

Development of an Instrument Measuring Middle School Students' Achievement of Pressure

Seçil BOZ

Aksaray University, Aksaray-Turkey

Hasan ÖZCAN

Aksaray University, Aksaray-Turkey

Ayberk BOSTAN SARIOĞLAN

Balıkesir University, Balıkesir-Turkey

Article History

Submitted: 21.09.2022

Accepted: 23.03.2023

Published Online: 05.04.2023

Keywords

Pressure
Achievement test
Test development

Abstract

Purpose: This study aims to develop a valid and reliable achievement test for measuring pressure.

Design & Methodology: Eighty seventh-grade students constitute the sample of the study. A convenience sampling strategy was used in the selecting of the participants. In the first stage of the research, the objectives were determined, and then an item pool of 30 items was created. The content validity of the test was established with the help of specification table. In the selection of questions, exams and objective comprehension tests administered by the Ministry of National Education in the past years were used. In the pilot study, the test was administered to 20 students and two items were deleted based on pilot study data. After the actual implementation of the achievement test, overall test difficulty, overall test discrimination, item difficulty, and item discrimination values were calculated.

Findings: The overall test difficulty value was .56, the overall test discrimination value was .47, and the KR-20 internal consistency coefficient was calculated to be .85. Five items were removed from the test based on the item difficulty and item discrimination values. As a result of the study, a valid and reliable test consisting of 23 multiple-choice questions was obtained.

Implications & Suggestions: The developed achievement test can be used by researchers and science teachers to measure students' conceptual understanding of pressure.



Ortaokul Öğrencilerinin Basınç Konusu ile İlgili Bilgilerini Ölçmeye Yönelik Bir Başarı Testinin Geliştirilmesi

Seçil BOZ

Aksaray Üniversitesi, Aksaray-Turkey

Hasan ÖZCAN

Aksaray Üniversitesi, Aksaray-Turkey

Ayberk BOSTAN SARIOĞLAN

Balıkesir Üniversitesi, Balıkesir-Turkey

Makale Geçmişi

Geliş:21.09.2022

Kabul: 23.03.2023

Online Yayın: 05.04.2023

Anahtar Sözcükler

Basınç
Başarı testi
Test geliştirme



DOI:10.29129/inujse.1178479

Öz

Amaç: Bu araştırma ile basınç konusuna yönelik geçerli ve güvenilir bir başarı testinin geliştirilmesi amaçlanmaktadır.

Yöntem: Yedinci sınıfta öğrenim görmekte olan 80 öğrenci, çalışmanın örneklemini oluşturmaktadır. Örneklem seçimi uygun örnekleme yoluyla yapılmıştır. Araştırmanın ilk aşamasında kazanımlar belirlenmiş, ardından belirtke tablosuyla kapsam geçerliğini sağlayacak 30 maddelik bir soru havuzu oluşturulmuştur. Soruların seçiminde geçmiş yıllarda Millî Eğitim Bakanlığı tarafından uygulanmış sınavlardan ve kazanım kavrama testlerinden yararlanılmıştır. Pilot olarak 20 öğrenciye uygulanan testten iki maddenin çıkarılmasına karar verilmiştir. Başarı testinin asıl uygulanmasından sonra ortalama güçlük, ortalama ayırt edicilik, madde güçlük ve madde ayırt edicilik değerleri hesaplanmıştır.

Bulgular: Testin ortalama güçlük değeri 0,56, ortalama ayırt edicilik değeri 0,47, KR-20 iç tutarlık katsayısı ise 0,85 olarak hesaplanmıştır. Madde güçlük ve ayırt edicilik değerlerine bakılarak beş madde daha testten çıkarılmıştır. Çalışma sonucunda 23 çoktan seçmeli sorudan oluşan geçerli ve güvenilir bir test elde edilmiştir.

Sonuçlar ve Öneriler: Geliştirilen başarı testi, basınç konusunda öğrencilerin kavramsal anlamalarını ölçmede araştırmacılar ve fen bilimleri öğretmenleri tarafından kullanılabilir.

GİRİŞ

Okullarda değerler ve yetkinlikler ile bütünleşen bilgi, beceri ve davranışlara sahip öğrenciler yetiştirmek amaçlanmaktadır (Millî Eğitim Bakanlığı [MEB], 2018). Bu amaca yönelik olarak da öğretim düzenlenmektedir. Verilen öğretimin sonuçlarının alındığı ve anlam kazandığı süreç, ölçme ve değerlendirme olarak ifade edilmektedir (Brown, Nidumolu, McConnell, Hecker ve Grierson, 2019). Ölçme ve değerlendirme süreci sonunda hem öğretimin etkililiđi hem de öğrencilerin bilişsel yapıları hakkında bilgi sahibi olunabilmekte ve gerekli noktalara müdahale edilebilmektedir. Bu amaca yönelik çok çeşitli başarı testleri kullanılmaktadır. Akademik başarıların ölçülmesi amacıyla klasik, boşluk doldurma, doğru-yanlış, eşleştirme ve çoktan seçmeli gibi pek çok ölçme aracı bulunmaktadır (Gullickson, 1985; Özsevgeç, Çepni ve Demirciođlu, 2004; Wise, 2015). Kullanılan bu ölçme araçlarının faydalarının yanında bazı sınırlıkları da mevcuttur. Klasik yazılı sorular, açık uçlu veya kapalı uçlu doğrudan cevabın yazıldığı sorular olup kapsam geçerliliđini sağlama noktasında sınırlı testlerdir (Couch, Hubbard ve Brassil, 2018; Timur, Dođan, İmer Çetin, Timur ve Işık, 2019). Bununla birlikte klasik yazılı sorular uygulayan kişinin değerlendirmesi açısından uzun zaman isteyen ve objektifliđi sağlamanın oldukça zor olduđu bir test türüdür (Tekindal, 2000). Cevabın sadece “Dođru” veya “Yanlış” şeklinde işaretlenmesi istenen test türü, doğru-yanlış testleri olarak ifade edilmektedir. Doğru-yanlış test grubunda kapsam geçerliliđi soru sayısı ile beraber artırılabilirken doğru cevap verme ihtimalinin %50 olması şans faktörünü artırmakta ve testi bu bakımdan sınırlı duruma getirmektedir (Timur ve diđerleri, 2019). Boşluk doldurma sorularında verilen boşluklara uygun kelimelerin yerleştirilmesi istenmektedir. Ancak bu testlerin büyük ölçüde bilgi düzeyini ölçmeye yönelik olması, testin öne çıkan sınırlılıđı olarak ifade edilebilir. Yine aynı düzeye hitap eden başka bir test grubu da eşleştirme sorularıdır. Eşleştirme soruları genellikle yeni konu veya kavramların öğrenildiđi ilk süreçte tercih edilmektedir. Bahsedilen tüm bu ölçme araçlarının içerisinde çoktan seçmeli testler en yaygın kullanılanlar arasındadır (Saraç, 2018; Tamir, 1971; Treagust, 1986; Yaman, 2016).

Gerekli ve yeterli soru sayısı ile geçerlik türlerinden biri olan kapsam geçerliliđinin (Keleciođlu ve Göçer Şahin, 2014) sağlanabildiđi çoktan seçmeli başarı testlerinde kullanılan seçenek sayısına göre şans faktörü de azaltılabilmektedir (Çardak ve Selvi, 2018; Temizkan ve Sallabaş, 2011). Öğrencinin yaratıcılıđını kısıtlaması ve kendi cümleleriyle ifade özgürlüđünün olmaması bu test grubunun olumsuz yönü olsa da (Üstüner ve Şengül, 2004) sayılan bu ölçme araçları arasında en öne çıkan test türü olduđu görülmektedir. Ayrıca puanlanması kolay testler arasında yer almaktadır (Akkuş ve Baykul, 2001). Aynı zamanda çoktan seçmeli testleri kullanmanın bazı avantajları olarak; aynı anda çok fazla bilgiyi test etmesi, değerlendirmenin kısa zaman alması, değerlendirmede objektif olunması nedeni ile güvenilirliđi yüksek ölçümler yapılması ve her sorunun zorluk seviyesinin tespit edilmesinin kolay olması olarak sıralanabilir (Kuechler ve Simkin, 2003; Simkin ve Kuechler, 2005; Yaman, 2016). Bununla birlikte çoktan seçmeli testlerde belirli bir sınav süresinde çok sayıda soru sorulabildiđi için (Karataş, Köse ve Coştı, 2003) çok sayıda kazanımı ölçmede kullanabilir. Çoktan seçmeli testler bu açılarından öğretmenlere ve araştırmacılara kullanımda hız ve pratiklik sağlamaktadır. Bu nedenler ile de öğrencilerin bilişsel yapılarını araştırmaya yönelik birçok çalışmada çoktan seçmeli testler geliştirilmektedir (Bolat ve Karamustafaöđlu, 2019; Karslı, Karamustafaöđlu ve Kurt, 2019; Kızıkan ve Bektaş, 2018; Saraç, 2018; Üçüncü ve Sakız, 2020).

Öğrencilerin fen bilimleri dersinde bazı konuları öğrenmede zorlandıkları ve çeşitli kavram yanlışlıklarına yönelik bilişsel yapıya sahip oldukları bilinmektedir. Basınç konusu da bu konular arasında yerini almaktadır (Can, 2019; Kaya ve diđerleri, 2018; Özcan ve Koca, 2019). Basınç konusunun öğrenciler tarafından anlaşılması zor olduđu gibi öğrencilerde bu kavrama ilişkin çeşitli kavram yanlışlıkları ile de karşılaşmaktadır. Bu kavram yanlışlıklarından bazıları şu şekilde sıralanabilir; basınç sadece ağırlığa bağlıdır (Can, 2019), basınç ile güç kavramı aynıdır (Can, 2019; Özdemir Benli, 2021), katı basıncı temas ettiđi yüzeyin cinsine bağlıdır (Can, 2019; Özdemir Benli, 2021); sıvı basıncı sıvının miktarına bağlıdır

(Bostan Sariođlan ve Can, 2021; Özdemir Benli, 2021); rakım günlük basınç deđişikliklerini etkiler (Akbas ve Gencturk, 2011), gazlar basınç uygulamaz çünkü gazlar yerinde durmaz (Özdemir Benli, 2021). Öğrencilerin basınç konusu ile ilgili bu kadar kavram yanılıđına sahip olması bu kavrama ilişkin kavram yanılıđlarının tespit edilmesinin ve bu kavram yanılıđlarına yönelik öğretimin önemini göstermektedir. Bu nedenle uygun yöntem ve tekniklerle desteklenecek bir öğretimin yanı sıra nitelikli bir başarı testi önemlidir. Basınç konusunun öğrencilerin anlamakta zorlandıkları bir konu olduđu için de bu konuya ilişkin öğrencilerin bilişsel yapılarını araştırmaya yönelik çalışmalarda yürütölmüştür. Alanyazın incelendiđinde ölkemizde ortaokul düzeyinde farklı birçok konuda geliştirilen başarı testlerinin olduđu görölmektedir. Bu testler çoktan seçmeli (Divarcı ve Kaya, 2019; Özcan, Koca ve Söğüt, 2019; Uçar ve Aktamış, 2019) olabileceđi gibi iki aşamalı testlerden (Ayvacı ve Durmuş, 2016) ve üç aşamalı testlerden (Çetinkaya ve Taş, 2016) de oluşmaktadır. Bunlardan basınç konusu ile ilgili olan çalışmalar açıklamaları ile Tablo 1’de paylaşılmaktadır.

Tablo 1

Basınç Konusunda Yapılmış Çalışmalarda Kullanılan Başarı Testleri

Baytok (2007)	Kazanımlar belirlenerek, ders kitaplarından faydalanılarak 50 maddelik soru havuzu oluşturulmuş, uzman görüşü alınarak bu maddeler 33’e düşürölmüş, pilot uygulama sonucu öğrenci görüşleri alınarak son hali verilmiş ve 20 maddelik bir başarı testi elde edilmiş. Çalışmada güçlük, ayırt edicilik, KR-20 iç tutarlık katsayısı gibi deđerlere yer verilmediđi görölmektedir.	Ortaokul 7. sınıf öğrencileri
Kirişciođlu (2007)	Kazanımlar belirlenerek kazanımlara uygun 51 çoktan seçmeli soru havuzu oluşturulmuş, uzman görüşüyle maddelerin düzenlenmesi yapılmıştır. Oluşturulan test 35 kişiye uygulanarak SPSS ile madde analizleri yapılmıştır. Araştırmacı tarafından başarı testinin KR-20 deđeri 0,92 olarak belirlenmiştir.	Ortaokul 7. sınıf öğrencileri
Taşkoyan (2008)	Kazanımlar belirlenerek 45 soruluk bir madde havuzu oluşturulmuş, uzman görüşü alınarak pilot uygulaması yapılmış ve analiz sonucunda testin ortalama güçlük indeksi 0,55, KR-21 0,841 olarak hesaplanmıştır.	Ortaokul 7. sınıf öğrencileri
Çakalliođlu (2008)	Kazanımlar belirlenerek ve kapsam geçerliđi dikkate alınarak 50 maddelik bir havuz oluşturulmuştur. Uzman görüşü alınarak pilot uygulaması yapılmış ve SPSS programında madde analizleri yapılarak uygun olmayan sorular atılmıştır. Madde analizleri yapılmış ancak ortalama deđerlere yer verilmediđi görölmektedir.	Ortaokul 7. sınıf öğrencileri
Bozan (2008)	Kazanımlar doğrultusunda açık uçlu problemler belirlenerek uzman görüşüne başvurulmuştur.	Ortaokul 7. sınıf öğrencileri
Şenyıl (2009)	Belirlenen hedef ve davranışlar doğrultusunda 40 maddelik havuz oluşturulmasının ardından uzman görüşüyle 35 maddeye indirilmiştir. Cronbach alfa katsayısı 0,8709 olarak tespit edilmiştir.	Ortaokul 7. sınıf öğrencileri
Bayri (2014)	Kazanımlar belirlenerek uzman görüşü alınmıştır.	Ortaokul 8. sınıf öğrencileri
Öztürk (2014)	Kazanımlar belirlenerek madde havuzu oluşturulmuş uzman görüşü alınarak pilot uygulama yapılmıştır. Madde analizler yapılan testteki	Ortaokul 8. sınıf öğrencileri

	verilerin analizi sonucu güvenilirlik katsayısı 0,81 olarak tespit edilmiştir.	
Divarcı (2016)	Kazanımların belirlenmesinin ardından madde havuzu oluşturularak uzman görüşü alınmış, pilot uygulama sonucu madde analizi yapılarak ortalama ayırt edicilik 0,41, ortalama güçlük 0,50 ve KR-20 güvenilirlik katsayısı 0,69 olarak tespit edilmiştir.	Ortaokul 8. sınıf öğrencileri
Özcan, Koca ve Söğüt (2019)	Kazanımların belirlenmesiyle madde havuzu oluşturulmuş ve belirtke tablosundan faydalanılmıştır. Ortalama madde ayırıcılık indeksi 0,404, ortalama madde güçlük indeksi 0,711, hesaplanan KR-20 iç tutarlık katsayısı da 0,73 olarak tespit edilmiştir.	Ortaokul 7. sınıf öğrencileri

Araştırmanın Amacı ve Önemi

Geçerlik ve güvenilirlik özelliklerini yüksek düzeyde sağlayabilen testler eğitim dahil her alanda güven verici şekilde kullanılabilir (Ercan ve Kan, 2004). Alanyazın incelendiğinde basınç konusuna yönelik geçerlik ve güvenilirlik analizi yapılan çalışmaların azlığı dikkat çekmektedir. Çalışmaların birçoğunun 2013 yılı öncesinde yapıldığı ve basınç konusu ile ilgili kazanımların değişmesinden dolayı bu başarı testlerinin kavram yanlışlarının tespitinde ve öğretimin bu kavram yanlışları doğrultusunda düzenlenmesinde yetersiz kalacağı görülmektedir. Yukarıda da bahsedildiği üzere, öğrenciler basınç konusunu anlamada zorlandıkları gibi öğrencilerde bu kavrama ilişkin çeşitli kavram yanlışları da mevcuttur. Bu nedenler ile öğrencilerin zorlandığı ve çeşitli kavram yanlışlarına sahip olduğu bir konuda öğretmenlerin kullanımına sunulmak üzere daha fazla sayıda nitelikli başarı testi ihtiyacı duyulduğu ifade edilebilir (Altay ve Balım, 2021). Bu sayede de öğretmenler öğrencilerde karşılaşılan kavram yanlışlarını tespit edebilir (Ecevit ve Şimşek, 2017) ve öğretimi bu kavram yanlışlarının giderilmesi yönünde düzenleyebilir (Demirciođlu, Aydın ve Demirciođlu, 2012). Geçerlik ve güvenilirlik çalışmaları yapılmış çok sayıda başarı testi bu açıdan öğretmenlerin kullanımına sunulması yararlı olacaktır. Aynı zamanda bu testler ile öğrencilerin ön bilgileri belirlenebileceği (Ercan, Bilen ve Ural, 2016) gibi öğretim sonrası değerlendirme amacı ile de kullanılabilir (Saraç, 2018). Bu bağlamda basınç konusunda çalışmalar yürütmeyi planlayan araştırmacılar da geçerlik ve güvenilirlik çalışmaları yapılmış bir başarı testi kullanabilirler. Basınç konusunda kullanılmak üzere çok sayıda başarı testi öğretmenler ve araştırmacılar tarafından farklı amaçlar ile kullanılabilir. Bu ihtiyacı karşılamak üzere bu çalışmada basınç konusuna yönelik geçerlik ve güvenilirlik çalışmaları yapılmış bir başarı testinin geliştirilmesi amaçlanmaktadır. Bu bağlamda yapılan çalışmanın sınırlı sayıdaki basınç başarı testi geliştirilmesi çalışmalarına katkı sağlayacağı düşünülmektedir. Bu çalışmada geliştirilmesi amaçlanan başarı testi çoktan seçmeli sorulardan oluşmakta olup, MEB tarafından yayınlanan sorulardan oluşmaktadır. Böylece öğretmenler tarafından daha sıklıkla kullanılabilir. Aynı zamanda bu başarı testi ile öğrencilerde basınç konusu ile ilgili sahip oldukları hatalı fikirleri belirlenebilecek ve öğretmenler bu kavram yanlışlarını gidermeye yönelik öğretim düzenleyebileceklerdir.

Bu amaç doğrultusunda aşağıda belirtilen sorulara cevap bulunmaya çalışılacaktır:

1. Basınç konusunda geliştirilen başarı testi yapı geçerlik düzeyi ne seviyededir?
2. Basınç konusunda geliştirilen başarı testindeki maddelerin madde güçlük ve madde ayırt edicilik değeri nedir?
3. Geliştirilen başarı testinin ortalama güçlük ve ayırt edicilik düzeyi ne durumdadır?
4. Geliştirilen testin güvenilirlik düzeyi nedir?

YÖNTEM

Araştırma Deseni

Bu araştırma ortaokul düzeyinde kullanılabilecek bir başarı testi geliştirmeye yönelik yürütülmüştür. Araştırmada ölçeğın kapsam geçerliğinin belirlenmesinde uzman görüşü alınırken nitel araştırma desenlerinden (Shuttleworth, 2016), ölçek maddelerinin geçerlik ve güvenilirlik oranlarını belirlemede nicel araştırma desenlerinden genel tarama deseninde yararlanılmıştır.

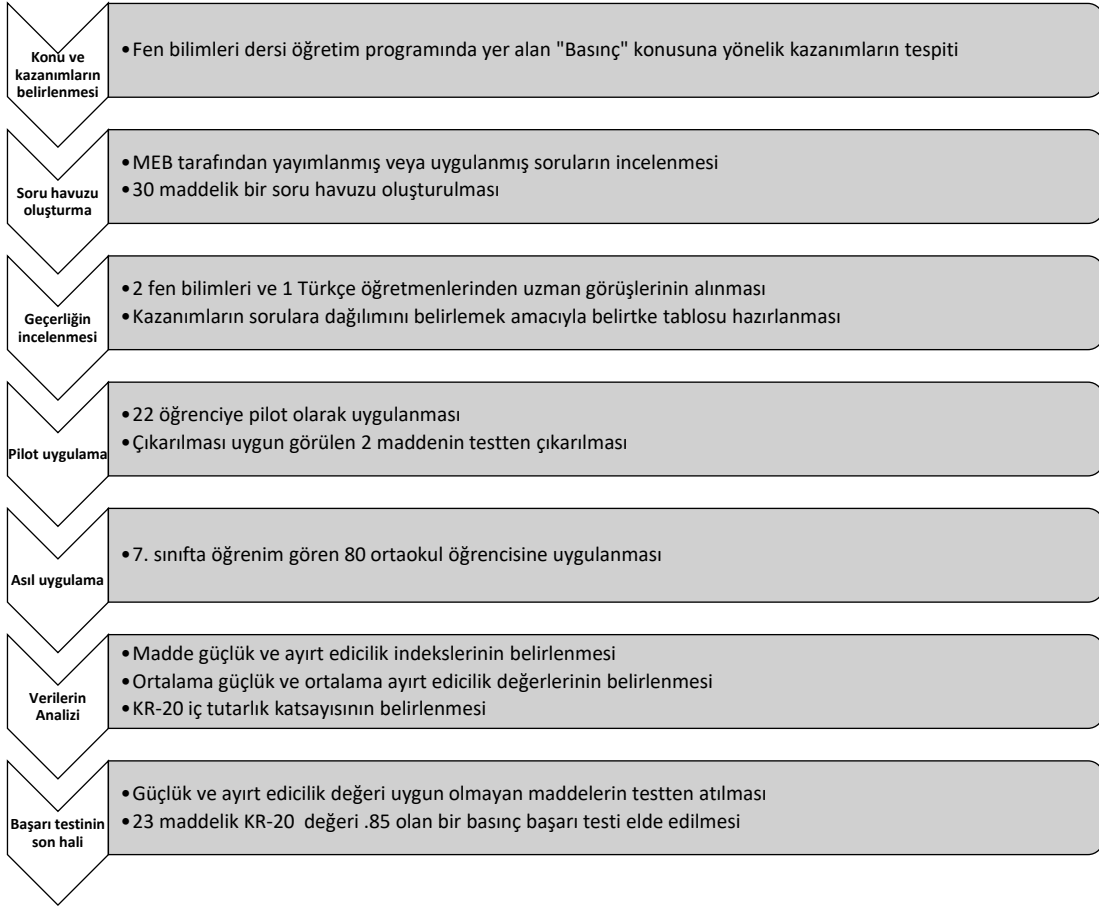
Katılımcılar / Evren ve Örneklem / Çalışma Grubu

Çalışma grubunu, Aksaray ilinde ortaokula devam etmekte olan 80 yedinci sınıf öğrencisi oluşturmaktadır. Test geliştirmede 50-100 arası örneklem büyüklüğü zayıf olarak tanımlanmasına rağmen test geliştirmede kabul edilebilir bir sayı olduğu belirtilmektedir (Netemeyer, Bearden ve Sharma, 2003). Yedinci sınıf öğrencileri basınç konusu ile ilgili öğretimi ölçek uygulanmadan hemen önce aldıkları için tercih edilmiştir. Örneklem seçiminde seçkisiz olmayan örnekleme türlerinden biri olan uygun örneklem yöntemi kullanılmıştır. Araştırmacıardan biri belirlenen okulda görev yapmaktadır. Pratik ve katılımcılara ulaşılması kolay olduğu için bu örnekleme yöntemi tercih edilmiştir (Balci, 2021).

Veri Toplama Araçları

Bu çalışma 2018-2019 eğitim öğretim döneminde gerçekleştirilmiştir. Bu çalışmada basınç konusu ile ilgili geçerli ve güvenilir ölçümler yapan bir başarı testi geliştirmek amaçlandığı için burada ilk olarak geçerlik ve güvenilirlik kavramlarından bahsedilmektedir. Geçerlik, sadece ölçülmesi amaçlanan özelliđi, başka özellikleri karıştırmadan ölçmek olarak ifade edilebilir (Büyüköztürk, 2019; Tekin İftar, 2018). Bir testin geçerliđi varlık veya yoklukla deđil “yüksek”, “orta” ve “düşük” gibi derecelerle belirtilmektedir. Geçerliđi yüksek testler amaca yüksek oranda hizmet etmektedir. Güvenirlik ise testin ölçmek istediđi özelliđi ölçüm kaç kez tekrarlanırsa tekrarlansın aynı şartlarda kararlı bir şekilde ölçmesi olarak belirtilebilir (Büyüköztürk, 2019; Karasar, 2008). Tesadüfi hatalardan arınık bir ölçüm, güvenilirliđin bir diđer ifade şeklidir. Güvenirliđi özetleyen üç anahtar kelime kararlılık, tutarlılık ve duyarlılıktır (Büyüköztürk, Kılıç Çakmak, Akgün, Karadeniz ve Demirel, 2013).

Test geliştirme süreciyle ilgili adımlar aşağıda Şekil 1’de yer almaktadır.



Şekil 1. Test Geliştirme Süreci

Başarı testinde yer alan maddeler seçilirken, kazanımlar ve maddeler arasındaki uyumun belirlenmesinde Webb'in (1997) geliştirdiđi ölçütlerden üç tanesi göz önüne alınmıştır. Bu ölçütler ve kullanımı şu şekilde açıklanmaktadır:

- Kategorik birlik ölçütü:** Bu ölçüt ölçme aracında yer alan maddelerin kazanımların tamamını kapsama durumu hakkında genel bir bilgi vermeyi amaçlar. Bu amaçla ölçme değerlendirme için uygulanan testin soruları ile öğretim programında yer alan kazanımlar eşleştirilir. Uzmanların yaptığı bu eşleştirmenin uyum yüzdesi incelenir.
- Bilgi aralık uygunluğu ölçütü:** Bu ölçüt kazanımlar ile soruların uyumunu içermektedir. Öğrencilerin kazanım ile edinmeleri gereken davranış ve soruda yoklanmak istenen davranışın uyumunun ne derecede olduğunu açıklamayı amaçlamaktadır.
- Temsilin dengesi ölçütü:** Bu ölçüt soruların kazanımlara dağılımını içermektedir. Hazırlanan formülle birlikte soruların kazanımlara dağılımı hesaplanmaktadır.

Kategorik birlik ölçütünü test etmek için başarı testinde yer alan soruların kazanımlara göre dağılımı ve yenilenmiş Bloom taksonomisinde hangi basamaklarda yer aldıklarına ilişkin belirtke tablosu aşağıda açıklanmaktadır. Başarı testi geliştirme sürecinin ilk basamağında yedinci sınıf "Basınç" konusuyla ilgili 2013 yılı ve 2018 yılları fen bilimleri dersi öğretim programlarında yer alan kazanımlar belirlenmiştir. Tespit edilen kazanımlar ve hangi madde numaralarının hangi kazanımı ölçmeye yönelik sorulduđu Tablo 2'de yer almaktadır (Millî Eğitim Bakanlığı [MEB], 2013). 2018 yılı fen bilimleri dersi öğretim programındaki "Basınç" konusunda yer alan 3 kazanımın da 2013 yılı fen bilimleri programında yer alan

aşağıda yer alan kazanımlara benzer olduđu ve sadece konunun 8. sınıf düzeyinde işlendiđi görölmektedir.

Temsilin dengesi ölçütünde kazanımlara soru dağılımı aşağıda Tablo 2’de gösterilmektedir.

Tablo 2

Fen Bilimleri Dersi Basınç Konusuna Ait Kazanımlar

Kazanımlar	Madde Numarası
7.2.2.1. Katı basıncını etkileyen deđişkenleri deneyerek keşfeder ve bu deđişkenler arasındaki ilişkiyi analiz eder.	2,3,5,6,12,13,14,15,17,18,29
7.2.2.2.Sıvı basıncını etkileyen deđişkenleri deneyerek keşfeder ve bu deđişkenler arasındaki ilişkiyi analiz eder.	
a. Gazların da sıvılara benzer şekilde basınç uyguladıkları vurgulanır.	4,8,9,20,21,23,26,28,30
b. Sıvı ve gaz basıncını etkileyen deđişkenlere ve matematiksel bağıntılara girilmez.	
7.2.2.3.Katı, sıvı ve gazların basınç özelliklerinin günlük yaşam ve teknolojideki uygulamalarına örnekler verir.	1,7,10,11,16,19,22,24,25,27

Tablo 3’te her bir kazanım için basınç başarı testinde yer alan maddelerin yenilenmiş Bloom taksonomisine göre dağılımı yer almaktadır.

Tablo 3

Basınç Başarı Testi Belirtke Tablosu

Kazanımlar	Hatırlama	Yenilenmiş Bloom Taksonomisi Bilişsel Süreçleri				
		Anlama	Uygulama	Analiz Etme	Deđerlendirme	Yaratma
7.2.2.1.		2, 3, 13	5,6,12,18,29	14, 17	15	
7.2.2.2.		9, 20, 21,28	4, 8, 23, 30	26		
7.2.2.3.		11,16,19,22,27	1, 7, 24, 25		10	

Çoktan seçmeli olarak hazırlanan basınç başarı testinde yer alan maddelerin yenilenmiş Bloom taksonomisine göre en alt beceri olan hatırlatma ve en üst beceri olan yaratma basamaklarında yer almadığı görölmektedir. Bunun nedeni olarak da soruların çoktan seçmeli olarak hazırlanmasının öğrencilerin orijinal bir ürün elde etme becerilerini ölçemediđi için bu basamakta yer almadığı söylenebilir.

Belirlenen kazanımlara uygun ve kapsam geçerliğini sağlayacak şekilde MEB tarafından yayınlanmış ve uygulanmış sınavlardan 30 madde seçilmiştir. Bu soruların seçilmesinde öğrencilerin bu soru türleri ile sıklıkla karşılaşmaları ve öğretmenlerinde rahatlıkla uygulayabilecek olmalarıdır. Kazanımlar çoktan seçmeli maddeler ile oluşturulmak istenmiş ve bu maddeler ulusal düzeyde farklı yıllarda uygulanan sınavlardan (SBS, TEOG) ve MEB’in yayımladığı kazanım kavrama testlerinden (KKT) seçilmiştir. Özellikle bu soruların tercih edilmesinin sebebi soruların niteliğinin artırılmaya çalışılması olmuştur. Tablo 4’de belirtildiđi şekilde soru seçimi yapılmıştır.

Tablo 4

Basınç Başarı Testi Soru Havuzunun Oluşturulduğu Kaynaklar

Madde	SBS 2009	SBS 2010	SBS 2011	2014 TEOG (NİSAN)	2014 TEOG (MAYIS)	2016 TEOG (MAYIS)	MEB KKT (2015/ 2016 (7.sınıf	MEB KKT (2015/ 2016 (8.sınıf	MEB KKT (2016/ 2017 (8.Sınıf	MEB KKT (2017/ 2018 (8.Sınıf
1	*									
2		*								
3			*							
4				*						
5					*					
6						*				
7							*			
8							*			
9							*			
10							*			
11							*			
12								*		
13								*		
14								*		
15								*		
16								*		
17								*		
18								*		
19								*		
20								*		
21								*		
22									*	
23										*
24									*	
25									*	
26									*	
27									*	
28									*	
29										*
30										*

Ardından sorular bilgi aralık uygunluğu ölçütünün kontrol edilmesi için alan uzmanları tarafından incelenmiştir. Kapsam geçerliği ve görünüş geçerliği açısından incelenmesi için iki fen eğitimi alan uzmanı ile en az on yıl çalışmış iki fen bilimleri öğretmeninden ve dil yapısı ve anlaşılabilirliğinin kontrol edilmesi için bir Türkçe öğretmeninden görüşleri alınmıştır. Türkçe alan uzmanı testteki maddeleri dil ve anlatım biçimi yönünden kontrol etmiş ve maddelerin Türkçe yazım kurallarını uygun biçimde yazılıp yazılmadığını değerlendirmiştir. Fen eğitimi alan uzmanları başarı testindeki maddelerin tüm kazanımları kapsayıp kapsamadığını ve soruların programda yer alan ilgili kazanıma uygun olup olmadığını

deđerlendirmişlerdir. Deđerlendirme sonucunda alan uzmanları başarı testindeki maddelerin ilgili kazanıma yönelik hazırlandığı görüşünü belirtmişlerdir.

Uzman görüşleri doğrultusunda soruların kapsam ve görünüş geçerliğini sağlamada ve dilinin öğrenciler açısından anlaşılır olduğu kanaatine varılarak pilot uygulamaya başlanılmıştır. Uzman görüşlerine dayalı olarak testin kapsam geçerliğini sağladığı belirlenmiş ve sorularda deđişiklik yapılmamıştır. Pilot çalışmada 20 kişiye uygulama yapılmış ve neticesinde tekrarlayan seçenekler bulundurduğu için 28. madde testten çıkarılmıştır. Ayrıca bire bir aynı olduğu tespit edilen 9. ve 23. maddelerden 9. maddenin testten çıkarılmasına karar verilmiştir. Pilot uygulamada öğrencilerin soruları cevaplayabilmeleri için bir ders saati yani 40 dakikanın yeterli olduğu görülmüştür.

Verilerin Analizi

80 öğrenciye uygulanan basınç başarı testi dört seçenekli 28 çoktan seçmeli maddeden oluşmaktadır. Veri analizi SPSS 24.0 paket programı ve Microsoft Excel kullanılarak yapılmıştır. SPSS ve Excel’de soruları doğru cevaplayan öğrencilerin cevapları 1.00, yanlış cevaplayan ve cevap vermeyen öğrencilerin cevapları ise .00 olacak şekilde kodlanmıştır. Elde edilen verilerin betimsel analizi yapılmıştır. Maddelerin güçlük ve ayırt edicilik indeksleri hesaplanarak yorumlanmış ve ortalama deđerleri bulunmuştur. Madde güçlük indeksi ile bir testteki soruların doğru cevaplanma yüzdesini gösterirken, madde ayırt edicilik indeksi testte yer alan soruların istenilen özelliđi ölçme durumunu belirlemektedir. Güvenirlik analizi için KR-20 deđeri hesaplanarak deđerlendirilmiştir.

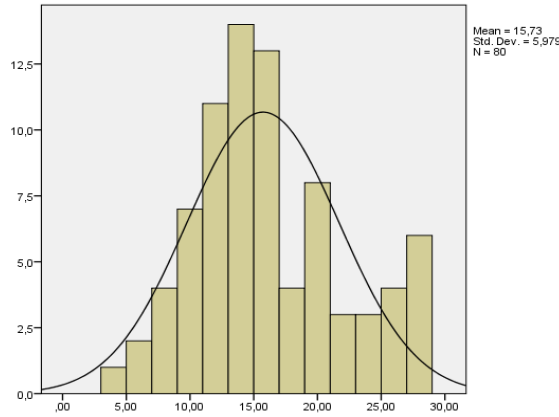
BULGULAR

Bu bölümde basınç kavramı ile ilgili başarı testi geliştirilme sürecine yönelik bulgulara yer verilmektedir. İlk olarak teste ait betimsel analizden elde edilen bulgular Tablo 5’te sunulmuştur.

Tablo 5
Basınç Başarı Testine Yönelik Betimsel Analiz

N	80
Ortalama	15.725
Medyan	15.00
Mod	15.00
Standart Sapma	5.979
Varyans	35.746
Çarpıklık	.535
Basıklık	-.372
Ranj	24.00
Minimum puan	4.00
Maksimum puan	28.00

Tablo 5 incelendiğinde testten alınan en düşük puan 4, en yüksek puan 28 olarak belirlenmiş olup ranjin 24.00 deđeri aldığı görülmektedir. Çarpıklık ve basıklık deđerlerinin +1.00 ile -1.00 arasında yani normal dağılım sınırları içerisinde kaldığı belirlenmiştir (Karaman, 2020). Aynı zamanda ortalama, mod ve medyan deđerlerinin birbirine yakın olması dağılımın normal olduğunun bir diđer göstergesidir (Kan, 2009). Şekil 2 incelendiğinde de dağılımın normal olduğu tespit edilmiştir.



Şekil 2. Normal Dağılım Grafiđi

Bir başarı testinde analiz edilmesi ve sonuçlarının incelenmesi gereken diđer özellikler ise madde güçlük ve madde ayırt edicilik değerleridir. Madde güçlük indeksi maddeyi doğru cevapların toplam katılımcı sayısına bölünmesiyle bulunabilir (Başol, 2019; Bayrakçeken, 2012; Kan, 2009). En yüksek 1 (bir) değerini alır, bu durum maddeyi tüm katılımcıların doğru cevapladığı anlamına gelir. Madde kimse tarafından doğru cevaplanmadığında ise bu değer 0 (sıfır) olur. Kısacası madde güçlük indeksi 1'e yaklaştıkça madde kolaylaşır, 0'a yaklaştıkça zorlaşır şeklinde yorumlanabilir (Bayrakçeken, 2012). Genellikle maddeler için güçlük indeksinin .50 olması beklenilmektedir (Hasańebi, Terzi ve Küçük, 2020) ve bu durum testte yer alan maddelerin orta güçlükte olduğunu göstermektedir. Çok kolay ya da çok zor maddelerden oluşan testlerin güvenilirlik düzeyleri de düşük olmaktadır (Hasańebi ve diđerleri, 2020) ve bu nedenle de bu maddelerin kullanılması tercih edilmektedir. Tablo 6'daki değerler incelendiğinde madde güçlük indeksi değerlerinin yorumlanması ifade edilmiştir (Başol, 2019).

Tablo 6

Madde Güçlük İndeks Deđerlerinin Yorumu

.85 - 1.00 arası	Çok kolay madde	Kullanılması tercih edilememektedir
.61 - .84 arası	Kolay madde	Kullanılabilir
.40 - .60 arası	Orta güçlükte madde	En uygun maddedir
.39 - .16 arası	Zor madde	Kullanılabilir
.15 - .00 arası	Çok zor madde	Kullanılması tercih edilmemektedir

Madde ayırt edicilik değeri, maddenin kazanımın anlaşıldığı kişilerce yapılması, anlaşılmayan kişilerce yapılmaması olarak ifade edilebilir. Testten yüksek puan alanların düşüğe doğru sıralanmasıyla üst ve alt gruplar arasındaki fark belirlenebilir. Üst gruptaki katılımcıların maddeye daha fazla doğru cevap vermesiyle maddenin ayırtıcılık özelliđi artacaktır. Madde ayırtıcılık indeksi +1 ile -1 arasında değer alabilmektedir (Bayrakçeken, 2012). Tablo 7'de ifade edildiđi şekliyle madde ayırt edicilik değerleri yorumlanmıştır (Başol, 2019).

Tablo 7

Madde Ayırt Edicilik Deđerlerinin Yorumu

0,40 ve üzeri	Çok iyi derecede ayırt edici maddedir
0,30 – 0,39 arası	Ayırt edici maddedir
0,20 – 0,29 arası	Düzeltilerek kullanılabilir
0,19 ve altı	Testten çıkarılmalıdır

Tablo 8
Basınç Başarı Testi İstatistik Deđerler

Basınç Başarı Testi	Sayısal Veriler
Soru sayısı	28
Uygulanan kiři sayısı	80
Madde güçlük indeksi ortalama deđeri	0,56
Madde ayırt edicilik ortalama deđeri	0,47
KR-20 güvenilirlik katsayısı	0,85

Tablo 8'deki sayısal veriler incelendiđinde ortalama güçlük indeksi deđerinin 0,56 olması maddelerin orta güçlükte olduđu şeklinde yorumlanabilir. Ortalama ayırt edicilik deđeri ise 0,47 bulunmuştur ve bu testin genel anlamda ayırt ediciliđinin çok iyi düzeyde olduđunu göstermektedir. Ancak tüm bunlara rađmen maddelerin tek tek güçlük ve ayırt edicilik deđerlerinin incelenmesi varsa çıkarılması gereken maddelerle ilgili fikir verecektir.

KR-20 iç tutarlılık katsayısı, testin maddelerinin ayrı ayrı güçlük indekslerinin hesaplanabildiđi testlerde uygulanır (Kan, 2009). Geliştirilmek istenen başarı testi, cevapların dođru (1.00) veya yanlış (.00) olarak kodlandıđı bir ölçme aracı olduđu için güvenilirlik düzeyinin tespitinde KR-20 güvenilirlik katsayısı hesaplanarak yorumlanmıştır (Tan, 2020). Güvenilir bir başarı testinin KR-20 deđeri minimum 0,70 olmalıdır (Büyükcöztürk, 2019). Testin KR-20 deđeri tabloda görüldüđu gibi 0,85 olarak hesaplanmıştır.

Tablo 9
Madde Güçlük ve Ayırt Edicilik Deđerleri

Madde	Madde ayırt edicilik indeksi	Madde güçlük indeksi
1	0,37	0,76
2	0,45	0,75
3	0,20*	0,89*
4	0,61	0,59
5	0,37	0,68
6	0,69	0,58
7	0,49	0,73
8	0,24**	0,59
9	0,16*	0,75
10	.0, 3	0,33
11	0,53	0,75
12	0,33	0,48
13	0,61	0,36
14	0,57	0,68
15	0,57	0,60
16	0,61	0,39
17	0,45	0,33
18	0,69	0,38
19	0,29**	0,46
20	0,57	0,53
21	0,53	0,51

22	0,33	0,75
23	0,41	0,56
24	0,73	0,59
25	0,29**	0,44
26	0,41	0,31
27	0,65	0,41
28	0,45	0,59

*Testten çıkarılması gereken maddeler

**Düzeltilerek kullanılabilir maddeler

Basınç başarı testinde üçüncü madde neredeyse ayırt edici bir madde değildir (0,20) ve çok kolay bir soru (0,89) olarak görülmektedir. Bu madde testten çıkarılmalıdır. Dokuzuncu madde ise 0,16 ayırt edicilik değerine sahiptir ve kesinlikle testten çıkarılmalıdır. Sekizinci, 19. ve 25. maddelerin ayırt edicilik değerleri 0,20 ile 0,29 arasında olmasından dolayı düzeltilerek kullanılmaya uygundur. Ancak bu başarı testinin amacı çıkmış sınav sorularından ve MEB tarafından yayımlanmış kazanım kavrama testlerinden geçerliği ve güvenilirliği sağlanmış bir başarı testi elde etmektir. Bundan dolayı bu maddelerin düzeltilmesi mümkün değildir. Bu nedenle bahsi geçen üç maddenin de testten çıkarılmasının daha uygun olacağı düşünülmektedir.

TARTIŞMA, SONUÇ VE ÖNERİLER

Bu çalışmada başarı testi geliştirme süreci sonunda 2013 yılı fen bilimleri dersi öğretim programındaki basınç konusu kazanımlarına yönelik geçerliği ve güvenilirliği yüksek çoktan seçmeli bir test elde edilmiştir. Başarı testinin kapsam geçerliği ve görünüş geçerliği için uzman görüşleri alınmıştır. Belirtke tablosu ile her kazanım için yeterli ve yaklaşık aynı sayıda soru kullanılmıştır. Başarı testini oluşturan maddeler MEB tarafından uygulanan merkezi sınavlar (SBS, TEOG) ve MEB tarafından yayınlanan KKT soruları arasından seçilmiştir. Geçerlik ve güvenilirlik çalışmaları sonucunda başarı testi basınç konusunda yer alan üç kazanımı ölçen 23 maddeden oluşmuştur. Bu duruma benzer olarak alanyazında başarı testi geliştirme sürecinde ulusal düzeyde uygulanan sınavlardan hazırlanan başarı testlerinin mevcut olduğu görülmektedir (Dumanoglu ve Bezir Akçay, 2018; Özcan ve diğerleri, 2019; Saylan Kırmızıgül ve Kaya, 2019). Böylece geliştirilen başarı testi öğretmenlerin kullanabilecekleri bir yapıda geliştirilmiş olmaktadır.

Başarı testinin elde edilen ortalama güçlük değeri testin güçlük değerinin orta seviyede olduğunu göstermektedir. Elde edilen değer testin güçlük değerinin istenilen düzeyde olduğunu göstermektedir. Bu sayede testin uygulandığı öğrenciler ortalama güçlükteki bir başarı testi ile karşılaşmış olacak ve bu test öğrencilerin kavramsal anlamalarını ölçmede amaca ulaşabilecektir. Alanyazında başarı testi geliştirmeye yönelik çalışmalarda da yakın ortalama güçlükte testler geliştirildiği görülmektedir (Akbulut ve Çepni, 2013; Bolat ve Karamustafaođlu, 2019; İlhan ve Hoşgören, 2017; Nacaroglu, Bektaş ve Kızıkan, 2020; Saraç, 2018; Sontay ve Karamustafaođlu, 2017). Geliştirilen başarı testinin ayırt edicilik değeri ise 0,47 olarak belirlenmiştir. Bu değer testin uygulandığı grupta kazanıma sahip olan ve olmayan öğrencileri tespit etme noktasında oldukça iyi düzeydedir. Bu test ölçme ve değerlendirmenin amaçlarından biri olan ayırt etme özelliğini sağlamaktadır. Alanyazında ortalama ayırt edicilik değeri 0,65 ve üzeri çalışmalar da göze çarpmaktadır (Divarcı ve Kaya, 2019; Saraç, 2018). Bu çalışmanın alanyazında basınç konusuyla ilgili geliştirilen bir diğer çalışmayla kıyaslandığında daha orta seviyede ve daha ayırt edici olduğu görülmektedir (Özcan ve diğerleri, 2019). Bu nedenle de bu başarı testi öğretmenlerin ve araştırmacıların kullanımı için uygun bir testtir.

Testte yer alan her bir maddenin elde edilen güçlük ve ayırt edicilik indekslerine bakılarak bu değerler yorumlanmıştır. Testteki maddelerin ayırt edicilik katsayıları incelendiğinde düşük ayırt edicilik katsayısına sahip maddelerin düzeltilerek kullanılması gerekliliğinden ancak bu maddeler çıkmış

sorulardan seçildiğinden dolayı düzeltilmesi tarafımızca mümkün olmadığı için düşük ayırt edicilik katsayısına sahip beş maddenin testten çıkarılması uygun görülmüştür. Sonuç olarak elde edilen testte yer alan maddelerin ayırt edicilik düzeyleri orta düzeyde yer almaktadır. Testin son halinde testte yer alan maddelerden altı zor, yedi kolay ve on orta güçlük düzeyinde madde yer almaktadır. Geliştirilen başarı testinde en fazla sayıda orta güçlük derecesinde madde yer aldığı görülmektedir ve orta güçlük düzeyindeki maddelerden oluşan ölçme araçlarının güvenilirlik düzeyleri daha yüksektir (Hasançebi ve diğerleri, 2020). Bu durumla birlikte testte olması gerekli olan kolay ve zor madde sayıları da birbirine yakındır ve soruların zorluk derecesinin oransal olarak dağıldığı görülmektedir. Bu durumda testte yer alan maddelerin güçlük indekslerinin belirlenen oransal dağılıma sahip olduğunu göstermektedir.

Başarı testinin KR-20 güvenilirlik katsayısı hesaplanmış ve testin güvenilirliğinin yüksek olduğu bulunmuştur. Bu değer testin güvenilirlik düzeyinin iyi bir seviyede olduğunu göstergesidir. Bu orandaki bir güvenilirlik katsayısı öğrencilerin testten aldıkları puanların oldukça güvenilir olduğu anlamına gelmektedir (Aydođdu, Tatar, Yıldız ve Buldur, 2012; Demir, Kızılay ve Bektaş, 2016). Alanyazında fen alanında geliştirilen başarı testleri incelediğinde KR-20 iç tutarlılık katsayısının bu değere yaklaşık değer aldığı çalışmalar tespit edilmiştir (Eren ve diğerleri, 2020; İlhan ve Hoşgören, 2017; Karanlı ve diğerleri, 2019; Nacarođlu ve Bektaş, 2019; Sontay ve Karamustafaođlu, 2017; Sontay ve Karamustafaođlu, 2020; Soylu, Karamustafaođlu ve Karamustafaođlu, 2020; Yanar, Saylan Kırmızıgöl ve Kaya, 2019).

Bu çalışmadan elde edilen sonuçlar ışığında şu önerilerde bulunulmuştur. Geliştirilen başarı testi 2013 yılı fen bilimleri dersi öğretim programında yer alan yedinci sınıf basınç konusu ile ilgili üç kazanıma yönelik olmakla birlikte 2018 yılı fen bilimleri dersi öğretim programında basınç sekizinci sınıf düzeyinde yer almasına rağmen basınç konusundaki üç kazanım önceki program ile benzerdir (MEB, 2018). Başarı testindeki sorular ile öğretim programındaki kazanımlar ilişkili olduğunda öğrencilerin başarısı da artmaktadır (Sontay ve Karamustafaođlu, 2017). Bu çalışma sonucunda geliştirilen başarı testinin 2018 yılı fen bilimleri dersi öğretim programında da öğretmenler ve araştırmacılar tarafından kullanılabilir bir yapıda olup olmadığı sonraki çalışmalarda araştırılabilir. Geliştirilen başarı testi ortaokul 7. sınıf düzeyindedir ve sonraki çalışmalarda 8. sınıf düzeyi için de aynı testin geçerlik ve güvenilirlik çalışmaları yapılabilir.

Başarı testi geliştirme sürecinde MEB tarafından uygulanan merkezi sınavlarda sorulan sorular ile MEB tarafından yayınlanan KKT soruları kullanılmıştır. Bu nedenle de geliştirilen başarı testi çoktan seçmeli sorulardan oluşan bir testtir. Çoktan seçmeli bir test olması değerlendirmede objektiflik ve kolaylık sağlamaktadır. İleriki çalışmalarda basınç konusunda farklı soru türlerini ihtiva eden başarı testleri de geliştirilebilir. Örneğin üç aşamalı veya son zamanlarda sıklıkla kullanılan dört aşamalı testler (Kiray ve Simsek, 2021) geliştirilebilir. Ayrıca PISA ve TIMSS gibi uluslararası sınavlardaki sorular test geliştirme sürecinde kullanılabilir ve bu sayede sorularda çeşitlilik sağlanmış olunabilir.

Geçerlik ve güvenilirlik çalışmaları yapılan basınç başarı testi öğretmenler ve bu konuda çalışma yürütecek araştırmacılar tarafından öğrencilerin ön bilgilerini belirlemede kullanılabileceği gibi öğretim sonrası öğrencilerin bilişsel yapılarını belirlemede de kullanılabilir. Bu sayede araştırmacılara bu konuda geçerliği ve güvenilirliği sağlanmış başarı testleri kullanma olanağı sunulabilir.

Bu çalışmanın sınırlılıklarında biri olarak test geliştirme sürecinin 80 ortaokul öğrencisi ile yürütülmesi gösterilebilir. Bu sayıdaki öğrenci ile yürütülen test geliştirme sürecinde elde edilen veriler geçerlik ve güvenilirlik çalışmalarını yürütmek için yeterli olmuştur. Ancak bundaki sonraki araştırmalarda daha fazla sayıda örnekleme kapsayan çalışmalar yürütülmesi uygun olacaktır.

KAYNAKÇA

Akbas, Y., & Gençturk, E. (2011). The effect of conceptual change approach to eliminate 9th grade high school students' misconceptions about air pressure. *Educational Sciences: Theory and Practice*, 11(4), 2217-2222.

- Akbulut, H. İ., & 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.
- Akkuş, O., & Baykul, Y. (2001). Çoktan seçmeli test maddelerini puanlamada, seçenekleri farklı biçimlerde ağırlıklandırmanın madde ve test istatistiklerine olan etkisinin incelenmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 20, 9-15.
- Aktamış, H., & Uçar, R. (2019). Astronomi'ye yönelik tutum ölçeđi ve 7. sınıf "Güneş Sistemi ve Ötesi" ünitesine yönelik başarı testi geliştirme çalışması. *Batı Anadolu Eğitim Bilimleri Dergisi*, 10(1), 57-78.
- Altay, E., & Balım, A. G. (2021). Maddenin tanecikli yapısı ile ilgili kavram yanılgılarının tespiti. *Batı Anadolu Eğitim Bilimleri Dergisi*, 12(2), 576-592.
- Aydođdu, B., Tatar, N., Yıldız, E., & Buldur, S. (2012). The science process skills scale development for elementary school students. *Journal of Theoretical Educational Science*, 5(3), 292-311.
- Ayvacı, H. Ş., & Durmuş, A. (2016). Bir başarı testi geliştirme çalışması: Isı ve sıcaklık başarı testi geçerlik ve güvenilirlik araştırması. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi*, 35(1), 87-103.
- Balci, A. (2021). *Sosyal bilimlerde araştırma yöntem, teknik ve ilkeler (15. Baskı)*. Ankara: Pegem Akademi.
- Başol, G. (2019). *Eğitimde ölçme ve değerlendirme (6. Baskı)*. Ankara: Pegem Akademi.
- Bayrakçeken, S. (2012). Test Geliştirme. İçinde E. Karip (Ed.), *Ölçme ve değerlendirme*, (5. Baskı), (ss. 294-324), Ankara: Pegem Akademi.
- Bayri, N. G. (2014). *Sekizinci sınıf öğrencilerinin basınç konusuyla ilgili gösterim türleri arasında geçiş yapabilme durumlarının incelenmesi* (Yayınlanmamış yüksek lisans tezi). Kastamonu Üniversitesi Fen Bilimleri Enstitüsü, Kastamonu.
- Baytok, H. (2007). *Yapılandırmacı öğrenme kuramına dayalı öğretimin ilköğretim 7. sınıf basınç konusunda öğrenci başarısı ve tutumuna etkisi* (Yayınlanmamış yüksek lisans tezi). Balıkesir Üniversitesi Fen Bilimleri Enstitüsü, Balıkesir.
- Bolat, A., & Karamustafaođlu, S. (2019). Vücutumuzdaki sistemler ünitesi başarı testi geliştirme: geçerlik ve güvenilirlik. *Gazi Eğitim Bilimleri Dergisi*, 5(2), 131-159.
- Bostan Sariođlan, A., & Can, Y. (2021). Effect of open inquiry based learning approach on the conceptual understanding of secondary school students. *International Online Journal of Educational Sciences*, 13(2), 432-446.
- Bozan, M. (2008). *Problem çözme etkinliklerinin 7. sınıf öğrencilerinin basınç konusu ile ilgili başarı, tutum ve üstbilis becerilerinin gelişimine etkisi* (Yayınlanmamış doktora tezi). Balıkesir Üniversitesi Fen Bilimleri Enstitüsü, Balıkesir.
- Brown, A., Nidumolu, A., McConnell, M., Hecker, K., & Grierson, L. (2019). Development and psychometric evaluation of an instrument to measure knowledge, skills, and attitudes towards quality improvement in health professions education: the beliefs, attitudes, skills, and confidence in quality improvement (BASIC-QI) scale. *Perspectives on Medical Education*, 8(3), 167-176.
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2013). *Bilimsel araştırma yöntemleri (15. Baskı)*. Ankara: Pegem Akademi Yayınları.
- Büyüköztürk, Ş. (2019). *Sosyal bilimler için verilerin analizi el kitabı (26. Baskı)*. Ankara: Pegem Akademi.
- Can, Y. (2019). *Sorgulamaya dayalı öğrenme yaklaşımının 7. sınıf öğrencilerinin basınç kavramı ile ilgili kavramsal anlamalarına etkisi* (Yayınlanmamış yüksek lisans tezi). Balıkesir Üniversitesi Fen Bilimleri Enstitüsü, Balıkesir.
- Couch, B. A., Hubbard, J. K., & Brassil, C. E. (2018). Multiple–true–false questions reveal the limits of the multiple–choice format for detecting students with incomplete understandings. *BioScience*, 68(6), 455-463.

- Çakaliođlu, S. N. (2008). *Proje tabanlı öğrenme yaklaşımına dayalı fen bilgisi öğretiminin akademik başarı ve tutuma etkisi* (Yayınlanmamış yüksek lisans tezi). Çukurova Üniversitesi Sosyal Bilimler Enstitüsü, Adana.
- Çardak, Ç. S., & Selvi, K. (2018). Öğretim ilke ve yöntemleri dersi için başarı testi geliştirme süreci. *Akdeniz Eğitim Araştırma Dergisi*, 12(26), 379-406. doi: 10.29329/mjer.2018.172.19.
- Çetinkaya, M., & Taş, E. (2016). "Vücudumuzda Sistemler" ünitesine yönelik üç aşamalı kavram tanı testi geliştirilmesi. *Ordu Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Araştırmaları Dergisi*, 6(15), 317-330.
- Demir, N., Kızılay, E., & Bektaş, O. (2016). 7. sınıf çözümler konusunda başarı testi geliştirme: Geçerlik ve güvenilirlik çalışması. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 10(1), 209-237.
- Demirciođlu, G., Aydın, M. A., & Demirciođlu, H. (2012). Kavramsal deđişim metninin ve üç boyutlu modelin 7. sınıf öğrencilerinin atomun yapısını anlamalarına etkisi. *Bayburt Eğitim Fakültesi Dergisi*, 7(2), 70-96.
- Divarçı, Ö.F., & Kaya, H. (2019). 8. sınıf "Maddenin Halleri ve Isı" ünitesine yönelik geçerliliđi ve güvenilirliđi sağlanmış bir akademik başarı testi geliştirme çalışması. *Eğitim Kuram ve Uygulamaları Araştırmaları Dergisi*, 5(2), 214-238.
- Divarçı, Ö.F. (2016). *Multimedya destekli probleme dayalı öğrenme yaklaşımının 8. Sınıf öğrencilerinde akademik başarıya, tutuma ve kalıcılıđa etkisi: Basınç konusu* (Yayınlanmamış yüksek lisans tezi). Amasya Üniversitesi Fen Bilimleri Enstitüsü, Amasya.
- Dumanođlu, F., & Bezir Akçay, B. (2018). Elektrik enerjisi başarı testinin geliştirilmesi. *E-Kafkas Eğitim Araştırmaları Dergisi*, 5(2), 20-39.
- Ecevit, T., & Şimşek, P. Ö. (2017). Öğretmenlerin fen kavram öğretileri, kavram yanlışlarını saptama ve giderme çalışmalarının deđerlendirilmesi. *İlköğretim Online*, 16(1), 129-150.
- Ercan, O., Bilen, K., & Ural, E. (2016). 'Earth, Sun and Moon': Computer assisted instruction in secondary school science - achievement and attitudes. *Issues in Educational Research*, 26(2), 206-224.
- Ercan, İ., & Kan, İ. (2004). Ölçeklerde güvenilirlik ve geçerlik. *Uludağ Üniversitesi Tıp Fakültesi Dergisi*, 30 (3) 211-216.
- Eren, A. A., Tanık Önal, N., & Büyük, U. (2020). Elementler ve bileşikler konusu için geçerli ve güvenilir bir başarı testi geliştirme çalışması. *Pearson Journal of Social Sciences & Humanities*, 6(6), 152-167.
- Gullickson, A. R. (1985). Student evaluation techniques and their relationship to grade and curriculum. *The Journal of Educational Research*, 79(2), 96-100.
- Hasançebi, B., Terzi, Y., & Küçük, Z. (2020). Madde güçlük indeksi ve madde ayırt edicilik indeksine dayalı çeldirici analizi. *Gümüşhane Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 10(1), 224-240.
- İlhan, N., & Hoşgören, G. (2017). Fen bilimleri dersine yönelik yaşam temelli başarı testi geliştirilmesi: Asit baz konusu. *Fen Bilimleri Öğretimi Dergisi*, 5(2), 87-110.
- Kan, A. (2009). Ölçme araçlarında bulunması gereken nitelikler, Hakan Atılğan (Ed.). *Eğitimde ölçme ve deđerlendirme*. Anı Yayıncılık.
- Karaman, P. (2020). Temel kavramlar. İçinde F. Orçan (Ed.), *Sosyal bilimlerde istatistik Excel ve SPSS uygulamaları* (ss. 27-52). Ankara: Anı Yayıncılık.
- Karasar, N. (2008). *Bilimsel araştırma yöntemi (18. Baskı)*. Ankara: Nobel Yayınları.
- Karataş, F. Ö., Köse, S., & Coştu, B. (2003). Öğrenci yanlışlarını ve anlama düzeylerini belirlemede kullanılan iki aşamalı testler. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 13(13), 54-69.
- Karslı, G., Karamustafaođlu, S., & Kurt, M. (2019). Yenilenen fen bilimleri dersi öğretim programına yönelik 7. sınıf "hücre ve bölünmeler" ünitesi başarı testi: Geçerlik ve güvenilirlik. *Fen Bilimleri Öğretimi Dergisi*, 7(1), 68-98.

- Kaya, D., Bozdađ, H. C., & Ok, G. (2018). Yedinci sınıf öđrencilerinin basınç konusundaki kavramsal anlamaları ve kavram yanılgılarının matematiksel hatalar açısından incelenmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18(1), 321-341.
- Keleciođlu, H., & Göçer Şahin, S. (2014). Geçmişten günümüze geçerlik. *Eđitimde ve Psikolojide Ölçme ve Deđerlendirme Dergisi*, 5(2), 1-11.
- Kızıkan, O., & Bektaş, O. (2018). Fen eğitiminde başarı testi geliştirilmesi: hücre bölünmesi ve kalıtım örneđi. *Maarif Mektepleri Uluslararası Eğitim Bilimleri Dergisi*, 2(1), 1-18.
- Kiray, S. A., & Simsek, S. (2021). Determination and evaluation of the science teacher candidates' misconceptions about density by using four-tier diagnostic test. *International Journal of Science and Mathematics Education*, 19(5), 935-955.
- Kirişciođlu, S. (2007). *İlköđretim 7. Sınıf fen bilgisi dersi "basınç" konusunun yapılandırmacı öğrenme yaklaşımına dayalı öğretiminin akademik başarıya etkisi* (Yayınlanmamış yüksek lisans tezi). Pamukkale Üniversitesi Fen Bilimleri Enstitüsü, Denizli.
- Kuechler, W. L., & Simkin, M. G. (2003). How well do multiple choice tests evaluate student understanding in computer programming classes? *Journal of Information Systems Education*, 14(4), 389-400.
- Milli Eğitim Bakanlığı [MEB], (2013). *İlköđretim kurumları (ilkokullar ve ortaokullar) fen bilimleri dersi öğretim programı (3, 4, 5, 6, 7 ve 8. Sınıflar)*. Ankara: Talim ve Terbiye Kurulu Başkanlığı.
- Milli Eğitim Bakanlığı [MEB], (2018). *İlköđretim kurumları (ilkokullar ve ortaokullar) fen bilimleri dersi (3., 4., 5., 6., 7. ve 8. sınıflar) öğretim programı*. Ankara: Talim ve Terbiye Kurulu Başkanlığı.
- Nacarođlu, O., Bektaş, O., & Kızıkan, O. (2020). Madde döngüleri ve çevre sorunları konusunda başarı testi geliştirme: geçerlik ve güvenilirlik çalışması. *Kastamonu Eğitim Dergisi*, 28(1), 36-51.
- Nacarođlu, O., & Bektaş, O. (2019). Fen bilimleri dersindeki madde ve deđişim ünitesine yönelik geçerli ve güvenilir başarı testi geliştirme: BİLSEM örneđi. *Maarif Mektepleri Uluslararası Eğitim Bilimleri Dergisi*, 3(2), 67-88.
- Netemeyer, R.G., Bearden W.O., & Sharma S. (2003). *Scaling Procedures: Issues and Applications*. Sage Publications, Inc Thousand Oaks, California.
- Özcan, H., Koca, E., & Söđüt, M. (2019). Ortaokul öđrencilerinin basınç kavramıyla ilgili anlayışlarını ölçmeye yönelik bir test geliştirme çalışması. *Eđitim ve Toplum Araştırmaları Dergisi*, 6(1), 130-144.
- Özcan, H., & Koca, E. (2019). STEM yaklaşımı ile basınç konusu öğretiminin ortaokul 7. sınıf öđrencilerinin akademik başarılarına ve STEM'e yönelik tutumlarına etkisi. *Eđitim ve Bilim*, 44(198), 201-227.
- Özdemir Benli, E. (2021). The impacts of STEM supported Science teaching on 8th grade students' elimination of misconceptions about "solid, fluid and gas pressure", and their attitudes towards science and stem. *International Online Journal of Education and Teaching*, 8(1), 205-228.
- Özsevgeç, T., Çepni, S., & Demirciođlu, G. (2004). *Fen bilgisi öğretmenlerinin ölçme deđerlendirme okuryazarlık düzeyleri*. 4. Ulusal Fen ve Matematik Eğitimi Kongresi, Marmara Üniversitesi, İstanbul.
- Öztürk, M. (2014). *8. sınıf kuvvet ve hareket ünitesine yönelik bilgisayar destekli öğretim uygulamalarının etkililiđinin araştırılması* (Yayınlanmamış yüksek lisans tezi). Karadeniz Teknik Üniversitesi Eğitim Bilimleri Enstitüsü, Trabzon.
- Saraç, H. (2018). Fen bilimleri dersi 'maddenin deđişimi' ünitesi ile ilgili başarı testi geliştirme: geçerlik ve güvenilirlik çalışması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18(1), 416-445.
- Saylan Kırmızıgöl, A., & Kaya, H. (2019). 'Aynalarda yansıma ve ışığın sođrulması' konusunda geçerliđi ve güvenilirliđi sađlanmış başarı testi geliştirme çalışması. *Kuramsal Eğitimbilim Dergisi*, 12(2), 474-493.
- Simkin, M. G., & Kuechler, W. L. (2005). Multiple-choice tests and student understanding: What is the connection?. *Decision Sciences Journal of Innovative Education*, 3(1), 73-98.

- Sontay, G., & Karamustafaođlu, S. (2020). Fen bilimleri dersi “güneş, dünya ve ay” ünitesine yönelik başarı testinin geliştirilmesi. *Gazi Üniversitesi Gazi Eğitim Dergisi*, 40 (2), 511-551.
- Sontay, G., & Karamustafaođlu, S. (2017). 5. sınıf fen bilimleri dersi “yer kabuđunun gizemi” ünitesine yönelik başarı testi geliştirme. *Fen Bilimleri Öğretim Dergisi*, 5(1), 62-86.
- Soylu, Ü. İ., Karamustafaođlu, S., & Karamustafaođlu, O. (2020). 6. sınıf “madde ve ısı” ünitesi başarı testi geliştirme: Geçerlik ve güvenilirlik. *Ihlara Eğitim Araştırmaları Dergisi*, 5(2), 271–293.
- Shuttleworth M. (2016). *Content validity*. <https://explorable.com/contentvalidity.html> adresinden 23.08.2022 tarihinde elde edilmiştir.
- Şenyıl, M. (2009). *İlköğretim fen bilgisi dersinde buluş yoluyla öğretim stratejisinin öğrencilerin üst düzey düşünme becerilerine etkisi* (Yayınlanmamış yüksek lisans tezi). Selçuk Üniversitesi Fen Bilimleri Enstitüsü, Konya.
- Tamir, P. (1971). An alternative approach to the construction of multiple choice test items. *Journal of Biological Education*, 5(6), 305-307.
- Tan, Ş. (2020). *Öğretimde ölçme ve değerlendirme KPSS el kitabı (14. Baskı)*. Ankara: Pegem Akademi.
- Taşkoyan, S. N. (2008). *Fen ve teknoloji öğretiminde sorgulayıcı öğrenme stratejilerinin öğrencilerin sorgulayıcı öğrenme becerileri, akademik başarıları ve tutumları üzerindeki etkisi* (Yayınlanmamış yüksek lisans tezi). Dokuz Eylül Üniversitesi Eğitim Bilimleri Enstitüsü, İzmir.
- Tekindal, S. (2000). Klasik yazılı sınavla ve çok sorulu testle elde edilen ölçümlerin güvenilirlik ve geçerliği. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 8(8), 38-46.
- Tekin İftar, E. (2018). *Eğitim ve davranış bilimlerinde tek denekli araştırmalar (2. Baskı)*. Ankara: Anı Yayıncılık.
- Temizkan, M., & Sallabaş M. E. (2011). Okuduđunu anlama becerisinin değerlendirilmesinde çoktan seçmeli testlerle açık uçlu yazılı yoklamaların karşılaştırılması. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 30, 207-220.
- Timur S., Dođan F., İmer Çetin N., Timur B., & Işık R. (2019). 6. sınıf hücre konusuna ilişkin başarı testi geliştirilmesi: geçerlik ve güvenilirlik çalışması. *Çukurova Üniversitesi Eğitim Fakültesi Dergisi*, 48(2), 1202-1219.
- Treagust, D. (1986). Evaluating students' misconceptions by means of diagnostic multiple choice items. *Research in Science Education*, 16(1), 199-207.
- Üçüncü, G., & Sakız, G. (2020). Başarı testi geliştirme süreci: İlkokul dördüncü sınıf maddeyi tanıyalım ünitesi örneđi. *Kastamonu Eğitim Dergisi*, 28(1), 82-94.
- Üstüner, A., & Şengül, M. (2004). Çoktan seçmeli test tekniđinin Türkçe öğretimine olumsuz etkileri. *Fırat Üniversitesi Sosyal Bilimler Dergisi* 14(2), 197-208.
- Webb, N. L. (1997). *Determining alignment of expectations and assessments in mathematics and science education*. NISE brief 1(2). Madison, WI: University of Wisconsin-Madison, National Institute for Science Education.
- Wise, S. L. (2015). Effort analysis: Individual score validation of achievement test data. *Applied Measurement in Education*, 28(3), 237-252.
- Yaman, S. (2016). Çoktan seçmeli madde tipleri ve fen eğitiminde kullanılan örnekleri. *Gazi Eğitim Bilimleri Dergisi*, 2(2), 151-170.
- Yanar, S., Saylan Kırmızıgöl, A., & Kaya, H. (2019). 6. sınıf ışık ve ses konusuna yönelik başarı testi geliştirme çalışması. *SDU International Journal of Educational Studies*, 6(2), 53-72. doi: 10.33710/sduijes.57741