, N. (2014). Modern Fizik Konuları ile İlgili Kavram Testi Geliştirilmesi ve Uygulanması: Modern Fizik Kavram Testi (MKFT). *Journal Of European Education*, 4(2).
- Demir, N., Kızılay, E. ve Bektaş, O. (2016). 7. Sınıf Çözümler Konusunda Başarı Testi Geliştirme: Geçerlik ve Güvenirlik Çalışması. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 10(1).
- Demirci, N. ve Efe, S. (2007). İlköğretim öğrencilerinin ses konusundaki kavram yanlışlarının belirlenmesi. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*. 1(1). 23-56.
- Erdoğan, M. Y. ve Kurt, F. (2012). Öğretmenlerin Ölçme ve Değerlendirme Yeterlik Algılarının Bazı Değişkenler Açısından İncelenmesi. *Eğitim Bilimleri Dergisi*, 1(2), 23-36.

- Gömlüksiz, M. ve Erkan, S. (2010). *Eğitimde ölçme ve değerlendirme (2. Baskı)*. Ankara: Nobel Yayın Dağıtım.
- 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*, VIII (I), 40-57.
- Kan, A. (2014). *Ölçme Aracı Geliştirme, Eğitimde Ölçme ve Değerlendirme* (ed. Satılmış Tekindal), Ankara, Pegem A.
- 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.
- Karasar, N. (2009). *Bilimsel araştırma yöntemleri*. Ankara: Nobel yayınları
- Karataş, F. Ö., Köse S. ve Coştu, B. (2003). Öğrenci yanılgılarını ve anlama düzeylerini belirlemede kullanılan iki aşamalı testler. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 13(1), 54-69.
- Kenan, O. ve Özmen, H. (2014). Maddenin tanecikli yapısına yönelik iki aşamalı çoktan seçmeli bir testin geliştirilmesi ve uygulanması. *Journal of Research in Education and Teaching*, 3(3), 371-378.
- Küçükahmet, L. (2002). *Öğretimde planlama ve değerlendirme (13. Baskı)*. Ankara: Nobel Yayın Dağıtım
- MEB. (2013). *İlköğretim kurumları (ilkokullar ve ortaokullar) fen bilimleri dersi (3, 4, 5, 6, 7 ve 8. sınıflar) öğretim programı*. Ankara: MEB Yayınevi
- Mintzes, J. J., Wandersee, J. H. & Novak, J. D. (2001). Assessing understanding in biology. *Journal of Biological Education*, 35(3), 118-125.
- Ogan Bekiroğlu, F. (2004). *Ne kadar Başarılı?, Klasik ve Alternatif Ölçme-Değerlendirme Yöntemleri ve Fizikte Uygulamalar* (1. baskı). Ankara: Nobel Yayın Dağıtım.
- Özçelik, D. A. (1998). *Ölçme ve Değerlendirme*. Ankara: ÖSYM Yayınları
- 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.
- Şener, N., & Taş, E. (2017). Developing Achievement Test: A Research for Assessment of 5th Grade Biology Subject. *Journal of Education and Learning*, 6(2), 254.
- Şimşek, Ü. (2007). *Çözümler ve kimyasal denge konularında uygulanan jigsaw ve birlikte öğrenme tekniklerinin öğrencilerin maddenin tanecikli yapıda öğrenmeleri ve akademik başarıları üzerine etkisi*. Doktora tezi, Atatürk Üniversitesi.
- Taşkın, (2008). *Fen ve teknoloji öğretiminde yeni yaklaşımlar*. Ankara: Pegem Akademi
- Tekin, H. (2010). *Eğitimde ölçme ve değerlendirme (20. Baskı)*. Ankara: Yargı Yayınevi.



- Tosun, C. ve Taşkesenligil, Y. (2011). Revize edilmiş Bloom'un taksonomisine göre çözümler ve fiziksel özellikleri konusunda başarı testinin geliştirilmesi: Geçerlik ve güvenirlik çalışması. *Kastamonu Eğitim Dergisi*, 19(2), 499-522.
- Turgut, M. F. ve Baykul, Y. (2012). *Eğitimde ölçme ve değerlendirme* (4. Baskı). Ankara: Pegem Akademi.
- Yeşilyurt, E. (2012). Fen ve Teknoloji Dersinde Kullanılan Ölçme-Değerlendirme Yöntemleri ve Karşılaşılan Güçlükler. *TurkishStudies*, 7(2), 1183-1205.

## EXTENDED ABSTRACT

### 1. Introduction

It is an important necessity to determine whether the education, which is defined as the process of bringing changes in human behavior, is achievement full in the process or changes of behavior (Karaca, 2008; Erdogdu and Kurt, 2012). This requirement in the training process is met by assessment and evaluation. For this reason, the measurement and evaluation emerge as an indispensable element of the educational process (Yesilyurt, 2012).

Measurement is to observe any quality and to express the result of the observation in numbers or adjectives (Turgut and Baykul, 2012). Evaluation is expressed as a result of comparing the results obtained with the measurement to a criterion (Bahar, Nartgun, Durmus and Bıcak, 2012).

To determine the realization levels of the goals and behaviors identified in science education and to determine the student achievements, there is a need for qualified and well-prepared measurement tools covering all relevant teaching objectives. Also, the validity and reliability of these measurement tools must be high enough to be qualified (Gonen, Kocakaya and Kocakaya, 2011). Measuring instruments used in education and training to measure and evaluate student achievement are; questionnaires, oral tests, correct miss tests, multiple choice tests, paired tests, void filling tests, scales, short response tests, written examinations, open ended questions, two step tests (Simsek, 2007; Ogan Bekiroglu, 2004).

Multiple choice tests are limited in determining students' critical thinking skills and creativity. Nonetheless, multiple-choice tests are particularly preferred measurement tools for determining students' conceptual misconceptions and achievements. (Kucukahmet, 2002).

The purpose of this study is to develop an achievement test consisting of multi-choice questions with the validity of points and validity of the fifth-grade science course 'Change of Matter' unit.

Science education achievement tests are being developed relating to subject areas developed especially when looking to determine the concept of illusion, it is seen that a lot of the achievement test. In addition, there was no study of the achievement test development for the fifth grade 'Change of Matter' unit within the scope of science research in the field. Therefore, it is considered that your work will contribute to the field of writing. In addition, for the researchers, the achievement of the students in this regard can also be used in the research process.

### 2. Method

This research is an achievement test development work. Examination of achievement test development studies in educational research is followed by various stages of test development process preparation, application and reporting stages (Gomleksiz and Erkan, 2010).

This study was conducted using screening design from quantitative research methods. The scan pattern allows quantitative representation of the trend, attitude, or views in the universe through studies on a sample selected from within a universe (Creswell and Plano Clark, 2011). In the scan pattern, science has made observations, establishing relationships between events, and reaching generalizations on the controlled invariant relations (Karasar, 2009).

The universe of this work consists of all sixth-grade students who are studying in the district of Umraniye, Istanbul province in the academic year of 2014-2015. The sample is composed of 230 students who are in the sixth grade in a secondary school affiliated to Umraniye district of Istanbul province.

It is stated that as the size of the study group increases in research, more accurate estimates can be made to approach the actual scores more. It is recommended that the number of items should be at least two times, preferably 10 times, considering factors such as factor analysis, item analysis, etc. in determining the size of the working group (Buyukozturk, Kılıc Cakmak, Akgun, Karadeniz and Demirel, 2012). Within the scope of this research, the achievement test, which was prepared in 30 items, was applied approximately 8 times the number of items. In this context, it can be said that the number of samples is sufficient in the validity and reliability analyzes of the developed test. It was effective for the students who participated in the research to be informed about the 'Change of Matter' unit within the science course of these students when they were selected from the sixth-grade students of the secondary school.

### 3. Findings, Discussion and Results

Some steps have been observed while the test developed for purpose. These steps are; to determine the aim and the subject of the test, to write the items and to form the pool of questions, to make the spelling check and to take the expert opinion, to examine the validity of the scope of the test and to carry out the application, to make the item difficulty and discrimination analysis after the application and then to carry out the reliability analysis of the test, Is the preparation of the revised Bloom Taxonomy table, which is a classification table according to the scientific process dimension and the knowledge accumulation dimension of each item.

The statistical results of the achievement test developed for the 'Change of Matter' unit are generally as follows;

- MCU-ST Question Number; 28
- Number of persons employed (N); 230
- Kuder Richardson-20 reliability coefficient (r); 0.714
- Average substance strength (P<sub>jx</sub>); 0.501
- Average substance discrimination (R<sub>jx</sub>); 0.488

As a result of the analyzes made, it can be said that an achievement test was developed in which the validity and scores of the fifth-grade science course 'Change of Matter' unit were reliable.

As a result of the research, the following suggestions can be made; Fifth-grade science course achievement test, which consists of more questions, can be developed to meet the

achievements of the 'Change of Matter' unit. Interview with some students can be conducted within the scope of the pilot study so as to be able to make a full decision on the related questions after the field search. The students who will participate in similar studies can be selected from the same level but from different education and training institutions and the number of samples can be increased while validity reliability studies are carried out. Two-stage, three-stage multiple-choice tests or open-ended tests may be developed using alternative assessment techniques to measure students' achievements towards the 'Change of Matter' unit at a higher level.

**Ek-1: Maddenin Değişimi Ünitesi Başarı Testi**

**SORULAR**

1. Çaydanlıkta ısıtılan su ( I ).....olayı ile su buharına dönüşür. Su buharı soğutulduğunda ise tekrar su hâline geçer. Bu olaya ( II ).....denir.

Yukarıda verilen paragraftaki **I ve II** numaralı boşluklara aşağıdakilerden **hangisi gelmelidir?**

<b>I</b>	<b>II</b>
A) Erime	Yoğuşma
B) Buharlaşma	Donma
C) Buharlaşma	Yoğuşma
D) Donma	Erime

2. İçinde bir miktar hava bulunan ağzı kapalı esnek bir balon içinde sıcak su bulunan bir kaba konuluyor.

Buna göre balonda aşağıdakilerden hangisi **gerçekleşir?**

- A) Büzülme
- B) Hacminde artma
- C) Hâl değişimi
- D) Hacminde azalma

3. Aşağıdaki olaylardan hangisi **hacim artışı** sonucunda gerçekleşir?

- A) Kışın su borularının patlaması.
- B) Kışın elektrik tellerinin gerginleşmesi.
- C) Şişirilmiş balonun soğukta büzülmesi.
- D) Termometrenin cıva seviyesinin düşmesi.



Günlük hayatta sıkça karşılaştığımız **yoğusma olayları** nelerdir?

4.

Fen Bilimleri öğretmeni Elif Hanım'ın sorusuna bazı öğrenciler aşağıdaki yanıtları veriyor.

Kışın evimizdeki camların buğulanması



Yusuf

Buzdolabından çıkan şişenin dışında su damlacıkları oluşması



Azra



Mina

Yazın göllerdeki su seviyesinin azalması

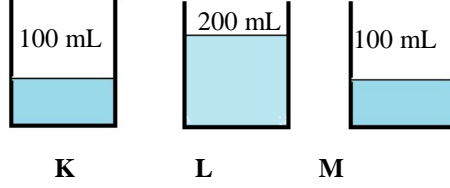
Buna göre hangi öğrencilerin verdiği yanıtlar doğrudur?

- A) Yusuf ve Mina                      B) Azra ve Mina  
C) Yusuf ve Azra                      D) Yusuf, Azra ve Mina

5. Isı ve sıcaklık kavramları ile ilgili aşağıdaki ifadelerden hangisi **yanlıştır**?

- A) Sıcaklıkları farklı maddeler temas ettiğinde aralarında ısı alış verişi olur.  
B) Sıcaklıkları yüksek maddeler temas ettikleri cisimleri ısıtır.  
C) Isı akışı sıcaklığı düşük olan maddeden yüksek olan maddeye doğrudur.  
D) Aynı maddeye çok ısı verdiğimizde çok, az ısı verdiğimizde az ısınır.

6.



*Yukarıdaki kaplarda aynı sıcaklıkta saf su bulunmaktadır. Kaplar özdeş ısıtıcılar kullanılarak 10 dakika ısıtılıyor. Buna göre termometre yardımıyla sıvıların sıcaklıkları ölçüldüğünde aşağıdakilerden hangisi **yanlış olur?***

- A) Sıcaklığı en fazla olan L'dir.
- B) K ve M'nin sıcaklığı eşittir.
- C) K, L ve M eşit ısı almıştır.
- D) K'nın sıcaklığı L'den fazladır.

7.



*Sıcaklıkları farklı olan A ve B tuğlaları birbirine temas ettiğinde A tuğlasının sıcaklığı artarken, B tuğlasının sıcaklığı azalmaktadır.*

Buna göre aşağıdakilerden hangisi **yanlıştır?**

- A) Başlangıçta A ve B tuğlalarının sıcaklıkları eşittir.
- B) A ve B tuğlalarının son sıcaklıkları eşittir.
- C) Isı akışı B'den A'ya doğrudur.
- D) A tuğlası ısı almıştır.

8.



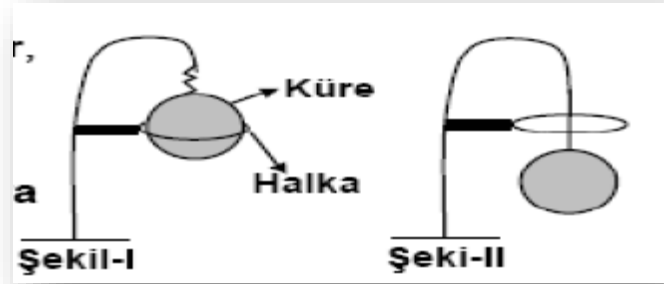
Yukarıdaki olay ile ilgili,

- I- Dondurma ısı vererek hal değiştirmiştir.
- II- Dondurma katı halden sıvı hale geçmiştir.
- III- Dondurmanın yukarıdaki değişimi erime olayına örnektir.

Yargılarından hangisi veya hangileri doğrudur?

- E) Yalnız-I    B) Yalnız-II    C) I ve II    D) II ve III

9.

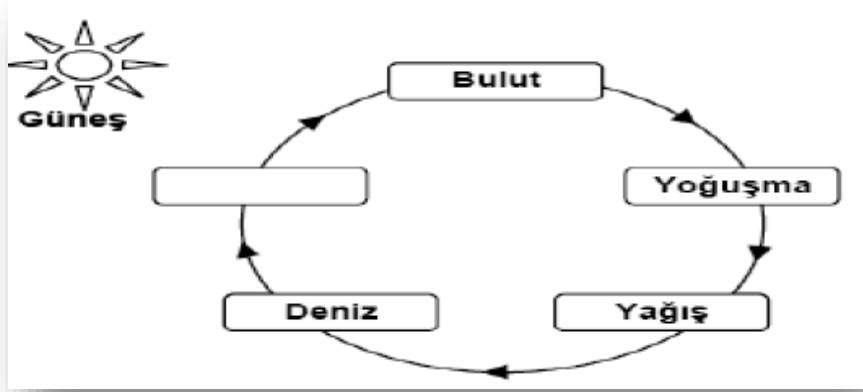


Özdeş metal küreler sıcaklıkları farklı sıvılarda eşit sürelerde bekletilip çapları eşit halkalara bırakıldığında şekil-I ve şekil-II' deki durum gözleniyor. Buna göre yukarıdaki deney seti ile aşağıdaki sorulardan hangisi cevaplanamaz?

- A) Maddeler arasında ısı alış-verişi olur mu?
- B) Sıcaklık değişimi maddelerin hacmini etkiler mi?
- C) Isının akış yönü maddelerin sıcaklığına bağlı mıdır?
- D) Isı alan maddenin şekli değişir mi



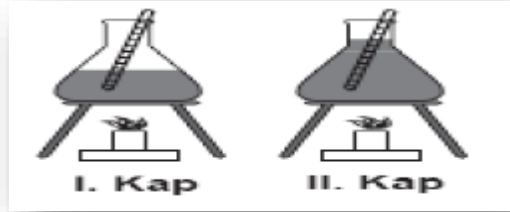
10.



Şekilde verilen su döngüsünde boş bırakılan kutuya aşağıdaki olaylardan hangisi gelmelidir?

- A) Buharlaşıma B) Erime C) Donma D) Sis

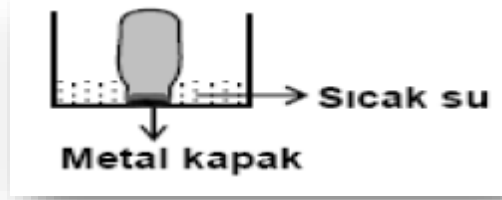
11. Neşe öğretmen, aşağıda şekilde özdeş kaplarda bulunan aynı sıcaklıktaki suları özdeş ısıtıcılarla eşit süre ısıtıyor. Daha sonra öğrencilerden suyun sıcaklıkları hakkında yorum yapmalarını istiyor.



Öğrencilerin yaptığı yorumlardan hangisi doğrudur?

- F) I. Kaptaki suyun sıcaklığı daha düşüktür.  
G) II. Kaptaki suyun sıcaklığı daha düşüktür.  
H) Kaplardaki maddeler aynı olduğundan sıcaklıkları eşittir.  
İ) Kaplardaki suyun miktarları farklı olduğundan bir şey söylenemez.

12.



Yukarıda kapağı zor açılan bir kavanoz, şekildeki gibi bir süre sıcak suda bekletildikten sonra kapak açılıyor. Bu sonuca göre aşağıdakilerden hangisi yada hangileri doğrudur?

- I- Metalin genişmesi, camın genişmesinden büyüktür.
  - II- Metalin genişmesi, camın genişmesinden küçüktür.
  - III- Metalin genişmesi, camın genişmesine eşittir.
- A) Yalnız-I B) Yalnız-II C) I ve II D) II ve III

13.

1	2	3
Erime	Yoğuşma	Donma
4	5	6
Buharlaştırma	Kaynama	Isınma

Buhardan buza dönüşüm sırasında yukarıda verilen olaylardan hangileri sırasıyla gerçekleşir?

- A) 1-4-6 B) 2-3-5 C) 4-5 D) 2-3

**14.** Erime ve Donma olayı ile ilgili aşağıdaki ifadelerden hangisi ya da hangileri **doğrudur?**

**I.** Katı maddelerin ısı alarak sıvı hâle geçmesine Erime olayı adı verilir.

**II.** Donmakta olan bir sıvının tamamı katılaşıncaya kadar sıcaklığı değişmez.

**III.** Donma ısı alarak, Erime ısı vererek gerçekleşir.

A) Sadece 2

B) 1 ve 2

C) 1 ve 3

D) 1, 2 ve 3

**15.** Maddelerin **ısı alış veriş**i ile ilgili olarak;

**I.** Fırından yeni çıkmış ekmek ısı kaybederek soğur.

**II.** Dondurucudan yeni çıkmış dondurma ısı alarak erir.

**III.** Islak çamaşırlar ısı vererek kurur.

Yukarıdaki ifadelerden hangisi ya da hangileri **yanlıştır?**

A) 1 ve 3

B) Sadece 3

C) 2 ve 3

D) 1, 2 ve 3

**16.** 1) Sıvının her sıcaklığında olur. 2) Sadece sıvının yüzeyinde gerçekleşir.

3) Sadece belli sıcaklıklarda gerçekleşir. 4) Sıvının her tarafında gerçekleşir.

5) Oluşurken sıcaklık değişmez.

6) Oluşurken sıcaklık değişebilir.

Yukarıda **kaynama** ve **buharlaşıma** olayları ile ilgili ifadeler verilmiştir. Bu ifadeler **kaynama** ve **buharlaşıma**ya ait olarak hangi seçenekte **doğru şekilde** gruplanmıştır?

**Kaynama**

**Buharlaşıma**

A) 3, 4, 5

1, 2, 6

B) 1, 2, 3

4, 5, 6

C) 3, 4, 6

1, 2, 5

D) 1, 2, 4

3, 5, 6

**17.**

**Isıtılan maddelerin hacimleri artar yani genişirler, soğutulan maddelerin hacimleri ise azalır yani büzüşürler.**

Aşağıdaki olaylardan hangisi yukarıda verilen bilgi ile **izah edilemez?**

A) Yazın elektrik tellerinin sarkması, kışın ise tellerin gerilmesi.

B) Raylar arasındaki boşlukların yazın kapanıp, kışın açılması.

C) Isıtılan içi hava dolu şişenin belli bir süre sonra patlaması.

D) Buzlukta unutulmuş su dolu şişenin, donarak çatlaması.

18. Aşağıda verilen ifadelerden hangisi **buharlaştırma** olayının bir sonucudur?

- A) Kar yağarken havanın ılık olması.  
B) Sıcaklığı artırılan suyun kaynamaya başlaması.  
C) Islak çamaşırların rüzgârlı havalarda daha hızlı kuruması.  
D) Dolaptan çıkarılan dondurmanın bir süre sonra erimesi.

19. Maddelerin hâl değişimleri ile ilgili olarak aşağıda verilen ifadelerden hangisi **doğrudur?**

- A) Maddelerin ısı alarak katı hâlden sıvı hâle geçmesi olayına donma adı verilir.  
B) Yoğuşma olayının tersi erime olayıdır.  
C) Kaynama aynı zamanda bir buharlaştırma olayıdır.  
D) Bir ortamda sıcaklığın artması yoğuşma olayını, düşmesi ise buharlaştırma olayını hızlandırır.

20. Her maddenin kendine özgü bir takım özellikleri vardır. Bu özellikler her madde için farklıdır. Bunlara **maddenin ayırt edici özellikleri** denir

Buna göre aşağıdakilerden hangisi ayırt edici **özelliklerdendir?**

- A) Renk  
B) Kütle  
C) Hacim  
D) Erime noktası

21.

Madde	Erime sıcaklığı (°C)	Kaynama sıcaklığı (°C)
Su	0	100
Alkol	- 115	78
Cıva	- 39	357

Yukarıdaki tabloda bazı saf maddelerin **erime ve kaynama sıcaklıkları** verilmiştir.

Buna göre, aşağıdaki sıcaklıkların hangisinde her üç madde de **gaz hâldedir?**

- A) 45 °C  
B) - 7 °C  
C) 90 °C  
D) 360 °C

22.

1. Termometre  
2. Yangın alarm sistemleri  
3. Konserveler  
4. Sıcak hava - seyahat balonları

Yukarıda verilen madde ve olaylardan hangisi ya da hangileri **maddelerin genleşme özelliğinden** yararlanılarak yapılmıştır?

- A) 1 ve 2  
B) 2 ve 3  
C) 1, 2 ve 3  
D) 1, 2 ve 4



25.

**Kışın bazı bölgelerimizde hava sıcaklığı – 25 °C'ye kadar inmektedir. Bu şartlarda u donduğu hâlde benzin donmaz.**

Bu durum aşağıdakilerden hangisi ile **açıklanabilir?**

- A) Benzinin daha yoğun olması
- B) Her sıvının donma sıcaklığının farklı olması
- C) Benzin ile suyun donma noktasının aynı olması
- D) Suyun daha yoğun olması

26. Havada bulunan su buharının çok soğuk ilkbahar ve sonbahar gecelerinde sıvı hale geçmeden donarak yeryüzünde ince bir kar tabakası oluşturur. Bu olaya .....denir.

- A) Dolu
- B) Kar
- C) Kırağılaşma
- D) Çiy

27. Bir bardak kolanın içerisine bir buz parçası atılmıştır. Buna göre aşağıdakilerden hangisi **yanlıştır?**

- A) Kola ısı verir.
- B) Her iki madeninde sıcaklığı artar.
- C) Buz ısı alır.
- D) Kola soğur.

28. Katı bir maddenin doğrudan doğruya **gaz haline geçmesi olayına** ne denir?

- A) Erime
- B) Yoğuşma
- C) Süblimleşme
- D) Kırağılaşma

**Cevap Anahtarı**

M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M 10	M 11	M 12	M 13	M 14
C	B	A	C	C	A	A	D	A	A	B	A	D	B
M 15	M 16	M 17	M 18	M 19	M 20	M 21	M 22	M 23	M 24	M 25	M 26	M 27	M 28
B	A	D	C	C	D	D	D	D	A	B	C	B	C