

SINDIRIM SİSTEMİ ÜNİTESİNE YÖNELİK BAŞARI TESTİ GELİŞTİRME: GEÇERLİK VE GÜVENİRLİK ÇALIŞMASI

DEVELOPMENT OF A ACHIEVEMENT TEST FOR THE DIGESTIVE SYSTEM UNIT: A VALIDITY AND RELIABILITY STUDY

Araştırma Makalesi

Gürcü Ebru ÖZKILIÇ¹, Oktay BEKTAŞ² ve Melek KARACA³

Makale gönderim tarihi: 19 Mart 2023

Makale kabul tarihi : 24 Mayıs 2023

Özet

Bu çalışmanın amacı, ortaokul öğrencilerinin altıncı sınıf "Sindirim Sistemi" konusundaki öğrenme düzeylerini ölçmeye yönelik geçerlik ve güvenilirlik kontrolleri yapılmış, çoktan seçmeli sorulardan oluşan bir başarı testi geliştirmektir. Çalışma nicel araştırmanın tarama desenine göre yapılandırılmıştır. 2020-2021 eğitim-öğretim yılı ikinci döneminde Şanlıurfa İli Siverek ilçesinde altıncı sınıfta öğrenim gören 229 öğrenci, araştırmanın örneklemini oluşturmaktadır. Testin kapsam geçerliği için kazanımlar belirlenmiş, kazanımlardan 44 adet çoktan seçmeli sorudan oluşan soru havuzu oluşturulmuştur. Test bir ölçme değerlendirme uzmanına, beş fen bilimleri öğretmenine, bir biyoloji uzmanına, üç Türkçe öğretmenine, bir fen eğitimi alanında akademisyen olmak üzere toplamda 11 uzmanın görüşüne sunulmuş ve kapsam geçerliği kontrol edilmiştir. Testteki soruların madde güçlük ve madde ayırt edicilik indeksleri hesaplanmış; soruların güçlüğü kolay, orta, zor olduğu, ayırt ediciliğinin ise çok iyi olduğu görülmüştür. Veri analizinde Factor programı, MS Excel 2020 ve SPSS.25 paket programları kullanılmıştır. Yapı geçerliği için AFA ve DFA aynı örneklem üzerinde uygulanmıştır. Geliştirilen başarı testi, KR-20 (Kuder-Richardson 20) güvenilirlik katsayısı 0,84 olan 16 soruluk geçerlik ve güvenilirliği kontrol edilmiş bir başarı testidir. Geliştirilen teste açık uçlu sorular eklenip iki aşamalı şekilde kullanılabilir.

Anahtar Kelimeler: Başarı testi, Geçerlik, Güvenirlik, Sindirim sistemi

Abstract

The aim of this study is to develop an achievement test consisting of multiple-choice questions with validity and reliability in order to measure the learning levels of middle school students in the sixth grade "Digestive System". The study was structured according to the screening design of quantitative research. In the second semester of the 2020-2021 academic year, 229 sixth grade students in the Siverek district of Şanlıurfa constitute the sample of the research. For the content validity of the test, the gains were determined, and a question pool consisting of 44 multiple-choice questions was created from the objectives. Content validity was ensured by submitting the test to a total of 11 experts, including one assessment and evaluation expert, five science teachers, one biology expert, three Turkish teachers, and one academician in the field of science education. The item difficulty and item discrimination indexes of the questions in the test were calculated, and it was seen that the difficulty of the questions was easy, medium, difficult, and the discrimination was very good. Factor program, MS Excel 2020 and SPSS.25 package programs were used in data analysis. For construct validity, EFA and CFA were applied on the same sample. The test KR-20 (Kuder-Richardson 20) is an achievement test of 16 questions with a reliability coefficient of 0.84 with validity and reliability. Open-ended questions can be added to the developed test and it can be used in two stages.

Keywords: Achievement test, Validity, Reliability, Digestive system

¹ Öğretmen, Milli Eğitim Bakanlığı, ebru.fnc.91@gmail.com, ORCID ID: 0000-0001-6797-3497

² Prof. Dr., Erciyes Üniversitesi, obektas@erciyes.edu.tr, ORCID ID: 0000-0002-2562-2864

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

GİRİŞ

Ölçme ve değerlendirme kavramları eğitim öğretimin olduğu her dönemde ihtiyaç duyulan ayrılmaz bir parçası, dolayısıyla üzerinde durulması gereken önemli kavramlar olarak karşımıza çıkmaktadır. Ölçme ve değerlendirmenin olmadığı bir eğitim öğretim sürecinden bahsetmek mümkün değildir (Başol, 2015). Örneğin, bireylere uygulanan testteki bilgilerin o bireyde olup olmadığı ancak bir ölçme aracı sayesinde öğrenilebilir. Öğrencinin fen bilimleri dersindeki başarısını gözüne, yüzüne bakarak sadece fiziksel olarak doğrudan anlamak mümkün değildir. Akademik anlamda başarılı öğrenciyi başarısız öğrenciden ayırırken de yine mutlaka bir ölçme aracına ihtiyaç duyarız (Baştürk, 2014). Doğrudan ölçülemeyen durumların varlığı, ölçme ve değerlendirmenin eğitim öğretim sürecinin ayrılmaz bir parçası olduğunu göstermektedir. Ölçme; sayı ve semboller kullanılarak kişilerin ya da nesnelerin belirli özelliklere sahip olup olmadığını gözlemlemek, eğer bu özelliklere sahipse ne kadarına sahip olduğunu, sahiplik ölçüsünü ifade etme durumunu gösterir (Taşkaya ve Meydan, 2012). Değerlendirme ise ölçme sürecindeki amaçlara eğitim öğretim sürecinde ne kadar ulaşıldığının bir nişanesi, süreçte kullanılan öğelerin işleyip işlemediğinin bir ölçüsü varsa işlemeyen noktaların açığa çıkarılmasını sağlayan karar verme sürecidir (Temizkan ve Sallabaş, 2011). Ölçme ve değerlendirme kavramları birbirinden bağımsız düşünülemez olup ölçmeyi anlamlı kılan öğe değerlendirmedir (Baştürk, 2014). Ölçme ve değerlendirme ders kapsamında düşünüldüğünde dersin farklı bölümlerinde de kullanılır. Örneğin; okullarda derste anlatılacak konuya geçmeden önce öğrencilerin hazırbulunmuşluklarını belirlemek amacıyla kullanılabilir. Ayrıca dersin başlangıcında, konunun girişinde ön bilgilerini görebilmek için kullanılabilir. Öğretim süreci devam ederken öğrencinin süreç boyunca o dersten ne kadar verim aldığını görmek amaçlı da kullanılabilir. Eğitim öğretimin her basamağında yapılan farklı tür sınavlarda kullanım alanlarındandır (Özçelik, 1998). Öğrenciye, “Verilen bilginin ne kadarını öğrettik?” sorusunun cevabı, ölçme ve değerlendirme kapsamında ele alınır (Temizkan ve Sallabaş, 2011). Ölçme kavramı hayatın her aşamasında var olmakla birlikte aynı zamanda uzun yıllardır kullanılan köklü bir kavram olarak karşımıza çıkmaktadır (Gönen vd., 2011).

Ölçme işleminde bireylerin veya nesnelerin belirli özelliklere sahip olup olmadığını gözlemlerken, gözlemlere bazen hatalar karışabilir ve bu hata nedeniyle gözlenen bir özelliğin gerçek durumu yansıtılmayabilir. Bu durumda gerçek değer tahmin edilmeye çalışılır (Çelen, 2008). Klasik test teorisi rehber alınarak bu gerçek değerden sapabilecek hatalar için Spearman-Brown yöntemi kullanılmıştır (Baykul, 2015). Bu noktada tahminleri yapabilmek için bazı istatistiksel kuramlar karşımıza çıkmaktadır (Baykul, 2015). İki önemli kuram mevcuttur (Çelen, 2008). Bu kuramlardan biri “*Klasik Test Kuramı*” iken, diğeri “*Madde Tepki Kuramı*”dır. (Çelen ve Aybek, 2013). Klasik test kuramında ölçülmek istenen özelliğin geçerli ve güvenilir bir şekilde ölçülmesi söz konusudur (Çelen, 2008). Aynı zamanda klasik test kuramında öğrencinin başarısı ölçülürken, her soruya verilen doğru cevaplar toplanır. Madde tepki kuramında ise toplama işlemi yapılmaz. Sebebi her soru bazında düşünüldüğünde, soruların her birinin hesaplanan toplam puan için katkısının eşit olmamasıdır. Klasik test kuramında var olan çeldiriciler eşit oranda etkileme sağlarken, madde tepki kuramında çeldiriciler doğru cevapla ilgili, doğru cevaba yakınlığı ile alakalı sıralama mevcuttur. İki kuram için de ortak varsayımlar bulunmaktadır. Örneğin, iki kuramda da bir maddeye verilen cevap diğer maddeleri etkilemez. Ya da iki kuram da testin ölçmeye çalıştığı özelliği ölçer. İki kuram da normallik varsayımını karşılar. Kısaca iki kuramın özellikleri ele alınıp, çalışmada referans alınması bakımından değerlendirildiğinde bazı sonuçlar ortaya çıkmıştır. Test geliştirme süreçlerinde klasik test kuramının madde tepki kuramına göre daha yaygın bir şekilde kullanıldığı görülmüştür. Literatürde birbirine üstünlükleri konusunda net bir karar bulunmamaktadır (Çelen ve Aybek, 2013; Çıtak Gözen, 2010). Mevcut çalışmada klasik test kuramı esas alınarak bir başarı testi geliştirilmeye çalışılmıştır. Sebebi, bütün çeldiriciler için doğru cevaptan eşit oranda farklılaşma kabul edilmiş olup taksonomi kavramı klasik test teorisine göre ortaya çıkarak çoktan seçmeli testler için özellikler belirlenmiştir.

Öğrenciye verilen bilgilerin anlaşılıp anlaşılmadığını, anlaşıldıysa ne kadarının anlaşıldığını belirlemek için bazı ölçme araçları kullanılır. Bu ölçme araçlarından bazıları yazılı yoklamalar, boşluk doldurma, doğru yanlış testleri, kısa cevaplı testler ve çoktan seçmeli testlerdir (Çalışkan ve Yiğittir, 2008). Çoktan seçmeli testler, öğrenci başarısını ölçmek amaçlı öğrenciye belli bir süre verilerek, verilen bilgilerin ne kadarının anlaşılıp anlaşılmadığını görmek amaçlı seçeneklerden oluşan ve bu

seçeneklerden öğrencinin doğru olan cevabı bulması istenen test şeklidir. Çok sayıda soru sormaya uygundur. Günümüzde ekonomikliği nedeniyle çok fazla tercih edilen etkili bir şekilde yoklamaya olanak sağlayan önemli ölçme araçlarından biridir (İpek Akbulut ve Çepni, 2013). Ayrıca çoktan seçmeli testler, testler arasında en çok kullanılan özellikle geniş katılımcı kitlesinin olduğu durumlarda, bilişsel alanın farklı düzeylerinden soru sorulabilmesi, kısa sürede ölçülecek özelliğin ölçülebilmesi, objektif değerlendirmenin olması bakımından önemsenmektedir. Ayrıca hazırlanmasının kolay olması ve kapsam geçerliğinin yüksek olması açısından dünyada çok fazla tercih edilen bir ölçme aracı olarak görülmektedir (Anderson, 2003). Çoktan seçmeli testler Türkiye de dahil olmak üzere birçok ülkede özellikle başarı testleri şeklinde sıklıkla kullanılmaktadır (İpek Akbulut ve Çepni, 2013; Akhter ve Bahoo 2015; Demir, vd., 2016; Sontay ve Karamustafaoğlu, 2020; Narlı ve Baser 2008; Saraç, 2018; Saylan Kırmızıgül ve Kaya, 2019; Şerabatır, 2008). Özellikle uluslararası sınavlarda başarılı olan ülkeler ve hatta uluslararası sınav olan PISA da dahil birçok sınavda çoktan seçmeli testler kullanılmaktadır (PISA, 2018). Çoktan seçmeli testlerin tercih edilme nedenleri arasında özellikle çok fazla kişinin katıldığı ulusal düzeydeki sınavlarda uygulanabilirliğinin kolay olması gelmektedir. Ayrıca değerlendirmenin objektif olması, hızlı bir şekilde ölçülmek istenen özellikleri ölçme imkanı vermesi de tercih edilmesini artırmaktadır (Baştürk, 2014). Çoktan seçmeli testlerin avantajlarının yanında sınırlılıkları da bulunmaktadır. Örneğin üst düzey bilişsel becerileri ölçme konusunda yetersiz kalmaktadır (Haladyna, 1997). Ayrıca şans başarısı ihtimali de mevcuttur. Hazırlanması sırasında uzmanlıkta gerektirmektedir (Baykul, 2015). Ancak sınırlılık durumları kontrol altına alınıp en aza indirgenebilir. Örneğin, kapsam geçerliğini sağlamak adına belirtke tablosu oluşturulabilir. Testin bir kaynaktan direkt alınması önenebilir. Seçeneklerin birbirine ipucu vermemesine dikkat edilebilir. Seçenekler birbirinden bağımsız oluşturulabilir. Maddelerin açık ve net bir şekilde açıklaması yapılabilir (Baykul, 2015). Bu çalışma kapsamında çoktan seçmeli testlerle ilgili yukarıda belirtilen bilimsel veriler göz önünde bulundurularak uygulanabilirliğinin kolay olması, ölçülmek istenen davranışı hızlı ve kolay bir şekilde ölçebilmesi, değerlendirmenin objektif olması gibi amaca hizmet etmesi bakımından çoktan seçmeli bir başarı testi geliştirilmesi amaçlanmıştır. Çoktan seçmeli testler ilk bakışta hazırlanması basit, yalın gibi gözükse de zor, karmaşık ve uzun bir sürece sahiptir. Test soruları hazırlanırken hedeflerin iyi tanımlanması kadar iyi bir ölçme değerlendirme bilgisine sahip olmak da önemlidir (Çolak ve Demircioğlu, 2010). Öğrencilerin başarılarını belirlemek için geliştirilen iyi bir çoktan seçmeli testte, soruların nitelikli olması, ölçülmek istenen hedef davranışı ölçmesi, testteki soruların bütün kazanımları kapsamaması, her kazanımdan soru sorulması, bu doğrultuda soruların iyi bir şekilde organize edilmesi gerekir (Gönen vd., 2011). Soruların iyi hazırlanması ve soruların gruplandırılması sürecinde mevcut taksonomiler içerisinde en iyi bilinen, en çok faydalanılan diğer taksonomilere göre daha çok tercih edilen, Benjamin Bloom'un kendi adıyla 1956'da geliştirilen ve daha sonrasında gelişen, değişen eğitim koşulları sebebiyle güncellenen Bloom taksonomisidir (Anderson, 2005 ve Çakıcı vd., 2012).

Bloom taksonomisinde, bilgiler basitten karmaşığa doğru, alt düzey bilgilerden üst düzey bilgilere şeklinde altı boyutta sınıflandırılmaktadır. Bu boyutlar bilgi, kavrama, uygulama, analiz, değerlendirme ve sentez basamaklarıdır. Bilgi basamağında genelde tanım cümleleri yer alır. Öğrenci üzerine bilgi eklemeyen aynı şekilde aldığı bilgiyi aktarır. Kavrama basamağında bilgiyi kendi cümleleri doğrultusunda açıklaması beklenir. Uygulama basamağında problem çözer. Analiz basamağında bir bütünü parçalara ayırıp bunlar arasında ilişki kurup, bunları yorumlaması önemsenir. Sentez basamağında çok yönlü bakış açısıyla düşünme, orijinal bir ürün oluşturması beklenir. Son olarak değerlendirme basamağında ise ölçütler karşılaştırılıp bir karara varma söz konusudur (Birgin, 2016). Alt düzey bilgiler, daha çok bilgiyi tanıma ve hatırlama düzeyinde öğrendiğini kavrama, anlama ve bildiğini uygulama gibi daha az karmaşık ve temel olan bilgileri içerirken; üst düzey bilgiler öğrencinin öğrendiği tüm bilgileri düşünüp analiz etmesini, verilen bilgilerin hepsini düşünüp karar vermesini, ortaya yeni farklı özgün bir ürün çıkarmasını sağlar (Çakıcı vd., 2012). Güncellenen Bloom taksonomisinde bilgi basamağı "*Hatırlama*", kavrama basamağı ise "*Anlama*" olarak belirtilmiştir. Ayrıca sentez basamağı, değerlendirme basamağı ile yer değiştirip en üst basamağa çıkarılmıştır (Tosun ve Taşkesenligil, 2011). Bloom taksonomisine göre her basamak birbiriyle ilişkili ve birbirinin tamamlayıcısıdır. Bir basamakta bilgi tam olarak öğrenilip kavranmadığı sürece diğer basamaklarda bundan etkilenecek öğrenmede eksiklik yaşanacak ve etkili bir öğrenme

süreci gerçekleştirilemeyecektir (Çakıcı, vd., 2012). Bloom taksonomisinde bulunan her bir basamak farklı bir bilişsel süreci kullanmayı gerektirdiğinden, öğrencilerin başarılarının ölçüleceği çoktan seçmeli testlerde her basamakla ilgili soruların sorulması, soruların bütün konuyu kapsamamasını, bütün konuyu yansıtmamasını sağlayacağından önemlidir (Linn ve Gronlund, 1995). Bu nedenle de bu taksonomi öğretmenler, eğitimciler için önemli ve gerekli görülmektedir (Çolak ve Demircioğlu, 2010).

Bloom'a göre basamaklar arasında değerlendirme basamağındaki sorular önemli bir yer oluşturur. Süreçteki bilgilerin tamamının düşünülüp karar verilmesi, derin öğrenmelerin sağlanması nedeniyle öğretmene, öğrencinin konuyu genel olarak anladığı düşüncesini verir (Çakıcı vd., 2012). Yine Bloom'a göre sentez basamağındaki bilgiler çoktan seçmeli testlerle ölçülemeyeceğinden sentez basamağı soruları için açık uçlu sorular kullanılabilir (Birgin, 2016). Her basamaktan soru sorulması (sentez hariç), taksonominin temel aldığı düşüncelerin çalışma açısından uygunluğu, Bloom'un kullandığı ölçütlerin testin geliştirme amacına hizmet etmesi önemli görülmektedir. Geliştirilen başarı testlerinin basamaklar doğrultusunda hazırlandığında öğretim programı kazanımları ile daha uyumlu olacağı öngörülmektedir. Basamaklar sayesinde soruların programda yer alan içeriği tamamen kapsayıp yansıtacağı gibi düşüncelerden dolayı Bloom taksonomisi bu çalışmanın altyapısını, kuramsal çerçevesini oluşturmaktadır.

Fen eğitimi ile ilgili literatür incelendiğinde, özellikle ortaokul biyoloji konularına yönelik öğrenci başarısını ölçme amaçlı testlerin sınırlı sayıda olduğu görülmektedir. Biyoloji konularına yönelik bazı çalışmalar; Genetik (Çakır ve Aldemir, 2011); Solunum ve Fotosentez (Tekkaya ve Balcı, 2003); Hücre bölünmesi ve Kalıtım (Kızıkan ve Bektaş, 2018); Hücre (Timur vd., 2019) şeklindedir. Sindirim sistemi konusu ile ilgili test geliştirme çalışmaları da yine literatür tarandığında karşımıza çıkmaktadır (Kaya ve Gül, 2020; Güngör ve Özkan, 2017 ve Kahyaoğlu, 2005).

Fen eğitimiyle ilgili literatürdeki çalışmalar incelendiğinde altıncı sınıf "*Sindirim Sistemi*" konusu ile ilgili kazanımların tamamını içeren, her kazanımın ölçülmeye çalışıldığı, Bloom taksonomisinin her basamağından (sentez hariç) soruların sorulduğu geçerli ve güvenilir bir başarı testine rastlanmamıştır. Ayrıca bir önceki yani 2013 yılındaki fen bilimleri öğretim programındaki vücudumuzdaki sistemler ünitesi ile 2018 yılındaki fen bilimleri öğretim programındaki vücudumuzdaki sistemler ünitesinin kapsamı farklılaşmaktadır. 2013'te konu yedinci sınıflarda verilirken, 2018'de altıncı sınıflara çekilerek kazanımlarda ve kapsamda değişiklikler yapılmıştır. Ayrıca kazanımların açıklamalarında da birtakım genişletmeler mevcuttur (MEB, 2013; 2018). Bu sebeple 2018 yılından önce sindirim sistemi konusunda geliştirilen testler bu bağlamda düşünüldüğünde yetersiz kalmaktadır. Açıklanan sebeplerden yola çıkıldığında, güncellenen 2018 öğretim programına uygun, kazanımları ve kapsamı değişen sindirim sistemine yönelik, bilişsel düzeyde Bloom taksonomisine uygun hazırlanmış aynı zamanda üst düzey becerileri de yoklayan teste ihtiyaç bulunmaktadır. Ayrıca araştırmacılara ve öğretmenlere kullanım açısından elverişli olan ve öğretmenler arasında çokça tercih edilen çoktan seçmeli bir ölçme aracının gerekliliği uygun görülmüştür. Mevcut çalışma kapsamında, ortaokul altıncı sınıf öğrencilerinin "*Vücudumuzdaki Sistemler*" ünitesine ait "*Sindirim Sistemi*" konusundaki akademik başarılarını ölçmeye yönelik geçerli ve güvenilir bir ölçme aracı geliştirilmesi amaçlanmaktadır. Bu amaç doğrultusunda, aşağıda verilen alt problemlere cevap aranmıştır:

1. Altıncı sınıf öğrencilerinin "*Sindirim Sistemi*" konusundaki başarılarını belirlemek amacıyla geliştirilen çoktan seçmeli test geçerli midir?
2. Altıncı sınıf öğrencilerinin "*Sindirim Sistemi*" konusundaki başarılarını belirlemek için geliştirilen çoktan seçmeli test güvenilir midir?

YÖNTEM

Çalışma Deseni

Araştırma, altıncı sınıf "*Vücudumuzdaki Sistemler*" ünitesindeki "*Sindirim Sistemi*" konusu kapsamında geçerli ve güvenilir çoktan seçmeli bir başarı testi geliştirme çalışmasıdır. Nicel araştırma desenlerinden tarama deseni kullanılmıştır. Tarama deseni, var olan durumu ortaya çıkarmak amaçlı

yapılan bir araştırma türüdür (Çepni, 2010). Tarama deseni kullanılarak yapılan çalışmalarda bir evrenin içinden seçilen örneklem grubu üzerinde yapılan çalışmalar vasıtasıyla evrenin genelindeki durum, davranış veya düşünceler istatistiksel olarak tasvir edilir (Creswell, 2013). Çalışmada, altıncı sınıfların sindirim sistemi konusu ile ilgili genel durumunu geniş gruplar üzerinden ortaya çıkaracak, başarı düzeyi hakkında bilgi verecek, geçerli ve aynı zamanda güvenilir ölçme aracı geliştirilmek istendiğinden dolayı tarama deseni tercih edilmiştir.

Araştırma ve Yayın Etiği

İnönü Üniversitesi Etik Kurulunun 2021/10-12 tarih ve 10 oturum sayılı yazısı ile etik açıdan uygun görülmüştür.

Örneklem / Çalışma grubu

Araştırmanın ulaşılabilir evrenini Şanlıurfa Siverek ilçesinde öğrenim gören ortaokul altıncı sınıf öğrencileri, örneklemi ise; rastgele örneklem türünden küme örnekleme ile belirlenen 2020-2021 eğitim öğretim yılının ikinci döneminde Şanlıurfa Siverek ilçesinde seçilen ortaokullarda altıncı sınıfta öğrenim gören 229 öğrenci oluşturmaktadır. Küme örnekleme kullanılarak büyük bir evren kendisinden daha küçük ve çalışılabilir parçalara ayrılmaktadır (Patton, 2005). Bu örneklemin seçilme nedeni; evrenin büyük olmasından kaynaklı gruplama yapılarak daha küçük gruplarla çalışılmak istenmesidir. Çalışma grubu büyüklüğü belirlenirken analiz işlemleri düşünülerek soru sayısının 10 katı ve evrende bulunan kişi sayısının %10'u olarak belirlenmiştir (Bernard, 2011). Bu çalışma kapsamında 25 maddenin yer aldığı başarı testi, soru sayısının dokuz katı örnekleme uygulanmıştır. Bu açıdan teste bakıldığında geçerlik ve güvenilirlik analizi için örneklem sayısının yeterli olacağı düşünülmektedir.

Veri Toplama Araçları

Altıncı sınıf "*Vücudumuzdaki Sistemler*" ünitesinde "*Sindirim Sistemi*" konusunda geçerli ve güvenilir bir başarı testi oluşturulurken sorulara karar vermek için öncelikle 2018 Fen Bilimleri Dersi Öğretim Programı temel alınmıştır. Sorular, bu öğretim programında yer alan "*Canlılar ve Yaşam*" konu alanındaki "*Vücudumuzdaki Sistemler*" ünitesinde "*Sindirim Sistemi*" konusundaki kazanımlar baz alınarak oluşturulmuştur. Daha sonra öğretim programındaki kazanımlara ek olarak soru bazlı kazanım yazılmıştır. Bu kazanımlar yazılırken literatürden yararlanılmıştır (Baladın Duman, 2019; Kurtcuoğlu, 2007; Kaya ve Gül, 2020). "*Sindirim Sistemi*" ile ilgili yapılan literatürde bulunan ortaokul ve lise olmak üzere tüm çalışmalar incelenmiştir (Bozdağ, 2017; Güngör ve Özkan, 2017; Kaya ve Gül, 2020). Kazanımlara uygun olan sorular seçilip gerekli düzenlemeler yapıldıktan sonra başarı testi kapsamına alınmıştır. Örneğin, özellikle lise düzeyindeki sorular ortaokul öğrencileri için anlaşılır hale getirilmiş olup seçenekler dört seçenek olmak üzere değiştirilmiştir. "*Sindirim Sistemi*" başarı testi için birinci yazar tarafından oluşturulan soru bazlı kazanımlar da kapsam geçerliği kısmında ayrıntılı olarak belirtilmiştir. Sorular oluşturulurken özellikle Bloom taksonomisinde yer alan tüm basamaklardan kazanımları ölçen sorular olmasına özen gösterilmiştir. Bu çalışma kapsamında altıncı sınıf sindirim sistemi ünitesiyle ilgili olarak düzey belirleyici (summative) olma özelliği taşıyan çoktan seçmeli 25 soru içeren geçerli ve güvenilir başarı testi geliştirilmesi amaçlanmıştır. Bu test Google forms uygulaması ile düzenlenmiş ve düzenlenmiş hali Siverek ilçesinde bulunan devlet ortaokullarına yine Google form aracılığıyla gönderilmiştir. Öğrenciler tarafından soruların tamamı işaretlenmiştir. Ölçüt test olarak literatürden geçerliği ve güvenilirliği kanıtlanmış, kazanımlarla uyumlu bir test kullanılmıştır. Her iki test için de geçerlik ve güvenilirlik çalışması yapılmış ve bulgular kısmında bu çalışmaya ait veriler sunulmuştur.

Verilerin Analizi

Çalışmada elde edilen verilerin analizi için başarı testinin geçerlik ve güvenilirlik çalışmaları yapılmıştır. Bu kapsamda betimsel istatistikler, madde analizi (güçlük, ayırt edicilik), tetrakorik faktör analizi (AFA ve DFA) incelenmiştir. AFA ve DFA aynı örneklem üzerinde uygulanmıştır. Analizlerde

FACTOR, MS EXCEL 2020 ve SPSS.25 paket programları kullanılmıştır. SPSS.25'te yapılan betimsel istatistiksel analizlerin sonuçları ile ölçme aracından elde edilen verilerin normal dağılım gösterip göstermediği incelenmiştir. Bunun için öncelikle öğrencilerin verdiği yanıtlar "1" ve "0" olarak puanlanmış doğru yanıt için "1" yanlış yanıt için "0" kodu girilmiştir. Bu testte 25 soru yer aldığı için testten alınabilecek en yüksek puan "25" iken en düşük puan ise sıfır olarak belirlenmiştir. Doğru cevapları yanlış cevaplar etkilememektedir. Sırasıyla veriler normallik analizi, geçerlik güvenirlik indeksleri kapsamında madde güçlük ve madde ayırt edicilik indeksi ilişkisiz örneklem t testi, güvenirlik analizi, açımlayıcı faktör analizi (AFA) ve doğrulayıcı faktör analizi (DFA) yapılmıştır. Öğrencilerin aldıkları puanlar alt ve üst gruplar şeklinde Excel programında puan sıralamalarına göre sıralanmış olup SPSS.25 paket programında ilişkisiz örneklem t testi analizi yapılmıştır. Çalışmada güvenirlik analizini gerçekleştirmek için KR-20 değerine bakılmıştır. Başarı testlerinde KR-20 değeri yaygın olarak kullanılan yöntemdir (Barchard ve Hakstian, 1997). KR-20 katsayısı testte yer alan soruların homojen bir şekilde bir bütün olup olmadığını anlamamızı yardımcı olur. KR-20 değeri iki değerli ölçümlenmiş maddeler için uygundur (Crocker ve Algina, 1986).

BULGULAR

Normallik Analizi Bulguları

Yapılan analizler sonucunda, 229 öğrencinin almış olduğu puanların normal dağıldığı belirlenmiştir. Testte yer alan sorular için basıklık ve çarpıklık değerleri bulunmuş olup bu değerler genel olarak -2 ile +2 değerleri arasında dağılım gösterdiğinden 229 öğrencinin testte yer alan her bir sorudan almış olduğu puanların normal dağıldığı söylenebilir (George ve Mallery, 2010). Testte yer alan sorulardan beşinci, sekizinci, 16., 18. ve 22. soruların basıklık değerleri -2'den çok az bir farkla fazla olduğu görülmekte olup bu soruların aritmetik ortalama mod medyan değerlerinin birbirine yakın olduğu ve diğer analizlerinde problem olmadığı için sıkıntı teşkil etmediği söylenebilir.

Geçerlik Analizi Bulguları

Kapsam Geçerliği

Testin Amacının Belirlenmesi

Geliştirilen başarı testi, güncellenen fen bilimleri öğretim programı kapsamında, altıncı sınıf "*Vücudumuzdaki Sistemler*" ünitesi "*Sindirim Sistemi*" konusuna yönelik, öğrencilerin öğrenme düzeyleri açısından uygun, konu bitiminde düzey belirleme yani "*Summative*" amacıyla kullanılabilir. Bu bağlamda bu test bu konuda araştırma yapacak araştırmacılara, konu hakkında soru arayacak öğrenciye, dersin sonunda dersteki bilgilerin ne kadarının öğrenilip öğrenilmediğini anlamak için ölçme yaparken öğretmene faydalı bir içerik sunması amaçlı hazırlanmıştır.

Konunun Belirlenmesi

Günlük hayatta içinde bulunduğumuz ve daha önce geliştirilen başarı testlerinde yer bulmayan "*Sindirim Sistemi*" konusuna ait alt konu ve kavramlar bu çalışma kapsamında testin içeriğine alınmıştır. Bu konuda öğrencilerin sindirim sistemine ait yapı ve organlara ilişkin bilgi beceri kazanmaları, bu konu kapsamında verilen kavramları öğrenme, kendi yaşamları için çok elzem olan sindirim sisteminin vücutta nasıl gerçekleştiğini anlamaları, vücutlarına ve yaşamlarına karşı bir farkındalık oluşturulması amaçlanmaktadır. (MEB, 2018). Sindirim sistemi başarı testinin kapsam geçerliği kapsamında öncelikle 2018 Fen bilimleri öğretim programında yer alan kazanımlar belirlenmiştir. Bu kazanımlar aşağıda belirtilmiş olup kazanımların hepsi dikkate alınmıştır.

F.6.2.2.1. Sindirim sistemini oluşturan yapı ve organların görevlerini modeller kullanarak açıklar.

F.6.2.2.2. Besinlerin kana geçebilmesi için fiziksel (mekanik) ve kimyasal sindirime uğraması gerektiği çıkarımını yapar.

F.6.2.2.3. Sindirime yardımcı organların görevlerini açıklar.

Daha sonra 2018 Fen Bilimleri Öğretim programında altıncı sınıf sindirim sistemi konusunda konunun tüm detaylarıyla ölçülmesi, bu konuda programdaki kazanımların yetersiz olması ve Bloom taksonomisindeki tüm basamaklara uygun kaynaklı kazanımların olmamasından dolayı ana kazanımlara bağlı kalınarak soru bazında ek kazanımlar yazılmıştır.

Daha sonra kazanımların ve soruların hangi düzeyleri ölçtüğüne yönelik bir belirtke tablosu hazırlanmıştır. Bu belirtke tablosu başarı testi uygulanmadan önce tüm konu kapsamının görülmesi ve kazanımlarla soru sayısının, taksonomik basamağın uyuşup uyuşmadığının görülmesi adına hazırlanmıştır (Tablo 1).

Tablo 1. Sindirim sistemi konulu başarı testi soru havuzuna ait belirtke tablosu

Kazanımlar	Bilgi	Kavrama	Uygulama	Analiz	Değerlendirme
Vücuda alınan besinlerin sindirimi sırasında izlediği yolu sıraya koyar.	1.Soru				
Sindirim sistemini oluşturan yapı ve organların görevlerini açıklar.		2.Soru			
Sindirim sisteminde sindirimin gerçekleştiği organları model üzerinde gösterir.			3.Soru		
Sindirim sisteminin görevlerini açıklar.		4.Soru			
Sindirim sistemi organlarının görevleriyle ilgili analiz yapar.				5.Soru	
Ağız, mide ve ince bağırsağın ortak yönlerini açıklar.		6.Soru			
Tükürük sıvısının görevlerini ifade eder.		7.Soru			
Kimyasal ve fiziksel sindirimin gerçekleştiği organları şekil üzerinde gösterir.			8.Soru		
Safra kesesinin yağların sindirimi için çok önemli olduğunu ayırt eder.				9.Soru	
Fiziksel sindirimin gerçekleştiği olayları açıklar.		10.Soru			
Sindirimde görevli organları ve görevlerini tablodan çıkarım yapar.		11.Soru			
Vücudun organlarını ve görevlerini model üzerinde gösterir.			12.Soru		
Sindirim sisteminde sindirimin gerçekleştiği yerleri bilir.	13.Soru				
Sindirim sisteminin amacı konusunda en doğru tanıma karar verir.					14.Soru
Besinlerin kana geçebilmesi için fiziksel (mekanik) ve kimyasal sindirime uğraması gerektiğini analiz eder.				15.Soru	

Karbonhidrat ile ilgili öğrendiği bilgileri, farklı örnekler üzerinden analiz eder.	16.Soru
Sindirimle ilgili olayları kavrar.	17.Soru
Şekil üzerinde gösterilen sindirime yardımcı organların, görevine en doğru şekilde karar verir.	18.Soru
Besinlerin sindiriminde enzimlerin ve yüzey alanlarının etkisini analiz eder.	19.Soru
Besinlerin kana geçebilmesi için sindirime uğraması gerektiğini ve sindirime uğramayan besini analiz eder.	20.Soru
Kimyasal sindirimin gerçekleştiği organları analiz eder.	21.Soru
Besinlerin sindirilme sürelerini, yüzey alanlarına bakarak karar verir.	22.Soru
Sindirim sisteminde görevli organ ve yapıların görevlerini analiz eder.	23.Soru
Kimyasal sindirimin hangi organlarda ve hangi salgı veya enzimlerle gerçekleştiğini ayırt eder.	24.Soru
Sindirilmeden kana geçen besinleri farklı besin grupları üzerinde uygular.	25.Soru

Belirtke tablosu hazırlanırken bu süreçte bir ölçme uzmanından, iki fen bilgisi öğretmeni ve bir de biyoloji uzmanından dönütler alınarak, düzeltmeler yapılarak tablo oluşturulmuştur. Böylece testin kapsam geçerliğini artırmak amaçlanmıştır. Testin kapsamının düşük olması geçerliği ve güvenilirliği tehlikeye sokacağından bu konunun üzerinde durulması gereken önemli bir nokta olduğunu söyleyebiliriz. Çoktan seçmeli test kapsamında soruların seçiminde Bloom taksonomisine göre kazanımlara uygun olacak şekilde her düzeyden mutlaka soru sorulmuştur (sentez hariç). Özellikle testte üst düzey soruların olduğu uygulama ve analiz sorularına daha fazla yer verilmiştir. Kazanımlara uygun olarak sorular basamaklar açısından dengeli bir şekilde oluşturulmuştur.

Soru Havuzunun Oluşturulması

Sindirim sistemi başarı testi soruları oluşturulurken öncelikle 2018 fen bilimleri öğretim programındaki kazanımlar dikkate alınarak literatürde bulunan konuyla ilgili başarı testleri araştırılmıştır. Bulunan testler incelenmiş olup çoktan seçmeli testler belirlenmiştir. Ortaokul düzeyinde çoktan seçmeli sorulardan oluşan sindirim sistemi ile ilgili çok sınırlı çalışma bulunduğundan lise düzeyinde bu konuyla ilgili literatürdeki çalışmalar incelenmiştir. İncelenen testler içerisinde 2018 fen bilimleri öğretim programında yer alan sindirim sistemi konusuyula ilgili kazanımların paralel olduğu sorular alınmıştır. Bu sorular daha sonra ortaokul düzeyine uyarlanıp testte kullanılmaya uygun hale getirilmiştir. Daha sonra fen bilimleri dersi sindirim sistemi ile ilgili oluşturulan testler incelenerek kazanımlara uygun konuyla ilgili konuyu kavratacak sorulardan yararlanılmıştır. Araştırmacı tarafından da hazırlanan soruların bulunduğu testte testin ilk halindeki soru sayısı 44'tür.

Uzman Görüşünün Alınması

Sindirim sistemiyle ilgili 44 soruluk hazırlanan başarı testi beş Fen bilgisi öğretmenine ve bir fen eğitimi alanında Dr. Öğretim üyesine, soruların kazanımlara ve öğretim programına uygunluğuna bakılması için sunulmuştur. Bir tane Biyoloji uzmanına soruların kazanımlara ve ortaokul seviyesine uygunluğu ve kavram yanlışlığı oluşturabilecek soruların olup olmadığını anlamak için gösterilmiştir. Bir tane Ölçme değerlendirme uzmanına testin mümkün olduğu kadar geçerlik ve güvenilirliğini olumsuz etkileyecek etmenlerden arındırılmasını sağlamak ve soruların bilimsel açıdan doğruluğunu basamaklarla kazanımların uyuşup uyuşmadığını kontrol etmek için başvurulmuştur. Son olarak üç tane Türkçe öğretmenine soruların açık net anlaşılır olup olmadığı, sorularda imla yazım kurallarının kontrolü açısından değerlendirmesi amacıyla gösterilmiştir. Alınan dönütlere göre test kapsamında değişiklikler yapılmıştır. Bazı soruların kazanıma göre basamağı değiştirilmiştir. Bazı sorularda Türkçe açısından cümlede anlatım bozukluğuna dikkat edilerek eklemeler veya çıkarmalar yapılmıştır. Bir tane soru öğrenci düzeyine uygun olmaması kafa karıştırıcı olması sebebiyle testten çıkarılmıştır. Uzman görüşleriyle ilgili dönüt düzeltmeler yaklaşık olarak üç hafta gibi bir süreçte gerçekleştirilmiştir. Ayrıca uzman görüşleri için kontrol listesinden yararlanılmıştır. Tüm bu dönüt düzeltmeler sonucunda sindirim sistemi başarı testi, öğrencilerin bulunduğu yaş grubu özellikleri de dikkate alınarak 25 soruya indirilmiştir.

Ölçme ve Değerlendirme Uzmanının Önerileri

Ölçme uzmanı tarafından alınan dönütler aşağıda belirtilmiştir.

- ✓ Değerlendirme basamağı ile bilgi basamağı sorularının soru sayısının birbirine yakın olması gerektiği,
- ✓ Soru sayısının 20' den 25 ya da 30'a çıkarılması gerektiği,
- ✓ Sorulara ait bazı kazanımlarda eksiklik olduğu ve kazanımlara ekleme yapılması gerektiği,
- ✓ Belirtke tablosunda bazı eksik kısımların olduğu ve onların düzeltilmesi gerektiği,
- ✓ Görsel olarak verilen sorularda görsele başlık ekleme ve soru kökünün doğru ifade ederek sorma gerekliliği,

SORU 16: Ayşe fen bilimleri dersinde makarna, börek, ekmeğ gibi besinlerin karbonhidrat olduğunu ve karbonhidratların da sindirimini ağızda başladığını öğrenmiştir. Ayşe'nin bu öğrenmiş olduğu bilgileri aşağıdakilerden hangisiyle ilişkilendirmesi bilgiyi transfer ettiği anlamına gelmektedir?

- A) Pirinç ve balık
- B) Süt ve yumurta
- C) Patates ve poğaç
- D) Zeytin ve peynir

Bilgiyi transfer etmek? Bu soru kökü yerine hangisiyle ilişkilendirirse konuyu anladığı anlaşılır şekilde düzenlenebilir.

Şekil 1. Fen bilimleri öğretmeninden gelen dönüt örneği

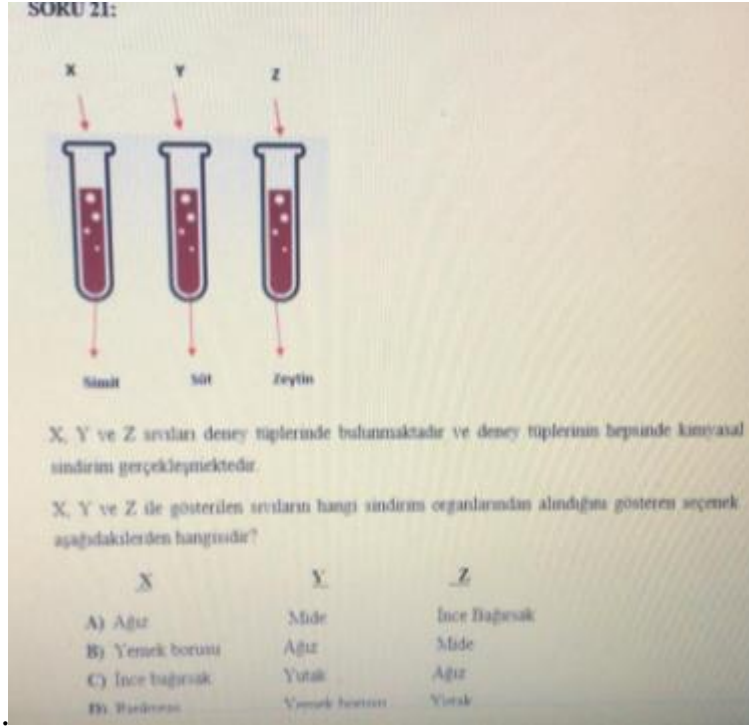
- ✓ Bazı basamakların kazanıma uygun olmadığı, basamağın değiştirilmesinin gerekliliği, (5., 6., 13., 15. sorular gibi). Örnek,

SORU 13: Aşağıdakilerden ifadelerden hangisi yanlıştır?
A) Anüste sindirim gerçekleşmez.
B) Sindirimin başladığı yer midedir.
C) Yağların sindirimi ağzın içinde başlar.
D) Suyun sindirimi ince bağırsakta başlar.

Hocam bu soru bilgi basamağı C şıkında ağızda başlar diyelim.

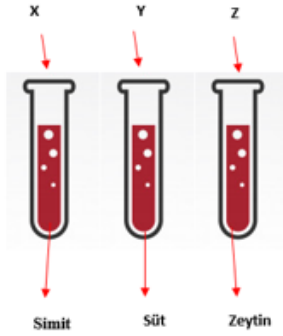
Şekil 2. Fen bilimleri öğretmeninden gelen dönüt örneği

- ✓ Bazı kazanımların değiştirilmesi gerektiği (örnek: 13. Soru gibi)
- ✓ Görsellerde vurgu yapılması gereken yerlerde düzeltme yapılması, (Örnek, soru 21)



Şekil 3. Sorunun ilk hali

SORU 21:



Görsel – 5: Kimyasal sindirimin gerçekleştiği deney tüpleri

Görsel – 5 'te verilen deney tüplerinde sırasıyla simit, süt ve zeytin bulunmaktadır. Farklı sindirim organlarından alınan X, Y ve Z sıvıları görseldeki gibi deney tüplerine sırasıyla ilave edilmektedir. Deney tüplerinin hepsinde kimyasal sindirim gerçekleştiği gözlemlenmektedir.

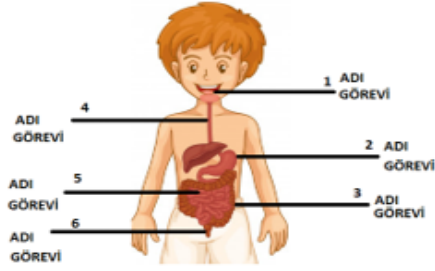
Buna göre X, Y ve Z ile gösterilen sıvıların hangi sindirim organından alındığı, aşağıda verilen seçeneklerden hangisinde doğru verilmiştir?

<u>X</u>	<u>Y</u>	<u>Z</u>
A) Ağız	Mide	İnce Bağırsak
B) Yemek borusu	Ağız	Mide
C) İnce bağırsak	Yutak	Ağız
D) Pankreas	Yemek borusu	Yutak

Şekil 4. Fen bilimleri öğretmeni tarafından verilen dönüt düzeltme sonucunda teste alınan sorunun son hali

- ✓ Bazı sorularda vurgulanması gereken yerlerin altının çizilip koyu yazılması gerektiği,
- ✓ Görsel kullanımı gerektirmeyen sorularda görselin çıkarılması gerektiği,
- ✓ Yazı tipinin farklı olduğu yerde yazı tipinin düzeltilmesi gerektiği,
- ✓ Görünüşte teste uygun olmayan sorular için görünüşün düzeltilmesi gerektiği, (Örnek, 12. soru)

SORU 12:



Görsel - 3

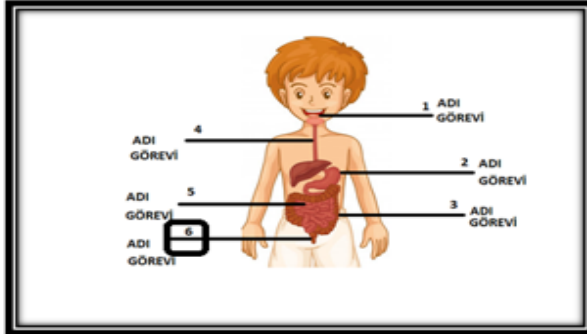
Görsel – 3’de vücudun organları ve görevleri gösterilmiştir.

Görselde 6 numarayla gösterilen organ ve görevi aşağıdakilerden hangisinde doğru verilmiştir?

- A) Anüs-Besin atıklarının dışarı atıldığı yerdir.
- B) Pankreas-Karbonhidrat, protein, yağların sindirimini sağlar.
- C) İnce bağırsak-Bütün besinlerin sindiriminin tamamlanmasını sağlar.
- D) Kalın bağırsak-Fazlalık su, vitamin ve minerallerin emilimi gerçekleşir.

Şekil 5. Sorunun testteki ilk hali

SORU 12:



Görsel – 3: Vücudun organları ve görevleri

Görsel – 3’de vücudun organları ve görevleri gösterilmiştir.

Görsel – 3 ‘ te 6 numarayla gösterilen organ ve görevi aşağıdakilerden hangisinde doğru verilmiştir?

- A) Anüs-Besin atıklarının dışarı atıldığı yerdir.
- B) Pankreas-Karbonhidrat, protein, yağların sindirimini sağlar.
- C) İnce bağırsak-Bütün besinlerin sindiriminin tamamlanmasını sağlar.
- D) Kalın bağırsak-Fazlalık su, vitamin ve minerallerin emilimi gerçekleşir.

Şekil 6. Testin dönüt ve düzeltmelere uygun olarak değiştirilmiş son hali

- ✓ Anlaşılır olmayan soruların düzeltilmesi gerektiği,



Şekil 7. Biyoloji konusunda uzmandan gelen dönüt örneği

Düzeltilme yapılmayan sorularda kazanım ve sorunun uyumlu olduğu ve taksonomik sürecin uyumlu olduğu, gibi konularda dönütler alınmıştır.

Testin Ölçüt Geçerliğinin İncelenmesi

Ölçüt geçerliği için, iki tane teste ihtiyaç vardır. Bu testlerden biri uygulanacak olan asıl test iken diğeri daha önceden oluşturulmuş geçerliği ve güvenilirliği belirlenmiş ölçüt alınan testtir. Burada asıl test ile ölçüt alınan test soruları farklı olmalı, kazanımlar aynı olmalıdır. Bunun sebebi ileriye yönelik tahminlerde ölçütün, geçerli ve güvenilir olduğu konusunda şüphe oluşmasını. Ayrıca yapılan testler eş zamanlı uygulanabilir (Çakmur, 2012).

Bu araştırma kapsamında yapılacak olan ölçüt geçerliği testi için literatür taranmıştır. Taramalar sonucunda daha önce ortaokul yedinci sınıf kapsamında "7. Sınıf Sindirim Sistemi Konusunda İki Aşamalı Test Geliştirilerek Kaoram Yamlıklarının Tespit Edilmesi" isimli uygulanan geçerli ve güvenilirliği belirlenmiş olan test referans alınmıştır. Sindirim sistemi ile ilgili olan 20 soruluk ölçüt test kullanılmıştır. Bu testte yer alan kazanımlar araştırmacının testindeki kazanımlarla uyumlu, sorular birbirinden farklı, ölçülen kademenin ikisinin de aynı olduğu gibi gerekçelerle bu test ölçüt test olarak belirlenmiştir. Ölçüt testi Ek 2'de sunulmuştur.

Madde İndeksleri

Başarı testinde yer alan sorular için geçerlik indekslerine bakılmıştır. Bu kapsamda her soru bazında ayrı ayrı madde güçlük indeksleri ve madde ayırt edicilik indeksleri hesaplanmıştır. Bu indeksleri hesaplamak için öncelikle Excel programında veriler büyükten küçüğe doğru sıralanmış ve teste katılan öğrencilerin %27'lik alt grubu ve %27'lik üst grubu belirlenmiştir. Yüksek puan alanlar üst grubu, düşük puan alanlar ise alt grubu oluşturmaktadır. Burada toplam alt grup ve üst grup belirlenirken 229 öğrenci arasından puanı en yüksek ve en düşük olan %27'lik öğrencilerin puanları alınmıştır. Bu değerlerin dışında kalan %46'lık kısım dikkate alınmamıştır (Hasançebi, 2020). 229 öğrenci için %27'si hesaplandığında 61,8 karşılık gelmekte olup 62 kişi olarak alınmıştır. Bu nedenle 62 tane en iyi puan alan, 62 tane de en düşük puan alan öğrenci seçilerek toplam 124 kişinin verileri ile madde güçlük ve madde ayırt edicilik indeksleri hesaplanmıştır. Arada kalan 105 öğrencinin verileri bu hesaplamalar için kullanılmamıştır. Her bir soru için madde güçlük indeksi (p) ve madde ayırt edicilik indeksi (r) hesaplamaları yapılırken bu çalışmada SPSS.25 programı kullanılmıştır. Taslak başarı testine ait olarak soruların geçerlik indeksleri formülle hesaplanmış olup geçerlik indeksi değerleri Tablo 2'de sunulmuştur.

Tablo 2. Taslak başarı testi için geçerlik değerleri

Sorular	Gruplar	Doğru Yanlış Sayısı		Güçlük (p)	Ayırt Edicilik (r_{jx})
		Doğru	Yanlış		
S1	Üst Grup (%27= 62)	62	0	0,70 kolay	0,68 çok iyi
	Alt Grup (%27= 62)	25	37		
S2	Üst Grup (%27= 62)	62	0	0,58 kolay	0,87 çok iyi
	Alt Grup (%27= 62)	10	52		
S3	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S4	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S5	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S6	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S7	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S8	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S9	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S10	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S11	Üst Grup (%27= 62)	62	0	0,52 kolay	1,00 çok iyi
	Alt Grup (%27= 62)	3	59		
S12	Üst Grup (%27= 62)	62	0	0,50 orta	0,96 çok iyi
	Alt Grup (%27= 62)	0	62		
S13	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		

S14	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S15	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S16	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S17	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S18	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S19	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S20	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S21	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S22	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S23	Üst Grup (%27= 62)	62	0	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		
S24	Üst Grup (%27= 62)	62	0	0,36 zor	0,77 çok iyi
	Alt Grup (%27= 62)	0	62		
S25	Üst Grup (%27= 62)	45	17	0,50 orta	1,00 çok iyi
	Alt Grup (%27= 62)	0	62		

Tablo 2 incelendiğinde testte yer alan her bir soru için üst grup ve alt gruptan teste katılan sayısı, sorulara üst grup ve alt gruptan doğru yanıt verenlerin sayısı görülmektedir. Genel olarak madde güçlüğünün orta güçlükte olduğu, kolay ve zor soruların da testte yer aldığı görülmektedir. Madde ayırt edicilik değerlerinin ise 0,77 üstü yani ayırt etme gücünün çok iyi olduğu görülmektedir (Hasançebi, 2020).

Testin ayırt ediciliğine kanıt sunmak amacıyla ayırt edicilik indeksi hesaplamalarına ek olarak ilişkisiz örneklem t testi de yapılmıştır. İlişkisiz örneklem t testinin kullanılabilmesi için gerekli olan

varsayımlar vardır. Bu varsayımlardan biri normal dağılım sağladıktan sonra yeterli örnekleme ulaşılmış ulaşılmadığıdır. Bu çalışmadaki soruların hücre bazında basıklık ve çarpıklık değerleri incelendiğinde değerler +2 ile -2 aralığında olduğundan testten alınan puanların normal dağılım gösterdiği görülmekte olup varsayım karşılanmıştır (George ve Mallery, 2010). Yeterli örnekleme varsayımına bakmak için KMO değerleri hesaplanmış ve yorumlanmıştır. Bu çalışmada KMO değerinin 0,73 olduğu görülmüştür. KMO değeri 0,60 ve üzeri olduğu zaman yeterli örnekleme sağlanmış olacağından bu test için bu durum da sağlanmıştır (Büyüköztürk, 2012). Barlett testinin de anlamlı çıktığı görülmektedir. Burada önemli olan iki farklı grubun ortalamaları arasında anlamlı fark olup olmadığının tespit edilmesidir (Büyüköztürk, 2012). Burada iki farklı grup olarak üst ve alt grup arasında anlamlı farklılık olup olmadığına bakılmıştır. Bu anlamlı farklılığının üst grubun lehine olması önem arz etmektedir (Özçelik, 2010). Normallik varsayımı sağlandıktan sonra varyansların homojenliğine bakılmış ve homojenliği sağladığı tespit edilmiştir. Varyansların homojenliği kısmında ise Levene İstatistiğine bakılır. Alt grup ve üst grup varyanslarının eşit olup olmadığı tespit edilir. Eğer burada anlamlılık değeri (p) 0.05'ten büyük olursa alt grup ve üst grup arasında anlamlı farklılık yoktur (Kalaycı, 2010). Sonuç itibarıyla üst grupta olan öğrencilerin her bir soru bazında ortalamalarının alt grupta olan öğrencilerden anlamlı olarak yüksek çıktığı görüldüğünden taslak testte yer alan her bir madde için ayırt edicidir, yorumu yapılabilir.

Yapı Geçerliği Bulguları

Yapı geçerliği, testin gerçekten ölçülmek istenilen yapıyı ölçüp ölçmemesine odaklandığı için diğer geçerlik türleri de bu geçerliğin bir parçasını oluşturmaktadır (Kelecioğlu ve Göçer Şahin, 2014). Madde analizleri yapıldıktan sonra faktör analiz programıyla hem Açımlayıcı faktör analizi (AFA) hem Doğrulamalı faktör analizi (DFA) aynı program ile aynı örnekleme üzerinden yürütülmüştür.

Açımlayıcı Faktör Analizi

Açımlayıcı faktör analizi aynı yapıyı ya da değişkenleri ölçen soruları bir araya getirerek ölçme işlemini daha az faktörle açıklama işidir (Akbaş vd., 2019). Faktör döndürmesinde iki farklı yöntem kullanılır olup en yaygın olanı dik döndürme yöntemidir. Faktörlerin birbiriyle bağımlı olduğu durumda kullanılır (Saraçlı, 2011). Yapılan bu çalışmada faktörlerin birbiriyle bağımlı olduğunun düşünülmesinden kaynaklı dik döndürme (Raw Warimax) yöntemi kullanılmıştır. Bu çalışmada AFA sonuçlarına bakıldığında testteki soruların hangi faktörler altında toplandığı incelenmiş olup yedi, dört, üç, iki ve bir faktörlü analizler yapılmış ve en iyi analizin iki faktörlü analizde olduğu belirlenmiştir. Her faktörün altında en az üç soru olması, açıklanan toplam varyans yüzdesinin 0,40'ın altında olmaması, soruların binişik olmaması gibi durumlar faktör analizinde önemlidir (Çakır, 2014). Özellikle birbiriyle alakasız olan değişkenlerin sayısının azaltılması, ilgili soruların bir faktör altına girmesi bunları açıklayabilme yorumlama açısından bakıldığında diğer analizlerde bu husus sağlanmamış olup iki faktörlüde sağlanmıştır. Üç, dört ve yedi faktörlü yapıların açıklanan varyansın yüksek olmasına rağmen faktör altına düşen soruların sayısının yetersiz olması ve bazı soruların hiçbir faktörün altına girmediği görüldüğünden dolayı iki faktörlü yapı işleme alınmıştır. Fakat iki faktörlü yapıda da dokuzuncu, onuncu ve 19. sorunun herhangi bir faktör altına girmediği görülmektedir. Bu üç soru çıkarılıp tekrar faktör analizi yapılmıştır. Faktör analizi sonucunda her sorunun bir faktörün altında yer aldığı görülmüştür. Daha sonra testte yer alan soruların binişikliğine bakılmıştır. Binişiklik için faktör yükleri arasında en az 0,10'luk farkın olması durumu göz önünde bulundurulur (Çakır, 2014). Testte yer alan soruların binişiklik durumları incelendiğinde 7,15,16,20 ve 22. soruların binişik olduğu görülmüştür ve o sorular çıkarılıp tekrar faktör analizi yapılmıştır. Daha sonra güvenilirlik analizi yapıldığında 24. Sorunun güvenilirlik analizinde, 24. sorunun testteki diğer soruların toplamından oluşan bütün arasındaki korelasyonun 0,02 çıkmasından kaynaklı o soru da çıkarılıp tekrar faktör analizi yapılmıştır. Tüm bu analizler sonucunda, toplam varyansın 0,55 olduğu görülmüştür. Faktör yükü 0,40'ın üzerindeyse faktör sayısı ona göre belirlenip faktörlerin altındaki sorulara göre doğrulama yapılır (Çakır, 2014). Burada faktör 1'in altında 13 soru (1, 2, 3, 4, 5, 6, 11, 12, 13, 14, 18, 21, 25) toplanmış olup, faktör 2'nin altında ise üç sorunun (8, 17, 23) toplandığı görülmektedir. Bu çalışma faktörlerin altında en az üç soru olduğu için iki faktörlü yapı olarak alınabileceğini ifade eder. Bu kısımda taslak başarı testine ait olarak KMO değerleri, çarpıklık ve basıklık değerleri, faktör değerleri, binişiklik durumları ve hangi sorunun hangi faktör altında yer aldığı incelenmiş olup bu kapsamda hesaplanan değerler Tablo 3, Tablo 4, Tablo 5'te sunulmuştur.

Tablo 3. Başarı testi için KMO değeri

Kaiser-Meyer-Olkin (KMO) test örneklem yeterliği	0,73
Küreselliğin Bartlett Testi	P 0,00

Tablo 3 incelendiğinde KMO değeri ve Bartlett Testi sonucu verilmiştir. KMO değeri tablodan görüldüğü üzere 0,73 kabul edilebilir uygun olarak görülmektedir. KMO değeri 0,60'dan büyükse faktör analizi gerçekleştirilebilir (Pallant, 2016). Küreselliğin Bartlett testinin de anlamlılık değerinin (p) 0,00 olduğu görülmektedir. Bu değerler, faktör analizi yapılabileceğini ve yeterli örnekleme sahip olduğunu göstermektedir.

Tablo 4. Başarı testi faktör değerleri

Faktör	Öz değer	Varyansların Yüzdesi	Toplam Varyans Yüzdesi
1	7,30	0,45	0,45
2	1,50	0,09	0,55
3	1,16	0,07	
4	0,95	0,05	
5	0,83	0,05	
6	0,69	0,04	
7	0,62	0,03	
8	0,61	0,03	
9	0,54	0,03	
10	0,41	0,02	
11	0,36	0,02	
12	0,28	0,01	
13	0,26	0,01	
14	0,20	0,01	
15	0,09	0,00	
16	0,07	0,00	

Tablo 4 incelendiğinde toplam varyans yüzdelerinin verilmiş olduğu ve açıklanan toplam varyansın değerinin 0,55 olduğu görülmektedir. Açıklanan toplam varyansın 0,50'nin üstünde olması geçerli bir analiz olarak kabul edileceğinden bu testten elde edilen değerlerin de istenen düzeyde olduğu söylenebilir (Çakır, 2014). Öz değer (Eigen Value), faktör sayısına karar verirken dikkate alınır. Aynı zamanda faktörler tarafından açıklanan varyansı hesaplar. Her ne kadar öz değer (eigenvalues) istatistikinde birden büyük olan değerler anlamlı olarak değerlendirilse de (Çakır, 2004) burada elde

edilen analizler, factor programı ile yapıldığından anlamlı olan öz değerlerin karşısında açıklanan toplam varyans değeri çıkmaktadır (Lorenzo-Seva ve Van Ginkel, 2016). Ayrıca output dosyası incelendiğinde bu test için “*Advised number of dimension:1*” bir faktörlü yapıyı önermektedir. Yazarlar, soru içeriklerini ve iki faktörlü yapıyı karşılaştırmalı olarak inceleyerek iki faktörlü yapının bu test için uygun olduğuna karar vermişlerdir.

Tablo 5. Başarı testi faktörleri ve maddeleri

Sorular	Faktör 1	Faktör 2
S1	0,68	
S2	0,58	0,37
S3	0,56	0,42
S4	0,44	
S5	0,62	
S6	0,53	
S11	0,87	
S12	0,63	
S13	0,55	
S14	0,76	
S18	0,65	
S21	0,64	
S25	0,64	0,32
S8		0,67
S17		0,50
S23		0,57

Yapılan tekrarlı faktör analizi sonucunda bir faktörlü analiz yapılmış olup (10. ve 24. soru hariç çünkü bu iki soru güvenilirlik analizinde değerleri 0,20'nin altında çıktığı için faktör analizine alınmamıştır) toplam varyansın 0,38'ini açıkladığı görülmüştür. Açıklanan toplam varyansın 0,50 üzerinde olması faktör analizi için önemli bir kıstas (Yaslıoğlu, 2017) olduğundan tek faktörlü yapı tercih edilmemiştir. İki faktörlü yapının analizi yapıldığında ise açıklanan toplam varyansın 0,55 olduğu görüldüğünden ve bu değer çok daha iyi bir şekilde testi açıkladığından uzman önerisi ile iki faktörlü yapı tercih edilmiştir. Tablo 5'te soruların hangi faktör altında hangi değerlere sahip olduğu ve binişiklik durumları görülmektedir. Tablo incelendiğinde, faktör yükleri arasında en az 0,10'luk farkın olduğu yani binişikliğin olmadığı ve her sorunun bir faktör altında olduğu görülmektedir. İki faktörden oluşan testin birinci faktörünün altında 13 soru olduğu görülmektedir. Soru içerikleri incelendiğinde birinci faktörün ismi “*Sindirimde yer alan organlar ve görevleri*” olarak adlandırılmıştır. İkinci faktörün altında ise üç soru bulunmaktadır. İkinci faktörün ismi ise “*Sindirim çeşitleri ve sindirimin gerçekleştiği yerler*” şeklinde adlandırılmıştır.

Doğrulayıcı faktör analizi

DFA'da geçerliği değerlendirmek amacıyla, Ki-kare Uyum Testi (χ^2), Yaklaşık Hataların Ortalama Karekök (RMSEA), Karşılaştırmalı Uyum İndeksi (CFI), İyilik Uyum İndeksi (GFI), Düzeltilmiş İyilik Uyum İndeksi (AGFI), Normlaştırılmamış Uyum İndeksi (NNFI) gibi değerlere bakılmıştır. Literatür tarandığında ki-karenin örneklem büyüklüğüne karşı hassas olması nedeniyle, normlaştırılmış ki-kare olarak isimlendirilen χ^2/sd oranının kullanılması önerilmekte olup örneklemin büyük olduğu durumlarda bu oranın üçün altında olması mükemmel uyumu sağladığını, beşin altında çıkması ise orta düzeyde uyumun gösterimi olarak kabul edilmektedir. (Kline, 2005; Şimşek, 2007; Yılmaz ve Çelik, 2009). RMSEA ve RMR değerlerinin 0,05'ten küçük veya eşit olmasının istenilen uyumun iyi olduğunu, 0,05 ile 0,08 arasında olması yeterli bir uyumun olduğunu, 0,08 ile 0,10 arasında olması ise vasat bir uyumun olduğunu göstermektedir (Şimşek, 2007; Yılmaz ve Çelik, 2009). CFI değerinin 0,95'ten, GFI değerinin 0,90'dan, AGFI değerinin ise 0,85'ten daha büyük olması ise uyumun kabul edilebilir olduğunu göstermektedir (Yılmaz ve Çelik, 2009). Bu çalışmadaki bu değerlere bakıldığında Ki-kare Uyum Testine bakıldığında χ^2/sd oranının yaklaşık olarak 1,59 olduğu yani mükemmel uyum gösterdiği yorumu yapılabilir. RMSEA değerinin 0,05 olduğu için uyumun iyi olduğu söylenebilir. Yine bu çalışmada CFI, GFI, AGFI ve NNFI değerlerinin bire yaklaşmış olması uyumun gayet iyi olduğunu göstermektedir (Şimşek,2007; Yılmaz ve Çelik, 2009).

Doğrulayıcı faktör analizi kapsamında uyum indeksleri hesaplanmış olup hesaplanan değerler Tablo 6'da sunulmuştur. Ki kare değeri ve uyum ölçüsü değerleri verilmiştir. (χ^2/sd) değeri sıfır ile üç arasında olduğunda iyi bir uyum şeklinde nitelendirilirken üç ile beş arasında kabul edilebilir uyum olarak nitelendirilir (Hooper vd., 2008).

$$\chi^2/ sd = 141.710 / 89$$

χ^2/ sd değerinin 1,59 olduğu ve bu değerinde mükemmel uyum gösterdiği söylenebilir.

Tablo 6. DFA analizi için uyum indeksleri değerleri (Seçer,2017)

Uyum kriteri	Hesaplanan değerler	Mükemmel uyum sınırı	Kabul edilebilir uyum sınırı	Uyum durumu
RMSEA	0,05	0,00 - p<0,05	0,50-0,80	Mükemmel uyum
CFI	0,98	0,97 ve üzeri	0,95 ve üzeri	Mükemmel uyum
GFI	0,98	0,90 ve üzeri	0,85 ve üzeri	Mükemmel uyum
AGFI	0,97	0,90 ve üzeri	0,85 ve üzeri	Mükemmel uyum
NNFI	0,98	0,90 ve üzeri	0,85 ve üzeri	Mükemmel uyum

Tablo 6 incelendiğinde DFA analizi için uyum indeks değerleri incelenmiş olup değerlerin mükemmel uyuma sahip oldukları söylenebilir.

Ölçüt Geçerliği

Bu çalışma kapsamında literatürde geçerliği ve güvenilirliği sağlanmış, taslak başarı testinin kazanımlarına uygun, altıncı sınıf sindirim sistemi başarı testi ölçüt test olarak kullanılmıştır. Yeterli örnekleme ulaşılabilmesi için soru sayısının beş katı örnekleme ulaşılması gerekir (Büyüköztürk, 2012). Yeterli örnekleme ölçüt testin soru sayısının yedi katına uygulanması ile ulaşılmıştır. Hem taslak başarı testi hem ölçüt testi için mod, medyan ve aritmetik ortalama değerlerinin birbirine yakın olması ve basıklık ve çarpıklık değerlerinin +1 ve -1 aralığında olması sonucunda normal dağılım

varsayımı karşılanmıştır (Morgan, Leech, Gloeckner, ve Barrett, 2004). Hesaplanan değerler Tablo 7’de sunulmuştur.

Tablo 7. Ölçüt Testi ve Başarı Testine İlişkin Değerler

		Ölçüt Testi	Başarı Testi
N	Katılımcı Sayısı	194	229
Ortalama		7,81	13,44
Medyan		7,00	14,00
Mod		6,00	16,00
Standart Sapma		3,64	5,54
Varyans		13,27	30,74
Çarpıklık		0,759	-0,006
Basıklık		0,064	-1,062
Ranj		17,00	23,00
Minimum		2,00	2,00
Maksimum		19,00	25,00

Tablo 7 incelendiğinde ölçüt testinden alınan ve başarı testinden alınan değerler ortalama, mod, medyan, standart sapma, ranj, varyans, minimum ve maksimum değerleri görülmektedir. Ölçüt testinin ve başarı testinin çarpıklık ve basıklık katsayılarının -1 ve +1 aralığında olduğu görülmektedir.

Güvenirlilik Analizi Bulguları

Güvenirlilik testte yer alan soruların kendi arasında tutarlığın olmasıdır (Yaşar, 2014). Bu çalışma kapsamında KR-20 güvenirlilik katsayısına bakılmış olup 0,84 olarak bulunmuştur. Güvenirliliğin 0,60 ile 0,90 arasında değer alması testten alınan puanların güvenirliliğinin çok iyi olduğunu gösterir (Can, 2014). Bu çalışma da bu değer karşılanmış olup testten elde edilen puanların güvenilir olduğu yorumu yapılabilir. Geliştirilen başarı testine ait olarak soruların güvenirlilik indeksleri hesaplanmış olup güvenirlilik indeksi değeri Tablo ‘de sunulmuştur.

Tablo 8. Taslak başarı testinin KR-20 değeri

KR-20	Standardize öğelere göre KR-20 değeri	Soru sayısı
0,84	0,84	16

Tablo 8 incelendiğinde testten elde edilen puanları güvenirliliği 0,84 olarak bulunmuş olup soru sayısının 16 tane olduğu görülmektedir.

Tartışma ve Sonuç

Bu çalışmada, altıncı sınıf fen bilimleri dersinin ikinci ünite ikinci bölümü olan “Sindirim sistemi” konu kazanımlarını içeren geçerliği ve güvenilirliği sağlanmış başarı testi geliştirilmesi amaçlanmıştır. Bu çalışmayı literatürdeki çoğu çalışmadan üstün kılan yönleri mevcuttur. Klasik test teorisine göre test geliştirme basamaklarına uyumlu şekilde yürütülerek hata oranının en az düzeye indirilmesine dikkat edilmiştir. Taksonomisi belli, soruları temel aldığı Bloom taksonomisine uygun, taksonomik olarak basamaklara dikkat edilmiş ve uzman görüşleri alınarak titizlikle hazırlanmış bir başarı testi geliştirme çalışmasıdır. Literatürde yer alan başarı testi geliştirme çalışmaları incelendiğinde bu özellikler açısından eksikler olduğu görülmektedir. Örneğin; Demir ve Dindar’ın (2006) çalışmasında çoktan seçmeli testte değerlendirme basamağından soru sorulmamıştır. Gülmez Güngörmez ve Akgün’ün (2018) çalışmasında ise çoktan seçmeli testte sentez basamağından sorunun yer alması gibi taksonomik hatalar görülmektedir. Demir ve Öner Armağan’ın (2019) çalışmasında dayandığı felsefenin net olarak belli olmaması, Kargın ve Gül’ün (2021) çalışmasında ise uzman görüşlerinin az ve uzmanlık alanlarının açık ve net bir şekilde ifade edilmemesi gibi faktörler bu çalışmanın önemini göstermektedir.

Bu çalışmada birinci yazar tarafından literatür taranarak her soru için soru bazında kazanım yazılmış, taksonomik anlamda doğruluğu için titiz çalışılmış ve başlangıçta soru havuzu geniş tutulmuştur. Bloom Taksonomisi basamaklarına sorular uzman kontrolünde doğru şekilde yerleştirilmiştir. Çoktan seçmeli testlerde sentez basamağından soru sorulması çoktan seçmeli testin doğasına uygun değildir (Birgin, 2016). Değerlendirme basamağı ile ilgili soruların çoktan seçmeli testlerde yer alması üst düzey bilgileri ölçme açısından önemlidir (Çakıcı vd., 2012). Buna karşın literatürde yer alan başarı testleri incelendiğinde taksonomik hatalarla ilgili örnekler karşımıza çıkmaktadır. Örneğin Kaya ve Gül’ün (2020) “Sindirim Sistemi” konusunda geliştirmiş olduğu başarı testinde değerlendirme basamağına dair soru bulunmamaktadır. Demir ve Dindar (2006); Gündüz (2009); Gülmez Güngörmez ve Akgün’ün (2018) yaptıkları çalışmalarda ise çoktan seçmeli testte sentez düzeyinden soru sorma, uygulama basamağından soru sormama gibi taksonomik hatalar görülmektedir. Kahyaoğlu’nun (2005) ön test ve son test şeklinde çalıştığı, dolaşım ve sindirim sisteminin birlikte yer aldığı çalışmasında da sindirim sistemi ile alakalı açık uçlu soruların sığ kaldığı düşünülmektedir.

Madde güçlük indeksi testi cevaplayıcılara soruların ne kadar zor ya da ne kadar kolay olduğuyla ilgili bilgi verir ve buradaki indekslerin aldığı değerler 0 ve 1 aralığında değişir. Buradaki değerler 0’a yaklaşırsa zorlaşırken, 1’e doğru kolaylaşır. Genel olarak testin güçlük indeksinin 0,50 olması beklenir (Hasançebi, 2020). Başarı testinde madde güçlük indeksinin 0,40 ile 0,60 arasında olması tercih edilir. (Hotiu, 2006). Yine güçlük için soruların çok zor ya da çok kolay olması testlerin güvenilirliğini düşürür (Hasançebi, 2020). Bu değerlerin 0,30 ile 0,70 aralığında olması soruların orta ve kolay güçlükte kabul edilebilir olduğundan bu çalışmada madde güçlük indekslerine bakıldığında genel olarak 0,30 ile 0,70 aralığında olduğu görülmekte olup madde güçlük indekslerinin kabul edilebilir olduğu söylenebilir. Maddelerin ayırt edicilik indeksi ise ilgili sorunun soruyu bilen öğrenciyle bilmeyen öğrenciyi ne kadar ayırdığını gösterir. İstenilen başarılı öğrencinin soruyu doğru yapması, başarısız öğrencinin soruyu yanlış yapmasıdır (İpek Akbulut ve Çepni, 2013). Ayırt ediciliği yüksek sorulardan oluşan testin güvenilir olduğu söylenebilir. Madde ayırt edicilik indeksi için değer aralığının -1 ve +1 arasında olması gerekir (Hasançebi, 2020). Madde ayırt edicilik indeksinin 0’a yaklaşması sorunun bilen ve bilmeyen öğrencileri ayırması bakımından iyi olmadığını, +1’e yaklaşması sorunun bilen ve bilmeyen öğrencileri ayırması bakımından çok iyi olduğunu ifade eder (Hasançebi, 2020). Burada bakılan değerlerin 0,40 ve daha büyük olduğu durumlarda ayırt ediciliğin çok iyi, 0,30 ve 0,39 aralığında olursa iyi ve geliştirilebilir olduğunu, 0,20 ve 0,29 aralığında düzeltilmeli ve geliştirilmeli olduğunu ve son olarak 0,19 ve daha düşükse madde ayırt ediciliğinin çok düşük ve mutlaka çıkarılması gerektiğini ifade eder (Hasançebi, 2020). Bu çalışma kapsamında ayırt ediciliğe bakıldığında her soru bazında ayırt ediciliğin çok iyi 0,40 üzerinde olduğu ve çoğu sorunun +1 olduğu görülmektedir. Ayrıca test sorularının ayırt edici olup olmadığını belirlemek için ilişkisiz örneklem t testi yürütülmüştür. Bu analizler sonucunda soruların genel olarak güçlüğü 0,50 olduğu zor ve kolay sorularında olduğu görülmektedir. Yapılan çalışmanın literatürde yer alan başarı testi geliştirme çalışmaları ile madde güçlük ve madde ayırt edicilik indeks değerlerinin çok iyi olması bakımından

benzerlikleri olduğu görülmektedir (Bolat ve Karamustafaoğlu, 2019; Baladın Duman, 2019; Kurtcuoğlu, 2007; Kaya ve Gül, 2020; Uğur, 2010). Örneğin, Baladın Duman (2019) çalışmasındaki soruların güçlüğüyle mevcut çalışma paralellik gösterirken, ayırt edicilik olarak bakıldığında bu çalışmada testten atılması gereken, kullanılmaması gereken soru bulunması olumsuz bir durumdur. Mevcut çalışmada ise sindirim sisteminde geliştirilen bu testin ayırt ediciliğinin yüksek olduğu söylenebilir. Kapsam geçerliği açısından farklı alanlarda uzman kontrolünün sağlanması, uzmanlardan gelen dönütlerin dikkate alınıp hassasiyetle incelenmesi, Bloom Taksonomisinin her basamağından (sentez hariç) soru sorulması, kazanımların soruya özel yazılması, literatürün taranması, soru havuzunun geniş olması, soruların açık net anlaşılır olması için düzeye uygun olarak kelimelerin seçilmesi, varsa yanlış anlaşılacak kelimelerin yanına anlaşılacak şekilde yazılması ve sürecin her aşamasında titizlikle davranılması bu çalışmayı özgün kılan hususlardır. Geliştirilen başarı testinin bu yönleriyle kapsam geçerliğini sağlandığı söylenebilir.

Çalışmada, yapı geçerliğini sağlamak adına Açımlayıcı faktör analizi (AFA) ve doğrulayıcı faktör analizi (DFA) yürütülmüştür. AFA ve DFA aynı örneklem üzerinde uygulanmış olup AFA ve DFA FACTOR 11.04.02 programıyla hesaplanmıştır. Yine literatürdeki çalışmalar incelendiğinde AFA ve DFA'nın aynı örneklem (Nacaroğlu, vd., 2020; Uçar ve Aktamış, 2019) veya farklı örneklem (Metin, vd., 2021; Özer Özkan ve Acar Güvendir, 2018) üzerinde uygulandığı çalışmaların olduğu görülmektedir. AFA sonucunda toplam varyansın 0,55'ini açıklayan, 16 sorudan oluşan, iki faktörlü yapı elde edilmiş ve bu yapı DFA ile doğrulanmıştır. Literatür tarandığında yapı geçerliği açısından bu çalışmanın literatürde bulunan diğer çalışmalarla benzerlikleri ve farklılıkları olduğu görülmektedir. Örneğin; Bektaş ve Nacaroğlu (2019), Saylan Kırmızıgül ve Kaya'nın (2019) çalışmalarında KMO değerine yani yeterli örnekleme sağlama durumuna bakıldığında bu çalışmayla benzer değerlerde yeterli örnekleme sağlandığı görülmektedir. Buna karşın, Bolat ve Karamustafaoğlu (2019), Gül ve Kargın (2021), Keçeci vd. (2019) gibi çalışmalarda AFA ve DFA ile ilgili gerekli açıklamaların yapılmaması, Demir ve Öner Armağan (2019) çalışmasında sadece AFA'ya dair sınırlı açıklamanın yer alması gibi yapı geçerliği analizlerinin yeterli olmadığı çalışmalar da bulunmaktadır. Bu çalışmada ise AFA ve DFA ile ilgili gerekli ve yeterli bilgiler verilmiştir. DFA sonucunda ise değerlerin mükemmel uyum gösterdiği görülmektedir.

Başarı testi geliştirilirken literatürde yapılan önceki çalışmalar incelenmiştir. Geçerliği ve güvenilirliği yüksek olan aynı kazanımları ölçen daha önce yapılmış bir teste bakılmıştır (Özkan, 2017). Ölçüt olarak alınan Özkan (2017) testinin güvenilirliği incelendiğinde 0,74 olduğu, bu yapılan testin güvenilirliğinin ise 0,85 olduğu görülmekte olup iki test karşılaştırıldığında yapılan bu testin daha güvenilir olduğu ortaya çıkmaktadır. Yine Özkan'ın testiyle bu çalışma ayırt edicilik kapsamında karşılaştırıldığında Özkan'ın(2017) testinin ayırt ediciliğinin 0,45 çıkması bu testin ayırt ediciliğinin bire yakın çok yüksek düzeyde çıkması ile bu yapılan testin ayırt ediciliğinin daha fazla olduğu anlaşılmaktadır. Özkan'ın (2017) testindeki soruların güçlüğüyle bu testin güçlüğü karşılaştırıldığında Özkan'ın testinin madde güçlüğü 0,59 iken bu testinde genel olarak 0,50 civarında olduğu görülmektedir. Bu açıdan testler benzerlik göstermektedir. Ayrıca Özkan'ın(2017) testi 2018 yılında güncellenen fen bilimleri öğretim programındaki değişiklikler yönünden eksik kalmaktayken bu çalışma güncellenen programa uygundur.

Güvenirlik çalışmaları sonucunda KR 20 güvenilirlik katsayısı 0,85 olarak hesaplanmıştır. Bu değer 0,60'dan büyük olduğu için (Büyüköztürk, 2002) sindirim sistemi konusu kazanımları ile geliştirilen başarı testinden elde edilen puanların güvenilir olduğu sonucuna ulaşılmıştır. Literatürde geliştirilen başarı testlerinde güvenilirliği yüksek testlere rastlanmakta ve bu bakımdan test literatürle uyum sağlamaktadır (Baladın Duman, 2019; Kurtcuoğlu, 2017; Kaya ve Gül, 2020). Literatürde yer alan başarı testlerinde test güvenilirliği analizleri için KR-20 güvenilirlik katsayısını kullanan birçok çalışma bulunmaktadır (Güneş ve Serdaroğlu, 2018; Nacaroğlu, vd., 2019; Nacaroğlu ve Bektaş, 2019; Saraç, 2018; Saylan Kırmızıgül ve Kaya, 2019). Buna karşın, başarı testi geliştirme çalışmalarında Cranbach's Alpha güvenilirlik katsayısı kullanan çalışmalar da bulunmaktadır (Ayvacı ve Durmuş, 2016; Demir ve Öner Armağan, 2019; Keçeci, vd., 2019; Kızılkapan ve Bektaş, 2018; Şen ve Eryılmaz, 2011). Her iki güvenilirlik katsayısı da eşdeğer olduğu için, mevcut çalışmada kullanılan. KR-20 güvenilirlik katsayısının kabul gören bir analiz türü olduğu söylenebilir (Crocker ve Algina, 1986).

Bu başarı testi literatürde yer alan geçerliği ve güvenilirliği sağlanmış testlere benzerliğiyle destek çalışma olurken, dayandığı felsefesi olmayan ya da kuramsal temelleri açıklanmayan, kapsam geçerliği sıkıntılı olan, taksonomi bilgisi verilmeyen ya da taksonomik olarak yanlışlıkları olan, güvenilirliği düşük, ayırt ediciliği ve gücü düşük, analizleri doğru şekilde yapılmamış diğer çalışmalara ve yeni araştırmalara da örnek olması, eksikliklerin görülmesi açısından önemli olduğu ifade edilebilir (Bolat ve Karamustafaoğlu, 2019; Demir ve Dindar, 2006; Demir ve Öner Armağan, 2019; Gündüz, 2009; Gülmez Güngörmez ve Akgün, 2018; Gül ve Kargın, 2021; Keçeci vd., 2019).

Sonuç olarak altıncı sınıf seviyesinde geliştirilmiş olan "Sindirim Sistemi" konusuna ilişkin 2018 fen bilimleri öğretim programında yer alan kazanımlara ve amacına uygun, yeterli şekilde örneklemin bulunduğu, geçerlik ve güvenilirliği kontrol edilmiş, öğrencilerin öğrenme düzeylerini ölçebilecek bir başarı testi geliştirilmiştir.

Öneriler

Öğretmenler, öğrencilerinin "Sindirim Sistemi" konusundaki öğrenme düzeylerini belirlemek amacıyla geliştirilen testi kullanılabilirler. Araştırmacılar tarafından veri toplama aracı olarak kullanılabilir. Bu çalışma ortaokul altıncı sınıf "Sindirim Sistemi" konusunda geçerliği ve güvenilirliği sağlanmış bir başarı testi olarak literatüre katkı sağlasa da, testin çoktan seçmeli sorulardan oluşması sebebiyle öğrencilerin sindirim sistemi konusundaki kavramlarını derinlemesine belirleme konusunda sınırlı kaldığı söylenebilir. Bu sebeple ileride yapılacak olan çalışmalara bu testin iki aşamalı olarak geliştirilmesi ve uygulanması önerilebilir. Böylece ilk aşamada çoktan seçmeli sorulardan oluşan, ikinci aşamada ise açık uçlu sorudan oluşacak test, araştırmacılara öğrencilerin sindirim sistemi konusuna yönelik kavramlarını detaylı olarak inceleme fırsatı sunacaktır. Bu açıdan bu test revize edilip tekrar kullanılabilir. Test için daha fazla katılımcıyla uygulamalar yapılabilir. Testin uygulanmasının ardından öğrencilerle yapılacak olan nitel görüşmelerle, elde edilen sonuçlar analiz edilerek test güçlendirilebilir.

KAYNAKÇA

- Akbaş, U., Karabay, E., Yıldırım Seheryeli, M., Ayaz, A. ve Demir, Ö. O. (2019). Türkiye ölçme araçları dizininde yer alan açılımlayıcı faktör analizi çalışmalarının paralel analiz sonuçları ile karşılaştırılması. *Journal of Theoretical Educational Science*, 12 (3), 1095-1123. doi: 10.30831/akukeg.453786
- Aker, S., Dünder, C. ve Pekşen, Y. (2005). Ölçme araçlarında iki yaşamsal kavram: Geçerlik ve güvenilirlik. *O.M.Ü. Tıp Dergisi*, 22(1), 50-60.
- Akhter, N. ve Bahoo, R. (2015). *Development of a semi standardized achievement test of education for intermediate level. Journal of educational research* (1027-9776). 2015, Vol. 18 Issue 1, Dept of Education IUB, Pakistan
- Anderson, L.W. (2003). *Classroom assessment: Enhancing the quality of teacher decision making*. New Jersey: Routledge.
- Anderson, L. W. (2005). Objectives evaluation and the improvement of education. *Studies in Educational Evaluation*, 31(2), 102- 113. doi:10.1016/j.stueduc.2005.05.004
- Ayvacı, H. Ş. ve Durmuş, A. (2016). Bir başarı testi geliştirme çalışması: Isı ve sıcaklık başarı testi geçerlik ve güvenilirlik araştırması. *Ondokuz Mayıs University Journal of Education Faculty*, 35(1), 87-103.
- Baladın Duman, B. (2019). *Besin içerikleri ve sindirim sistemi konularında tga yöntemine dayalı olarak geliştirilen etkinliklerin değerlendirilmesi*. (Yayımlanmamış Yüksek Lisans Tezi). Trabzon Üniversitesi, Lisansüstü Eğitim Enstitüsü, Trabzon.

- Baltacı, A. (2018). Nitel araştırmalarda örnekleme yöntemleri ve örnek hacmi sorunsalı üzerine kavramsal bir inceleme. *Bitlis Eren Üniversitesi Sosyal Bilimler Dergisi*, 7 (1), 231-274.
- Barchard, K. A. ve Hakstian, A. R. (1997). Alfa katsayısı ile örnekleme modelinin çıkarsamaya etkisi. *Eğitimsel ve Psikolojik Ölçüm*, 57(6), 893-905. <https://doi.org/10.1177/0013164497057006001>
- Başol, G. (2015). *Eğitimde ölçme ve değerlendirme*. (3. Baskı). Pegem Akademi. Ankara
- Baştürk, S. (2014). Ölçme araçlarının taşınması gereken nitelikler. S. Baştürk (Ed.), *Eğitimde Ölçme ve Değerlendirme*. Nobel Akademik Yayıncılık, Ankara.
- Baykul, Y. (2015). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Pegem Akademi. Ankara.
- Bernard, H. R. (2011). *Research Methods in Anthropology: Qualitative and Quantitative Approaches* (5th edition), Rowman Altamira.
- Birgin, O. (2016). *Bloom Taksonomisi*. E. Bingölbali, S. Arslan ve İ. Ö. Zembat (Edit.) Matematik Eğitiminde Teoriler içinde (ss. 839-860). Pegem Akademi, Ankara.
- Bolat, A. ve Karamustafaoğlu, S. (2019). Vücudumuzdaki sistemler ünitesi başarı testi geliştirme: Geçerlik güvenilirlik. *Gazi Eğitim Bilimleri Dergisi*, 5(2), 131-159. doi: <https://Dx.doi.Org/10.30855/Gjes.2019.05.02.008>
- Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Kuram ve Uygulamada Eğitim Yönetimi*, 32(32), 470-483. Retrieved from <https://dergipark.org.tr/tr/pub/kuey/issue/10365/126871>
- Büyüköztürk, Ş. (2012). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem Yayıncılık.
- Can, A. (2014). *SPSS ile bilimsel araştırma sürecinde veri analizi*, (3. Basım). Pegem Akademi Yayıncılık, Ankara.
- Can, A. (2017). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi*. (5. Baskı). Pegem Akademi, Ankara.
- Creswell, J. W. (2013). *Araştırma Deseni* (Çev. Bursal, M: Ed.: Demir, S. B.). Eğiten Kitap.
- Crocker, L. ve Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth: Holt, Rinehart and Winston.
- Çakıcı, Y., Ürek, H. ve Dinçer, E. O. (2012). İlköğretim öğrencilerinin soru oluşturma becerilerinin incelenmesi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 8(1), 43-68.
- Çakır, A. (2014). Faktör analizi. Doktora programı. <https://docplayer.biz.tr/38884504-T-c-istanbul-ticaret-universitesi-sosyal-bilimler-enstitusu-isletme-anabilim-dali-isletme-doktora-programi-faktor-analizi-ayhan-cakir-1250d91213.html>
- 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), 335-353. Retrieved from <https://dergipark.org.tr/tr/pub/mkusbed/issue/19554/208352>
- Çakmur, H. (2012). Araştırmalarda ölçme-güvenilirlik-geçerlilik. *Taf Prev Med Bull*, 11(3), 339-344. doi:10.5455/Pmb.1-1322486024
- Çalışkan, H. ve Yiğittir, S. (2008). *Sosyal bilgilerde ölçme ve değerlendirme*. Özel öğretim yöntemleriyle sosyal bilgiler öğretimi içinde ss. 217-281. Tay, B. ve A. Öcal (Ed.). Pegem A Yayıncılık, Ankara.
- Çelen, Ü. (2008). Klasik test kuramı ve madde tepki kuramı yöntemleriyle geliştirilen iki testin geçerlilik ve güvenilirliğinin karşılaştırılması. *Elementary Education Online*, 7(3), 758-768.

- 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.
- Çepni, S. (2010). *Araştırma ve proje çalışmalarına giriş* (5. Baskı). Celepler Matbaacılık, 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.
- Çolak, K. ve Demircioğlu, İ. H. (2010). Tarih dersi sınav sorularının Bloom taksonomisi'nin bilişsel alan düzeyi açısından sınıflandırılması. *Millî Eğitim Dergisi*. 40 (187) , 160-171. Retrieved from <https://dergipark.org.tr/tr/pub/milliegitim/issue/36197/407037>
- Demir, M. ve Dindar, H. (2006). Beşinci sınıf öğretmenlerinin fen bilgisi dersi sınav sorularının Bloom taksonomisine göre değerlendirilmesi. *Gazi Eğitim Fakültesi Dergisi*, 26(3), 87-96.
- Demir, N. ve Öner Armağan, F. (2019). Astronomi başarı testi geliştirme: Geçerlik ve güvenilirlik çalışması. *Uluslararası Eğitim Bilimleri Dergisi*, 3(1), 52-70.
- Demir, N., Kızılay, E. ve Bektaş, O. (2016). 7. Sınıf çözümler konusunda başarı testi geliştirme: geçerlik ve güvenilirlik çalışması. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi (Efmed)*, 10 (1), 209-237.
- Erkorkmaz, Ü., Etikan, İ., Demir, O., Özdamar, K. ve Sanisoğlu, S. Y. (2013). Doğrulayıcı faktör analizi ve uyum indeksleri. *Türkiye Klinikleri Journal of Medical Sciences*, 33(1), 210-223.
- George, D. ve Mallery, M. (2010). *Windows adım adım SPSS: Basit bir kılavuz ve referans*, 17.0 güncelleme. (10. Baskı). Pearson, Boston.
- Gömlüksiz, M. ve Erkan, S. (2010). *Eğitimde ölçme ve değerlendirme* (2. Baskı). Nobel Yayın Dağıtım, Ankara.
- 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 Vuu(I)*, 40-57.
- Gülmez Güngörmez, H. ve Akgün, A. (2018). Ortaokul öğrencilerinin fen bilimleri dersindeki kuvvet ve enerji ünitesine yönelik akademik başarı testi geliştirme çalışması. *Diyalektolog Ulusal Sosyal Bilimler Dergisi*, 18, 85-99.
- Gündüz, Y. (2009). İlköğretim 6, 7 ve 8. sınıf fen ve teknoloji sorularının ölçme araçlarına ve Bloom'un bilişsel alan taksonomisine göre analizi. *Yüzüncü Yıl Üniversitesi, Eğitim Fakültesi Dergisi*, VI(II), 150-165.
- Güneş, M. H. ve Serdaroğlu, C. (2018). Bitki ve hayvanlarda üreme, büyüme ve gelişme ünitesinde geliştirilen başarı testinin geçerliliği ve güvenilirliği. *Bilim, Eğitim, Sanat ve Teknoloji Dergisi (BEST Dergi)*, 2(1), 35-40.
- Güngör, S. N. ve Özkan, M. (2017). Fen bilgisi öğretmen adaylarına ağızda nişasta sindiriminin TGA (tahmin-gözlem-açıklama) yöntemiyle öğretimi: Amilaz Örneği. *Atatürk Üniversitesi Kazım Karabekir Eğitim Fakültesi Dergisi*, 34, 30-54.
- Haladyna, T. M. (1997). *Writing test items to evaluate higher order thinking*. Allynand Bacon: Needham Heights, MA.
- Hasançebi, B., Terzi, Y. ve Küçük, Z. (2020). Madde güçlük indeksi ve madde ayırt edicilik indeksine dayalı çeldirici analizi. *Gümüşhane Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 10(1), 224-240.

- Hooper, D., Coughlan, J. ve Mullen, M. R. (2008). Structural equation modelling: guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), 53-60.
- Hotiu, A. (2006). *The relationship between item difficulty and discrimination indices in multiple-choice tests in a physical science course. Unpublished master's thesis*, Florida Atlantic University.
- İpek Akbulut, H. 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.
- Kahyaoğlu, M. (2005). *İlköğretim 6. sınıf "dolaşım ve sindirim sistemi" konusunda şarkı ve şiirin başarıya, hatırlamaya ve derse karşı tutuma etkisi. (Yüksek Lisans Tezi)*. Marmara Üniversitesi, Eğitim Bilimleri Enstitüsü, İstanbul.
- Kalaycı, Ş. (2010). *Faktör analizi*. Ş. Kalaycı (Edit.), *SPSS uygulamalı çok değişkenli istatistik teknikleri*, (5. Baskı), Asil Yayın Dağıtım, Ankara.
- Kargın, P. D. ve Gül, Ş. (2021). Altıncı sınıf vücudumuzdaki sistemler ve sağlığı ünitesine yönelik bir başarı testi geliştirilmesi, *İhlara Eğitim Araştırmaları Dergisi*, 6(1), 1-26.
- Kaya, S. (2020). *11. sınıf öğrencilerine "sindirim sistemi" konusunun react stratejisi ile öğretimi. (Yayımlanmamış Doktora Tezi)*. Atatürk Üniversitesi, Eğitim Bilimleri Enstitüsü, Erzurum.
- Kaya, S. ve Gül, Ş. (2020). 11. sınıflar için 'Sindirim sistemi' konusuna yönelik başarı testi geliştirme çalışması. *Uluslararası Sosyal ve Eğitim Bilimleri Dergisi*, 7(13), 72-97. <https://doi.org/10.20860/İjoses.701311>
- Keçeci, G., Yıldırım, P. ve Kırbag Zengin, F. (2019). Sistemler akademik başarı testi: Geçerlik ve güvenilirlik çalışması. *Ulusal Eğitim Akademisi Dergisi (UEAD)*, 3(1), 96-114.
- Kelecioğlu, H. ve Göçer Şahin, S. (2014). Geçmişten günümüze geçerlik. *Journal of Measurement and Evaluation in Education and Psychology*, 5(2), 1-11. <https://doi.org/10.21031/epod.41706>
- Kelloway E. K. (1998). *Assessing model fit. Using lisrel for structural equation modeling*. 3rd ed. USA: Sage Publications.
- Kızıkan, O. ve Bektaş, O. (2018). Yedinci sınıf maddenin yapısı ve özellikleri ünitesi başarı testi geliştirilmesi: Geçerlik ve güvenilirlik çalışması. *The Journal of International Lingual, Social and Educational Sciences*, 4 (2), 186-202.
- Kline, T. (2005). *Psychological Testing: A Practical Approach to Design and Evaluation*.
- Lorenzo-Seva, U., ve Van Ginkel, J. R. (2016). Multiple imputation of missing values in exploratory factor analysis of multidimensional scales: estimating latent trait scores. *Anales de Psicología/Annals of Psychology*, 32(2), 596-608.
- Linn, R. L. ve Gronlund, N. E. (1995). *Measurement and assessment in teaching*. (7th edition) Englewood Cliffs, NJ: Prentice-Hall.
- McKillup, S. (2011). *Statistics explained: An introductory guide for life scientists*. Cambridge University Press.
- 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.

- Metin, M., Çevik, A. ve Gürbey, S. (2021). Öğretmenlerin uzaktan eğitime ilişkin görüşlerini belirleme ölçeği: Geçerlilik ve güvenilirlik çalışması. *Maarif Mektepleri Uluslararası Sosyal ve Beşerî Bilimler Dergisi*, 4(1), 15-35. <https://doi.org/10.47155/mamusbdd.911344>
- Morgan, G. A., Leech, N. L., Gloeckner, G. W. ve Barrett, K. C. (2004). *SPSS for introductory statistics: Use and interpretation* (2. ed.). London: Lawrence Erlbaum Associates.
- Nacaroğlu, O. ve Bektaş, O. (2019). Fen bilimleri dersindeki madde ve değişim ünitesine yönelik geçerli ve güvenilir başarı testi geliştirme: BİLSEM örneği. *Maarif Mektepleri Uluslararası Eğitim Bilimleri Dergisi*, 3(2), 67-88.
- Nacaroğlu, O., Bektaş, O. ve Kızılkapan, O. (2020). Madde döngüleri ve çevre sorunları konusunda başarı testi geliştirme: Geçerlik ve güvenilirlik çalışması. *Kastamonu Education Journal*, 28(1), 36-51. <https://doi.org/10.24106/kefdergi.3374>
- Narlı, S. ve Baser, N. (2008). "Küme, bağıntı, fonksiyon" konularında bir başarı testi geliştirme ve bu test ile üniversite matematik bölümü 1. sınıf öğrencilerinin bu konulardaki hazırbulunuşluklarını betimleme üzerine nicel bir araştırma. *Dokuz Eylül Üniversitesi Buca Eğitim Fakültesi Dergisi*, 24, 147-58.
- Özcan, H., Koca, E. ve Söğüt, M. (2019). Ortaokul öğrencilerinin basınç kavramıyla ilgili anlayışlarını ölçmeye yönelik bir test geliştirme çalışması. *Eğitim ve Toplum Araştırmaları Dergisi*, 6(1), 130-144.
- Özçelik, D. A. (2010). *Test hazırlama kılavuzu*. (4. Baskı). ss.111-29. Pegem Akademi, Ankara.
- Özer Özkan, Y. ve Acar Güvendir, M. (2018). Merkezi sınavların öğretmenler üzerindeki öğretimsel ve duyuşsal etkilerini belirlemeye yönelik öğretmen ölçeğinin geliştirilmesi. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 19(3), 189-204. <https://doi.org/10.17679/inuefd.394383>
- Özkan, F. (2017). *7.sınıf sindirim sistemi konusunda iki aşamalı test geliştirilerek kavram yanlışlarının tespit edilmesi*. (Yayınlanmamış yüksek lisans tezi). Erciyes Üniversitesi, Eğitim Bilimleri Enstitüsü, Kayseri.
- Pallant J. (2016). *SPSS kullanma kılavuzu SPSS ile adım adım veri analizi*. (S. Balcı ve B. Ahi, Çeviri). Ankara, Anı Yayıncılık.
- Patton, M. Q. (2005). *Qualitative Research*. John Wiley ve Sons.
- Saraç, H. (2018). Fen bilimleri dersi "Maddenin değişimi" ünitesi ile ilgili başarı testi geliştirme: Geçerlik ve güvenilirlik çalışması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18(1), 416-445.
- Saraçlı, S. (2011) Faktör analizinde yer alan döndürme metotlarının karşılaştırmalı incelenmesi üzerine bir uygulama. *Düzce Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi*, 1(3), 22-26.
- Saylan Kırmızıgül, A. ve Kaya, H. (2019). "Aynalarda yansıma ve ışığın soğrulması" konusunda geçerliği ve güvenilirliği sağlanmış başarı testi geliştirme çalışması. *Journal of Theoretical Educational Science*, 12(2), 474-493.
- Seçer, İ. (2017). *SPSS ve LISREL ile pratik veri analizi: analiz ve raporlaştırma*. (3. Baskı) Anı Yayıncılık, Ankara.
- Sipahi, B., Yurtkoru, E. S. ve Çinko, M. (2008). *Sosyal bilimlerde SPSS'le veri analizi*. Beta Basım Yayım Dağıtım, İstanbul.

- 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.
- 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.
- Şerabatır, S. (2008). *İlköğretim 6. ve 7. sınıf öğrencilerinin İngilizce dersinde kullandıkları kelime öğrenme stratejileri*. (Yayımlanmamış yüksek lisans tezi). Abant İzzet Baysal Üniversitesi, Sosyal Bilimler Enstitüsü, Bolu.
- Şimşek Ö. F. (2007). *Yapısal eşitlik modellemesine giriş, temel ilkeler ve LISREL uygulamaları*. Ekinoks, Ankara.
- Taşkaya, S. M. ve Meydan, A. (2012). Sınıf öğretmenliği anabilim dalında uygulanan ölçme yöntemleri üzerine bir inceleme. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 1(32), 23-34.
- Tavşancıl, E. (2002). *Tutumların ölçülmesi ve SPSS ile veri analizi*. (1.Baskı). Atlas Yayınları, Ankara.
- Tekkaya, C. ve Balcı, S. (2003). Öğrencilerin fotosentez ve bitkilerde solunum konularındaki kavram yanlışlarının saptanması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 24, 101-107.
- Temizkan, M. ve Sallabaş, M. E. (2011). Okuduğunu anlama becerisinin değerlendirilmesinde çoktan seçmeli testlerle açık uçlu yazılı yoklamaların karşılaştırılması. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 30, 207-220.
- Timur, S., Dogan, F., Cetin, N.I., Timur, B. ve Isik, R. (2019). Artirilmiş gerçeklik uygulamalarının 6. sınıf öğrencilerinin hücre konusundaki akademik başarılarına etkisi. *Akdeniz Eğitim Araştırmaları Dergisi*, 13(30), 126-138. <https://doi.org/10.29329/mjer.2019.218.8>
- Tosun, C., ve Taşkesenligil, Y. (2011). Revize edilmiş Bloom'un taksonomisine göre çözeltiler ve fiziksel özellikleri konusunda başarı testinin geliştirilmesi: geçerlik ve güvenilirlik çalışması. *Kastamonu Eğitim Dergisi*, 19(2), 499-522.
- Uçar, R. ve Aktamış, H. (2019). Astronomi'ye yönelik tutum ölçeği ve 7. sınıf "Güneş sistemi ve ötesi" ünitesine yönelik başarı testi geliştirme çalışması. *Batı Anadolu Eğitim Bilimleri Dergisi*, 10(1), 57-79.
- Uğur, K. U. (2010). *Lise öğrencilerin sindirim sistemi ile ilgili kavram yanlışlarının iki aşamalı testler ile tespit edilmesi*. (Yayımlanmamış yüksek lisans tezi). Eğitim Bilimler Enstitüsü, Selçuk Üniversitesi, Konya.
- Url1:https://acikders.ankara.edu.tr/pluginfile.php/129671/mod_resource/content/0/Veri%20Analizi%205.%20Hafta.pdf
- Url2:<https://medium.com/datarunner/veri-biliminde-normal-da%C4%9F%C4%B1%C4%B1m%C4%B1n-python-%C3%BCzerinden-g%C3%B6rselle%C5%9Ftirilmesi-ve-yorumlanmas%C4%B1-histogram-8381d12b85b9>
- Yaşar, M. (2014). İstatistiğe yönelik tutum ölçeği: Geçerlilik ve güvenilirlik çalışması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 36, 59-75.
- Yaşlıoğlu, M.M. (2017). Sosyal bilimlerde faktör analizi ve geçerlilik: Keşfedici ve doğrulayıcı faktör analizlerinin kullanılması, 46(Özel Sayı), 74-85. <http://dergipark.ulakbim.gov.tr/iuiiletme>
- Yılmaz, V. ve Çelik, H. (2009). *Lisrel ile yapısal eşitlik modellemesi*. Pegem Akademi, Ankara.

EKLER**Ek-1. Başarı testi****Adı Soyadı:****Sınıfı/No.su:**

Sevgili Öğrenciler,

Bu test "Vücudumuzdaki Sistemler ünitesindeki Sindirim Sistemi " konusuna yönelik öğrenme düzeylerini belirlemek amaçlı uygulanmaktadır. Her soru için dört seçenek bulunmaktadır. Dört tane seçenekten bir tanesini işaretlemeniz gerekmektedir.

Başarılar

6.SINIF SİNDİRİM SİSTEMİ KONULU BAŞARI TESTİ SORULARI**SORU 1: 1-Ağız**

- 2-Anüs
- 3-Mide
- 4-Yutak
- 5-İnce bağırsak
- 6-Yemek borusu
- 7-Kalın bağırsak

Vücuda alınan besinlerin sindirim sırası aşağıdakilerden hangisinde doğru olarak verilmiştir?

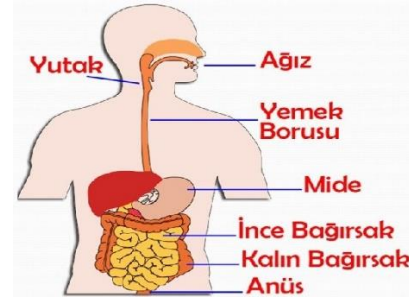
- A) 1-3-2-7-5-4-6
- B) 1-4-6-3-5-7-2
- C) 2-1-3-5-4-7-6
- D) 3-7-1-2-6-5-4

SORU 2:

1. Ağızdan gelen lokmaları yemek borusuna iletirim.
2. Yiyecekleri toplar, salgılarımla bulamaç haline getiririm.
3. Çorba kıvamındaki yiyecekleri kıvrımlı ve geniş yapımla kana ulaştırırım.

Aşağıdakilerden hangisinde organ – görev eşleşmesi doğru olarak verilmiştir?

	<u>Mide</u>	<u>Yutak</u>	<u>İnce bağırsak</u>
A) 1	1	2	3
B) 2	2	1	3
C) 1	1	3	2
D) 2	2	3	1

SORU 3:**Görsel - 1:** Sindirim sistemi organları

Sindirim sistemiyle ilgili Görsel - 1’de belirtilen organlardan hangilerinde sindirim gerçekleşmez?

- A) Ağız-Mide-Anüs
- B) Mide-Yutak-Kalın Bağırsak
- C) Yutak-Yemek Borusu-Anüs
- D) Yemek Borusu-Mide-İnce Bağırsak

SORU 4: Aşağıdakilerden hangisi sindirim sisteminin görevlerinden biridir?

- A) Hareket etmeyi kolaylaştırır.
- B) Kanın vücutta dolaşmasını sağlar.
- C) Dışarıdan gelen uyarılara yanıt verir.
- D) Büyük moleküllü besinleri küçük yapı taşlarına ayırır.

SORU 5:

- I. Anüste emilimin gerçekleşmez ve besin atıkları vücuttan anüs yoluyla uzaklaştırılır.
- II. Yemek borusu kaslı ve esnek yapısı sayesinde, besinlerin mideye iletilmesinde görev alır.
- III. Ağız sindirimin başladığı ilk organdır ve karbonhidratların kimyasal ve fiziksel sindirimi burada yapılır.
- IV. Kalın bağırsak fazlalık su, vitamin ve minerallerin emilimini gerçekleştirdiğinden sindirim sisteminde önemli bir yere sahiptir.

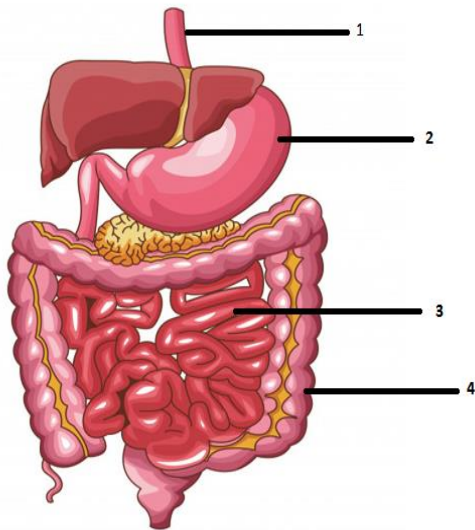
Verilen bilgilerden hangileri doğrudur?

- A) I-II-III
- B) I-II-IV
- C) II-III-IV
- D) I-II-III-IV

SORU 6: Aşağıda verilen özelliklerden hangileri ağız, mide ve ince bağırsak için ortakdır?

- A) Mide öz suyunu kullanma
- B) Yağların sindirimini yapması
- C) Kimyasal sindirimi gerçekleştirme
- D) Safra salgısı ile mekanik sindirim yapma

SORU 7: Aşağıda verilen görselde sindirim sistemine ait bazı yapılar gösterilmiştir.



Görsel - 2: Sindirim sistemine ait bazı yapılar

Buna göre Görsel - 2’de gösterilen numaralandırılmış organlarla ilgili olarak aşağıdakilerden hangisi yanlıştır?

- A) 2 numaralı organda proteinlerin kimyasal sindirimi başlar.
- B) 4 numaralı organ sindirimin en yoğun gerçekleştiği kısımdır.
- C) 3 numaralı organda besinlerin kimyasal sindirimi tamamlanır.
- D) 1 numaralı organda fiziksel (mekanik) ve kimyasal sindirim gerçekleşmez.

SORU 8: Sindirimde görevli organlar ve görevleri Tablo- 1’de verilmiştir.

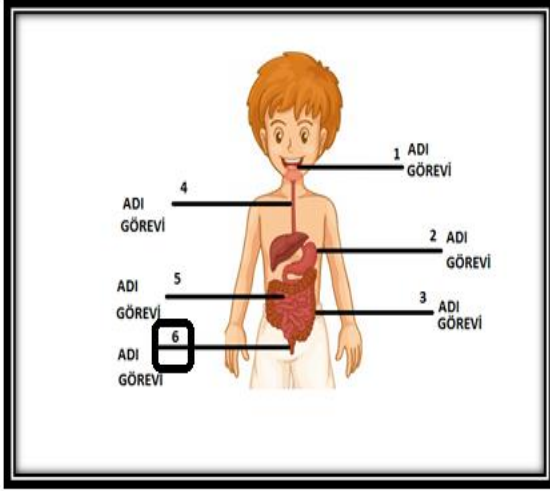
	1	2	3	4
Tükürükle besinlerin ıslatıldığı yerdir.	x			
Besinlerin emilerek kana karıştığı yerdir.		x		
Besinlerin bulamaç haline getirildiği yerdir.			x	
Sindirim sonucu oluşan atık maddelerin dışarı atıldığı yerdir.				x

Tablo - 1: Sindirim sisteminde görevli organlar

Tablo - 1’e göre özellikleri verilen sindirim organları hangisinde doğru olarak verilmiştir?

	1	2	3	4
A) Ağız	İnce Bağırsak	Mide	Anüs	
B) Mide	Ağız	Yutak	İnce Bağırsak	
C) Anüs	Mide	Yemek borusu	Kalın Bağırsak	
D) Yutak	Yemek borusu	Mide	İnce Bağırsak	

SORU 9:



Görsel - 3: Vücudun organları ve görevleri

Görsel - 3'de vücudun organları ve görevleri gösterilmiştir.

Görsel - 3 'te 6 numarayla gösterilen organ ve görevi aşağıdakilerden hangisinde doğru verilmiştir?

- A) Anüs-Besin atıklarının dışarı atıldığı yerdir.
- B) Pankreas-Karbonhidrat, protein, yağların sindirimini sağlar.
- C) İnce bağırsak-Bütün besinlerin sindirimini tamamlanmasını sağlar.
- D) Kalın bağırsak-Fazlalık su, vitamin ve minerallerin emilimi gerçekleşir.

SORU 10: Aşağıdakilerden ifadelerden hangisi doğrudur?

- A) Anüste sindirim gerçekleşmez.
- B) Yağların sindirimi ağızda başlar.
- C) Sindirimin başladığı yer midedir.
- D) Suyun sindirimi ince bağırsakta başlar.

SORU 11:

Öğretmen: Sindirimin amacı.....

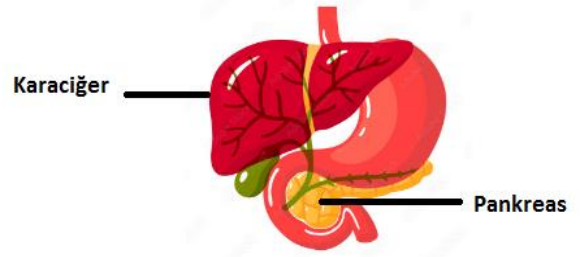
Aşağıdakilerden hangisi öğretmenin verdiği boşluğu en doğru şekilde tamamlar?

- A) Besinleri parçalamaktır.
- B) Besinlerin kana geçmesini sağlamaktır.
- C) Besinlerin parçalanıp hücrelere küçük geçmesini sağlamaktır.
- D) Vücuda alınan büyük moleküllü yiyeceklerin kana ve hücrelere geçebilecek kadar küçük parçalara ayrılmasını sağlamaktır.

SORU 12: Sindirim sistemi ile ilgili aşağıda verilen ifadelerden hangisi yanlıştır?

- A) Proteinlerin sindirimi midede başlar.
- B) Kalın bağırsakta sindirim gerçekleşir.
- C) İnce bağırsakta yağların sindirimi başlar.
- D) Karbonhidratların sindirimi ağızda başlar.

SORU 13:

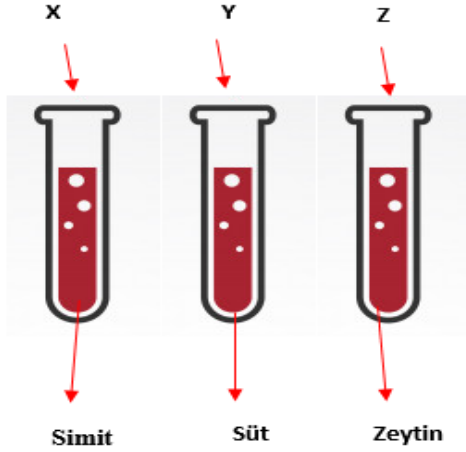


Görsel - 4: Pankreas ve Karaciğer görseli

Görsel - 4'te gösterilen organlardan pankreas ve karaciğer organlarının görevi en doğru şekilde hangisinde yazılmıştır?

- A) Sıvılar salgılayan organlardır.
- B) Salgıladıkları sıvıları ince bağırsağa gönderirler.
- C) Fiziksel (mekanik) ve kimyasal sindirimi sağlamada görev alırlar.
- D) Asıl görevleri sindirim olmayan fakat salgıladıkları sıvılar sayesinde sindirime yardımcı olan organlardır.

SORU 14:



Görsel - 5: Kimyasal sindirimin gerçekleştiği deney tüpleri

Görsel - 5 'te verilen deney tüplerinde sırasıyla simit, süt ve zeytin bulunmaktadır. Farklı sindirim organlarından alınan X, Y ve Z sıvıları görseldeki gibi deney tüplerine sırasıyla ilave edilmektedir. Deney tüplerinin hepsinde kimyasal sindirim gerçekleştiği gözlemlenmektedir.

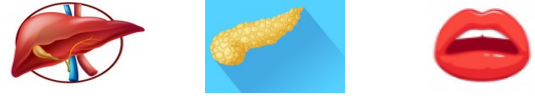
Buna göre X, Y ve Z ile gösterilen sıvıların hangi sindirim organından alındığı, aşağıda verilen seçeneklerden hangisinde doğru verilmiştir?

- | <u>X</u> | <u>Y</u> | <u>Z</u> |
|------------------|--------------|---------------|
| A) Ağız | Mide | İnce Bağırsak |
| B) Yemek borusu | Ağız | Mide |
| C) İnce bağırsak | Yutak | Ağız |
| D) Pankreas | Yemek borusu | Yutak |

CEVAP ANAHTARI

1	2	3	4	5	6	7	8	9
B	B	C	D	D	C	B	A	A
10	11	12	13	14	15	16		
A	D	B	D	A	B	C		

SORU 15:



Görsel - 5: Sindirim sistemine ait bazı yapı ve organlar

Sindirim sistemine ait bazı yapı ve organlar Görsel - 5 'te gösterilmiştir.

Görsel - 5 'te verilen sindirim sistemine ait bazı yapı ve organlar seçeneklerdeki bilgilerle eşleştirildiğinde hangi bilgi açıkta kalır?

- A) Tükürük sıvısı yardımıyla makarnanın kimyasal sindirimini başlatır.
- B) Salgıladığı mide öz suyuyla yumurtanın kimyasal sindirimini gerçekleştirir.
- C) Sindirime yardımcı organ olarak safra sıvısıyla zeytinin fiziksel sindirimini gerçekleştirir.
- D) Burada bulunan enzimler yulafın, yoğurdun, ayçiçeği yağının sindirimini gerçekleştirerek sindirime yardımcı olur.

SORU 16: Aşağıda verilen madde çiftlerinden hangileri sindirilmeden doğrudan kana karışır?

- A) Su - Kebap
- B) Tavuk- Pilav
- C) Mineral- Vitamin
- D) Çikolata- Baklava

Ek-2. Ölçüt başarı testi

Adı Soyadı:

Sınıfı/No.su:

Sevgili Öğrenciler,

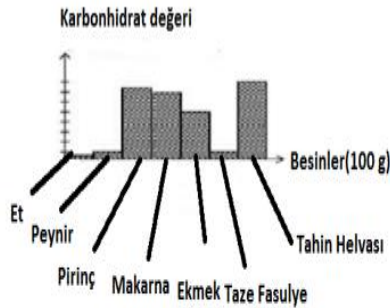
Bu ölçüt test "Vücudumuzdaki Sistemler ünitesindeki Sindirim Sistemi " konusuna yönelik öğrenme düzeylerini belirlemek amaçlı uygulanmaktadır. Her soru için dört seçenek bulunmaktadır. Dört tane seçenektan bir tanesini işaretlemeniz gerekmektedir.

Başarılar

SİNDİRİM SİSTEMİ ÖLÇÜT GEÇERLİĞİ TESTİ

SORU 1

1.



Doktor, Ayşe'ye fazla kilo aldığıni söyleyip beslenme uzmanına (diyetisyene) göndermiştir. Beslenme uzmanı, karbonhidratlı besinleri az yemesini önerip, bazı besinlerdeki karbonhidrat değerlerini yukarıdaki grafikte anlatmıştır. Ayşe, miktarları eşit olan, aşağıdaki yemeklerden hangisini yerse, beslenme uzmanının önerisine uymuş olur?

- A) Etili pirinç pilavı
- B) Etili taze fasulye
- C) Peynirli makarna
- D) Ekmek arasında tahin helvası

SORU 2:

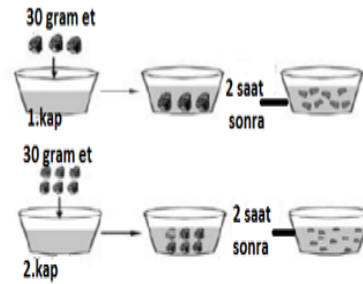
2. Aşağıda verilen birbiriyle bağlantılı cümleler arasındaki D "Doğru", Y "Yanlış" anlamındadır. Sindirim sistemi hakkında yeterli bilgiye sahip olan bir öğrenci kaçını çıkışa ulaşır?



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

SORU 3:

3.



Bir öğrenci, içinde eşit miktarda mide öz suyu bulunan I. ve II. kaplara et parçalarını koyuyor. Et parçalarında şekillerdeki gibi değişim gözleyen öğrenci, 2 saatlik bu gözlemi ile ilgili olarak hangi sonuca ulaşabilir?

- A) Büyük parçalı besinler, midede sindirime uğramaz
- B) İyi çiğnenen besinler, midede daha küçük parçalara ayrılır.
- C) Mide öz suyu, her türlü besinin kimyasal sindiriminde etkilidir.
- D) Hayvansal besin çeşitleri, midede fiziksel sindirime uğramaz

SORU 4:

4. Bir öğrenci kimyasal sindirimi gözlemlemek istiyor.

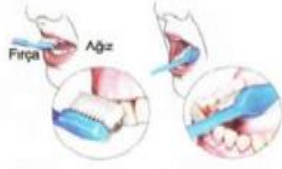


Enzimlerin etkili olabilmesi için ortamda belirli bir miktar su bulunması gereklidir. Buna göre öğrenci, şekildeki malzemelerle nasıl bir düzenek hazırlamalıdır?

- A) Besinleri bir deney tüpüne, suyu ve enzimi diğer deney tüpüne koymalıdır.
- B) Besinleri ve suyu bir deney tüpüne, enzimi diğer deney tüpüne koymalıdır.
- C) Besinleri ve suyu eşit miktarda her iki deney tüpüne, enzimi bu tüplerden sadece birine koymalıdır.
- D) Besinleri bir deney tüpüne, enzimi diğer deney tüpüne, suyu her iki deney tüpüne eşit miktarda koymalıdır.

SORU 5:

5. Sindirim sisteminin sağlığını korumak için yapılan bir etkinlik, şekilde gösterilmiştir.



Belirtilen etkinlik, hangi faaliyetin sürekliliğini sağlamak için yapılır?

- A) Yeterli ve dengeli beslenmeyi sağlamak
- B) Besinlerin fiziksel sindirimini yeterliliğini sağlamak
- C) Besinlerin kimyasal sindirimini sürekliliğini sağlamak
- D) Besinlerin tadının yeterli oranda algılanmasını sağlamak

SORU 6:

6. Öğrenciler nişastanın sindirimini, sindirim sisteminin farklı bölümlerinde şekildeki gibi göstermiştir.

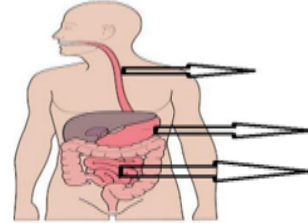


Nişastanın sindirimi ile ilgili olarak aşağıdakilerden hangisi söylenebilir?

- A) Ağızda sindirime uğramamıştır.
- B) Midede kimyasal sindirime uğramıştır.
- C) İnce bağırsakta kimyasal sindirimi tamamlanmıştır.
- D) Sindirilen nişastanın tamamı kalın bağırsaktan emilmiştir.

SORU 7:

7. Öğretmenin sorduğu bir soruya Ahmet, şekil üzerinde ki işaretlenmiş üç organı göstererek doğru cevabı vermiştir.



Buna göre, öğretmenin sorduğu soru aşağıdakilerden hangisidir?

- A) Hangi organlar sindirime yardımcı olur?
- B) Hangi organlar sindirim sistemi organıdır?
- C) Hangi organlarda sindirilen besinler kana geçer?
- D) Vitamin, su ve mineral hangi organ tarafından emilerek kana geçer?

SORU 8:

8. Kalın bağırsağın görevi aşağıdakilerden hangisidir?

- A) Sindirimin tamamlandığı yerdir.
- B) Besinlerin sindirilip kana geçişini sağlar
- C) Sindirime yardımcı olan pankreas öz suyunu salgılar
- D) Su, mineral ve vitaminin kana geçişini sağlandığı yerdir

9. Yutulan besinlerin yemek borusundan geçmesini, aşağıdakilerden hangisi sağlar?

- A) Dil
- B) Mide
- C) Yutak
- D) Gırtlak

SORU 9

9. Yutulan besinlerin yemek borusundan geçmesini, aşağıdakilerden hangisi sağlar?

- A) Dil
- B) Mide
- C) Yutak
- D) Gırtlak

Sindirim Sistemi Organları	Süre (Yaklaşık değerler)
Ağız	20 saniye
Yemek borusu	10 saniye
Mide	2-6 saat
İnce Bağırsak	5 saat
Kalın Bağırsak	24 saat

Yediğimiz bir besinin sindirim sistemi organlarında kaldığı yaklaşık değerler yukarıdaki tablodaki gibidir. Buna göre tablo için aşağıdakilerden hangisi yanlıştır?

- A) Kalın bağırsakta besin atıkları daha uzun süre kalırlar.
- B) Yediğimiz besinlerin atıkları bir gün sonra dışarı atılırlar.
- C) Yediğimiz besinler en kısa olarak yemek borusunda durmaktadır.
- D) Yediğimiz besinler mide de ve ince bağırsakta kaldığı süreler eşit olabilir.

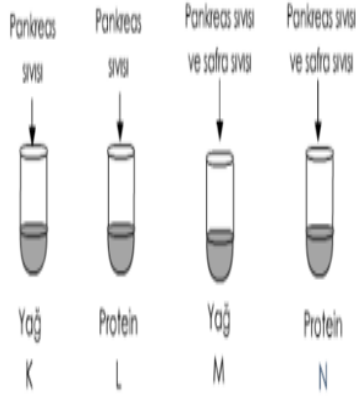
11.



Enzimler büyük moleküllü besinleri küçük moleküllere dönüştürmekte görevlidir. Buna göre enzimler yanda verilen gruplardan hangisine etki eder?

- A) Su, vitamin, mineral
- B) Vitamin, protein, yağ
- C) Su, protein, karbonhidrat
- D) Karbonhidrat, protein, yağ

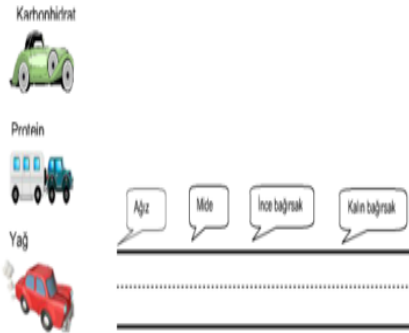
12.



Yukarıdaki deney düzeniği hazırlayan Can safra sıvısının sindirimi kolaylaştırdığı sonucuna ulaşmak istiyor. Can'ın istediği sonuca ulaşabilmesi için hangi düzenekleri seçmesi gerekir?

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

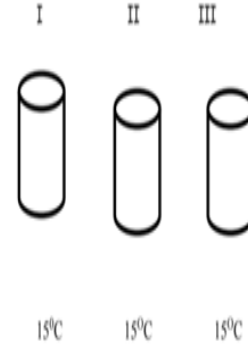
13.



Yukarıda karbonhidrat, protein ve yağ araçlarını görmekteyiz. Bu araçların sindirime başladıkları yerde benzin aldıkları, sindirimlerinin bittikleri yerde de yolculuklarını bitirdikleri bilinmektedir. Sizler araçların benzin aldıkları ve yolculuklarını bitirdikleri yerleri aşağıdaki seçeneklerden birini seçerek cevaplayınız?

Besinleri Aldıkları Yer	Yolculuklarını Bitirdikleri Yer
A) Ağız	İnce Bağırsak
B) Mide	İnce Bağırsak
C) İnce Bağırsak	İnce Bağırsak
D) Kalın Bağırsak	İnce Bağırsak

14. Şimdi de besinlerin sindirimini Mehmet arkadaşımız araştırmak ister. Aşağıdaki özdeş kaplarda deney düzeneklerini kurar. Mehmet aşağıdaki seçeneklerde belirtmiş olduğu durumlardan hangisini savunmazdınız?



15°C 15°C 15°C
20 gr enzim 15 gr enzim 20 gr enzim
250 gr protein 200 gr protein 200 gr protein

- A) Sıcaklık sindirim hızını etkiler.
- B) Besin miktarı sindirim hızını etkiler.
- C) Enzim miktarı sindirim hızını etkiler.
- D) Hangi kaptaki sindirim daha çabuk olur.

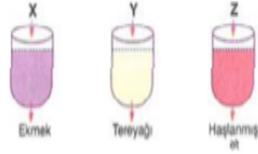
15.



Yukarıdaki organlardan kaç tanesi sindirim sisteminde bulunur?

- A) 4
- B) 5
- C) 6
- D) 7

16. Üç farklı organdan alınan X, Y, Z salgıları şekildedeki tüplere ekleniyor. Bir süre sonra, bu tüplerdeki besinlerin sindirimi gerçekleşiyor.



Buna göre, belirtilen salgıların alındığı organlar, hangisinde doğru eşleştirilmiştir?

- | X | Y | Z |
|------------------|---------------|---------------|
| A) Ağız | İnce Bağırsak | Mide |
| B) Ağız | Mide | İnce Bağırsak |
| C) Mide | Ağız | İnce Bağırsak |
| D) İnce Bağırsak | Mide | Ağız |

17. Bir beslenme uzmanı, beslenme ve sindirimle ilgili olarak şu açıklamayı yapıyor.

Beslenme Uzmanı: Yiyeceklerimiz ağızımızda ne kadar uzun süre çiğnersek o kadar kolay sindirilir ve daha hızlı kana geçer.

Buna göre, uzmanın açıklamasına gösterilebilecek neden, aşağıdakilerin hangisinde verilmiştir?

- A) Çiğneme kimyasal sindirimi sağlar. Kimyasal sindirim de besin emilimini artırır.
- B) Çiğneme fiziksel sindirimi sağlar. Fiziksel sindirim de enzim etkinliğini artırır.
- C) Çiğneme besinlerin yutulmasını sağlar. Hızlı beslenme kana besin geçişini hızlandırır.
- D) Çiğneme besinlerin tadının algılanmasını sağlar. Tadı algılanan besinler daha hızlı sindirilerek kana emilir.

18. İnsanın ince bağırsağında bulunan çok sayıda parmaklı yapıdaki villuslar, bir canlıya aşağıdaki özelliklerden hangisini kazandırır?

- A) Besinlerin daha hızlı sindirilmesi
- B) Besinlerin daha fazla sindirilmesi
- C) Sindirilen besinlerin daha fazla depolanması
- D) Sindirilen besinlerin kana daha hızlı emilmesi

19. Sağlıklı bir insan, bir günlük beslenmesinde, sadece 300 gram tereyağı tüketiyor. Bu bireyin sindirim sürecinde üç sindirim organındaki yağ miktarı nasıl değişim geçirir?

A)



B)



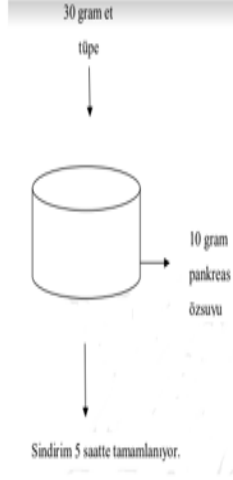
C)



D)



20. Bir uzman, şekildeki gibi düzeneği kurarak, etin sindirim durumunu araştırıyor ve belirtilen sonuca ulaşıyor.



Etin sindirim süresini azaltmak için;

1. Etin fiziksel olarak parçalanması
2. Kaba 10 gram daha pankreas özsu eklenmesi
3. Kaba 10 gram safra sıvısı eklenmesi

Değişikliklerinden hangileri yapmaya karar vermelidir?

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

CEVAP ANAHTARI

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
B	B	B	C	B	B	B	D	C	B	D	C	A	C	D	A	B	C	A	A

EXTENDED SUMMARY

When the literature on the units in the sixth grade science course is scanned, there are no valid and reliable achievement test development studies related to the "Digestive System". From this point of view, this study is important in that it is a valid and reliable achievement test, questions are asked from all levels of Bloom's steps (except synthesis), more questions are asked from the steps that measure higher level knowledge, and there is no other multiple-choice test in the literature that includes these features. offers.

It is not possible to talk about an education and training process without measurement and evaluation (Başol, 2015). While making observations whether it has a certain detail in the measurement process or in the future, sometimes errors can be confused with the observations and the actual state of an observed feature may not be reflected due to this error. In this case, the real value is tried to be given (Çelen guess, 2008). Spearman-Brown method was used for the errors that may deviate from the classical test assumption guidance guideline (Baykul, 2015). To fulfill the provisions of some statistical theories in order to make these center estimates (Baykul, 2015). There are two important coercion theories on the subject (Çelen, 2008). One of these theories is "Classical Test Theory". There is a valid and reliable audit of the feature to be measured (Çelen, 2008). Existing texts have also attempted to develop an achievement test based on the principles of classical test theory. Some loaded tools are used to determine whether the information given to the student has been understood, and if so, how much has been understood. One of these escape tools is elective tests (Çalışkan & Yiğittir, 2008). In the process of preparing the questions well and grouping the questions, the best known among the existing taxonomies is the Bloom taxonomy, which was developed in 1956 under Benjamin Bloom's own name (Çakıcı et al., 2012).

The survey design of the quantitative research method was used in the research. Survey design is a type of research conducted to reveal the existing situation (Çepni, 2010). In studies using survey design, the situation, behavior or thoughts in the population in general are described statistically through studies on a sample group selected from a population (Creswell, 2013).

Accessible population of the research is the sixth grade students of secondary school studying in Şanlıurfa Siverek, the sample is; The sample consists of 229 sixth grade students studying in secondary schools selected in Şanlıurfa Siverek district in the 2020-2021 academic year, which was determined by cluster sampling from the random sample type.

In order to decide on the questions, the 2018 Science Curriculum was taken as the basis. The questions were formed on the basis of the achievements in the "Digestive System" in the "Systems in Our Body" unit in the "Living Things and Life" subject area in this curriculum. Then, in addition to the gains in the curriculum, question-based learning outcomes were written. The literature was used while writing these achievements (Bamadan Duman, 2019; Kurtcuoğlu, 2007; Kaya and Gül, 2020). While creating the questions, care was taken to include questions measuring the achievements from all the steps in Bloom's taxonomy. Within the scope of this study, it was aimed to develop a valid and reliable achievement test consisting of 25 multiple-choice questions that have the feature of being summative regarding the sixth grade digestive system unit. This test was rearranged with the Google Forms application, and its edited version was sent to public secondary schools in Siverek through social media applications. As a criterion test, a test with proven validity and reliability in the literature was used.

Validity and reliability studies of the achievement test were carried out for the analysis of the data. In this context, descriptive statistics, item analysis (difficulty, distinctiveness), tetrachoric factor analysis (EFA and CFA) were examined. EFA and CFA were applied on the same sample. FACTOR, MS EXCEL 2020 and SPSS.25 package programs were used in the analysis. The results of the descriptive statistical analyzes made in SPSS.25 and whether the data obtained from the measurement tool showed normal distribution were examined. Data normality analysis, item difficulty and item discrimination index unrelated samples t-test, reliability analysis, exploratory factor analysis (EFA)

and confirmatory factor analysis (CFA) were performed within the scope of validity and reliability indices, respectively. The scores obtained by the students were listed in the Excel program in the form of upper and lower groups according to their score rankings, and unrelated samples t-test analysis was performed in the SPSS.25 package program. In the study, the KR-20 value was checked to perform the reliability analysis. The Kuder Richardson (KR-20) value is a widely used method in achievement tests (Barchard & Hakstian, 1997).

The aspect that makes the present study superior to most studies in the literature is that it is an achievement test development study based on classical test theory, with a certain taxonomy, in accordance with the Bloom's taxonomy on which questions are based, taxonomic attention was paid to the steps and meticulously prepared by taking expert opinions. While this achievement test is a supportive study with its similarity to the validated and reliable tests in the literature, the analyzes that do not have a philosophy on which it is based or whose theoretical foundations are not explained, whose content validity is problematic, taxonomy information is not given or taxonomic inaccuracies are low, reliability and difficulty are low. It can be stated that it is important in terms of setting an example for other studies and new research that have not been done correctly and seeing the deficiencies (Bolat and Karamustafaoğlu, 2019; Demir and Dindar, 2006; Demir and Öner Armağan, 2019; Gündüz, 2009; Gülmez Güngörmez and Akgün, 2018; Gül and Kargın, 2021; Keçeci et al., 2019). This developed test can be used by teachers to determine the learning levels of their students on the "Digestive System", and it can also be used by researchers as a data collection tool.