

MADDE ve ISI ÜNİTESİ BAŞARI TESTİ GELİŞTİRME ÇALIŞMASI

DEVELOPMENT of AN ACHIEVEMENT TEST ABOUT MATTER and HEAT UNIT

Ebru KAPLAN¹

Oktay BEKTAŞ²

Melek KARACA³

Başvuru Tarihi: 22.08.2021

Yayına Kabul Tarihi: 18.05.2022

DOI: 10.21764/maeuefd.985968

(Araştırma Makalesi)

Özet: Bu çalışma altıncı sınıf madde ve ısı ünitesine yönelik, geçerli ve güvenilir başarı testi geliştirmeyi amaçlamıştır. Testin geliştirilmesi süreci nicel araştırmanın tarama desenine uygun olarak yürütülmüştür. Araştırmanın örneklemini Kayseri ilinde bulunan 143 altıncı sınıf öğrencisi oluşturmaktadır. Veriler, SPSS, LISREL ve Factor programları kullanılarak analiz edilmiştir. AFA ile iki faktörlü 17 soruluk yapı belirlenmiş olup, testin KMO değeri 0,93 olarak hesaplanmış ve bu iki faktörlü yapının toplam varyansın %53'ünü açıkladığı tespit edilmiştir. Elde edilen bu yapı DFA ile doğrulanmıştır. Testin güçlük ve ayırt edicilik indekslerinin orta düzeyde olduğu tespit edilmiş olup ilişkisiz örneklemler t testi ile üst grup ile alt grup arasında anlamlı farklılıkla doğrulanmıştır. Testin güvenilirlik katsayısı 0,81 olarak hesaplanmıştır. Sonuç olarak, ortaokul öğrencilerinin “Madde ve Isı” ünitesine yönelik başarı düzeylerini belirlemek veya öğrenme eksikliklerini gidermek için hem araştırmacılar hem de öğretmenler tarafından kullanılabilir geçerli ve güvenilir bir ölçme aracı geliştirilmiştir.

Anahtar Sözcükler: ölçme ve değerlendirme, başarı testi, test geliştirme, madde ve ısı

Abstract: The study aimed to develop a valid and reliable achievement test for the sixth-grade matter and heat unit in accordance with Bloom taxonomy. The test's development process was conducted in accordance with quantitative research designs (survey) and applied to 143 sixth-grade students in Melikgazi District of Kayseri province. Data were analyzed using SPSS, LISREL, and Factor programs. A 17-question structure with two factors was determined with AFA, and the KMO value of the test was calculated as 0.93, and it was found that this two-factor structure explains 53% of the total variance. The resulting structure was verified by DFA. It was found that the difficulty and distinctiveness indices of the test were moderate, and unrelated samples proved that the upper group and the lower group differ significantly with independent sample t test. The reliability coefficient of the scores obtained from the final 17-question version of the test was calculated as 0.81. As a result, a valid and reliable measurement tool has been developed that can be used by both researchers and teachers to determine the achievement levels of Secondary School students for the “Matter and Heat Unit” or to address learning deficiencies.

Keywords: Measurement and assessment, achievement test, test development, matter and heat

Giriş

Eğitim sürecini yönlendiren öğretim programlarında ölçme ve değerlendirme konusu önemli yer tutmaktadır (MEB, 2018). Her birey birbirinden bağımsız, biricik, farklı bilgi düzeyi ve beceri eğilimlerine sahip olduğu için öğretim programları öğrencilerin ölçme ve değerlendirme faaliyetlerini, bu farklılığı en iyi bilen öğretmenlere bırakmıştır (MEB, 2018). Eğitimde öğretmenlerin öğrenci farklılıklarını gözeterik kullanabilecekleri farklı ölçme araçları mevcuttur.

¹ Millî Eğitim Bakanlığı, kplnebru38@gmail.com, ORCID: 0000-0003-2553-2209

² Doç. Dr., Erciyes Üniversitesi, Eğitim Fakültesi, obektas@erciyes.edu.tr, ORCID: 0000-0002-2562-2864

³ Sorumlu Yazar, Dr., Bağımsız Araştırmacı, melekkaraca38@gmail.com, ORCID: 0000-0002-6957-5932

Ölçme aracı geliştirmeye ilişkin alan yazın incelendiğinde klasik test ve madde tepki kuramı olmak üzere iki kuramın ön plana çıktığı görülmektedir (Çelen ve Aybek, 2013). Klasik test kuramında öğrenci başarısı her bir soruya verdiği doğru cevapların toplanması ile bulunurken, madde tepki kuramında her sorunun hesaplanan toplam puana katkısı eşit olmadığından klasik test kuramında olduğu gibi toplama işlemi yapılamaz. Klasik test kuramında tüm çeldiriciler eşit oranda etkili iken, madde tepki kuramında çeldiricilerin doğru cevaba yakınlık bakımından kendi içinde bir sıralaması vardır. Her iki kuramın da ortak varsayımları; testin normallik varsayımının karşılanması, testin ölçmek istediği özelliği ölçmesi ve bir maddeye verilen yanıtın diğer maddeleri etkilememesidir. Klasik test kuramı, madde tepki kuramına göre test geliştirmede yaygın olarak kullanılmakla beraber iki kuramın birbirine üstünlüğüne dair alan yazında net bir kabul yoktur (Çelen ve Aybek, 2013; Çıtak Gözen, 2010). Bu çalışma klasik test teorisine göre tasarlanmış, tüm çeldiricilerin eşit oranda doğru cevaptan farklılaştığı kabul edilmiştir. Klasik test teorisine göre taksonomi kavramı ortaya çıkmış olup çoktan seçmeli testlerin özellikleri belirlenmiştir.

Eğitimde kullanılan ölçme araçları, değerlendirme sürecinin başlangıç noktası kabul edildiğinden, ulaşılmak istenen hedeflerin sıralanması ihtiyacı ortaya çıkmış, böylece taksonomi kavramı alan yazına girmiştir (Bloom, vd., 1956). Bu taksonomilerden literatürde en çok kabul gören Bloom taksonomisi, alanyazına sunulduğu tarihten itibaren yarım asırdır yaygın olarak kullanılmakta olup, çok çeşitli dillere çevirisi yapılmış, soru hazırlamanın yanı sıra program geliştirme gibi farklı amaçlarla kullanılan bir taksonomidir (Tosun ve Taşkesenligil, 2011). Bloom taksonomisi öğrencilerin bilişsel düzeylerini belirlemek için uygun olan ve tercih edilen bir araçtır (Büyüköztürk, 2013).

Bloom taksonomisinde bilişsel düzeylerin altı boyutu vardır. Bunlar; bilgi, kavrama, uygulama, analiz, değerlendirme ve sentez basamaklarıdır. Bilgi basamağında yer alan bilgiler genellikle kitaplarda yer alan tanım cümleleridir ve bu bilginin öğrenci tarafından ekleme yapılmadan aynen aktarması beklenir. Kavrama basamağında ise öğrencinin bilgi basamağında aldığı bilgiyi özümsemesi, kendi cümlesi ile ifade etmesi beklenir. Öğrendiği bilgi birikimine dayanarak problem çözme ise uygulama basamağına ait becerilerdir. Analiz basamağında ise öğrencinin bir bütünün parçaları arasında ilişki kurması ve yorumlaması beklenmektedir. Bir üst basamak olan sentez basamağında çok yönlü düşünme becerisi ve yaratıcılık gereklidir. Bu aşamada öğrenciden orijinal bir ürün oluşturması beklenmektedir. Son olarak değerlendirme basamağında, ölçülen

özellik bir ölçütü karşılaştırılarak yargıya varma söz konusudur ki bu noktada eleştirel düşünme becerisi gereklidir (Birgin, 2016). Yenilenmiş Bloom taksonomisinde sentez ve değerlendirme basamaklarının yeri değiştirilerek sentez basamağı en üst düzeye çıkarılmıştır. Bilgi basamağı “Hatırlamak”, kavrama basamağı “Anlamak” olarak değiştirilmiştir (Tosun ve Taşkesenligil, 2011).

Eğitimde kullanılan geleneksel ölçme araçları; açık uçlu, çoktan seçmeli, doğru-yanlış, boşluk doldurma ve eşleştirme sorularıdır (Karip, 2012). Geleneksel ölçme araçları içerisinde eğitimde en çok kullanılan ölçme aracı çoktan seçmeli sorulardır (Demir, vd., 2016). Çoktan seçmeli sorular yaratıcılık gibi üst düzey bilişsel becerileri ölçmede yetersiz kalsa da genellikle öğrenci başarısını belirlemek veya kavram yanlışlarını tespit etmek amacıyla sıklıkla kullanılmaktadır (Haladyna, 1997). Kavram yanlışları, çoktan seçmeli soruların yanısıra görüşmelerle ya da farklı ek destek testleriyle de belirlenebilir. Alanyazında öğrencilerde var olan kavram yanlışlarını belirlemek amacıyla çoktan seçmeli test kullanan araştırmalar mevcuttur (Demirci ve Efe, 2007; Eryılmaz ve Tatlı, 2000).

Çoktan seçmeli testler; bir soru kökü ve bu soru köküne göre uygun doğru ve yanlış cevaplardan (seçenek) oluşan bir ölçme aracıdır. Geçerliği yüksek, puanlaması objektif ve uygulaması kolay olduğu için büyük grupların katılımı ile gerçekleşen merkezi sınavlar için kullanışlı bir test türüdür. Bununla beraber şans başarısı ihtimalinin varlığı ve hazırlanması için uzmanlık gerektirmesi çoktan seçmeli testlerin olumsuz özellikleridir (Baykul, 2015). Bu olumsuz özellikleri en aza indirebilmek amacıyla çoktan seçmeli testler hazırlanırken aşağıdaki özelliklere dikkat edilmelidir (Baykul, 2015):

- Testin kapsam geçerliğinin sağlanması için belirtke tablosu oluşturulmalıdır.
- Maddeler açık olmalı, yönlendirici olmamalıdır.
- Seçenekler mümkün olduğunca özdeş olmalıdır.
- Test içindeki maddeler birbirine kopya verecek şekilde olmamalıdır.
- Test bir kaynaktan olduğu gibi alınmamalıdır.
- Seçenekler birbirinden bağımsız olmalı, birbirine ipucu verilmemelidir.

Eğitimde en çok kullanılan ölçme aracı olması sebebi ile öğretmenler, uzmanlar tarafından hazırlanmış geçerli ve güvenilir çoktan seçmeli ölçme araçlarına ihtiyaç duymaktadırlar. Bu nedenle eğitim araştırmalarında çoktan seçmeli formda hazırlanmış olan başarı testlerinin

geliştirildiği birçok çalışma bulunmaktadır (Adıgüzel ve Özdoğru, 2013; Çardak ve Selvi, 2018; Elbay, 2020; Uzunöz ve Buldan, 2012; Ünal ve Çelikkaya, 2009). Benzer şekilde fen eğitimi alanında da çoktan seçmeli formda geliştirilen birçok başarı testi bulunmaktadır (Akbulut ve Çepni, 2013; Demir, Kızılay ve Bektaş, 2016; Divarcı, ve Kaya, 2019; Kargın ve Gül, 2021; Kızıkan ve Bektaş, 2018; Nacaroğlu, Bektaş ve Kızıkan, 2020; Saraç, 2018; Sontay ve Karamustafaoğlu, 2020; Üçüncü ve Sakız, 2020).

Alanyazında madde ve ısı ünitesi ile ilgili başarı testi geliştirme çalışmaları da bulunmaktadır (Altınsoy, 2011; Avcı, 2020; Çelik, 2010; İnal ve Aydın, 2015; Koç, 2014; Soylu, vd., 2020). Fen bilimleri dersinde, soyut kavram ve olaylar içermesi sebebiyle öğrencilerin anlamakta sıkıntı yaşadığı konulardan biri de “*Madde ve Isı*” ünitesidir (Avcı, 2020; İnal ve Aydın, 2015). Üniteye yer alan; tanecikli yapı, yoğunluk, ısı iletimi, ısı yalıtımı kavramlarının daha önce verilmemiş olması öğrencilerin bu soyut konuyu öğrenmelerini zorlaştırmakta ya da kavram yanlışlarına neden olmaktadır. Buna ek olarak, ilgili kavramlar bir üst öğrenim kurumunda merkezi sınav konularına, lise fizik konularına temel teşkil edecek olması bakımından anlamlı öğrenmenin gerçekleştirilmesi önem arz etmektedir.

Fen bilimleri 2018 öğretim programına göre altıncı sınıf madde ve ısı ünitesinin kapsamı değiştirilmiştir. Fen bilimleri 2013 öğretim programında altıncı sırada yer alan ünite, 2018 programında dördüncü sıraya getirilip yedi olan kazanım sayısı 13’e çıkarılarak toplam ünitelere göre ağırlık oranı %19,4 olarak hesaplanmıştır. Buradan da anlaşılacağı üzere madde ve ısı ünitesinin programda yeri ve önemi artırılmıştır (MEB, 2013; MEB, 2015). Yeni öğretim programına göre madde ve ısı ünitesine yönelik hazırlanmış başarı testi geliştirme çalışmaları incelendiğinde; üniteye ait tüm konuları kapsamadıkları (Avcı, 2020) ve soruların bilişsel düzey basamaklarında yanlışlıklar olduğu (Soylu vd., 2020) görülmüştür. Örneğin, Avcı’nın (2020) geliştirdiği başarı testinde; maddenin tanecikli yapısı, yoğunluk, yakıtlar ile ilgili kazanımlara yönelik soru sorulmadığı eksiklik olarak tespit edilmiştir. Soylu ve arkadaşlarının (2020) geliştirdikleri başarı testinde ise 15, 30, 35 ve 40. soruların sentez basamağında olduğu ifade edilmiştir. Ancak, çoktan seçmeli testlerde öğrencilerin vereceği cevaplar seçeneklerle sınırlı olduğu için sentez basamağına ait becerileri yoklamada yetersizdir (Kızılkaya ve Seven, 2017).

Bir önceki öğretim programı olan 2013 yılı fen bilimleri öğretim programında maddelerin ısı iletimi, ısı yalıtımı ve yakıtlar ile ilgili kazanımlar yer almakta olup, güncel programda yer alan

maddenin tanecikli yapısı ve yoğunluk kavramı ile ilgili kazanımlar yer almamaktadır. Dolayısıyla 2018 yılı öncesinde üniteye yönelik olarak geliştirilen testler kapsam bakımından yetersiz kalmaktadır. Tüm bu sebeplerle; 2018 öğretim programında kazanım sayısı ve önemi artırılan altıncı sınıf madde ve ısı ünitesinin tüm konu ve kazanımlarına yönelik, Bloom taksonomisine uygun bilişsel düzeylerde hazırlanmış ve üst düzey becerileri yoklamaya uygun, araştırmacılar ve öğretmenlerin kullanımına elverişli çoktan seçmeli bir ölçme aracı geliştirilmesi gerekli görülmüştür. Bu araştırmanın amacı, fen bilimleri dersi öğretim programı kapsamında, altıncı sınıf “*Madde ve Isı*” ünitesine yönelik geçerlik ve güvenilirlik kontrolleri yapılmış bir başarı testi geliştirmektir. Bu kapsamda, aşağıdaki alt problemlere yanıt aranacaktır;

1. Altıncı sınıf “*Madde ve Isı*” ünitesine yönelik olarak hazırlanmış başarı testi geçerli midir?
2. Altıncı sınıf “*Madde ve Isı*” ünitesine yönelik olarak hazırlanmış başarı testi güvenilir midir?

Yöntem

Çalışma Deseni

Ortaokul altıncı sınıf madde ve ısı ünitesine yönelik başarı testi geliştirmenin amaçlandığı bu çalışmada nicel araştırma desenlerinden kesitsel tarama tercih edilmiştir. Bir olgu ya da olaya açıklık getirmek ya da var olan bir sorunun çözümlenmesine ilişkin gerekli olan verilerin bilimsel araştırma yöntemlerine uygun olarak toplanıp sonuçların sayı ya da sembollerle ifade edilmesi nicel araştırma yöntemi olarak tanımlanmaktadır (Büyüköztürk, 2013). Tarama çalışmalarının temel amacı; geniş gruplar üzerinden, durumlarla ve olaylarla ilgili olarak nicel veriler ve çeşitli istatistikler üretmektir. Böylece olay ve olgular betimlenmeye çalışılır (Creswell, 2009). Kesitsel tarama çalışmalarında değişkenler tek seferde ölçülür (Büyüköztürk, 2013). Bu sebeple başarı testi geliştirme çalışması için tarama deseni tercih edilmiştir.

Çalışma Grubu

Bu çalışmanın hedef evrenini Kayseri ilinde bulunan ortaokul öğrencileri, ulaşılabilir evrenini Kayseri ili Melikgazi ilçesindeki dördüncü bölge ortaokul öğrencileri, örneklemini ise 143 ortaokul öğrencisi oluşturmaktadır. Ulaşım zorluğu, uygulama kolaylığı, maliyeti düşürme gibi çeşitli sebeplerle evrene uygulama yapmak zordur bu yüzden örneklem oluşturulur. Örneklem oluşturulurken ulaşılabilir evrenin %10’u, testteki madde sayısının 10 katı ya da G-power analizi

sonucunda ortaya çıkan örneklem sayısı baz alınır (Pallant, 2020). Bu çalışmada soru sayısının 10 katı katılımcıya ulaşmak hedeflenmiş olup 143 katılımcıya ulaşılmıştır.

Araştırma örnekleminin belirlenmesi tarama araştırmalarının en önemli özelliklerinden birisi olması sebebi ile örneklem seçilirken evreni iyi temsil edecek örneklem oluşturulmaya çalışılmıştır. Araştırma amacına uymayan ya da evreni temsil etmeyen örneklem seçimi araştırmaya önemli ölçüde zarar verebilir (Büyüköztürk, 2013). Örnekleme oluşturan 143 öğrenci, ulaşılabilir evrenden basit (seçkisiz) örnekleme yoluyla seçilmiştir. Basit seçkisiz örnekleme evrenden örneklem için seçim yapma işlemi seçkisizlik ilkesine uygun olarak yapılmaktadır. Burada kastedilen seçkisizlik evrenden alınan birimlerin örneklem olarak seçilme olasılıklarının eşit olmasıdır. İlave olarak örnekleme seçilebilecek birimlerin seçilme durumları birbirinden bağımsızdır (Büyüköztürk, 2013). Veriler online olarak toplanacağından örneklem içerisinde yer alacak kişilerle doğrudan iletişim kurma imkanı sınırlı olduğu için bu örnekleme türü kullanılmıştır.

Veri Toplama Araçları

Bu çalışmada veri toplama aracı olarak araştırmacı tarafından; alan yazın taraması (Avcı, 2020; İnal ve Aydın, 2015; Soylu vd, 2020) ve Milli Eğitim Bakanlığı (MEB) ders kitapları ve MEB kazanım testlerinde yer alan 62 soru arasından kazanımlara uygun olarak seçilen 25 soruluk (Tablo1) “*Madde ve Isı Taslak Başarı Testi*” (MITBT) kullanılmıştır. Test geliştirme süreci test geliştirme basamaklarına (testin amacının belirlenmesi; testin kapsamının belirlenmesi, soru tipi, sayı ve sürenin belirlenmesi; testin geçerliği, pilot uygulama, güvenilirlik ve madde analizi) uygun olarak yürütülmüştür (Güler, 2011). Bu amaçla geliştirilen testin amacı belirlenmiştir. Başarı testi fen bilimleri dersi kapsamında, altıncı sınıf “*Madde ve Isı*” ünitesine yönelik olarak, öğrencilerin öğrenme düzeyine uygun, konu sonunda düzey belirleme (Summative) amaçlı kullanılacak şekilde öğretmenlerin ve diğer araştırmacıların faydalanması amacı ile hazırlanmıştır. Amaç belirlendikten sonra testin kapsamı belirlenmiştir. Fen bilimleri dersi 2018 programı incelenmiş, üniteye ait konu ve kazanımlardan faydalanılmış, konuların tümünden sorulara yer verilmiştir. Soruların hazırlanması ve geliştirilmesi sürecinde fen bilimleri altıncı sınıf ders kitabında (Demirçalı ve Alkan, 2009) yer alan ünite sonu değerlendirme sınavı soruları ve MEB tarafından yayımlanan kazanım testleri incelenmiştir. Ünite kazanımlarına uygun olacak şekilde alan yazında (Avcı, 2020; İnal ve Aydın, 2015; Soylu vd, 2020) yer alan sorulardan da seçme işlemi gerçekleştirilmiştir. Bloom taksonomisine uygun olacak şekilde her basamaktan soru sorulmasına

ve daha çok üst düzey becerilerin yoklanmasına dikkat edilmiştir. Madde ve ısı başarı testi oluşturulurken yararlanılan kaynaklar, uzmanlardan gelen dönütler ve yapılan değişikliklere Tablo 1’de yer verilmiştir.

Tablo 1

Uzman dönütleri ve Madde ve ısı başarı testinde yapılan düzeltmeler

Soru numarası	Yararlanılan kaynak	Uzman dönütü	Yapılan düzenleme/ sorunun son hali
1, 2, 3	Soylu, vd, 2020	Seçenekler düzenlenmeli, “en iyi” gibi ifadeler vurgulanmalı	Seçenekler kıstadan uzuna doğru sıralanıp, vurgu ifadeleri koyu ve altı çizili hale getirilmiştir.
2	Soylu, vd, 2020	Kazanım soru kökü ile uyumlu değil, değiştirilmeli	Maddelerin; “tanecikli, boşluklu ve hareketli yapıda olduğuna karar verir” kazanımı “Grafikteki bilgilerden yola çıkarak gaz halindeki bir maddenin taneciklerinin titreşim, öteleme ve dönme hareketi yapacağını yorumlar” şeklinde değiştirilmiştir.
3	Soylu, vd, 2020	Kazanımda yer alan deney yaparak ifadesi çıkarılmalıdır.	Sorunun kazanımından “deney yaparak” ifadesi çıkarılmıştır.
4	MEB ders kitabı (Demirçalı ve Alkan, 2009)	Tüm halleri ifadesi açıklanmalı ve seçenekler sıralanmalı.	Soru köküne “tüm halleri (katı, sıvı, gaz)” ifadesi eklenmiştir. Seçenekler kıstadan uzuna doğru sıralanmıştır.
5	Soylu, vd, 2020	Soruda kökü ve seçeneklerde yer alan gereksiz ifadeler çıkarılmalı Görsel daha anlaşılır hale getirilmeli	Soru kökünde yer alan “katı haldeki buz” ifadesinde “katı haldeki” kısmı çıkarılmıştır. Seçeneklerden “Buz gaz haline geçince tanecikler arasındaki boşluk en fazladır” ifadesi “Buz gaz haline geçtiğinde tanecikler arasındaki boşluk en fazladır” olarak değiştirilmiştir. Sorunun görseli değiştirilmiştir.
6	Soylu, vd, 2020	Şık ifadesi değiştirilmeli	“Hangi şıkta” ifadesi “hangi seçenekte” olarak değiştirilmiştir. Seçenekler küçükten büyüğe olacak şekilde değiştirilmiştir.
8, 12, 15	Soylu, vd, 2020	Alıntı yapılan soru numaraları çıkartılmalıdır.	Alıntı yapılan soru numaraları görsellerin üzerinden silinmiştir.
9	Soylu, vd, 2020	Düzeyi tekrar değerlendirilmeli	Soru bilgi seviyesinden uygulama seviyesine çıkarılmıştır.
10	Soylu, vd, 2020	Maddelerin hangi halde bulunduğu soru köküne eklenmeli	Soru köküne “sıvı halde bulunan” ibaresi eklenmiştir.
11, 12	Soylu, vd, 2020	Sorular kavrama düzeyi olmalı.	Uygulama düzeyinde kalması gerektiğine karar verilmiştir.
13, 20	Soylu, vd, 2020	Görseller daha anlaşılır hale getirilmeli. Sorunun düzeyi tekrar değerlendirilmeli.	Sorunun değerlendirme basamağına uygun olacak şekilde, tüm seçeneklerin doğru cevap olabilmesi sağlanmıştır. Bunun için görseller tüm tavaların kulpları ısı yalıtkanı, gövdeleri ısı iletkeni olarak değiştirilmiştir.
14	Soylu, vd, 2020	Seçeneklerde daha yavaş, daha hızlı gibi ifadeler var, fakat ne ile	“Çelik ısıyı daha yavaş iletmiştir.” seçeneği “Çelik ısıyı tahtaya göre daha

		kıyaslandığında bu durum söz konusu net değil.	yavaş iletmiştir.” şeklinde tüm seçeneklerde ne ile kıyaslandığı belirtilerek değiştirilmiştir.
15	İnal ve Aydın, 2015	Verilere dayalı çıkarımda bulunması isteniyor ancak analiz düzeyi olarak belirlenmiş, uygun değil.	Analiz düzeyinde kalmasına karar verilmiştir.
16	Avcı, 2020	-	-
19	Soylu, vd, 2020	Soru maddesi yazılmamış	Soru kökü eklenmiş ve tanılayıcı dallanmış ağaçla ilgili bilgi verilmiştir
23, 24	Soylu, vd, 2020	Soru maddesi yazarken bütünlük olmalı, soru kökünün altını çizme, koyu yazma vb. Örnek olarak 23 ve 24. soruyu karşılaştırabilirsin.	Bütünlük sağlanmıştır.
24	Soylu, vd, 2020	-	-
25	MEB kazanım testi	Verilen özellikler olarak açıklama yapılmış, ancak sadece tanecik modeli verilmiş, oysaki ifade çoğul. Bu durum düzeltilmelidir.	Soru kökü “Buna göre kırmızı yüzeylerin aşağıda tanecik modelleri verilen maddelerden hangisi ile kaplanması en uygundur?” olarak değiştirilmiş, özellikleri verilen ifadesi çıkarılmıştır.

Geçerlik ve Güvenirlik

Geçerlik ve güvenilirlik çalışmaları bilimsel araştırmalar için oldukça önemli bir yere sahiptir. Yeterli örneklem sağlandıktan sonra ulaşılabilir evrene genelleme ihtimalini artıran en önemli kısım burasıdır.

Ölçme aracının doğru sonuçlar verebilmesi, farklı kişilerce farklı örneklemeler üzerinde yürütüldüğü zaman benzer sonuçlar vermesi istenir. Bu durumun sağlanması o ölçme aracının güvenilir bir ölçme aracı olduğu anlamına gelmektedir. Ölçme aracının güvenilir olup olmadığına her bir test maddesi ile yürütülen güvenilirlik analizi sonucunda karar verilebilir. Toplam puanlar üzerinden güvenilirlik analizi yürütülemez (Pallant, 2020). Farklı güvenilirlik belirleme yöntemleri vardır. Örneğin; KR-20, KR-21, Cronbach Alpha (iç tutarlılık katsayısı), Spearman Brown eşdeğer yarılar güvenirligi, paralel formlar güvenirligi bu yöntemlerden bazılarıdır (Tanrıoğen, 2009). Cronbach alpha, ölçeğin güvenilirliğini araştırmak için kullanabilecek kullanışlı ve esnek bir araçtır (Seçer, 2013). Güvenirlik katsayısının 0,70 ve üzerinde olması ölçme aracında yer alan maddelerin güvenilir olduğu anlamına gelmektedir (Pallant, 2020). Bu çalışma test geliştirme çalışması olduğu için KR-20 güvenilirlik katsayısı hesaplanmıştır. Güvenirlik analizi ve faktör analizi sonucunda maddeler çıkarıldıktan sonra kalan maddelerin güvenilirlik katsayısı yeniden hesaplanmıştır.

Geçerlik kavramı ölçülmek istenen davranışın başka özelliklerle karıştırılmadan, doğru bir şekilde ölçülmesi iken, güvenilirlik kavramı ise ardı ardına yapılan ölçümlerin tutarlı olması anlamına gelmektedir (Seçer, 2013).

Kapsam Geçerliği. Oluşturulan testin kapsam geçerliğini sağlamak için belirtke tablosu oluşturulmuştur (Tablo 2). Oluşturulan belirtke tablosu Bloom taksonomisine uygun olacak şekilde hazırlanmış olup her basamaktan soru sorulmuş ve üst düzey basamaklara daha çok ağırlık verilmiştir. Testin kapsam geçerliğini sağlamak açısından fen eğitimi ve kimya eğitimi alanından iki uzmanın görüşleri alınmıştır. Fen eğitimi uzmanının Zoom programı aracılığı ile, kimya eğitimi uzmanının mail yolu ile görüşleri alınmış olup düzenlemeler yapılmıştır. Birinci uzmanın incelemesi sonucunda alanyazından alınan soruların bilişsel düzeyinde yanlışlıklar olduğu yani analiz düzeyinde olan bir sorunun kavrama düzeyinde yazılmış olduğu görülmüş ve uygun basamağa alınmıştır. Benzer olarak alanyazından alınan bazı soruların yazılan kazanımlarla uyumlu olmadığı görülüp yeniden düzenlenmiştir. Görünüş geçerliğini sağlamak adına tüm seçenekler kısıdan uzuna doğru sıralanmış olup, olumsuz ifadeler koyu yazılarak altı çizilmiştir, testin tamamında seçenekler aynı formatta harflendirilmiştir. Soru kökünde yer alan fiiller, sorunun ait olduğu bilişsel düzeye uygun hale getirilmiştir. Tablo 1’de yapılan düzeltmelere yer verilmiştir. Uzman dönütleri sonucunda oluşturulan test bir Türkçe uzmanının incelemesinden geçirilerek dil ve anlatım bakımından uzman kontrolü sağlanmıştır. Ayrıca oluşturulan test bir fen bilgisi öğretmenine incelenilerek altıncı sınıf seviyesine uygun olduğundan emin olunmuştur. Aşağıda geliştirilen testin belirtke tablosu verilmiştir;

Tablo 2

Geliştirilen taslak başarı testinin belirtke tablosu

MEB Kazanımı	Soru No	Yazarlar tarafından oluşturulan kazanım	Bilgi	Kavrama	Uygulama	Analiz	Değerlendirme
F.6.4.1.1. Maddelerin; tanecikli, boşluklu ve hareketli yapıda olduğunu ifade eder.	1	Maddelerin; tanecikli, boşluklu ve hareketli yapıda olduğuna karar verir					+
	4	Maddenin katı, sıvı ve gaz halinin taneciklerden oluştuğunu bilir.	+				
F.6.4.1.2. Hâl değişimine bağlı olarak maddenin tanecikleri arasındaki boşluk ve taneciklerin hareketliliğinin değiştiğini deney yaparak karşılaştırır.	2	Grafikteki bilgilerden yola çıkarak gaz halindeki bir maddenin taneciklerinin titreşim, öteleme ve dönme hareketi yapacağını yorumlar					+
	3	Hâl değişimine bağlı olarak maddenin tanecikleri arasındaki boşluk ve taneciklerin hareketliliğinin değiştiğini karşılaştırır					+
	5	Suyun hal değişimi deneyinden hareketle taneciklerin durumunu yorumlar					+
F.6.4.2.2. Tasarladığı deneyler sonucunda çeşitli maddelerin yoğunluklarını hesaplar.	6	Maddelerin yoğunluğunu hesaplar			+		
	7	Belirli şekli olmayan katının yoğunluğunu hesaplar			+		
	8	Grafikteki verilere göre maddelerin yoğunluğunu hesaplar			+		
	9	Yoğunluğun ayırt edici özellik olduğunu günlük yaşamında uygular			+		
F.6.4.2.3. Birbiri içinde çözünmeyen sıvıların yoğunluklarını deney yaparak karşılaştırır.	10	Birbiri içinde çözünmeyen sıvıların yoğunluklarını deney yaparak karşılaştırır					+
F.6.4.3.1. Maddeleri, ısı iletimi bakımından sınıflandırır.	11 ve 12	Maddeleri, ısı iletimi bakımından sınıflandırır			+		
	13	Maddelerin ısı iletkenliğinin farklı olduğunu gözlemler					+
	14	Maddelerin ısı iletkenliğinin farklı olduğunu yorumlar					+
	15	Isı iletim grafiğinden yola çıkarak iletken ve yalıtkan maddelerin günlük yaşamdaki kullanımı alanlarını yorumlar					+
	16	Isı yalıtım malzemelerinin özelliklerini bilir	+				
F.6.4.3.2. Binalarda kullanılan ısı yalıtım malzemelerinin seçilme ölçütlerini belirler.	17	Isı yalıtım malzemelerinin özelliklerini kavrar		+			
	18	Isı yalıtımının yapılma amacını kavrar		+			
	21	Isı yalıtım malzemeleri arasından en uygun olanı seçer					+
F.6.4.3.4. Binalarda ısı yalıtımının önemini, aile ve ülke ekonomisi ve	25	Tanecik modeli verilen ısı yalıtım malzemelerinden en uygun olanı seçer					+

kaynakların etkili kullanımı bakımından tartışır.				
F.6.4.4.1. Yakıtları, katı, sıvı ve gaz yakıtlar olarak sınıflandırıp yaygın şekilde kullanılan yakıtlara örnekler verir. <i>Fosil yakıtların sınırlı olduğu ve yenilenemez enerji kaynaklarından biri olduğu belirtilir ve yenilenebilir enerji kaynaklarının önemi örnekler verilerek vurgulanır.</i>	20	Yenilenebilir enerji kaynaklarının ortak özelliğini açıklar		+
F.6.4.4.2. Farklı türdeki yakıtların ısı amaçlı kullanımının, insan ve çevre üzerine etkilerini tartışır.	23	Katı, sıvı ve gaz yakıtlara örnek verir		+
	24	Petrolden üretilen sıvı yakıtları bilir		+
F.6.4.4.3. Soba ve doğal gaz zehirlenmeleri ile ilgili alınması gereken tedbirleri araştırır ve rapor eder.	19	Doğal gaz zehirlenmelerine karşı alınması gereken tedbirleri kavrar		+
	22	Soba zehirlenmeleri ile ilgili alınması gereken tedbirleri analiz eder		+
	Toplam		3	5
			6	7
			4	

MEB öğretim programında bu üniteye ait 13 kazanım yer almaktadır. Yürütülen bu çalışmada bazı kazanımlar programdan aynen alınmış, bazı kazanımlar değiştirilmiş, bazı kazanımlar ise sentez basamağına ait beceriler gerektirdiği ve hazırlanan testle yoklanamayacağı için dahil edilmemiştir. Örneğin; “*Yoğunluğu tanımlar*” kazanımı “*Maddelerin yoğunluğunu hesaplar*” olarak genişletilmiş olup kazanıma uygun bilişsel seviyede sorular hazırlanmıştır. Öğretim programında yer alan “*Hâl değişimine bağlı olarak maddenin tanecikleri arasındaki boşluk ve taneciklerin hareketliliğinin değiştiğini deney yaparak karşılaştırır.*” kazanımının yerine “*Grafikteki bilgilerden yola çıkarak gaz halindeki bir maddenin taneciklerinin titreşim, öteleme ve dönme hareketi yapacağını yorumlar*” ve “*Hâl değişimine bağlı olarak maddenin tanecikleri arasındaki boşluk ve taneciklerin hareketliliğinin değiştiğini karşılaştırır*” kazanımları yazılmıştır. MEB öğretim programı kazanımından hareketle aynı beceriyi yoklayan birden fazla soru yazılmış, dolayısıyla sorunun testten atılması durumunda beceri yoklanmasının devamlılığı sağlanmaya çalışılmıştır. Sadece 11 ve 12. soruların kazanımları birbirinin aynısıdır. MEB öğretim programında yer alan “*F.6.4.3.3. Alternatif ısı yalıtım malzemeleri geliştirir.*” kazanımı geliştirilen çokta seçmeli testte yoklanamayacağı için yer almamıştır. Bu sebeple bu çalışmada belirlenen 24 kazanımın yoklanması hedeflenmiştir.

Yapı Geçerliliği. Bilimde yapı; nesneye ait özellikler, bu özellikleri ifade eden kavramlar ve bunlar arasında var olan ilişkileri içinde barındıran sistemlerdir (Tanrıoğen, 2009). Ölçme aracının yapısının belirlenmesinde faktör analizi en çok kullanılan tekniktir. Faktör analizi sayesinde ölçme

aracının ölçtüğü düşünülen niteliği ölçüp ölçmediğine ilişkin fikir sahibi olunur (Baykul, 2015). Faktör analizi, açımlayıcı faktör analizi ve doğrulayıcı faktör analizi olmak üzere iki çeşittir. Ölçekler eşit aralıklı, başarı testleri ise sıralı yapıdadır. Bu nedenle başarı testleri klasik test teorisine göre doğru cevaba “1” yanlış cevaba “0” puan verilerek kodlanan sıralı ölçekler olduğu için faktör yapısının belirlenmesinde ölçeklerden farklı olarak tetrakorik korelasyon analizi yapılması uygundur. Analiz Lorenzo-Seva ve Ferrando (2006) tarafından geliştirilen “Factor” programı ile gerçekleştirilmiştir. Bu program hem AFA hem de DFA sonuçlarını vermektedir. AFA kapsamında KMO değeri ve Barlett’s testinin sig değerini vermektedir. Faktör analizi sonucunda elde edilen KMO değeri hem örneklem yeterliğinin sağlandığının hem de normal dağılımın göstergesidir. KMO değerinin 0,60 ve üzerinde olması gerekmektedir (Seçer, 2013). KMO tablosunda bulunan sig. değerinin 0,05’ten küçük olması eldeki veri setinin faktör analizi yürütmek için uygun olduğu, verilerin anlamlı olduğu anlamına gelmektedir. Bu iki şart sağlanınca faktör analizi verileri yorumlanabilmektedir. Faktörlerin Eigenvalues değerinin “1” ve üzerinde olması ayrıca tüm faktörlerin toplam varyansın en az %40’ını açıklaması beklenmektedir (Seçer, 2013). DFA kapsamında ise Ki kare/ Serbestlik derecesi ve diğer uyum indeksi değerleri hesaplanmaktadır. AFA sonucunda elde edilen faktör yapısının DFA’da doğrulanabilmesi için “Ki kare/ Serbestlik derecesi” değerinin 5’ten küçük olması ve faktör yapısının toplam varyansın en az %40’ını açıklaması beklenir (Seçer, 2013).

Güçlük ve Ayırt edicilik İndeksleri. Ölçme aracında yer alan soruların güçlük ve ayırt edicilik indeksleri hesaplanarak soruların zorluk seviyesi ve bilen öğrenciler ile bilmeyen öğrencileri birbirinden ne oranda ayırt edebildiği ortaya çıkarılmaya çalışılır. Böylece ölçme aracının geçerliği ve güvenilirliği artırılır. Güçlük ve ayırt edicilik indeksleri her bir soru için ayrı ayrı hesaplanır. Katılımcıların üst grup ve alt grup olarak belirlenmesi gerekmektedir. Bunun için katılımcılar aldıkları puanlara göre sıralanır, en yüksek puanları alan katılımcıların %27’si üst grup, en düşük puanları alan katılımcıların %27’si alt grup olarak adlandırılır (Beuchertand ve Mendoza, 1979). Bu çalışmada katılımcı sayısı 143 olup ilk 39 katılımcı üst grup, son 39 katılımcı alt grup olarak belirlenmiştir.

Madde güçlük indeksi (p); “0” ile “1” arasında değer alır ve bu değer 1’e yakın olması soruların kolay, 0’a yakın olması soruların zor olduğu anlamına gelmektedir. Soruların orta güçlükte olması idealdir. Ölçme aracında yer alan maddelerin güçlük indeksi değerlerinin 0,3 ile 0,7 arasında olması

istenir (Büyüköztürk, 2013). Güçlük indeksi üst gruptan soruyu doğru cevaplayan katılımcı sayısı ile alt gruptan soruyu doğru cevaplayan katılımcı sayısının toplamının üst ve alt gruptaki toplam katılımcı sayısına bölünmesi ile hesaplanır.

Madde ayırt edicilik indeksi (r); “-1” ile “+1” arasında değer alır ve bu değer 1’e yakın olması sorunun ayırt ediciliğinin arttığını gösterir. Ayırt edicilik indeksinin negatif olması sorunun ters ayırt edici olduğu yani alt grupta yer alan öğrencilerin üst grupta yer alan öğrencilere göre daha başarılı olduğu anlamına gelir ki bu da sorunun çalışmadığını gösterir. Ayırt edicilik indeksinin 0,30 ve üzerinde olması gerekmektedir, 0,30’un altında olan sorular düzeltilmelidir, 0,20’nin altında ayırt edicilik düzeyine sahip olan ve negatif olan sorular ölçme aracından atılmalıdır (Büyüköztürk, 2013). Ayırt edicilik indeksi üst gruptan soruyu doğru cevaplayan katılımcı sayısından alt gruptan soruyu doğru cevaplayan katılımcı sayısının çıkarılarak üst veya alt grupta yer alan katılımcı sayısına bölünmesi ile hesaplanır.

İlişkisiz örneklem t testi. Farklı iki gruba uygulanan testin gruplar arasında fark oluşturup oluşturmadığını incelemek amacıyla ilişkisiz örneklem t testi analizi kullanılır (Pallant, 2020). Ölçme aracında yer alan soruların üst grupta ve alt grupta bulunan katılımcıları ne oranda birbirinden ayırt edebildiğini ortaya çıkarmak için ya da bir başka deyişle her bir soru için üst grupta yer alan katılımcılarla alt grupta yer alan katılımcılar arasında o sorudan aldıkları puanlar bakımından anlamlı fark olup olmadığını anlamak için ilişkisiz örneklem t testi analizi yapılmıştır. t testlerine başlamadan önce sağlanması gereken varsayımlar vardır,

- Örneklem yeterliği sağlanmalıdır.
- Normal dağılım sağlanmalıdır.
- Varyanslar homojen olmalıdır.

Varyansların homojenliğine analiz sonucunda hesaplanan Levene’s Testinin sig. değerine bakılarak karar verilir. Eğer sig. değeri 0,05’ten büyükse varyanslar homojendir. Elde edilen t testi kısmının sig değeri 0,05’ten küçükse gruplar arasında anlamlı fark vardır demektir (Pallant, 2020).

Veri Toplama Süreci. Pandemi sebebi ile doğrudan öğrencilere ulaşma imkanı olmadığından hazırlanan taslak başarı testi araştırmacı tarafından “*Google Formlar*” uygulamasına aktarılıp, öğrencilere çevrim içi ortamlarla ulaştırılmıştır. Formun ilk kısmına “*Gönüllü Oluru*”

bölümü eklenip katılımcıların bu çalışmaya gönüllü olarak katılımları sağlanmaya çalışılmıştır. Başarı testinde yer alan soruların boş bırakılmaması adına tüm soruların işaretlenmesi zorunlu hale getirilmiştir. Başarı testinin uygulama basamağına geçilmeden önce gerekli izin başvuruları yapılmış olup 297 başvuru numarası ile Erciyes üniversitesi sosyal ve beşeri bilimler etik kurulundan 29.06.2021 tarihinde onay alınmıştır. Veri toplama süreci yaklaşık olarak üç ay sürmüş, 20.06.2021 tarihinde tamamlanmış ve 143 katılımcıya ulaşılmıştır. Google formlardan “*Microsoft Office Excel*” formatında alınan veriler SPSS 26 paket programına aktarılmış ve analize hazır hale getirilmiştir.

Verilerin Analizi. Google formlardan “Microsoft Office Excel” formatında alınan veriler doğru cevaba “1” yanlış cevaba “0” kodlanarak SPSS 26 paket programına aktarılmıştır. Katılımcıların testten aldıkları puanlar toplanarak “total” değişkeni oluşturulmuştur. Total değişkeninde bulunan puanlar büyükten küçüğe sıralanarak 143 katılımcıdan 39’ar kişilik üst grup ve alt grup belirlenmiş ve üst grup “2” alt grup “1” olarak kodlanmıştır. Her bir soru için normal dağılım analizi, indeks hesaplamaları, güvenilirlik katsayısı hesaplama, AFA, DFA hesaplanmıştır.

Testte yer alan soruların normal dağılım gösterip göstermediğine mod, ortanca ve aritmetik ortalama değerlerinin birbirine yakın olup olmamasına, standart sapma değerinin büyüklüğüne, basıklık- çarpıklık değerlerinin -2 ile +2 arasında olmasına, histogram grafiğinde puanların çoğunun eğrinin altında kalmasına ve Q-Q Plot grafiğinde noktaların çizginin üzerinde olmasına bakılarak karar verilmiştir (Pallant, 2020; Seçer, 2013).

Faktör analizini gerçekleştirmek için katılımcıların testten aldıkları puanların “1” ve “0” olarak kodlandığı Excel dosyası oluşturulmuş olup Factor 11.02.04 programına içerik aktarılmıştır. Programda “Read Data” kısmına kişi sayısı ve değişken (soru) sayısı bilgileri girilerek “Single group analysis” sekmesi işaretlenmiştir. “Configure Analysis” sekmesinden “Polychoric corelations” işaretlenmiş olup faktör sayısı girilerek döndürme türü olarak faktörlerin ilişkili olduğu düşünüldüğünden “Direkt oblimun” seçilmiş “Compute” seçeneğinden analiz yapılmıştır.

Bulgular

Normal Dağılım Analizi Bulguları

Taslak testte yer alan soruların normal dağılım gösterip göstermediği incelenmiştir. Her bir soru için ortalama, ortanca ve mod değerlerinin birbirine yakın olduğu, standart sapma ve varyans değerlerinin de istenilen aralıkta olduğu görülmüştür. Basıklık ve çarpıklık değerlerinin ise -2 ile +2 değerleri arasında olduğu görülmüştür. Test sorularından 2, 3, 5, 7, 8, 9, 12, 22, 23 ve 24. soruların basıklık veya çarpıklık değerleri bu aralıkların dışında yer aldığı için normal dağılım gösterdiklerine histogram ve Q- Q plot grafikleri incelenerek karar verilmiştir. Ayrıca faktör analizi sırasında 5. ve 7. soruların analizden çıkarılmasına karar verilmiştir.

Madde İndekslerine Ait Bulgular

Test soruları için üst grup ve alt grupta bulunan öğrencilerin doğru cevap sayıları kullanılarak güçlük ve ayırt edicilik indeksleri hesaplanmıştır. Elde edilen bulgular aşağıda sunulmuştur.

Tablo 3

Taslak test sorularının indeks değerleri

Soru	Üst Grup Doğru Cevap Sayısı	Alt Grup Doğru Cevap Sayısı	Güçlük İndeksi	Ayırt Edicilik İndeksi
1	22	6	0,36	0,41
2	36	8	0,56	0,72
3	30	12	0,54	0,46
4	38	16	0,69	0,56
5	28	11	0,50	0,44
6	32	8	0,51	0,61
7	19	1	0,26	0,46
8	37	14	0,65	0,59
9	34	13	0,60	0,54
10	25	8	0,42	0,44
11	18	6	0,31	0,31
12	34	10	0,56	0,62
13	28	6	0,44	0,57
14	39	14	0,68	0,65
15	28	9	0,47	0,49
16	34	4	0,49	0,77
17	23	5	0,36	0,46
18	28	8	0,46	0,51
19	37	11	0,62	0,67
20	35	11	0,59	0,61
21	36	11	0,60	0,64
22	29	9	0,49	0,51
23	29	8	0,47	0,54
24	34	9	0,55	0,64
25	18	10	0,36	0,21

Tablo 3 incelendiğinde yedinci sorunun güçlük indeksi ve 25. sorunun ayırt edicilik değerinin 0,30'un altında olduğu istenilen aralıkların dışına çıktığı görülmüştür. Bu sorular faktör analizi esnasında taslak başarı testinden çıkarılmıştır. Bu iki sorunun çıkarılmasına hem indeks değerleri hem de faktör analizi sonuçları göz önünde bulundurularak karar verilmiştir. Testin ortalama

güçlük değeri 0,50 ve ortalama ayırt edicilik değeri 0,54 olarak hesaplanmış olup bu değerlerin 0,50 dolaylarında olması istenen bir durumdur.

İlişkisiz Örneklem t Testi Bulguları

Testin ayırt ediciliğini belirlemek amacı ile ayırt edicilik indeksi hesaplamalarına ek olarak ilişkisiz örneklem t testi analizi yapılmıştır. Elde edilen bulgular aşağıda sunulmuştur.

Tablo 4

İlişkisiz örneklem t testi sonucu

F	Sig.	t	df	Sig. (2-tailed)	
S1	30,635	,000	-4,124	76	,000
			-4,124	69,427	,000
S2	11,974	,001	-9,148	76	,000
			-9,148	65,821	,000
S3	2,309	,133	-4,553	76	,000
			-4,553	75,377	,000
S4	236,286	,000	-6,731	76	,000
			-6,731	45,765	,000
S5	,000	1,000	-4,222	76	,000
			-4,222	76,000	,000
S6	,323	,572	-6,810	76	,000
			-6,810	75,804	,000
S7	43,235	,000	-4,587	76	,000
			-4,587	41,447	,000
S8	87,058	,000	-6,886	76	,000
			-6,886	53,381	,000
S9	21,440	,000	-5,744	76	,000
			-5,744	68,507	,000
S10	9,113	,003	-4,285	76	,000
			-4,285	73,851	,000
S11	33,333	,000	-3,082	76	,003
			-3,082	69,238	,003
S12	8,838	,004	-6,898	76	,000
			-6,898	71,158	,000
S13	7,882	,006	-6,029	76	,000
			-6,029	72,570	,000
S14	439,669	,000	-8,238	76	,000
			-8,238	38,000	,000
S15	1,055	,308	-4,872	76	,000
			-4,872	75,673	,000
S16	,493	,485	-10,503	76	,000
			-10,503	75,295	,000

S17	37,012	,000	-4,784	76	,000
			-4,784	66,934	,000
S18	2,485	,119	-5,229	76	,000
			-5,229	75,125	,000
S19	46,325	,000	-8,200	76	,000
			-8,200	55,263	,000
S20	19,190	,000	-6,990	76	,000
			-6,990	66,630	,000
S21	29,425	,000	-7,556	76	,000
			-7,556	61,731	,000
S22	,272	,604	-5,210	76	,000
			-5,210	75,903	,000
S23	1,137	,290	-5,581	76	,000
			-5,581	75,540	,000
S24	5,801	,018	-7,347	76	,000
			-7,347	72,267	,000
S25	10,890	,001	-1,908	76	,060
			-1,908	74,704	,060

İlişkisiz örneklem t testi ilgili soru için üst ve alt grup arasında anlamlı fark olup olmadığını anlamak amacıyla yürütülür. Bu analiz sonucunda ilgili soru için üst grup lehine anlamlı bir fark çıkar ise o sorunun ayırt edici olduğu kabul edilir. Dolayısıyla, Tablo 4’te de görüldüğü gibi her soru için ilişkisiz örneklem t testi yürütülmüş ve soruların ayırt ediciliği incelenmiştir. Bu amaçla ilk önce Levene testine bakılarak ilişkisiz örneklem t testinin uygulanıp uygulanılmayacağına karar verilmiştir. Tablo 4’e göre testte yer alan 3, 5, 6, 15, 18, 22, 23 ve 24. sorular için varyanslar homojendir ($\text{sig} > 0,05$). Diğer sorular için varyansların homojen olmadığı satırın t testi kısmının sig. değerleri incelenmiş olup 25. soru hariç tamamında üst grup lehine anlamlı fark tespit edilmiştir. Ayırt edici olmadığı için 25. soru testten atılmıştır. Sonuç olarak üst grupta yer alan öğrencilerin her bir soru için ortalamaları alt grupta yer alan öğrencilerden anlamlı olarak yüksek çıkmıştır. Taslak test soruları ayırt edicidir.

Tetrakorik Korelasyon Analizi Bulguları

Taslak testte yer alan soruların faktör yapısını belirlemek amacı ile “Factor” programı ile analiz yapılmış hem AFA hem de DFA bulguları elde edilmiştir. Analiz ilk olarak 3 faktörlü yapı ile çalıştırılmış olup varyansın %45’ini açıklayan, KMO değeri 0,85 olarak hesaplanan Barlett’s testi de anlamlı olan bir yapı elde edilmiştir. Bu yapıya ilişkin öz (Eigen) değerleri Tablo 5’te verilmiştir.

Tablo 5

Üç faktörlü yapı için AFA bulguları

Değişken	Öz Değeri	Varyans Oranı	Toplam Varyans Oranı
1	7,87	0,314	0,314
2	1,73	0,069	0,384
3	1,64	0,065	0,450
4	1,44	0,057	
5	1,32	0,053	
6	1,21	0,048	
7	1,12	0,044	

Tablo 6

Üç faktörlü yapı için DFA bulguları

Uyum İndeksleri	Kabul edilebilir sınırlar	Mükemmel uyum sınırı	Ölçeğe ait değer	Ölçeğin uyum durumu
NNFI	0,90 ve üzeri	0,95 ve üzeri	0,91	Kabul edilebilir
GFI	0,85 ve üzeri	0,90 ve üzeri	0,96	Mükemmel
AGFI	0,85 ve üzeri	0,90 ve üzeri	0,95	Mükemmel
RMSEA	=0,05 ve =0,08 arası	=0,00 ve <0,05 arası	0,00	Mükemmel

DFA bulguları incelendiğinde ki kare/ serbestlik derecesi 0,7 olarak hesaplanmış olup 5'ten çok daha küçük bir değer elde edilmiştir. RMSEA değeri ise 0,00 olarak hesaplanmıştır. Buna ek olarak, diğer uyum iyiliği indekslerine bakılmıştır. Diğer uyum indeksleri de istenilen aralıklardadır (Seçer, 2013). Öte yandan, üç faktörlü yapı için binişiklik tablosunda 0,3 ve altında değer alan soruların testten çıkarılması gerekmektedir. Burada 5, 7, 11, 15, 18, 19 ve 25. soruların testten çıkarılması gerektiği belirlenmiştir. Tavsiye edilen faktör sayısı bir olduğu için madde çıkarmadan bir faktöre sınırlandırılarak analiz tekrarlanmıştır. Tek faktörlü yapının toplam varyansın %31'ini açıkladığı görülmüş olup binişik olan 5, 6, 7, 11, 15, 18, 19 ve 25.sorular testten çıkarılıp kalan 17 soru iki faktöre sınırlandırılarak analiz son kez tekrarlanmıştır. Yapılan analiz sonucunda varyansın %53'ünü açıklayan, KMO değeri 0,93 ve Barlett's testi de anlamlı olan bir yapı elde edilmiştir.

Tablo 7

İki faktörlü yapı için AFA bulguları

Değişken	Öz Değeri	Varyans Oranı	Toplam Varyans Oranı
1	7,33	0,431	0,431
2	1,69	0,099	0,530
3	1,40	0,082	
4	1,25	0,073	
5	1,12	0,065	
6	0,78	0,046	
7	0,65	0,037	

Tablo 8

İki faktörlü yapı için DFA bulguları

Uyum İndeksleri	Kabul edilebilir sınır	Mükemmel uyum sınırı	Ölçeğe ait değer	Ölçeğin uyum durumu
NNFI	0,90 ve üzeri	0,95 ve üzeri	0,96	Mükemmel
GFI	0,85 ve üzeri	0,90 ve üzeri	0,96	Mükemmel
AGFI	0,85 ve üzeri	0,90 ve üzeri	0,95	Mükemmel
RMSEA	=0,05 ve =0,08 arası	=0,00 ve <0,05 arası	0,08	Kabul edilebilir

DFA bulguları incelendiğinde ki kare/ serbestlik derecesi 1,20 olarak hesaplanmış olup 5'ten çok daha küçük bir değer elde edilmiştir. RMSEA değeri ise 0,08 olarak kabul edilebilir düzeyde hesaplanmıştır. Diğer uyum indeksleri de 0,95 ve üzerinde olup mükemmel uyuma sahiptir (Seçer, 2013). Buna ek olarak, binişiklik tablosu incelendiğinde 0,3 ve altında değer alan madde olmadığı belirlenmiştir. Bu durumda 17 maddelik iki faktörlü yapının doğrulandığı bulgusu elde edilmiştir.

Tablo 9

İki faktörlü yapıya ait sorular

Soru numarası	Faktör 1	Faktör 2
S1		0,773
S2	0,336	0,669
S3		0,593
S4	0,541	
S8	0,354	0,449
S9	0,559	
S10	0,571	
S12	0,582	
S13	0,461	
S14	0,543	0,410
S16	0,880	
S17	0,435	
S20	0,369	
S21		0,520
S22		0,462
S23	0,766	
S24	0,767	

Tablo 9 incelendiğinde birinci faktörün altına giren soruların; 4, 9, 10, 12, 13, 14, 16, 17, 20, 23 ve 24.maddeler ikinci faktörün altına giren soruların; 1, 2, 3, 8, 21 ve 22. sorular olduğu görülmektedir. Taslak testte yer alan soru içerikleri ve kazanımlar incelenerek birinci faktöre “*Maddenin yapısı ve ısı iletimi*”, ikinci faktöre ise “*Tanecik modeli ve yoğunluk*” ismi verilmiştir.

Güvenirlilik Analizi Bulguları

Taslak testin 25 soruluk ilk halinin güvenirlik analizi yapılmış olup elde edilen bulgular aşağıda verilmiştir.

Tablo 10

Güvenirlilik analizi ilk hali

KR-20 Güvenirlilik Katsayısı	Standartlaştırılmış Maddeler	Soru Sayısı
,836	,836	25

Pallant (2020)'a göre taslak ölçme aracında yer alan maddelerin güvenilir olduğu söylenebilir.

Tablo 11

Güvenirlilik analizi ilk hali korelasyon değerleri

Soru Numarası	Soru Çıkarılırsa Ortalama Değer	Düzeltilmiş Toplam Madde- Korelasyon	Soru Çıkarılırsa Güvenirlilik Katsayısı
S1	11,2587	,218	,835
S2	11,0909	,519	,824
S3	11,1049	,319	,832
S4	10,9371	,419	,828
S5	11,1399	,313	,832
S6	11,1469	,416	,828
S7	11,2238	,253	,836
S8	11,0839	,453	,827
S9	11,1259	,447	,827
S10	11,2657	,334	,831
S11	11,3497	,309	,832
S12	11,0909	,416	,828
S13	11,2098	,382	,830
S14	10,9371	,448	,827
S15	11,1538	,301	,833
S16	11,1399	,538	,823
S17	11,3077	,361	,830
S18	11,2098	,393	,829
S19	11,0210	,455	,827
S20	11,0140	,394	,829
S21	10,9930	,408	,829
S22	11,1119	,337	,831
S23	11,1329	,373	,830
S24	11,0699	,435	,827
S25	11,3147	,159	,837

Tablo 11 verileri incelendiğinde “Düzeltilmiş Toplam Madde Korelasyon” değeri 0,3’ün altında olan sorular 1, 7 ve 25. sorular olarak tespit edilmiş olup faktör analizi sonuçları da değerlendirilerek 7 ve 25. soruların çıkarılmasına, 1. sorunun tutulmasına karar verilmiştir. Ayrıca tabloda yer alan “Madde Çıkarılırsa Güvenirlilik Katsayısı” değeri incelendiğinde bu maddelerin

taslak testten çıkarılması durumunda güvenilirlik katsayısının artış göstereceği hatta 7. ve 25. soruların çıkarılması halinde güvenilirlik katsayısının 1. sorunun çıkarılması durumuna göre daha fazla artış sağlayacağı görülmektedir. Bu sebeple 1. soru testten atılmamıştır. İki faktörlü 17 maddelik yapıda yer alan maddelerle güvenilirlik analizi tekrarlanmıştır.

Tablo 12

Güvenirlik analizi ilk hali korelasyon değerleri

KR-20 Güvenirlik Katsayısı	Standartlaştırılmış Maddeler Soru Sayısı	
,806	,806	17

Pallant (2020)'a göre geliştirilen ölçme aracında yer alan maddelerin güvenilir olduğu bulgusuna ulaşılmıştır. Geliştirilen testin 17 soruluk son hali, başlangıçta 25 soruluk halinde sahip olduğu soru numarası değiştirilmeden EK-1'de verilmiştir. Böylece, özellikle test geliştirme çalışması yapmak isteyen yeni araştırmacıların analizleri sorularla daha net ilişkilendirecekleri düşünülmüştür.

Sonuç ve Tartışma

Fen bilimleri dersi öğretim programı kapsamında, altıncı sınıf “Madde ve Isı” ünitesine yönelik Bloom taksonomisine uygun olarak hazırlanmış, üst düzey basamaklarda yer alan bilgi ve becerileri ölçmeyi amaçlayan, geçerli ve güvenilir bir başarı testi geliştirmenin hedeflendiği bu çalışma, klasik test teorisine göre test geliştirme basamaklarına uygun olarak yürütülmüş olup hata oranı en aza indirilmeye çalışılmıştır. Bu bakımdan ilgili alan yazın taranarak soru havuzu oluşturulup, sorular uzman kontrolünden geçirildikten sonra taslak testte yer almıştır. Taslak testin öğrencilere uygulanmasının ardından, ölçme aracının geçerlik ve güvenilirlik analizleri yapılmıştır. Alan yazında var olan başarı testi geliştirme çalışmaları incelendiğinde bir kısmında ya belirtke tablosunun verilmediği ya da hatalı olduğu tespit edilmiştir (İnal ve Aydın, 2015; Soylu vd., 2020). Bu çalışmada ise belirtke tablosu titizlikle hazırlanmış ve iki uzmanın kontrolünden geçirilmiştir. Bu bakımdan bu çalışma alan yazınla farklılık göstermektedir.

Ölçme aracının geçerliğini test etmek amacı ile; kapsam geçerliği ve yapı geçerliği çalışmaları yürütülmüştür. Kapsam geçerliği kapsamında; soru havuzu oluşturma, havuzdan soru seçme, seçilen soruların düzenlenmesi ve belirtke tablosunun hazırlanması olmak üzere tüm süreç alanında uzman bir fen eğitimcisinin kontrolünden geçirilmiştir. Burada sürecin ayrıntılı olarak raporlanması kısmı alan yazında var olan çalışmalardan (Avcı, 2020; Soylu vd., 2020)

farklılaşmaktadır. Taslak test soruları alanında uzman bir kimya eğitimcisinin kontrolünden geçirilmiş olup “Yapı Geçerliği” başlığında anlatılan gerekli düzeltmeler yapılmıştır. Bloom taksonomisinde bulunan sentez hariç tüm basamaklara uygun bilgi ve becerilerin ölçülmesini sağlayacak soruların testte yer alması sağlanmıştır. Test geliştirme kapsamında, ölçüt geçerliği analizi yapılamamıştır. Alan yazında test geliştirme çalışmalarının birçoğunda da ölçüt geçerliğine ilişkin analizlere rastlanmamıştır (Saraç, 2018; Şener ve Taş, 2017; Üçüncü ve Sakız, 2020). Bu durumun olası nedenleri, geliştirilen testler ile uyumlu ölçüt çalışma bulunamaması ya da verilerine ulaşılamaması olabilir. Testin yapı geçerliğini sağlamak açısından tetrakorik korelasyon yöntemi ile AFA ve DFA analizleri yapılmıştır. Açıklayıcı faktör analizleri binişik maddelerin çıkarılması ile birçok kez yürütülmüş olup süreç boyunca yapılanlar ve çıkarılan maddeler şeffaf bir şekilde açıklanmıştır. Bu sayede yeni olan araştırmacılara yol göstereceği düşünülmektedir. İki faktörlü 17 soruluk testin KMO değeri 0,93 olarak hesaplanmış olup iyi düzeydedir. Bu iki faktörlü yapının toplam varyansın %53’ünü açıkladığı tespit edilmiştir. Açıklanan varyans oranının %40 ve üzerinde olması bakımından uygun olduğu görülmektedir (Seçer, 2013). Bu yapı DFA ile doğrulanmış olup hem ki kare/ serbestlik oranı hem de diğer uyum indeksleri bakımından istenilen aralıklarda tespit edilmiştir. Avcı (2020) ve Soylu vd. (2020) Madde ve Isı ünitesi üzerinde değişen öğretim programına uygun olarak hazırlanmış başarı testi geliştirme çalışmaları yapmışlardır. Bu çalışmalar incelendiğinde AFA ve DFA analizlerinin yapılmamış olduğu görülmektedir. Bu çok büyük bir eksikliktir. Bu bakımdan bu çalışmalar ile mevcut çalışma arasında farklılık tespit edilmiştir. Elde edilen yapıya ait birinci faktörde yer alan maddeler incelenerek “Maddenin yapısı ve ısı iletimi” ikinci faktöre “Tanecik modeli ve yoğunluk” ismi verilmiştir. Ölçeklerden farklı olarak başarı testlerinde analizlerin SPSS ve Lisrell programları yerine “Factor” programı kullanılarak analiz edilmesi faktör yapısının belirlenmesinde daha uygun bir yöntemdir (Lorenzo-Seva ve Ferrando, 2006). Bu sebeple bu çalışmanın süregelen bir uygulamanın alternatifi ve yeni yolu olması bakımından farklılık gösterdiği düşünülmektedir.

Testte yer alan tüm soruların tek tek normal dağılım gösterme durumu incelenmiştir. Bu kısım genellikle test geliştirme çalışmalarında bulunmayan kısımdır. Bu açıdan bundan sonraki çalışmalarda normal dağılım analizlerinin bulunmasının yaygınlaştırılacağı düşünülmektedir. Testte bulunan soruların indeks değerlerini hesaplamak amacıyla katılımcılar %27’lik üst ve alt gruplara ayrılmıştır. Her bir soru için güçlük ve ayırt edicilik değerleri hesaplanmış olup, istenilen aralıklarda olmayan 7. ve 25. sorular faktör analizi sonuçları da değerlendirilerek ölçme aracından

çıkarılmıştır. Testin ortalama güçlük değeri 0,50 ve ortalama ayırt edicilik değeri 0,54 olarak hesaplanmış olup bu değerlerin 0,50 dolaylarında olması istenen bir durumdur. Soylu, vd. (2020)'de güçlük ve ayırt edicilik katsayısı bakımından benzer sonuçlara ulaşmışlardır. Madde güçlük ve ayırt edicilik indeksleri hemen hemen tüm başarı testi geliştirme çalışmalarında hesaplanan ve raporlanan kısımlar arasında yer alırken, üst grup ve alt grupta yer alan öğrenciler arasında alınan puanlar bakımından anlamlı farkın olup olmadığının ilişkisiz örneklem t testi analizi ile belirlendiği çalışmaların sayısı azdır (Divarcı ve Kaya, 2019). Bu bakımdan da yürütülen çalışmanın diğer başarı testi geliştirme çalışmalarından farklılık gösterdiği düşünülmektedir. Yapılan t testi analizi sonucunda 25. soru hariç tamamında üst grup lehine anlamlı fark tespit edilmiştir. Sonuç olarak üst grupta yer alan öğrencilerin her bir soru için ortalamaları alt grupta yer alan öğrencilerden anlamlı olarak yüksek çıkmıştır. Bu duruma göre taslak test sorularının ayırt edici olduğu sonucuna ulaşılmıştır.

Ölçme aracından elde edilen puanların güvenilir olup olmadığını ortaya koymak adına hem test sorularının tamamının dahil edildiği ilk halinin hem de analizler tamamlandıktan sonra kalan son halinin KR-20 güvenilirlik katsayısı hesaplanmıştır. Testin 25 soruluk ilk halinden elde edilen puanların güvenilirlik katsayısı 0,84 olarak, 17 soruluk son halinden elde edilen puanların güvenilirlik katsayısı 0,81 olarak hesaplanmıştır. Elde edilen güvenilirlik katsayılarına bakılarak ölçme aracının güvenilir olduğu sonucuna varılmıştır. Soylu, vd. (2020) da güvenilirlik katsayısı bakımından benzer sonuçlara ulaşmışlardır.

Başlangıçta bütün kazanımları karşılamak üzere ve üst düzey basamakları da yoklayacak şekilde test geliştirilmesi yönünde sorular oluşturulmuştur. Testten soru atmamak için farklı analizler yapılmış, soruların testte kalması yönünde kanıtlar bulunmaya çalışılmıştır. Ancak analizler sonucunda testte 17 soru kalmıştır. Kalan soruların kazanımları ve bilişsel düzeyleri objektif bir şekilde incelenmiştir. Örneğin, analiz düzeyinde yer alan 5. soru ile ölçülmek istenen “*Maddenin taneciklerinin hareketliliğindeki değişim*” kazanımının 3. soru ile karşılanabildiği görülmüştür. Uygulama düzeyinde yer alan yoğunluk hesaplama ile ilgili 6. ve 7. soruların 8. soru ile karşılanabildiği tespit edilmiştir. Uygulama düzeyindeki “*Maddeleri ısı iletimi bakımından sınıflandırır*” kazanımına ait 11. sorunun hedef davranışlarının 12. soru ile karşılanabildiği düşünülmektedir. Son olarak değerlendirme düzeyinde yer alan 25. soruya ait kazanımın 21. soru ile karşılanabildiği sonucuna ulaşılmıştır. Buradan hareketle kazanım bakımından bir kayıp

olmadığı düşünülmektedir. Testten çıkarılan sorulardan 3'ü uygulama, 1'i analiz ve 1'i değerlendirme olmak üzere üst düzey bilişsel düzeylerdendir. Son durumda test sorularının bilişsel basamaklara göre dağılımı ise şu şekildedir; bilgi (3), kavrama (5), uygulama (3), analiz (6), değerlendirme (3). Test geliştirme çalışması yaparken, taslak testte her bir kazanıma ve bilişsel düzeye yönelik birden fazla soru yazılmasının/seçilmesinin araştırmacının amaçladığı düzeyde bir test elde etmesini kolaylaştıracağını belirtmek isteriz. Araştırmacıların taslak testte fazla soru olmasından kaçınmalarının olası nedeni, geçerliği ve normallığı sağlayabilmek için taslak testte yer alan madde sayısının en az 10 katı kişiye ulaşma konusunda kaygı taşımaları olabilir.

Sonuç olarak, ortaokul öğrencilerinin “*Madde ve Isı*” ünitesine yönelik başarı düzeylerini belirlemek veya öğrenme eksikliklerini gidermek için hem araştırmacılar hem de öğretmenler tarafından kullanılabilir geçerliği ve güvenilirliği kontrol edilmiş bir ölçme aracı geliştirilmiştir.

Öneriler

Geliştirilen başarı testinin öğrencilerin başarı düzeylerini belirlemek veya öğrenme eksikliklerini gidermek için hem araştırmacılar hem de öğretmenler tarafından kullanılması önerilmektedir.

Başarı testlerinin faktör analizlerinin tetrakorik korelasyon analizi ile yapılması önerilmektedir.

Test geliştirme çalışması yapılırken soru havuzu belirlenmeden önce ilgili örneklem grubunda bulunacak olası katılımcılarla ön görüşmeler yapılabilir.

Bu çalışmada geliştirilen başarı testi klasik test teorisini temel almıştır, aynı konu üzerinde madde tepki kuramını temel alan başarı testleri geliştirilebilir.

Bu çalışmada çoktan seçmeli sorulardan oluşan başarı testi geliştirildiği için sentez basamağında sorular bulunmamaktadır, açık uçlu soruların da yer aldığı ölçme araçları ile öğrencilerin sentez basamağına uygun kazanım düzeyleri yoklanabilir.

Yakıtlarla ilgili öğretim programında yer alan ünite kazanımı değerlendirme basamağındadır. Yürütülen çalışmada ise yakıtlarla ilgili daha düşük basamakta sorulara yer verilmiştir. Bu araştırmanın sınırlılığıdır. Araştırmacılara bu kazanıma yönelik değerlendirme basamağında sorularak test geliştirmeleri önerilir.

Geliştirilen testin son halinde üst düzey becerileri yoklayan soru sayısı azalmıştır. Soru havuzundan taslak teste soru seçerken her kazanımdan ve bilişsel düzeyde birden fazla soruya yer verilmesi araştırmacılara önerilir. Uygulamaya katılacak kişi sayısı ile ilgili endişe taşımadan, taslak testte ne kadar çok soru ile yola çıkılırsa araştırmacının istenilen düzeyde geçerli ve güvenilir bir test elde etmesi o kadar kolaylaşacaktır.

Kaynakça

- Adıgüzel, O. C. ve Özdoğru, F. (2013). Üniversitelerde ortak zorunlu yabancı dil I dersine yönelik bir akademik başarı testinin geliştirilmesi. *Trakya Üniversitesi Eğitim Fakültesi Dergisi*, 3(2), 1- 11.
- 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.
- Altınsoy, A.B. (2011). *Fen ve teknoloji dersinde çoklu zekâ kuramına dayalı öğretimin öğrencilerin başarılarına etkisi* (Yüksek lisans tezi). Selçuk Üniversitesi, Konya.
- Avcı, F. (2020). Yenilenmiş Bloom taksonomisine göre madde ve işi başarı testi: Geçerlik güvenilirlik çalışması. *Iğdir University Journal of Social Sciences*, 21, 263- 292.
- Başol, G., Çakan, M., Kan, A., Özbek, Ö. Y., Özdemir, D. ve Yaşar, M. (2013). Eğitimde ölçme ve değerlendirme. *Ankara: Pegem Akademi Yayıncılık*.
- Baykul, Y. (2015). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Pegem Akademi.
- Beuchert, A. K. & Mendoza, J. L. (1979). A Monte Carlo comparison of ten item discrimination indices. *Journal of Educational Measurement*, 16(2), 109-118.
- Birgin, O. (2016). *Bloom taksonomisi*. E. Bingölbali, S. Arslan, & İ. Ö. Zembat (Edit.), Matematik eğitiminde teoriler. Pegem Akademi.
- Bloom, B. S., Engelhart, M. D, Furst, E. J, Hill, W. H. ve Krathwohl, D. R. (1956). *Taxonomy of educational objectives*. Book I: Cognitive Domain. U. S: Longman.
- Büyüköztürk, Ş. (2013). *Sosyal bilimler için veri analizi el kitabı*. 8. Baskı. Ankara: Pegem Akademi.
- Creswell, J. W. (2009). *Research design, qualitative, quantitative, and mixed methods approaches (Third Edition)*. California: SAGE Publications.

- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Çardak, Ç. S. ve Selvi, K. (2018). Öğretim ilke ve yöntemleri dersi için bir başarı testi geliştirme süreci. *Akdeniz Eğitim Araştırmaları Dergisi*, 12(26), 379-406.
- Çelen, Ü. ve Aybek, E. C. (2013). Öğrenci başarısının öğretmen yapımı bir testle klasik test kuramı ve madde tepki kuramı yöntemleriyle elde edilen puanlara göre karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 4(2), 64-75.
- Çelik, E. (2010). *Fen eğitiminde probleme dayalı öğrenme yaklaşımının öğrencilerin akademik başarısına, tutumuna, akademik risk alma düzeyine ve kalıcılığa etkisi* (Yüksek lisans tezi). Gazi Üniversitesi, Ankara.
- Çıtak Gözen, G. (2010). Klasik test ve madde tepki kuramlarına göre çoktan seçmeli testlerde farklı puanlama yöntemlerinin karşılaştırılması. *İlköğretim Online (elektronik)*, 9(1), 170-187.
- Demir, N., Kızılay, E. ve Bektaş, O. (2016). Development of an achievement test about solutions for 7 th graders: a validity and reliability study. *Necatibey Faculty of Education Electronic Journal of Science & Mathematics Education*, 10(1), 209-237.
- 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.
- Demirçalı, S. ve Alkan, B. (2018). *Ortaokul ve imam hatip ortaokulu fen bilimleri 6. sınıf ders kitabı*. Milli Eğitim Bakanlığı Yayınları, ISBN 978-975-11-4692-2, 236 s., Demirçalı, S. (Ed).
- Divarcı, Ö. F. ve Kaya, H., (2019). 8. sınıf “maddenin halleri ve ısı” ünitesine yönelik geçerliliği ve güvenilirliği sağlanmış bir akademik başarı testi geliştirme çalışması. *Eğitim Kuram ve Uygulama Araştırmaları Dergisi*, 5(2), 214-238.
- Elbay, S. (2020). T.C. İnkılap Tarihi ve Atatürkçülük dersi 2. ünitesine yönelik kazanım odaklı başarı testi geliştirme çalışması. *E-Uluslararası Eğitim Araştırmaları Dergisi*, 11(1), 53-68, DOI: 10.19160/ijer.679934
- Eryılmaz, A., ve Tatlı, A. (2000). ODTÜ öğrencilerinin mekanik konusundaki kavram yanlışları. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 18(18), 93- 98.
- Güler, N. (2011). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem Akademi.
- Haladyna, T. M. (1997). *Writing test items to evaluate higher order thinking*. Allynand Bacon: Needham Heights, MA.

- İnal, Z. ve Aydın, A. (2015). Madde ve ısı ünitesinin öğretilmesinde model kullanımının akademik başarıya ve bilgilerin kalıcılığına etkisi. *Journal of Kirsehir Education Faculty*, 16(3), 13-37.
- Kargın, P. D. ve Gül, Ş., (2021). Altıncı sınıf “vücudumuzdaki sistemler ve sağlığı” ünitesine yönelik bir başarı testi geliştirilmesi. *Ihlara Eğitim Araştırmaları Dergisi*, 6(1), 1-26. <https://doi.org/10.47479/ihead.729412>
- Karip, E. (Ed.) (2012). *Ölçme ve değerlendirme*. Ankara: Pegem Akademi (5. Baskı).
- Kızılkaya, A., ve Seven, S. (2017). Fen öğretiminde jigsaw I tekniğinin öğrencilerin bloom taksonomisi'nin bilişsel alan alt ve üst düzey akademik başarılarına etkisi. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, 19(2), 250-270. <https://doi.org/10.17556/erziefd.334984>
- Kızılkapan, O. ve Bektaş, O. (2018). Fen eğitiminde başarı testi geliştirilmesi: Hücre bölünmesi ve kalıtım örneği. *Maarif Mektepleri Uluslararası Eğitim Bilimleri Dergisi*, 2(1), 1-18.
- Koç, Y. (2014). Okuma-yazma-uygulama ve öğrenci takımları başarı bölümleri yöntemlerinin öğrencilerin akademik başarıları üzerine etkisi: Madde ve ısı ünitesi. *Ekev Akademi Dergisi*, 58(1), 191-210.
- Lorenzo-Seva, U. ve Ferrando, P. J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavior research methods*, 38(1), 88-91.
- MEB. (2013). *İlköğretim kurumları fen bilimleri dersi öğretim programı*. Ankara: Talim ve Terbiye Kurulu Başkanlığı.
- MEB, (2018). *Fen bilimleri dersi taslak öğretim programı (ilkokul ve ortaokul 3, 4, 5, 6, 7, 8. sınıflar)*. Ankara: Temel Eğitim Genel Müdürlüğü.
- Nacaroğlu, O., Bektaş, O. ve Kızılkapan, O. (2020). Madde döngüleri ve çevre sorunları konusunda başarı testi geliştirme: Geçerlik ve güvenirlik çalışması. *Kastamonu Eğitim Dergisi*, 28(1), 36-51.
- Pallant, J. (2020). *SPSS kullanma kılavuzu: SPSS ile adım adım veri analizi*. Anı Yayıncılık.
- Sağlam Arı, G., Armutlu, C., Güneri Tosunoğlu, N. ve Yücel Toy, B., (2009). Pozitivist ve Postpozitivist paradigmlar çerçevesinde metodoloji tartışmalarının yönetim ve pazarlama alanlarına yansımaları. *Hacettepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 27(1), 113-141.

- Saraç, H. (2018). Fen bilimleri dersi ‘maddenin değişimi’ ünitesi ile ilgili başarı testi geliştirme: Geçerlik ve güvenirlik çalışması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18(1), 416-445.
- Seçer, İ. (2013). *SPSS ve LISREL ile pratik veri analizi*. Ankara: Anı Yayıncılık.
- Sontay, G. ve Karamustafaoğlu, O. (2020). Fen bilimleri dersi "Güneş, Dünya ve Ay" ünitesine yönelik başarı testinin geliştirilmesi. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 40(2), 511-551.
- Soylu, Ü. İ., Karamustafaoğlu, S. ve Karamustafaoğlu, O. (2020). 6. sınıf "madde ve ısı" ünitesi başarı testi geliştirme: geçerlik ve güvenirlik. *Ihlara Eğitim Araştırmaları Dergisi*, 5(2), 270-292. <https://doi.org/10.47479/ihead.800620>
- Şener, N. ve Tas, E. (2017). Developing achievement test: A research for assessment of 5th grade biology subject. *Journal of Education and Learning*, 6(2), 254-271.
- Tanrıöğen, A. (2009). *Bilimsel araştırma yöntemleri*. Anı yayıncılık.
- Tosun, C., ve Taşkesenligil, Y. (2011). Revize edilmiş Bloom'un taksonomisine göre çözeltiler ve fiziksel özellikleri konusunda başarı testinin geliştirilmesi: geçerlik ve güvenirlik çalışması. *Kastamonu Eğitim Dergisi*, 19(2), 499-522.
- Uzunöz, A. ve Buldan, İ. (2012). Ortaöğretim coğrafya dersi doğal sistemler konu alanı atmosfer ve iklim ünitesi başarı testi geliştirme çalışması. *Kastamonu Eğitim Dergisi*, 20(1), 291-312.
- Üçüncü, G. ve Sakız, G. (2020). Başarı testi geliştirme süreci: ilkokul dördüncü sınıf maddeyi tanıyalım ünitesi örneği. *Kastamonu Eğitim Dergisi*, 28(1), 82-94.
- Ünal, Ç. ve Çelikkaya, T. (2009). Yapılandırmacı yaklaşımın sosyal bilgiler öğretiminde başarı, tutum ve kalıcılığa etkisi (5. sınıf örneği). *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 13(2), 197-212.

EK 1_ MADDE VE ISI BAŞARI TESTİ

1. Ayşe öğretmen bir bardak su içerisine bir damla mürekkep damlatıp öğrencilerinin gözlem yapmalarını istiyor. Sizce yapılan gözlem sonucunda Ayşe öğretmenin bu deneyi yapmasını aşağıdakilerden hangisi **en iyi** açıklar?

- A. Suyun tanecikli yapıda olduğunu göstermek için yapmıştır.
- B. Mürekkebin tanecikli yapıda olduğunu göstermek için yapmıştır.
- C. Mürekkebin taneciklerinin öteleme hareketi yaptığını göstermek için yapmıştır.
- D. Suyun ve mürekkebin tanecikli ve boşluklu yapıda olduğunu kanıtlamak için yapmıştır.

2. Katı, sıvı ve gaz halde olduğu bilinen K, L ve M maddelerinin tanecikleri arasındaki boşluk ilişkisi grafikte verilmiştir.



Bu grafikten yola çıkarak sizce aşağıdakilerden hangisi **söylenemez**?

- A. K maddesi gaz halindedir.
- B. L halindeki maddede boşluk en azdır.
- C. M halindeki madde konulduğu kabın şeklini alır.
- D. K halindeki madde sadece titreşim ve dönme hareketi yapar.

3. Emre mutfağa annesinin yanına gittiğinde annesinin ocakta tereyağı erittiğini görmüştür.

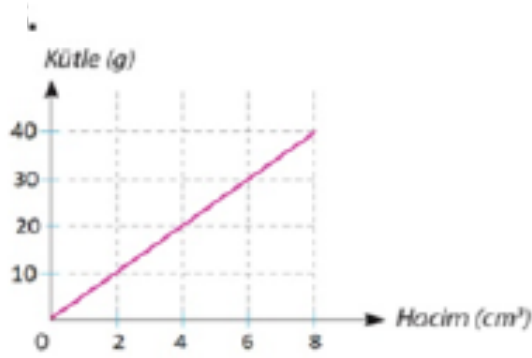
Tereyağıyla ilgili aşağıdaki seçeneklerden hangisini annesine **söyleyemez**?

- A. Tereyağı ısı aldıkça yoğunlaşma gerçekleşir.
- B. Tereyağı eridiğinde boşluklu yapısı değişir.
- C. Tereyağının katı haldeyken belli bir şekli vardır.
- D. Tereyağı hal değiştirdikçe öteleme ve dönme hareketi de yapar.

4. Aşağıda verilen özelliklerden hangisi maddenin tüm halleri (katı, sıvı, gaz) için ortaktır?

- A. Sıkıştırılabilir olması
- B. Taneciklerden oluşması
- C. Öteleme hareketi yapması
- D. Tanecikler arasında büyük boşluklar bulunması

8.



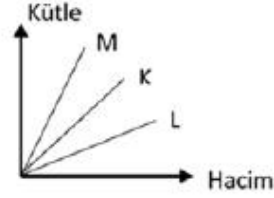
Yukarıdaki kütle- hacim grafiği verilen maddenin yoğunluğu kaç g/cm^3 'tür?

- A. 2
- B. 3
- C. 4
- D. 5

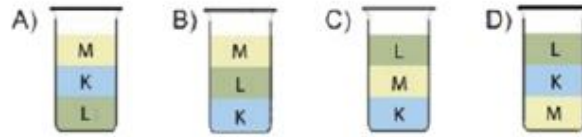
9. Elinde metal para ve silgi bulunan Hatice bu iki maddenin farklı maddeler olduğunu kanıtlamak istiyor. Hatice hangi işlemi yaparsa öğretmenini ikna etmiş olur?

- A. Metal para ve silginin kütesini hesaplamalıdır.
- B. Metal para ve silginin hacmini hesaplamalıdır.
- C. Metal para ve silginin sıcaklıklarını hesaplamalıdır.
- D. Metal para ve silginin yoğunluklarını hesaplamalıdır.

10. Özge'nin elinde sıvı halde bulunan üç ayrı madde vardır. Özge'nin elindeki bu üç madde birbiri içinde çözünmemekte ve birbiri ile karışmamaktadır. M, K ve L maddelerinin kütle ve hacimlerini ölçen Özge aşağıdaki kütle- hacim grafiğini elde etmiştir.



Bu sıvılardan eşit hacimde alınıp aynı kaba koyulduğunda görünüşleri nasıl olur?



12.



Zeynep hanım şekildeki tencere ile yemeğini yaparken aşağıdaki işlemleri gerçekleştirmiştir.

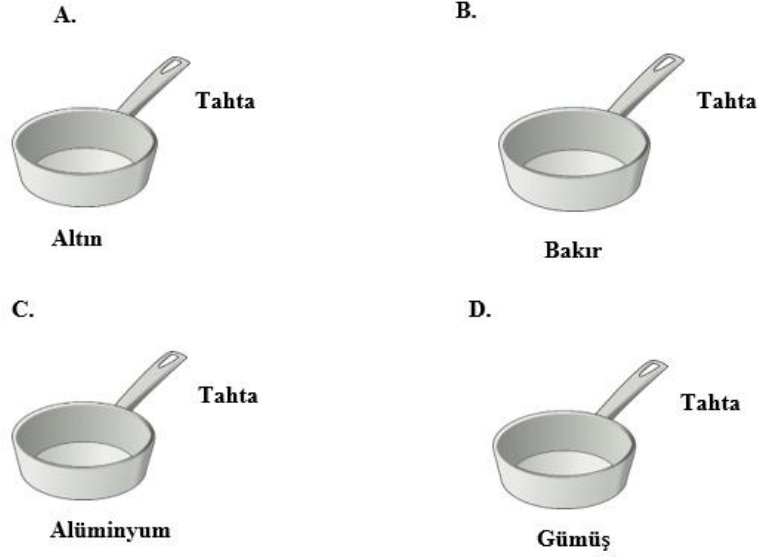
- I. Tencerenin metal kısmını ocağın üzerine koyarak yemeğini ısıtmaya başlamıştır.
- II. Yemeğini karıştırmak için tahta kaşık kullanmıştır.
- III. Tencerenin kapağını plastik kısmından tutarak açmıştır.
- IV. Tencerenin kulpundan tutarak tencereyi ocağın üstünden almıştır.

Zeynep hanımın yemek yaparken takip ettiği durumları ısı iletkeni ve ısı yalıtkanı olarak sınıflandırıldığında aşağıdakilerden hangisi doğru olur?

	<u>Isı İletkeni</u>	<u>Isı Yalıtkanı</u>
A.	III	I
B.	II	IV
C.	I	III
D.	IV	II

13. Bazı maddeler ısıyı daha iyi iletirler, bazıları ise iyi iletmez. Bu durum dikkate alındığında aşağıdaki tavalardan hangisinde yemek pişirmek **en uygundur**?

(Maddelerin ısı iletim sırası; Gümüş > Bakır > Altın > Alüminyum > Tahta)



14.



Sıcak bir yaz gününde Fatma parka oyun oynamaya gitmiştir. Oyun oynamaktan yorulup dinlenmek istediğinde şekilde gösterilen banka oturmuştur. Fatma bankın tahta kısmına oturmuş, metal kısmına kolunu koymuştur. Bir süre sonra kolunu birden çekmiş fakat bankta oturmaya devam etmiştir. Fatma'nın kolunu çekmesinin sebebi aşağıdakilerden hangisidir?

- A. Çelik ısıyı tahtaya göre daha yavaş iletmiştir.
- B. Çeliğin ısı iletkenliği tahtadan daha fazladır.
- C. Tahtanın ısı iletkenliği çelikten daha fazladır.
- D. Tahtaya dokunduğunda çeliğe dokunduğundan daha sıcak hissetmiştir.

16. Isı yalıtım malzemeleri seçilirken aşağıdakilerden hangisi tercih edilen bir özellik **değildir**?

Malzemenin;

- A. Kolay yanması
- B. Uzun ömürlü olması
- C. Düşük maliyetli olması
- D. İnsan ve çevre sağlığına uygun olması

17.



Ağlayan evin yüzünü güldürmek için bazı öğrenciler alternatif çözüm yolları önermişlerdir.

Ayşe: Isıyı iyi iletmeyen malzemeler kullanılmalıdır.

Ecrin: Ağır malzemeler kullanılmalıdır.

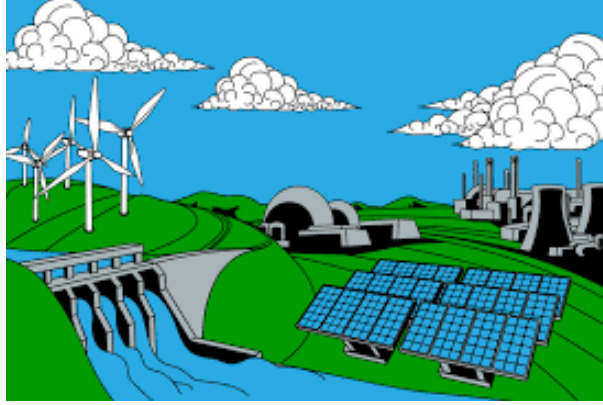
Nazlı: Kullanılan malzemenin iç kısımları boşluklu yapıda olmalıdır.

Ahmet: Evin her yeri aynı özellikteki malzeme ile kaplanmalıdır.

Hangi öğrencilerin önerileri doğrudur?

- A. Ayşe ve Ahmet
- B. Ecrin ve Nazlı
- C. Ayşe ve Nazlı
- D. Ecrin ve Ahmet

20.



Yukarıda rüzgar tribünleri, hidroelektrik santral, güneş panelleri ve jeotermal santrallere ait görsel verilmiştir. Verilen enerji türlerinin ortak özelliği aşağıdakilerden hangisidir?

- A. Atmosfere zehirli gaz vermeleri
- B. Canlı atıklardan elde edilmeleri
- C. Enerjiyi su sayesinde elde etmeleri
- D. Yenilenebilir enerji kaynağı olmaları

21.

Malzeme	Yanıcılık	Kullanım ömrü
Plastik köpük	Alev alır	Uzun ömürlü
Katran	Alev alır	Kısa ömürlü
Silikon yünü	Zor alev alır	Uzun ömürlü
Ahşap	Alev alır	Kısa ömürlü

Konutunun dış cephesine yalıtım yaptırmak isteyen Orhan Bey'in yukarıdaki tabloda verilen yalıtım malzemelerinden hangisini kullanması **en uygun** olur?

- A. Ahşap
- B. Katran
- C. Plastik köpük
- D. Silikon yünü

22.

- I. Soba yeteri kadar doldurulmalıdır.
- II. Sobalar yaşam alanlarına kurulmalıdır.
- III. Isı ve kurum biriken bacalar sık sık temizlenmelidir.
- IV. Sobanın bulunduğu oda havalandırılmalıdır.

Numaralandırılmış bu davranışlardan hangileri soba zehirlenmelerine karşı alınabilecek önlemlerdendir?

- A. I ve II
- B. I ve III
- C. I, III ve IV
- D. I, II, III ve IV

23. Aslı yakıtlara verilen örnekleri aşağıdaki gibi sıralamıştır.

- I. Gaz yağı, sıvı yakıtlara örnektir.
- II. LPG, gaz yakıtlara örnektir.
- III. Fuel- oil, binalarda ısınma amaçlı kullanılan gaz yakıtlara örnektir.
- IV. Odun, katı yakıtlara örnektir.

Yukarıda verilen öncüllerden hangileri doğrudur?

- A. I ve II
- B. I ve IV
- C. I, II ve III
- D. I, II ve IV

24. Petrolden üretilen aşağıdaki yakıtlardan hangisi sıvı yakıtlar arasında **yer almaz**?

- A. LPG
- B. Benzin
- C. Gaz yağı
- D. Mazot

Extended Abstract

Purpose

This study was designed according to classical test theory, assuming that all distractors differ from the correct answer equally. The characteristics of the multiple-choice tests, in which the concept of taxonomy emerged according to the classical test theory, were determined. Among these taxonomies, there are six dimensions of cognitive levels in Bloom's taxonomy, which is the most widely accepted in the literature. These; knowledge, comprehension, application, analysis, evaluation and synthesis. Bloom's taxonomy is a suitable and preferred tool to determine the cognitive levels of students. Among the traditional measurement tools, the most used measurement tool in education is multiple choice questions. Since it is the most widely used measurement tool in education, teachers need valid and reliable multiple-choice measurement tools prepared by experts.

"*Matter and Heat*" unit is one of the subjects that students have difficulty in understanding because it contains abstract concepts and events in science lessons. In the 2018 curriculum, it was deemed necessary to develop a multiple-choice measurement tool suitable for use by researchers and teachers, prepared at cognitive levels in accordance with Bloom's taxonomy, suitable for examining high-level skills, for all subjects and achievements of the sixth-grade matter and heat unit, the number and importance of which were increased. The aim of this research is to develop a valid and reliable achievement test for the sixth grade "*Matter and Heat*" unit within the scope of the science curriculum, prepared in accordance with the Bloom's taxonomy, aiming to measure the knowledge and skills at the higher levels. In this context, the sub-problems of the research are as follows:

1. Is the achievement test prepared for the sixth grade "*Matter and Heat*" unit valid?
2. Is the achievement test for the sixth grade "*Matter and Heat*" unit reliable?

Results

For the validity study, content and construct validity were performed. A specification table was created to ensure the content validity of the created test. Opinions of two experts from the fields of science education and chemistry education were received. The test, which was created as a result of expert feedback, was examined by a Turkish expert. In addition, the created test was examined by a science teacher and its suitability for the sixth-grade level was checked. Exploratory and confirmatory factor analysis was performed for construct validity. In addition, item indices and independent sample t-test were performed.

It was determined that each question in the test showed a normal distribution. When the item indices were examined, it was found that the difficulty index of the seventh question and the discrimination

index value of the 25th question were below the desired ranges, which were below 0.30. These questions were removed from the draft achievement test during factor analysis. According to the t-test results, all questions are distinctive for the upper group, except for the 25th question. In order to determine the factor structure of the questions in the draft test, analysis was made with the "Factor" program and both EFA and CFA findings were obtained. When the CFA findings were examined, the chi-square/degree of freedom was calculated as 1.20. The RMSEA value was calculated as 0.08 at an acceptable level. Other fit indices are 0.95 and above and have a perfect fit. In this case, it was found that the 17-item two-factor structure was confirmed. As a result of the reliability analysis, it was found that the developed test had a coefficient of .806.

Discussion and Conclusion

After the draft test was applied to the students, the validity and reliability analyzes of the measurement tool were made. When the achievement test development studies in the literature are examined, it has been determined that either the table of test specification is not given or it is incorrect. In this study, the table of test specification was meticulously prepared and checked by two experts. In this respect, this study differs from the literature.

In order to ensure the construct validity of the test, EFA and CFA analyzes were performed using the tetrachoric correlation method. Exploratory factor analyzes were carried out many times with the removal of overlapping items, and what was done and the items removed during the process were explained in a transparent way. In this way, it is thought that it will guide new researchers. The KMO value of the two-factor 17-question test was calculated as 0.93 and is at a good level. It was determined that this two-factor structure explained 53% of the total variance. It is seen that the explained variance rate is 40% and above (Secer, 2013). This structure was confirmed by CFA and it was determined in the desired ranges in terms of both the chi-square/freedom ratio and other fit indices.

As a result, a valid and reliable measurement tool that can be used by both researchers and teachers has been developed to determine the achievement levels of secondary school students for the "Matter and Heat" unit or to eliminate their learning deficiencies.

ETİK BEYAN: "MADDE ve ISI ÜNİTESİ BAŞARI TESTİ GELİŞTİRME ÇALIŞMASI" başlıklı çalışmanın yazım sürecinde bilimsel, etik ve alıntı kurallarına uyulmuş; toplanan veriler üzerinde herhangi bir tahrifat yapılmamıştır ve veriler toplanmadan önce Erciyes Üniversitesi Sosyal ve Beşeri Bilimler Etik Kurulu'ndan 29/06/2021 tarih ve 297 sayılı etik izin alınmıştır. Karşılaşılabilecek tüm etik ihlallerde "Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi Yayın Kurulunun" hiçbir sorumluluğunun olmadığı, tüm sorumluluğun Sorumlu Yazara ait olduğu ve bu çalışmanın herhangi başka bir akademik yayın ortamına değerlendirme için gönderilmemiş olduğunu taahhüt ederim.