
Kuram ve Uygulamada SOSYAL BİLİMLER DERGİSİ

Social Sciences: Theory & Practice

ISSN: 2619-9408

Geliş/Received: 26.02.2024 Kabul/Accepted: 12.12.2024

Makale Türü: Araştırma

Basınç Konusunda Geliştirilmiş Bir Başarı Testinin Güvenirliğinin Araştırılması*

Ali ÖZEN*

Nevzat YİĞİT**

Ebru MAZLUM GÜVEN***

ÖZ

Öğrencilerin basınç konusunu anlamakta zorlandıklarına dair yapılan çalışmalar, çeşitli veri toplama araçlarının kullanıldığını ortaya koymuştur. İncelenen çalışmalarda, temel veri toplama aracı olarak basınç başarı testlerinin tercih edildiği görülmektedir. Başarı testlerinin geliştirilmesi çalışmaları, başarı testlerinin bazı özelliklere sahip olması gerektiğini vurgulamaktadır. Önceden hazırlanan ve belirli örnekleme uygulanan başarı testlerine farklı çalışmalarda başvurulabilmesi ve araştırma sonuçlarının kullanılabilir olması için ilgili testlere yönelik güvenilirlik çalışmalarının yapılması, madde güçlük ve ayırıcılık indekslerinin incelenmesi önemlidir. Bu bağlamda bu çalışmanın amacı Özcan, Koca ve Söğüt (2019) tarafından geliştirilen başarı testinin, farklı bir araştırmada veri toplama aracı olarak kullanılabilirliğini araştırmaktır. Çalışmada nicel araştırma yöntemlerinden tarama deseni işe koşulmuştur. Çalışmanın örnekleme İstanbul ilinde, bir devlet okulunda öğrenim gören 149 sekizinci sınıf öğrencisidir. Öğrencilerin test maddelerine verdikleri yanıtların analizi, Test Analiz Programı (TAP) kullanılarak gerçekleştirilmiştir. Hesaplanan madde güçlük indeksinin 0,40 ile 0,77 aralığında değiştiği, madde ayırıcılık indeksinin ise 0,32 ile 0,74 arasında değiştiği görülmüştür. Başarı testinin ortalama madde güçlüğü 0,57 olarak hesaplanırken, ortalama madde ayırıcılık indeksi ise 0,55 olarak belirlenmiştir. Başarı testinin KR-20 güvenilirlik katsayısı 0,820 ve yarıya bölme yöntemiyle hesaplanan Spearman-Brown katsayısının 0,795 olduğu tespit edilmiştir. Bu psikometrik bulgular, enstrümanın güvenilirliğini doğrulamaktadır. Sonuç olarak, ilgili başarı testinin, basınç konusuna yönelik kavramsal anlayışı değerlendirmek için kullanışlı olduğu düşünülmektedir.

Anahtar Kelimeler: Basınç, ortaokul, başarı testi, güvenilirlik.

Investigation of the Reliability of an Achievement Test Developed on the Topic of Pressure

ABSTRACT

Studies have consistently highlighted the pedagogical challenge posed by students' difficulties in comprehending the concept of pressure, prompting the utilization of diverse methodological instruments for investigation. Within this context, pressure achievement tests have emerged as a principal tool for data collection in scholarly inquiries. The development of achievement tests

Atıf Bilgisi: Özen, A., Yiğit, N. & Mazlum Güven, E. (2024). Basınç konusunda geliştirilmiş bir başarı testinin güvenilirliğinin araştırılması. *Kuram ve Uygulamada Sosyal Bilimler Dergisi*, 8(2), 291-302. Doi: 10.48066/kusob.1443485

* Bu çalışma 5. Uluslararası Sosyal Bilimler Kongre'sinde sunulan bildirinin genişletilmiş versiyonudur.

*Öğretmen, Zuhur Ortaokulu, aliozen6199@gmail.com, ORCID: orcid.org/0009-0002-0233-1471.

** Prof. Dr., Trabzon Üniversitesi, Fatih Eğitim Fakültesi, Matematik ve Fen Bilimleri Eğitimi Bölümü, nyigit@trabzon.edu.tr, ORCID: orcid.org/0000-0001-7363-1637

***Dr. Öğr. Üyesi, Bayburt Üniversitesi, Eğitim Fakültesi, Matematik ve Fen Bilimleri Eğitimi Bölümü, ebruguyen@bayburt.edu.tr, ORCID: orcid.org/0000-0002-7758-4177.

emphasizes that these tests should possess certain characteristics. Conducting reliability studies on relevant tests is important for utilizing them in different studies and making research results available. In this context, the aim of this study is to investigate the usability of the achievement test developed by Özcan, Koca, and Söğüt (2019) as a data collection tool in a different research context. Employing a quantitative research paradigm, the investigation adopts a survey design methodology. The research cohort comprises 149 eighth-grade students enrolled in a public school situated in Istanbul. The analysis of students' responses to the test items is conducted utilizing the Test Analysis Program (TAP). The range of item difficult index spans from 0.40 to 0.77, with the item discrimination index oscillating between 0.32 and 0.74. The average item difficulty of the achievement test is computed at 0.57, while the average item discrimination index stands at 0.55. As a result of the item analysis, the KR-20 reliability coefficient of the achievement test was found to be 0.820, and the Spearman-Brown coefficient calculated using the split-half method was 0.795. These psychometric findings validate the robustness of the instrument. Consequently, the study ascertains the utility of the developed achievement test for assessing conceptual understanding of pressure across divergent student cohorts.

Keywords: Pressure, middle school, achievement test, reliability.

Giriş

Fen bilimleri kavramlarının etkili bir şekilde öğrenilmesi kişinin etrafındaki doğa olaylarını anlamlandırabilmesi, fen okuryazar ve bilimsel okuryazar birey olarak yetişmesi açısından önemli görülmektedir (Değirmenci, Karamustafaoğlu ve Kara Mustafaoğlu, 2021). Öğrencilerin fen bilimleri dersinde özellikle soyut kavramları anlamakta zorlandığı bilinmekte ve basınç konusu özellikle ortaokul öğrencileri tarafından anlaşılması güç bir konu olarak görülmektedir (Can, 2019). Basınç, kuvvet, yüzey alanı ilişkisi, sıvının yoğunluğu ve basınç ilişkisi ile kapalı kaplarda gaz basıncı öğrencilerin anlamakta zorlandığı ve kavram yanlışlığı yaşadığı kavramlara örnek olarak verilebilir (Yaman, 2016). Suprpto (2020) kavram yanlışlarının genellikle bilimsel kavramlarla çelişen önceden edinilmiş fikirlerden ve günlük deneyimlerden kaynaklandığını ileri sürmektedir. Geleneksel öğretim teknikleri ile öğrenciler aktif bir şekilde öğretim sürecine dahil edilemediği için aktarım şeklinde gerçekleştirilen öğretimin, öğrencilerin yanlışlığa oluşturmalarına sebep olduğu düşünülmektedir (Guerra-Reyes, Guerra-Dávila, Naranjo-Toro, Basantes-Andrade ve Guevara-Betancourt, 2024). Böylece öğrencilerin fen kavramlarına yönelik kavram yanlışları artmakta ve akademik başarıları düşmektedir (Sönmez, Geban ve Ertepinar, 2001).

Basınç konusuna ilişkin alanyazında,,öğrencilerin basınç çeşitlerini kavramada problem yaşadığına (Görkemli Taban, 2017; Kaya, Bozdağ ve Ok, 2017; Şahin ve Çepni, 2012; Yaman, 2016) öğrencilerin basınç ile ilgili matematiksel hesaplamaları yapmada zorlandığına (Kaya, Bozdağ ve Ok, 2017), basıncı etkileyen değişkenler arasındaki ilişkiyi kurmada sorun yaşadıklarına (Şahin, 2010) yönelik bulgular mevcuttur. İlgili kavram yanlışlarının giderilmesi ve etkili bir öğretimin sağlanabilmesi için araştırma-sorgulama tabanlı stratejilere başvurulması, etkileşime olanak veren yöntem-tekniklerin kullanılması (Suprato, 2020) ve bu yöntem tekniklerin etkililiğinin araştırılması gerekmektedir. Alanyazında, farklı yöntem-tekniklerin öğrencilerin basınç konusundaki akademik başarılarına etkisini belirlemek amacıyla birçok çalışmanın yapıldığı görülmektedir (Akgün, Tokur ve Özkara, 2013; Baytok, 2007; Can, 2019; Demirel, 2015; Gönen ve Andaç, 2009; Kirişçiöğlu, 2007; Özkara, 2011; Sadıç, 2016; Şahin ve Çepni, 2012; Tural, 2020; Turgutalp, 2021; Ünal ve Ergin, 2006; Yaman, 2016). Bu yöntem tekniklerin etkililiğini ölçmek için çeşitli veri toplama araçlarına başvurulmakla birlikte en sık kullanılan aracın başarı testleri olduğu görülmüştür.

Başarı testleri yapılan uygulamanın ne derece etkili olduğunu anlamak adına kullanılmaktadır. Başka bir deyişle başarı testleri ile öğretimin etkililiği ölçülerek öğrencilerin bilişsel yapıları ortaya koyulabilmektedir. Bu özelliği ile başarı testleri gerekli iyileştirmelerin yapılmasına olanak vererek ölçme değerlendirme sürecine katkı sağlamaktadır (Boz, Özcan ve Bostan-Sarioğlan, 2023). Akademik

başarının ölçülmesine yönelik çeşitli ölçme araçları kullanılmaktadır. Bunlardan en yaygın nesnel sonuçlar alınmasına olanak veren çoktan seçmeli maddelerden oluşan testlerdir (McMillan, 2018). Çoktan seçmeli testlerin üst düzey becerileri ölçmede yetersiz kaldığı ifade edilse de maddeler iyi yapılandırıldığında bu tür beceriler de ölçülebilmektedir (Davis, 1993). Çoktan seçmeli testlerin uygulanması ve değerlendirilmesinin kolay olması, birçok kazanımı ölçebilecek maddeyi barındırabilmesi ve nesnel sonuçlar vermesi gibi avantajları dolayısıyla tercih edildiği görülmektedir (Holder ve Mills, 2001).

Basınç Konusuna Yönelik Yapılan Çalışmalar

Yapılan çalışmalar incelendiğinde basınç konusu anlatımında kullanılan yöntem, teknik ve materyallerin, TGA (Atabey ve Çiftçi, 2019; Sadıç, 2016; Tural, 2020), argümantasyon (Demirel, 2015; Özkara, 2011), buluş yoluyla öğrenme (Ünal, 2005; Ünal ve Ergin, 2006) gözden geçirme stratejisi (Gönen ve Andaç, 2009) STEM yaklaşımı (Turgutalp, 2021), multimedya destekli probleme dayalı öğrenme (Saltan ve Divarcı, 2017), farklı yöntem ve tekniklerin bir arada kullanılması ile zenginleştirilmiş öğretim materyali (Şahin ve Çepni, 2012) ve sorgulamaya dayalı öğrenme (Can, 2019) olduğu görülmüştür.

Basınçla ilgili çalışmaların yapılış amaçları incelendiğinde, konunun öğretimine (Akgün, Tokur ve Özkara, 2013; Baytok, 2007; Demirel, 2015; Gönen ve Andaç, 2009; Kirişçioğlu, 2007; Özkara, 2011; Sadıç, 2016; Tural, 2020; Turgutalp, 2021; Saltan ve Divarcı, 2017; Ünal, 2005; Ünal ve Ergin, 2006; Şahin ve Çepni, 2012; Yaman, 2016) test geliştirmeye (Özcan, Koca ve Söğüt, 2019) ve problem çözümünde yapılan hataları belirlemeye (Atabey ve Çiftçi, 2019; Bozan ve Küçüközer, 2007) yönelik çalışmalar dikkat çekmektedir.

Basınç konusu ile ilgili yapılan çalışmalarda örneklem, ortaokul öğrencileri (Akgün ve ark., 2013; Baytok, 2007; Bozan ve Küçüközer, 2007; Can, 2019; Demirel, 2015; Gönen ve Andaç, 2009; Kirişçioğlu, 2007; Özkara, 2011; Ünal, 2005; Ünal ve Ergin, 2006; Sadıç, 2016; Saltan ve Divarcı, 2017; Şahin ve Çepni, 2012; Tural, 2020; Turgutalp, 2021; Yaman, 2016;) ve öğretmen adayları (Atabey ve Çiftçi, 2019; Pabuçcu, 2016; Ünsal, 2019) olarak belirlenmiştir.

Alanyazın incelemesinde araştırmalarda çeşitli veri toplama araçları kullanıldığı görülmüştür. Bunlar, basınç başarı testi (Akgün ve ark., 2013; Baytok, 2007; Gönen ve Andaç, 2009; Kirişçioğlu, 2007; Özcan ve Koca, 2019; Özcan ve ark., 2019; Özkara, 2011; Saltan ve Divarcı, 2017; Ünal, 2005; Ünal ve Ergin, 2006), gelişim dosyası (Kirişçioğlu, 2007), süreçte sınıf içi yapılan tartışmalar (Laçın Şimşek, Öztuna Kaplan ve Çorapçığıl, 2018), açık uçlu test (Bozan ve Küçüközer 2007; Pabuçcu, 2016; Ünal, 2005), mülakat (Baytok, 2007; Karamustafaoğlu ve Kandaz, 2006), anket (Karamustafaoğlu ve Kandaz, 2006; Yaman, 2016), kavram yanılışı testi (Ünsal, 2019) ve çalışma kağıtları (Atabey ve Çiftçi, 2019; Laçın Şimşek ve ark., 2018) şeklindedir. İncelenen bu çalışmalarda temel veri toplama aracı olarak basınç başarı testlerinin kullanıldığı görülmüştür.

Yapılan çalışmalar incelendiğinde basınç konusu öğretiminde kullanılan yöntem ve tekniklerin çeşitliliği dikkat çekmektedir. Çalışmaların büyük çoğunluğunda öğrenci başarısını artırmaya yönelik, ortak amacı öğrencilerin konuyu daha derin ve kalıcı öğrenmesini sağlamak olan çeşitli aktif öğrenme stratejilerinin kullanıldığı görülmektedir. İncelenen çalışmalarda değerlendirme araçları olarak başarı testleri, açık uçlu testler, mülakatlar ve anketler gibi veri toplama araçlarının kullanımı basınç konusunun çok yönlü bir yapısı olduğunu işaret etmektedir. Bununla birlikte başarı testlerinin, eğitimde öğrencilerin bilişsel gelişim düzeylerini, konuya dair sahip oldukları bilgiyi ve kavramsal anlayışlarını ölçmek amacıyla en sık başvurulan araçlardan biri olduğu görülmektedir. Basınç konusu gibi bilimsel kavramların öğretiminde, öğrencilerin konuyu ne kadar anladıklarını ve bu bilgileri nasıl uygulayabildiklerini ölçmek için başarı testleri önemli bir yol gösterici olabilir. Böylece öğretmenler,

öğrencilerin hangi alanlarda başarılı olduklarını, hangi kavramları anlamakta zorlandıklarını ve öğrenme süreçlerinde hangi noktalara daha fazla odaklanılması gerektiğini belirleyebilirler.

Alanyazında farklı konular üzerinde başarı testi geliştirme çalışmaları yapılmıştır (Ayvacı ve Durmuş, 2016; Kızılkapan ve Bektaş, 2018; Nacaroğlu, Bektaş ve Kızılkapan, 2020; Özkan ve Eryılmaz Muştı, 2018; Özcan ve ark., 2019; Saraç, 2018; Şen ve Eryılmaz, 2011). Başarı testleri, öğrencilere öğretilen materyali anlama ve hatırlama düzeylerini ölçmek için tasarlanmış olup, eğitim başarısının doğrudan bir göstergesi olarak hizmet eder (Hayati ve ark., 2020).

Bu testler, öğrenci performansındaki güçlü ve zayıf yönlerin belirlenmesini kolaylaştırarak, eğitimcilerin buna göre müdahaleler yapmasına olanak tanır (Tian ve Sun, 2018). Bu nedenle başarı testi puanları ile çalışılan araştırmalarda düzgün tasarlanmış, geçerliği ve güvenilirliği yüksek olduğu tespit edilmiş bir test akademik değerlendirmelerin bütünlüğünü korumada kritik bir öneme sahiptir (Topay ve Yılmaz, 2023).

Burada bahsi geçen geçerlik, değerlendirme sonucunda ortaya çıkan çıkarımların, kullanımların ve sonuçların uygunluğunu ifade eden bir özelliktir. Güvenirlik ise temel anlamda elde edilen puanların ne ölçüde hatadan arınmış olduğuyla ilgilidir (McMillan, 2018). Her ölçme aracı geliştirildiği yaş grubu ve popülasyon için geçerli ve güveniliridir. Araştırmacıların ölçme araçlarından edindiği puanların, geçerli ve güvenilir olması gerektiği için aynı yaş grubunda kullanılsa da farklı bir örneklemden veri alınacaksa geçerlik güvenilirlik çalışmalarının yeniden yapılması gerekmektedir (Yalçın, 2021). Dolayısıyla daha önceden hazırlanan ve belirli örnekleme uygulanan başarı testlerine farklı çalışmalarda başvurulabilmesi ve araştırma sonuçlarının kullanılabilir olması için de ilgili teste yönelik geçerlik ve güvenilirlik çalışmalarının yapılması gerekmektedir. Ancak geçerlik ve güvenilirlik çalışmaları testte yer alan maddelerin performansı ile ilgili bilgi vermeyebilir. Bu nedenle maddelerin güçlük indeksi ve ayırt edicilik derecelerinin belirlenmesi de önemli görülmektedir (Hasançebi, Terzi ve Küçük, 2020). Bu çalışma kapsamında ise Özcan ve arkadaşları (2019) tarafından geliştirilen başarı testinin, veri toplama aracı olarak işe koşulacağı farklı bir araştırma için kullanılabilirliği araştırılmıştır. İlgili ölçek Özcan ve arkadaşları (2019) tarafından 7. sınıf öğrencileri ile basınç konusuna yönelik öğrenci başarısını ölçmek üzere geliştirilmiştir. Uyarlanacak ölçeğin kullanılacağı çalışmada da aynı konu ele alınacaktır. Bununla birlikte ilgili ölçeğin geliştirildiği tarihten bu yana Fen Bilimleri Öğretim Programı'nda bir değişiklik olmamıştır. Bu nedenle bu çalışma kapsamında geçerlik çalışmaları yapılmamış ancak ölçek farklı bir bağlamda kullanılacağı için güvenilirlik ve madde güçlük/ayırıcılık indeksi hesaplarına yer verilmiştir. Bu bağlamda, bu çalışmanın amacı, öğrencilerin basınç konusundaki akademik başarılarını ölçmek için tasarlanmış basınç başarı testinin farklı bir örneklem üzerinde uygulanacak olması nedeniyle testin güvenilirlik durumunu araştırmaktır. Çalışmanın alt problemleri ise:

1. Basınç konusunda geliştirilmiş başarı testinin madde güçlük indeksi değerleri nasıldır?
2. Basınç konusunda geliştirilmiş başarı testinin madde ayırıcılık indeksi değerleri nasıldır?
3. Basınç konusunda geliştirilmiş başarı testi güvenilir midir?

Yöntem

Araştırmanın Modeli

Bu çalışmada nicel araştırma yaklaşımı çerçevesinde yer alan tarama yöntemi kullanılmıştır. Tarama araştırmaları evreni veya evreni temsil eden bir örneklem üzerinde çalışılarak evrenin görüş, tutum, bilgi gibi özelliklerini belirlemek amacıyla gerçekleştirilir (Sezgin Selçuk, 2019). Bu araştırmada güvenilirlik çalışması yapabilmek adına tarama deseni kullanılmıştır.

Örneklem

Bu çalışma İstanbul'da bir devlet okulunda gerçekleştirilmiştir. Güvenirlik çalışmalarında genellenebilir bir sonuç ortaya koyabilmek için örneklem sayısı önemlidir. Geçerlik güvenirlik çalışmalarında örneklem sayısının madde sayısının en az beş katı olması önerilmektedir (Büyüköztürk, 2002). Bu çalışma kapsamında ele alınan başarı testi 20 sorudan oluşmaktadır. Çalışmanın örneklemini 149 öğrencidir. Çalışmanın gerçekleştirildiği tarih itibariyle basınç konusunu öğrenmiş oldukları için 8. sınıf öğrencileri ile çalışılmıştır.

Veri Toplama Araçları

Çalışmada Özcan ve arkadaşları (2019) tarafından geliştirilen Basınç konusuna yönelik akademik başarı testi kullanılmıştır. Test 20 çoktan seçmeli sorudan oluşmaktadır. Testin kullanımı ile ilgili izinler yazardan elektronik posta yoluyla alınmıştır. Yazarlar, kapsam geçerliği çalışmaları için 2 fen alanında uzman ve 2 fen bilimleri öğretmeni ile çalışmıştır. Öğrenci seviyesi ve kapsam açısından maddelerin uygunluğuna dair alınan geri bildirimler, kapsam geçerliği çalışmalarının temelini oluşturmuştur. İlgili testin güvenirlik katsayısı ise yazarlar tarafından KR-20, 0.73 olarak hesaplanmıştır. Araştırmacılar elde ettikleri bu değerler sonucunda geliştirdikleri testin, öğretim programının işaret ettiği hedef ve davranışları ölçme ve değerlendirmede ülkemizdeki ortaokul öğrencilerine uygulanabilir nitelikte geçerli ve güvenilir bir ölçek olduğunu belirtmişlerdir. Başarı testindeki soruların kazanımlara göre dağılımı ise aşağıdaki tabloda gösterilmiştir.

Tablo 1. Kazanımların soru dağılımını gösteren tablo

Kazanımlar	Sorular
Katı basıncını etkileyen değişkenleri deneyerek keşfeder.	3,5,7,13,14,17,19
Sıvı basıncını etkileyen değişkenleri tahmin eder ve tahminlerini test eder.	1,2,4,6,8,10,11,12,15,16,20
Katı, sıvı ve gazların basınç özelliklerinin günlük yaşam ve teknolojideki uygulamalarına örnekler verir.	9,18

2018'de yürürlüğe giren Fen Bilimleri Öğretim Programı'nda katı, sıvı ve gaz basıncına yönelik üç kazanım bulunmaktadır. Gaz basıncı, sıvı basıncına ait kazanımın alt maddesi şeklinde verilmiştir. Bu nedenle hazırlanan tabloda (Tablo 1) gaz basıncına yönelik sorular sıvı basıncı kazanımını hücrelerinde verilmiştir. Güvenirliği araştırılmak istenen test incelendiğinde 2 ve 10 numaralı soruların gaz basıncına yönelik olduğu görülmektedir.

Verilerin Analizi

Verilerin analizi için test analiz programı kullanılmıştır. TAP, özellikle başarı testlerinde yapı geçerliği kapsamında maddelerin güçlük indeksleri, ayırt edicilik indeksleri ve güvenirlik hesaplamalarında kullanılan bir analiz programıdır. Bir testin güvenirliğini belirlemede birçok faktör etkili olabilmektedir. Bu araştırma kapsamında, güvenirliği yüksek olan maddelerden oluşan bir testin güvenirliğinin de yüksek olacağı öngörüldüğü için, madde zorluk indeksi, madde ayıricılık gücü indeksi ve iç tutarlılık hesaplamaları gerçekleştirilmiştir (Karaca, 2004).

Maddenin güçlük indeksi(p_j), bir maddeyi doğru yanıtlayanların tüm cevaplayıcı sayısına oranıdır. Bu indeks 0 ile 1 arasında eğer alır ve değeri 0'a yakın olan maddelerin zor 1'e yakın olan maddelerin ise kolay olduğu söylenebilir. Güçlük indeksi testteki sorunun doğru yapılma istatistiğini belirtir ve güçlük değerinin 0,50 civarında olması istenir. Ayrıca madde güçlüğü 0.29 ve altında ise maddenin zor, 0.30-0.49 aralığında ise orta güçlükte, 0.50-0.69 aralığında ise kolay ve 0.70-1 arasının çok kolay olduğu yorumu yapılabilir (Hasançebi ve ark., 2020).

Ayırt edicilik ise bilenle bilmeyeni ayırt edebilme gücüdür. Maddelerin ayırt edicilikleri (r_{jx}) %27 'lik alt-üst gruplar ile çift serili korelasyon katsayıları (r_{pb}) ile hesaplanmıştır. Bu değer 0,30 üzerinde ise ilgili maddenin iyi, 0,40 üzerinde ise çok iyi ayırt etme gücüne sahip olduğunu gösterir (Yiğit, 2008).

İç tutarlılıkla ilgili olarak başarı testlerinde Kuder-Richardson (KR) 20 ve 21 değerlerinin hesaplanması önerilmektedir (Tavşancıl, 2014). KR-21 madde analizi yapılmamış testlere uygulanmakta, KR-20 ise testteki her bir maddenin güçlük indekslerinin hesaplanması durumunda kullanılmaktadır. Bu nedenle bu çalışmada iç tutarlılığı hesaplamak üzere KR-20 değerine başvurulmuştur. KR-20 değeri 0 ile 1 arasında değişmektedir. KR-20 değerinin 0.80 ile 1.00 arasında hesaplanması testin yüksek derecede güvenilir olduğunu, 0.60 ve altında hesaplanması ise güvenilirliğin düşük olduğunu göstermektedir (Kalaycı, 2008). Bunun yanısıra yarı test güvenilirlik yöntemi ile Spearman-Brown katsayısı da hesaplanmıştır. Yarı test yönteminde test maddeleri yansız olarak ikiye bölünür ve yarı testler uygulandıktan sonra testin tamamı için Spearman-Brown katsayısı hesaplanır (Tavşancıl, 2014).

Bulgular

Bu çalışma kapsamında Özcan, Koca ve Söğüt (2019) tarafından geliştirilen başarı testinin güvenilirliğini incelemek üzere iç tutarlılığı hesaplanmıştır. Aynı zamanda başarı testinin maddelerinin güçlük değerleri ve ayırt edicilik indeksleri incelenmiştir. Yapılan hesaplamalar sonucu maddelerin güçlük indeksine yönelik bulgular Tablo 2'de verilmiştir.

Tablo 2. Madde güçlük indeksi değerleri

Madde	p_j	Madde	p_j	Madde	p_j	Madde	p_j
1	0,52	6	0,69	11	0,77	16	0,49
2	0,64	7	0,65	12	0,52	17	0,52
3	0,68	8	0,59	13	0,46	18	0,50
4	0,59	9	0,56	14	0,40	19	0,64
5	0,48	10	0,58	15	0,67	20	0,54

Tablo 2 incelendiğinde 5, 13, 14, 16. maddelerin güçlüklerinin 0,49 veya altı değerinde, 1, 2, 3, 4, 6, 7, 8, 9, 10, 12, 15, 17, 18, 19 ve 20. maddelerin ise güçlüklerinin 0,50 ve 0,69 aralığında ve 11. maddenin güçlük değerinin ise 0,77 olduğu görülmektedir. Maddelerin ortalama güçlük indeksleri hesaplandığında ise bu değer 0,57 olduğu görülmektedir.

Yapılan hesaplamalar sonucu maddelerin ayırt edicilik indeksine yönelik bulgular Tablo 3'te verilmiştir. Yapılan analizler testte yer alan maddelerin biri hariç madde ayırt edicilik indekslerinin 0,40'ın üzerinde olduğunu göstermektedir.

Tablo 3. Madde ayırt edicilik indeksi değerleri

Madde	r_{jx}	r_{pb}	Madde	r_{jx}	r_{pb}	Madde	r_{jx}	r_{pb}	Madde	r_{jx}	r_{pb}
1	0,63	0,43	6	0,67	0,60	11	0,33	0,37	16	0,55	0,49
2	0,41	0,43	7	0,49	0,45	12	0,53	0,45	17	0,74	0,58
3	0,56	0,50	8	0,63	0,48	13	0,64	0,51	18	0,48	0,41
4	0,41	0,36	9	0,62	0,52	14	0,50	0,44	19	0,52	0,47

5 0,46 0,42 10 0,67 0,57 15 0,58 0,55 20 0,63 0,51

Madde ayırt edicilik indeksine göre 0,40 üzeri değerler sorunun iyi düzeyde ayırt edicilik düzeyine sahip olduğunu göstermektedir. Buna göre 11 numaralı soru dışındaki maddeler çok iyi ayırt edicilik düzeyine sahiptir. 11 numaralı sorunun ayrıcalık indeksi 0,33 olarak bulunmuştur. Bu madde de iyi ayırt ediciliğe sahip soru olarak görülmektedir. Maddelerin ortalama ayırt edicilik değeri ise 0,55 olarak bulunmuştur. Çift serili korelasyon değerleri de 0,36 ile 0,60 arasında değişmekte olup, ortalaması 0,48 olarak hesaplanmıştır.

Testin iç tutarlık katsayısını hesaplamak üzere KR-20 ve Spearman-Brown katsayısı incelenmiştir. Tablo 4'te KR-20 ve Spearman-Brown katsayıları verilmiştir.

Tablo 4. Güvenirlik analiz değerleri

KR – 20	0,82
Spearman – Brown	0,79

Tablo 4 incelendiğinde KR-20 değerinin 0,82 olarak bulunduğu görülmektedir. Spearman – Brown katsayısı ise 0,79 olarak hesaplanmıştır.

Sonuç

Maddelerin ortalama güçlük değerinin 0,50 civarında olması gerekmektedir (Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel, 2012). Uygulama sonucunda testin ortalama güçlük indeksi 0,57 olarak bulunmuştur. Bu değere göre testin öğrenciler üzerinde orta güçlükte bir etkiye sahip olduğu söylenebilir.

Testte bulunan maddelerin ayırt edicilik indeksleri incelenmiştir. %27 'lik alt-üst gruplar ile yapılan ayırt edicilik bilenle bilmeyenin ayırt edilebilmesidir. Testteki maddelerin ayırt edicilik değerleri 0,33 ile 0,74 arasında değişmektedir. Ayırt edicilik değeri 0,30 üzerinde iyi, 0,40 üzerinde çok iyi düzey ayırt edicilik anlamına gelmektedir. Bununla birlikte soruların çift serili korelasyon katsayıları da 0,36 ile 0,60 arasında değişmektedir. Testteki maddelerin ayrı ayrı ayırt edicilik değerleri incelendiğinde tüm soruların iyi ve çok iyi düzeyde ayırt ediciliğe sahip olduğu görülmektedir. Testin ortalama ayırt edicilik indeksleri %27 'lik alt-üst gruplar yöntemiyle incelendiğinde bu değer 0,55 ve ortalama çift serili korelasyon değeri ise 0,48 olarak bulunmuştur. Ayırt edicilik değeri alanyazına göre iyi düzeyde ayırt edicilik anlamına gelmektedir (Büyüköztürk ve ark., 2012). Alanyazında benzer çalışmalarda da 0,40 üzerinde ayırt ediciliğe sahip çalışmaların (Bolat ve Karamustafaoğlu, 2019; Nacaroglu, Bektaş ve Kızıkan, 2020; Saraç, 2018) olduğu görülmüştür.

Başarı testinin analiz çalışmaları yapılırken ilk olarak madde güçlük indeksleri incelenmiştir. Hasaıbe ve diđerlerinin (2020) çalışmasındaki değerlerden yola çıkarak bu arařtırmadaki maddelerin 4'ünün orta güçlükte, 15'nin kolay ve 1'inin çok kolay olduđu söylenebilir. Ayrıca testin ortalama madde güçlük indeksi 0,57 olarak bulunmuştur. Bu değer testin orta güçlükte bir test olduğunu ve testin öğrenciler açısından uygun olduğunu göstermektedir. Nitekim alanyazında güçlük indeksi 0,50 civarında olan başarı testlerinin kullanıldığı benzer çalışmalar (Bolat ve Karamustafaoğlu, 2019; Nacaroglu, Bektaş ve Kızıkan, 2020) bulunmaktadır.

Testin KR-20 değeri incelendiğinde bu değer 0,82 olarak bulunmuştur. KR-20 değerinin 0.80 ile 1.00 arasında hesaplanması testin yüksek derecede güvenilir olduğunu gösterir (Kalaycı, 2008). Bu nedenle bu çalışma kapsamında ele alınan başarı testinin güvenilir sonuçlar verebileceđi söylenebilir. Güvenirlik katsayısı yarı test güvenirlik yöntemi ile hesaplanmış ve Spearman-Brown katsayısı 0.79 olarak bulunmuştur. 0.70 ve üstü güvenirlik değerleri ölçme aracının güvenilir olduğunu göstermektedir

(Tavşancıl, 2014). Bu durumda testin güvenilirliğine dair elde edilen KR-20 ve Spearman-Brown değerlerinin birbirlerini desteklediği görülmektedir.

Sonuç olarak bir başarı testinin farklı örneklem üzerindeki güvenilirliği araştırılmıştır. Ayırt edicilik indekslerine bakıldığında ayırt etme düzeyinin yüksek, güçlük indekslerine bakıldığında orta güçlükte bir test olduğu görülmüştür. Ayırt ediciliği yüksek ve orta güçlükteki testler alanyazına göre istenilen testlerdir. Bunun sonucunda testin uygulanabilir olduğu sonucuna varılmıştır. Güvenirlik için KR – 20 ve Spearman Brown katsayıları alanyazına göre istenilen değerlerin üzerinde çıkmıştır. Bu da uygulanacak testin güvenilir olduğunu göstermektedir. Elde edilen bulgulara göre basınç konusunda geliştirilmiş olan başarı testinin farklı bir örneklem üzerinde de çalışılabilir olduğu belirlenmiştir.

Yapılan çalışma neticesinde hesaplanan madde güçlük indeksine göre zor kategorisinde bir madde görülmemiştir. Gelecekte bu testin güvenilirlik çalışmalarının yapılacağı farklı araştırmalarda bu durum göz önünde bulundurularak madde düzenlenmesi yapılabilir. Basınç konusu son yıllarda çalışılan 2013 ve 2018 Fen Bilimleri Öğretim Programı'nda yer almakla birlikte 2024 programında bu konuya yönelik kazanımlara yer verilmediği tespit edilmiştir (Millî Eğitim Bakanlığı [MEB], 2024). Yeni öğretim programında öğrenciler basınç konusu ile 9.sınıf itibari ile tanışacaklardır. İlgili kazanımlar incelendiğinde bazılarının eski programdakiler ile ortak olduğu görülmektedir. Bu bağlamda ilgili test özellikle basıncı etkileyen değişkenler ve günlük hayatta kullanım üzerine yazılan sorular üzerinden genişletilip geçerlik ve güvenilirlik çalışmaları yapılarak ölçeğin kullanılabilirliği araştırılabilir.

Araştırma ve Yayın Etiği

Bu araştırmada, Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi'nde belirtilen tüm kurallara titizlikle uyulmuştur. Yönergede yer alan *Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler* başlığı altında bahsi geçen hiçbir eylem gerçekleştirilmemiştir.

Etik Kurul İzni

Bu araştırmayı gerçekleştirmek üzere Trabzon Üniversitesi Sosyal ve Beşerî Bilimler Bilimsel Araştırma ve Yayın Etik Kurulu'ndan 08.09.2023 tarihli 2023-9/1.1 sayılı etik kurul izni alınmıştır.

Yazarların Katkı Oranı

Bu çalışmaya tüm yazarlar eşit oranda katkıda bulunmuştur.

Çıkar Çatışması

Çalışmanın ortaya çıkmasında herhangi bir çıkar çatışması mevcut değildir.

Kaynakça

- Akgün, A. Tokur, F. & Özkara, D. (2013) TGA stratejisinin basınç konusunun öğretimine olan etkisinin incelenmesi. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, 2(2), 348-369.
- Atabey, N., & Çiftçi, A. (2019). Fen bilimleri öğretmen adaylarının gaz basıncı ile ilgili kavram yanılgılarının TGA yöntemiyle belirlenmesi. *Turkish Studies-Educational Sciences*, 14(2), 1-17.
- Ayvacı, H. Ş., & 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.
- Baytok, H. (2007). *Yapılandırmacı öğrenme kuramına dayalı öğretimin ilköğretim 7. Sınıf basınç konusunda öğrenci başarısı ve tutumuna etkisi*. Yayınlanmamış yüksek lisans tezi. Balıkesir: Balıkesir Üniversitesi, Fen Bilimleri Enstitüsü.

- Bolat, A., & Karamustafaoğlu, S. (2019). "Vücudumuzdaki sistemler" ünitesi başarı testi geliştirme: Geçerlik ve güvenilirlik. *Gazi Eğitim Bilimleri Dergisi*, 5(2), 131-159.
- Boz, S., Özcan, H., & Sarıođlan Bostan, A. (2023). Ortaokul öğrencilerinin basınç konusu ile ilgili bilgilerini ölçmeye yönelik bir başarı testinin geliştirilmesi. *İnönü Üniversitesi Eğitim Bilimleri Enstitüsü Dergisi*, 10(19), 14-29.
- Bozan, M., & Küçüközer, H. (2007). İlköğretim öğrencilerinin basınç konusu ile ilgili problemlerin çözümünde yaptıkları hatalar. *İlköğretim Online*, 6(1), 24-34.
- Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Kuram ve Uygulamada Eğitim Yönetimi Dergisi*, 32, 470-483.
- Büyüköztürk Ş., Çakmak, E. K., Akgün Ö. E., Karadeniz Ş. & Demirel F. (2012). *Bilimsel araştırma yöntemleri* (12. Baskı). Pegem Akademi.
- Can, Y. (2019). *Sorgulamaya dayalı öğrenme yaklaşımının 7. sınıf öğrencilerinin basınç kavramı ile ilgili kavramsal anlamalarına etkisi*. Yayımlanmamış yüksek lisans tezi. Balıkesir: Balıkesir Üniversitesi, Fen Bilimleri Enstitüsü.
- Davis, B. G. (1993). *Tools for teaching*. San Francisco, CA: Jossey-Bass.
- Değirmenci, S., Karamustafaoğlu, O., & Karamustafaoğlu, S. (2021). Fen bilgisi öğretmen adaylarının bazı mekanik kavramlarına yönelik imaj ve metaforları. *Türkiye Sosyal Araştırmalar Dergisi*, 25(2), 595-622.
- Demirel, R. (2015). Katı basınçta argümantasyon faaliyetlerinin uygulanması. *Sorgulamaya Dayalı Etkinlikler Dergisi*, 5(2), 70-90.
- Gönen, S. & Andaç, K. (2009). Gözden geçirme stratejisi ile desteklenmiş yapılandırmacı öğrenme yaklaşımının öğrencilerin basınç konusundaki erişilerine ve bilgilerinin kalıcılığına etkisi. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, (12), 28-40.
- Görkemli Taban, T. (2017). *Fen bilgisi öğretmen adaylarının sıvı basıncı konusundaki kavram yanlışlarının dört aşamalı tanı testi ile belirlenmesi*. Yayımlanmamış yüksek lisans tezi. Necmettin Erbakan Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Guerra-Reyes, F., Guerra-Dávila, E., Naranjo-Toro, M., Basantes-Andrade, A., & Guevara-Betancourt, S. (2024). Misconceptions in the learning of natural sciences: A systematic review. *Education Sciences*, 14(5), 497.
- Hasańebi, 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.
- Hayati, U., Ediyani, M., Maimun, M., Anwar, K., Fauzi, M. B., & Suryati, S. (2020). Test technique as a tool for evaluation of learning outcomes. *Budapest International Research and Critics Institute-Journal (BIRCI-Journal)*, 3(2), 1198-1205.
- Holder, W. W., & Mills, C. N. (2001). Pencils down, computer up--the new CPA exam. *Journal of Accountancy*, 191(3), 57-60.
- Kalaycı, Ş. (2008). *SPSS uygulamalı çok deđişkenli istatistik teknikleri*. Asil Yayın Dağıtım.
- Karaca, E. (2004). Seçme gerektiren, kısa cevaplı ve doğru- yanlış testlerinin madde ve test özelliklerinin karşılaştırılması. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi* (10).
- Kaya, D., Bozdađ, H. C. & Ok, G. (2018). Yedinci sınıf öğrencilerinin basınç konusundaki kavramsal anlamaları ve kavram yanlışlarının matematiksel hatalar açısından incelenmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18 (1), 321-341.
- Kızkapan, O., & Bektaş, O. (2018). Fen eğitiminde başarı testi geliştirilmesi: Hücre bölünmesi ve kalıtım örneđi. *Maarif Mektepleri Uluslararası Eğitim Bilimleri Dergisi*, 2(1), 1-18.
- Kirişciođlu, S. (2007). *İlköğretim 7. sınıf fen bilgisi dersi" basınç" konusunun yapılandırmacı öğrenme yaklaşımına dayalı öğretiminin akademik başarıya etkisi*. Yayımlanmamış yüksek lisans tezi. Pamukkale Üniversitesi, Fen Bilimleri Enstitüsü.

- Laçın-Şimşek, C., Öztuna-Kaplan, A., Çorapçığıl, A., & Mısır, M. E. (2018). Fen bilgisi öğretmenliği 3. sınıf öğrencilerinin basınç-kaynama noktası ilişkisine yönelik düşünceleri: Bir TGA uygulaması. *Kastamonu Eğitim Dergisi*, 26(5), 1679-1690.
- McMillan, J. H. (2017). *Principles and practice that enhance student learning and motivation* (7th Edition), Pearson.
- Millî Eğitim Bakanlığı (MEB) (2024, Eylül 16). *Fen Bilimleri Dersi Öğretim Programı* <https://mufredat.meb.gov.tr/>
- Nacaroğlu, O., Bektaş, O., & Kızılcapan, O. (2020). Madde döngüleri ve çevre sorunları konusunda başarı testi geliştirme: Geçerlik ve güvenirlik çalışması. *Kastamonu Eğitim Dergisi*, 28(1), 36-51.
- Özcan, H., & Koca, E. (2019). STEM yaklaşımı ile basınç konusu öğretiminin ortaokul 7. sınıf öğrencilerinin akademik başarılarına ve STEM'e yönelik tutumlarına etkisi. *Eğitim ve Bilim*, 44(198).
- Özcan, H., Esra, Koca, & Söğüt, M. (2019). Ortaokul öğrencilerinin basınç kavramıyla ilgili anlayışlarını ölçmeye yönelik bir test geliştirme çalışması. *Eğitim ve Toplum Araştırmaları Dergisi*, 6(1), 130-144.
- Özkan, E. B., & Eryılmaz Muştı, Ö. (2018). 8. sınıf basit makineler ünitesine yönelik başarı testi geliştirme: geçerlik ve güvenirlik çalışması. *Hitit Üniversitesi Sosyal Bilimler Enstitüsü Dergisi* 11(1), 737-754.
- Özkara, D. (2011). *Basınç konusunun sekizinci sınıf öğrencilerine bilimsel argümantasyona dayalı etkinlikler ile öğretilmesi*. Yayımlanmamış yüksek lisans tezi. Adıyaman Üniversitesi, Fen Bilimleri Enstitüsü.
- Sadiç, A. (2016). Açık hava basıncı konusunun öğretiminde kullanılacak örnek TGA etkinlikleri. *Araştırma Temelli Etkinlik Dergisi (ATED)*, 6(2), 63-79
- Saltan, F., & Divarci, O. F. (2017). Using blogs to improve elementary school students' environmental literacy in science class. *European Journal of Educational Research*, 6(3), 347-355.
- Saraç, H. (2018). Fen bilimleri dersi maddenin değişimi ünitesi ile ilgili başarı testi geliştirme: geçerlik ve güvenirlik çalışması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18(1), 416-445.
- Sezgin Selçuk, G. (2019). Tarama yöntemi. H. Özmen & O. Karamustafaoğlu (Ed.). *Eğitimde araştırma yöntemleri içinde* (ss.140-161). PegemA.
- Sönmez, G., Geban, Ö., & Ertepinar, H. (2001). Altıncı sınıf öğrencilerinin elektrik konusundaki kavramları anlamalarında kavramsal değişim yaklaşımının etkisi. *Yeni bin yılın başında Türkiye'de fen bilimleri eğitimi sempozyumu içinde* (35-38 ss.). Maltepe Üniversitesi.
- Suprpto, N. (2020). Do we experience misconceptions?: An ontological review of misconceptions in science. *Studies in Philosophy of Science and Education*, 1(2), 50-55.
- Şahin, Ç. (2010). *İlköğretim 8. sınıf "kuvvet ve hareket" ünitesinde zenginleştirilmiş 5E öğretim modeline göre rehber materyaller tasarlanması, uygulanması ve değerlendirilmesi*. Yayımlanmamış Doktora tezi. Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü.
- Şahin, Ç., & Çepni, S. (2012) 5E öğretim modeline dayalı öğretimin öğrencilerin gaz basıncı ile ilgili kavramsal anlamalarına etkisi. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 6(1), 220-264
- Şen, H. C., & Eryılmaz, A. (2011). Bir başarı testi geliştirme çalışması: Basit elektrik devreleri başarı testi geçerlik ve güvenirlik araştırması. *Van Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 8(1), 1-39.
- Tavşancıl, E. (2014). *Tutumların ölçülmesi ve SPSS ile veri analizi*. Nobel.
- Tian, H., Sun, Z., Tian, H., & Sun, Z. (2018). Historical development of academic achievement assessment. *Academic Achievement Assessment: Principles and Methodology*, içinde (pp 3-13), Springer.

- Topay, N., & Yılmaz, M. (2023). Biyoloji ve fen bilgisi öğretmenlerine yönelik tamamlayıcı ölçme ve değerlendirme teknikleri başarı testi geliştirilmesi. *Gazi Eğitim Bilimleri Dergisi*, 9(2), 214-240.
- Tural, G. (2020). Promoting students' understanding of the concept of pressure: Active learning environment versus traditional one. *Mimbar Sekolah Dasar*, 7(3), 284-303.
- Turgutalp, E. (2021). 8. sınıf basınç konusunda STEM öğretme-öğrenme modelinin uygulanmasının öğrenci başarısına ve girişimcilik becerisine etkisinin araştırılması. Yayınlanmamış Doktora tezi. Bursa Uludağ Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Ünal, G. (2005). *Fen öğretiminde derinliğine öğrenme basınç konusunda modelleme*. Yayınlanmamış Yüksek lisans tezi. Dokuz Eylül Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Ünal, G., & Ergin, Ö. (2006). Buluş yoluyla fen öğretiminin öğrencilerin akademik başarılarına, öğrenme yaklaşımlarına ve tutumlarına etkisi. *Journal of Turkish Science Education*, 3(1), 36-52.
- Yalçın, S. (2021). Ölçek geliştirme ve uyarlama süreci bilgi notu. Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü Ders Notları. <https://egitim.ankara.edu.tr/wp-content/uploads/sites/347/2021/06/Olcek-Gelistirme-ve-Uyarlama.Bilgi-notu.-Doc.-Dr.-Seher-Yalcin.pdf>
- Yaman, E. (2016). *Ortaokul 8.sınıf öğrencilerinin basınç konusunda kavramsal anlamalarının incelenmesi*. Yayınlanmamış yüksek lisans tezi. Balıkesir: Balıkesir Üniversitesi, Fen Bilimleri Enstitüsü.
- Yiğit, N. (2008). Test geliştirme. G. Başol (Ed.) *Eğitimde ölçme değerlendirme içinde* (153-171). Lisans Yayıncılık.

Extended Abstract

Introduction

It is recognized within educational discourse that students encounter challenges in comprehending abstract scientific concepts, with the topic of pressure notably posing difficulties for middle school students (Can, 2019). This situation can lead to an increase in students' conceptual misconceptions and a decrease in academic achievement (Sönmez, Geban, & Ertepinar, 2001).

The literature indicates that students experience difficulties in understanding various types of pressure (Görkemli Taban, 2017; Kaya, Bozdağ & Ok, 2017; Şahin & Çepni, 2012; Yaman, 2016) and struggle with making mathematical calculations related to pressure (Kaya, Bozdağ & Ok, 2017), as well as establishing relationships between variables affecting pressure (Şahin, 2010; Yaman, 2016). To address these conceptual misconceptions and ensure effective teaching, different methods and techniques should be employed. In the literature, numerous studies have been conducted to determine the impact of various methods and techniques on students' academic achievements in the topic of pressure (Akgün et al., 2013; Baytok, 2007; Can, 2019; Demirel, 2015; Gönen & Andaç, 2009; Kirişçioğlu, 2007; Özkara, 2011; Sadıç, 2016; Şahin & Çepni, 2012; Tural, 2020; Turgutalp, 2021; Ünal & Ergin, 2006; Yaman, 2016). When considering the objectives of research endeavors concerning pressure, it is notable that investigations are directed towards various facets including teaching the subject (Akgün et al., 2013; Baytok, 2007; Demirel, 2015; Gönen & Andaç, 2009; Kirişçioğlu, 2007; Özkara, 2011; Sadıç, 2016; Tural, 2020; Turgutalp, 2021; Saltan & Divarcı, 2017; Ünal, 2005; Ünal and Ergin, 2006; Şahin and Çepni, 2012; Yaman, 2016), the development of assessment tools (Özcan, Koca, and Söğüt, 2019), and the identification of errors encountered in problem-solving endeavors (Atabey and Çiftçi, 2019; Bozan and Küçüközer, 2007). Various data collection tools have been utilized in studies related to the topic of pressure. These include pressure achievement tests (Akgün et al., 2013; Baytok, 2007; Gönen & Andaç, 2009; Kirişçioğlu, 2007; Özcan & Koca, 2019; Özcan et al., 2019; Özkara, 2011; Saltan & Divarcı, 2017; Ünal, 2005; Ünal & Ergin, 2006), portfolios (Kirişçioğlu, 2007), in-class discussions recorded during the process (Laçın Şimşek, Öztuna Kaplan & Çorapçigil, 2018),

open-ended tools (Bozan and Küçüközer, 2007; Pabuçcu, 2016; Ünal, 2005), interviews (Baytok, 2007; Karamustafaoğlu & Kandaz, 2006), surveys (Karamustafaoğlu & Kandaz, 2006; Yaman, 2016), misconceptions tests (Laçın Şimşek et al., 2018; Ünsal, 2019), and worksheets (Atabey & Çiftçi, 2019). In the examined studies, pressure achievement tests have been observed as the primary data collection tool. Achievement tests can be used to assess the effectiveness of teaching. In other words, they allow for the measurement of instructional effectiveness and the revelation of students' cognitive structures. With this feature, achievement tests facilitate necessary improvements and contribute to the assessment and evaluation process (Boz, Özcan, & Bostan-Sarioğlu, 2023). Various assessment tools are utilized to measure academic achievement, with multiple-choice items being the most common due to their ability to yield objective results (McMillan, 2018). While it is acknowledged that multiple-choice tests may fall short in measuring higher-order skills, when well-structured, they can also assess such skills (Davis, 1993). Validity and reliability studies related to the relevant test need to be conducted for previously prepared achievement tests to be applied to specific samples to and utilized in different studies. However, validity and reliability studies may not provide information regarding the performance of the items in the test. Therefore, determining the difficulty index and discrimination levels of the items is also considered important (Hasağebi, Terzi & Küçük, 2020). In this context, the purpose of this study is to investigate the reliability status of the pressure achievement test designed to measure students' academic achievements in the topic of pressure, as it will be administered on a different sample group.

Method

This study employed a survey method within the framework of quantitative research approach. The achievement test that was developed by Özcan, Koca, and Söğüt (2019) for the topic of Pressure was utilized in the study. The test comprises 20 questions, and the sample of the study consists of 149 8th grade students. The analysis of students' responses to the test items is conducted utilizing the Test Analysis Program (TAP).

Results and Discussion

The range of item difficulty index spans from 0.40 to 0.77, with the item discrimination index oscillating between 0.32 and 0.74. The average item difficulty of the achievement test is computed at 0.57, while the average item discrimination index stands at 0.55. As a result of the item analysis, the KR-20 reliability coefficient of the achievement test was found to be 0.820, and the Spearman-Brown coefficient calculated using the split-half method was 0.795. These psychometric findings validate the robustness of the instrument. As a result of the study, no item was observed to fall under the difficult category according to the calculated item difficulty index. In future research where the reliability of this test will be investigated, item adjustments can be made considering this situation. The topic of pressure, which was included in the 2013 and 2018 Science Education Curricula, is notably absent from the objectives outlined in the 2024 curriculum (Ministry of National Education [MoNE], 2024). According to the updated curriculum, students will first encounter the concept of pressure in the 9th grade. A review of the new objectives reveals some overlap with those in the previous curriculum. In this context, the test could be refined and expanded by incorporating questions that emphasize variables influencing pressure and its real-world applications. Future studies could then evaluate the scale's validity and reliability to ensure its continued applicability.