

Kocaeli Üniversitesi

Eğitim Dergisi

E-ISSN: 2636-8846

2025 | Cilt 8 | Sayı 1

Sayfa: 73-105



Kocaeli University
Journal of Education

E-ISSN: 2636-8846

2025 | Volume 8 | Issue 1

Page: 73-105

Ortaokul öğrencileri için Blockly ile programlama
başarı testi geliştirilmesi

Developing an achievement test of programming
with Blockly for middle school students

Asım KANMAZ,  <https://orcid.org/0000-0002-8310-3560>

Hatay Mustafa Kemal Üniversitesi, Fen Bilimleri Enstitüsü, asim.kanmaz@gmail.com

Hamide ŞAHİNKAYASI,  <https://orcid.org/0000-0001-8787-8163>

Hatay Mustafa Kemal Üniversitesi, Eğitim Fakültesi, hamide@mku.edu.tr

Bu çalışma, Asım KANMAZ tarafından hazırlanmakta olan "Bilgisayarsız Ortamda Blok Temelli Programlama Söz Dizim Yapısını Öğrenmeyi Destekleyen Öğretim Materyali Geliştirilmesi ve Etkililiğinin İncelenmesi" başlıklı doktora tezinden üretilmiştir.

ARAŞTIRMA MAKALESİ

Gönderim Tarihi	Düzeltilme Tarihi	Kabul Tarihi
11 Ağustos 2024	13 Aralık 2024	16 Aralık 2024

Önerilen Atıf

Recommended Citation

Kanmaz, A., & Şahinkayası, H. (2025). Ortaokul öğrencileri için Blockly ile programlama başarı testi geliştirilmesi. *Kocaeli Üniversitesi Eğitim Dergisi*, 8(1), 73-105. <http://doi.org/10.33400/kuje.1531434>

ÖZ

Ortaokul öğrencilerine kodlama becerileri temellerinin kazandırılması, Bilişim Teknolojileri ve Yazılım dersinin esas amaçlarından biridir. Bu derslerde kodlama eğitimi için Scratch, Blockly, Code.org, Kodu Game Lab gibi blok temelli programlama araçları kullanılmaktadır. Milli Eğitim Bakanlığı ile Google arasındaki iş birliği protokolü kapsamında 2017-2018 eğitim öğretim yılından itibaren bu derste Blockly blok kodlama uygulaması kullanılmaktadır. Ancak, alan yazınında Blockly ortamında yapılan kodlama öğretiminde kullanılabilecek bir başarı testi mevcut değildir. Bu çalışmanın amacı, ilgili dersin “Problem Çözme ve Programlama” ünitesinin programlama ile ilgili kazanımlarını Blockly uygulaması bağlamında ölçen geçerli ve güvenilir bir başarı testi geliştirmektir. Çalışmada, klasik test teorisine dayanan test geliştirme aşamaları izlenmiştir. Söz konusu ünitenin programlama ile ilgili 12 kazanımı için dörder madde hazırlanarak toplam 48 soruluk madde havuzu oluşturulmuş ve kapsam geçerliği için uzman görüşleri alınmıştır. Uzmanlardan gelen görüşlere göre bazı maddelerde değişiklikler yapılmıştır. Araştırmada ölçüt örnekleme tercih edilmiştir. Deneme testi, 48 soruluk madde havuzundan oluşturulmuş ve ilgili dersi bir önceki yıl almış olan 151 6. sınıf öğrencisine uygulanmıştır. Testin 24 maddelik nihai formu ise 72 5. sınıf öğrencisine ünite işlendikten sonra uygulanmıştır. Nihai testin ortalama madde güçlük indeksi 0.63, ortalama madde ayırtıcılık indeksi 0.48 ve güvenirlik katsayısı (KR-20) 0.86 olarak hesaplanmıştır. Geliştirilen test, orta derece zorlukta ve test güvenirliği yüksektir. Bu testin, Blockly ile programlama öğretimi yapan öğretmenlerin ve araştırmacıların ölçme ve değerlendirmelerine katkı sunması beklenmektedir.

Anahtar Sözcükler: blok tabanlı programlama, Blockly, başarı testi

ABSTRACT

Teaching the basics of coding skills is among the main objectives of the Information Technologies and Software (ITS) course in middle schools in Türkiye. Block-based programming tools such as Scratch, Blockly, and Code.org are used for coding education. Blockly application has been used in ITS courses since the 2017-2018 academic year based on the cooperation protocol between the Ministry of National Education and Google. However, no achievement test has been encountered for coding in the Blockly environment in the literature. This study aims to develop a valid and reliable test in the context of Blockly for programming-related learning objectives in ITS's Problem Solving and Programming unit. Test development stages based on classical test theory were implemented. An item pool of 48 questions, measuring the programming-related 12 learning objectives of the unit, was created and expert opinions were applied for content validity. Some items were improved based on expert opinions. Criterion sampling was used. The trial test was applied to 151 6th-grade students who had taken the course the previous year. The final form of the test, having 24 items, was administered to 72 5th-grade students after the unit was taught. The final test's item difficulty index mean was 0.63, the item discrimination index mean was 0.48, and the reliability coefficient (KR-20) was 0.86. The test is of medium difficulty and is highly reliable. This test is expected to contribute to measuring and evaluating learning outcomes for teachers and researchers who teach programming with Blockly.

Keywords: block-based programming, Blockly, achievement test

GİRİŞ

Öğretimin temel hedefi, öğretim programı doğrultusunda öğrencilere öğretilen bilgi ve becerilerin kalıcı hale getirilmesidir (Demirtaşlı, 2017). Öğrencilerin, öğrenme süreçlerinde değişimleri ölçme ve değerlendirme yoluyla anlaşılır (Tekin, 2019). Ölçme, bir öğrencinin öğrenme çıktı düzeyini belirleme veya elde edilen öğrenme çıktısını nicel verilere dönüştürme sürecidir. Değerlendirme ise, ölçülen özelliğin ölçme sonucunu bir ölçüt ile karşılaştırma ve bu özellik için bir karara varma işlemidir (Özçelik, 2010).

Öğrencilerin başarılarını ölçmek ve değerlendirmek için sözlü sınavların yanı sıra yazılı sınavlar, kısa yanıtli testler, doğru-yanlış testleri, ölçekler, açık uçlu sorular, boşluk doldurmalı testler, eşleştirmeli testler ve çoktan seçmeli testler kullanılır (Turgut & Baykul, 2014; İpek Akbulut & Çepni, 2013; Balcı, 2011). Başarı testleri, öğrencilerin akademik seviyesini ölçmek için eğitimciler tarafından sıklıkla tercih edilen araçlardır (Haladyna, 1997). Bu testler, öğrencilerin öğrenme çıktılarının ölçülmesini sağlar (Best & Khan, 2006). Özçelik (2013) çoktan seçmeli soruları içeren başarı testlerinin etkili bir ölçme aracı olduğunu belirtmiştir. Alan yazınına bakıldığında öğrenme sürecini ölçmek için genellikle çoktan seçmeli sorulardan oluşan başarı testleri tercih edildiği görülür (Akdeniz, Öztürk & Bakırcı, 2017; Korkmaz & Konukaldı, 2015; Açıkgöz & Karşı, 2015; İpek Akbulut & Çepni, 2013; Can Şen & Eryılmaz, 2011; Gönen, Kocakaya & Kocakaya, 2011; Çalık & Ayas, 2003).

Başarı testleri programlama eğitiminde de önemli bir ölçme değerlendirme aracıdır. Dünyada programlama eğitimi incelendiğinde bazı ülkelerde anasınıfından liseye kadar farklı kademelerde programlama eğitiminin yer aldığı görülür (Demirer & Sak, 2016). Türkiye’de ise Bilişim Teknolojileri ve Yazılım (BTY) dersi 2012 yılında 5. ve 6. sınıflara zorunlu 7. ve 8. sınıflara seçmeli ders olarak müfredata girmiştir (MEB-TTKB, 2015). Programlama eğitimi sürecinde Scratch, Blockly, Code.org, Kodu Game Lab gibi blok tabanlı programlama araçları kullanılır. Örneğin, Scratch; görsel ara yüze sahip, sürükle-bırak yöntemi kullanılarak kodlama yapılan bir programlama platformudur. (Arslan & Akçelik, 2019). Scratch programının yanı sıra Millî Eğitim Bakanlığı (MEB), 5. sınıflar için Google tarafından geliştirilen, çevrimiçi ve çevrimdışı kullanılabilen Blockly Blok Tabanlı Programlama Uygulaması’nı (BBTPU) önerir. BBTPU görsel bir şekilde sürükle-bırak yöntemiyle kodlama temellerinin öğretilmesini sağlar (Aytekin, Sönmez Çakır, Kulaöz & Yücel, 2018). BTY dersi için Temel Eğitim Genel Müdürlüğü ile Google arasındaki iş birliği protokolü kapsamında 2017-2018 eğitim-öğretim yılında BBTPU resmi ve özel okullarda derslerde kullanılmaya başlanmıştır (MEB, 2017). BTY dersinde kazanımlar uygulamalı olarak ya da yazılı olarak ölçülebilmektedir. Blok tabanlı programlama ortamları ile ilgili kazanımlar uygulamalı olarak dereceli puanlama anahtarı ile ölçülmektedir. Başarı testleri ise yazılı olarak kazanımların ölçülmesini sağlayan ölçme araçlarından biridir. Şenel (2019), BTY dersi öğretmenlerinin ölçme yöntemi olarak en fazla çoktan seçmeli maddeleri içeren ölçme araçları kullandıklarını belirtmiştir.

Alan yazınında genellikle Scratch uygulama ortamı bağlamında blok tabanlı programlama başarı testi geliştiren ve kullanan bazı çalışmalar vardır. Oluk, Korkmaz ve Oluk (2018), 5. sınıf öğrencilerinin bilgi-işlemsel düşünme becerilerini incelemek için Scratch programlama ortamına yönelik dört seçenekli 27 maddeden oluşan bir başarı testi geliştirmiştir. Bakırcı (2019), 6. sınıflar için algoritma, değişken, koşul, kontrol, döngü gibi temel programlama kavramlarını kapsayan çoktan seçmeli 25 soru içeren bir başarı testi geliştirmiştir. Şahin (2019), programlama öğretiminde ters yüz sınıf uygulamalarının ortaokul öğrencilerinin akademik başarısına etkisini ölçmek için 6. sınıflara yönelik bir başarı testi hazırlamıştır. Aydın (2019), algoritma öğretiminde Scratch kullanımının ortaokul öğrencilerinin başarısına etkisini incelemiştir. Daha önce algoritma ve blok tabanlı programlama eğitimi almamış 5. ve 7. sınıf öğrencilerine sekiz soruluk açık uçlu bir başarı testi uygulamıştır. Alagöz ve Bilgiç (2023) ortaokul öğrencilerine yönelik sabit, değişken, döngü, koşul, algoritma gibi temel programlama kavramlarını ölçen; açık ve kapalı uçlu toplam 14 sorudan oluşan bir başarı testi hazırlamıştır. Piedade ve Dorotea (2023) Scratch tabanlı etkinliklerin 4. sınıf öğrencilerinin bilgi-işlemsel düşünme becerileri üzerine etkisini incelemek için Zapata-Cáceres, Martín-Barroso ve Román-

González (2020) tarafından geliştirilen BCTT (Beginners Computational Thinking Test) isimli çoktan seçmeli testi uygulamıştır. Grover, Jackiw ve Lundh (2019), 6-7-8. sınıflara yönelik olan çalışmasında VELA isimli etkinlikte Scratch blok tabanlı programlamaya yönelik sorulara da yer verdiği açık uçlu ve çoktan seçmeli sorulardan oluşan dokuz soruluk bir test kullanmıştır. Demir ve Seferoğlu (2019) ise Scratch programlama platformu bağlamında liselere yönelik 18 maddelik bir başarı testi hazırlamıştır.

Alan yazınında programlama eğitime yeni başlayanlar için geliştirilen ve özellikle BBTPU'nun öğrenme süreci çıktılarını ölçen kapsamlı bir başarı testi mevcut değildir. Bu çalışma, öğretmenlerin ve araştırmacıların BBTPU kullanılan programlama eğitiminde 5. sınıf öğrencilerinin öğrenme çıktılarını geçerli ve güvenilir bir şekilde ölçmesini sağlayan bir başarı testi geliştirerek alan yazınına katkı sunmaktadır.

YÖNTEM

Araştırma tarama modelindedir. Nicel araştırma yöntemlerinden biri olan tarama araştırması, bir örneklem grubunun görüşlerini ve davranışlarını açıklamak için verilerin toplandığı araştırma desenidir (Creswell, 2017).

Katılımcı Bilgisi

Araştırmanın örnekleme, seçkisiz olmayan amaçsal örnekleme yöntemlerinden ölçüt örneklemedir. Ölçüt örneklemede, çalışma grubu bir dizi ölçütü karşılar (Yıldırım ve Şimşek, 2016). Bu çalışmada deneme uygulaması katılımcıları için kullanılan ölçüt, çalışmaya katılan öğrencilerin 5. sınıfta BTY dersini almış 6. sınıf öğrencileri olmasıdır. Nihai formun uygulandığı grubun ölçütü ise "Problem Çözme ve Programlama" (PÇP) ünitesini derste işlemiş 5. sınıf öğrencileri olmasıdır. Araştırmanın deneme uygulamasının yapıldığı grup, 2021-2022 yılında Hatay ili Antakya ilçesinde bulunan Beyhan Gencay Ortaokulu (n=69) ve İnönü Ortaokulunda (n=82) öğrenim gören toplam 151 (71 Kız, 80 Erkek) ortaokul 6. sınıf öğrencisidir. Nihai testin uygulandığı grup ise Hatay ili Arsuz ilçesi Madenli Ortaokulunda öğrenim gören 72 (34 Erkek, 38 Kız) 5. sınıf öğrencisidir. Öğrencilerin cinsiyet oranlarının yakın olduğu söylenebilir.

Veri Toplama Aracı

Yazarlar tarafından geliştirilen başarı testinin 48 maddeden oluşan deneme testi öğrencilere Google Formlar uygulaması üzerinden dijital ortamda uygulanmıştır. Deneme uygulaması verileri üzerinde yapılan analizler sonucu seçilen maddelerden oluşturulan ve Ek 1'de sunulan 24 maddelik nihai test ise öğrencilere kâğıt üzerinde uygulanmıştır.

Test Geliştirme Süreci

Beşinci sınıf BTY dersinin PÇP ünitesinin programlama ile ilgili kazanımlarını, özellikle ikinci kısmında yer alan blok tabanlı programlama konusunun kazanımlarını ölçmek için test geliştirilirken klasik test teorisine dayanan ve aşağıda açıklanmış olan Özçelik (2013)'in başarı testi geliştirme adımları izlenmiştir.

1. Testin amacı ve kapsamının belirlenmesi

Testin amacı, BTY dersi 5. sınıf öğretim programında (MEB, 2018) yer alan PÇP ünitesinin programlama ile ilgili kazanımlarını (Tablo 1) BBTPU ortamından Labirent ve Kuş senaryolarındaki görselleri ve senaryoları kullanarak ölçmektir.

2. Madde havuzunun oluşturulması

Araştırmada 48 maddelik madde havuzu ikinci yazarın danışmanlığında birinci yazar tarafından alan yazını taranarak, görüş alınarak ve öğretmenlik bilgi birikimine dayanarak oluşturulmuştur. Hazırlanan testte üniteye öğretilecek her yeni davranışı yoklayacak en az bir soru bulunmalıdır (Özçelik, 2013). Deneme formunun hazırlanan soruların kapsam geçerliğini koruması için nihai testin en az üç katı kadar madde içermesi önerilir (Özüdoğru & Aksu, 2019). Bu çalışmada, Tablo

1’de yer alan 12 kazanımın her birinden 4 soru olmak üzere BBTPU’ya uygun 48 soru hazırlanmıştır. Oluşturulan maddelerin kazanım ifadelerinin hangi davranış grubunu ölçtüğünü gösteren bir belirtke tablosu hazırlanmıştır (Tablo 1). Bu tablo aynı zamanda kapsam geçerliğini de destekler (Can Şen & Eryılmaz, 2011).

Tablo 1*Deneme Testi için Belirtke Tablosu*

Ünite Bölümü	Kazanımlar	Hatırlama	Anlama	Uygulama	Deneme testindeki madde numaraları	Toplam
Problem Çözme Kavramları ve Yaklaşımları	BT.5.5.1.6. Problemi çözmek için gerekli değişken, sabit ve işlemleri açıklar.		4		M1, M2, M3, M4	4
	BT.5.5.1.8. Problem çözümünde ifade ve eşitliklere örnek verir.	1	3		M9, M10, M11, M12	4
	BT.5.5.1.7. Problem çözümünde kullanılabilecek operatörlere örnek verir.	3	1		M5, M6, M7, M8	4
	BT.5.5.1.11. Verilen bir problemde ifade ve eşitlikleri kullanarak çözüm üretir.			4	M37, M46, M47, M48	4
	BT.5.5.1.10. Verilen bir problemin çözümünde operatörleri kullanır.			4	M42, M43, M44, M45	4
	BT.5.5.1.17. Matematik ve bilgisayar bilimi arasındaki ilişkiyi tespit eder.		4		M35, M36, M38, M39	4
Programlama	BT.5.5.2.4. Doğrusal mantık yapısını açıklar.	3	1		M13, M14, M15, M16	4
	BT.5.5.2.5. Doğrusal mantık yapısını kullanan algoritmalar geliştirir.			4	M17, M18, M19, M20	4
	BT.5.5.2.6. Karar yapısını ve işlevlerini açıklar.	2	2		M21, M22, M25, M26	4
	BT.5.5.2.7. Karar yapıları içeren algoritmalar geliştirir.			4	M27, M28, M40, M41	4
	BT.5.5.2.8. Döngü yapısını ve işlevlerini açıklar.	2	2		M23, M24, M29, M30	4
	BT.5.5.2.9. Döngü yapısı içeren algoritmalar oluşturur.			4	M31, M32, M33, M4	4
Toplam		11	17	20		48

3. Ölçme biçiminin belirlenmesi

PÇP ünitesinin blok tabanlı programlama ile ilgili kazanımlarını ölçmek için hedef kitlenin yaşına uygun olarak dört seçenekli çoktan seçmeli test maddeleri hazırlanmıştır. Seçme gerektiren maddeler objektif puanlama sağlar. Bu tür maddelerde yanıtların belirlenmesi ve puanlanması işlemlerine daha az yorum karışır. Açık uçlu maddelere göre daha güvenilir puanlama sağlar. Seçme gerektiren maddelerin en büyük sınırlılığı şans etmeninden etkilenmesidir (Gültekin, 2017). Bu sınırlılığı aşmak amacıyla çeldiricilerin güçlü olması için onların soru ile ilişkili başka soruların doğru cevabı olmasına dikkat edilmiştir (Özçelik, 2013). Deneme testinde öğrencilerin

doğru cevapladığı maddeler dikkate alınarak madde analizi gerçekleştirilmiştir. Öğrencilerin boş bıraktığı veya yanlış yaptığı maddeler dikkate alınmamıştır. Veriler analiz edilmeden önce Microsoft Excel ortamında öğrencilerin verdiği cevapların doğru yanıtları için bir (1), yanlış verdikleri veya boş bıraktıkları yanıtları için sıfır (0) puan verilerek puanlama matrisi hazırlanmıştır (Haladyna, 2011). Her öğrencinin toplam puanı, öğrencinin testteki toplam doğru cevap sayısıdır. Nihai testte de puanlama matrisi aynı yöntemle hazırlanmıştır. Testin nihai formu ileriki çalışmalarda kullanılırken şans başarısından arınmış puanı hesaplamak için üç yanlış bir doğruyu götürecek şekilde puanlama yapılabilir.

4. Testin geçerliğinin sağlanması

Deneme testinde yer alacak maddelerin kapsam geçerliği için uzman görüşüne başvurulmuştur. Kırk sekiz sorudan oluşan madde havuzu için 14 uzmandan görüş alınmıştır (Şekil 1). Uzmanlar, bilgisayar bilimleri alanı ile Bilgisayar ve Öğretim Teknolojileri Eğitimi (BÖTE) alanından 8 öğretim üyesi/görevlisi, ölçme ve değerlendirme alanından 1 öğretim üyesi, Türkçe alanından 1 öğretim üyesi ve 4 bilişim teknolojileri öğretmeninden oluşmaktadır (Tablo 2). Ölçme ve değerlendirme uzmanı test maddelerinin uygunluğunu, Türkçe uzmanı anlam, dil bilgisi ve imla kurallarını, bilgisayar bilimleri ile BÖTE alanı uzmanları ise kazanımları, ilgili soruları ve öğrencilere uygunluğunu incelemiştir. Uzmanlara madde havuzu ve görüşlerini belirteceği uzman görüş formu e-posta yolu ile gönderilmiştir. Uzmanlardan her bir madde için “Uygun”, “Düzeltilmeli” ve “Çıkartılmalı” derecelendirmelerinden birini seçmeleri ve yorum yapmaları istenmiştir (Alisinanoğlu & Şimşek, 2013).

Şekil 1

Kapsam Geçerliğinin Sağlanması için İzlenen Adımlar



Tablo 2

Uzmanların Görevleri ve Uzmanlık Alanları

Uzman	Unvanı	Uzmanlık Alanı
U1	Prof. Dr.	BÖTE/Yazılım Mühendisliği
U2	Doç. Dr.	BÖTE
U3	Dr. Öğr. Üyesi	BÖTE
U4	Dr. Öğr. Üyesi	BÖTE
U5	Dr. Öğr. Üyesi	BÖTE
U6	Dr. Öğr. Üyesi	Ölçme ve Değerlendirme
U7	Dr. Öğr. Gör.	BÖTE/Bilgisayar Programcılığı
U8	Dr. Öğr. Gör.	BÖTE/Bilgisayar Programcılığı
U9	Öğr. Gör.	Bilgisayar Programcılığı
U10	Prof. Dr.	Türkçe
U11	Öğretmen	Bilişim Teknolojileri
U12	Öğretmen	Bilişim Teknolojileri
U13	Öğretmen	Bilişim Teknolojileri
U14	Öğretmen	Bilişim Teknolojileri

Her bir maddenin Kapsam Geçerlik Oranı'nı (KGO) elde etmek için Lawshe (1975) tekniği kullanılmıştır (Formül 1). Ayre ve Scally (2014), kapsam geçerliğinin tespitinde uzmanların sayısını en az 5 ve en fazla 40 olacak şekilde Kapsam Geçerlik Ölçütü'nün (KGÖ) minimum

Asım KANMAZ, Hamide ŞAHİNKAYASI

Developing an achievement test of programming with Blockly for middle school students

değerlerini belirlemiştir (Tablo 3). Testte yer alacak maddelerin kapsam geçerliğini test etmek için, uzman görüşlerinden alınan verilerin KGO ve kapsam geçerlik indeksi (KGİ) oranları hesaplanarak nicel verilere dönüştürülmüş ve KGÖ oranı ile karşılaştırılarak anlamlılık düzeyine bakılmıştır.

Formül 1

KGO Hesaplama Yöntemi (Lawshe, 1975)

$$KGO = N_G / (N/2) - 1$$

N_G : Gerekli diyen uzmanların sayısı, N : Araştırmaya katılan uzmanların toplam sayısı

Tablo 3

0.05 Anlamlılık Düzeyinde KGO'ların Minimum/Kritik Değerleri (Ayre & Scally, 2014, ss. 82)

Uzman Sayısı	Minimum Değer	Uzman Sayısı	Minimum Değer
5	1.000	23	0.391
6	1.000	24	0.417
7	1.000	25	0.440
8	0.750	26	0.385
9	0.778	27	0.407
10	0.800	28	0.357
11	0.636	29	0.379
12	0.667	30	0.333
13	0.538	31	0.355
14	0.571	32	0.375
15	0.600	33	0.333
16	0.500	34	0.353
17	0.529	35	0.314
18	0.444	36	0.333
19	0.474	37	0.297
20	0.500	38	0.316
21	0.429	39	0.333
22	0.455	40	0.300

KGO, (-1) ve (+1) arasında bir değere sahiptir. Formül 1'de görüldüğü gibi uzmanların madde havuzuna verdiği dönütlere göre KGO hesaplanır. Uzmanların yarısı test maddesine "Gerekli" dönütünü veriyse KGO=0, yarısından fazlası "Gerekli" dönütünü veriyse KGO>0, yarısından fazlası "Gerekli" dönütü vermemişse KGO<0 olacaktır. KGO değerinin KGO=0 veya KGO<0 olması test maddesinin kapsam geçerliliğinin olmadığını gösterir (Lawshe, 1975; Ayre & Scally 2014). KGİ değeri, ölçekteki maddelerin kapsam geçerlik oranı değerlerinin ortalamasıdır (Lawshe, 1975).

5. Testin güvenirliği

Testin güvenirliği için Kuder-Richardson-20 (KR-20) analizi yapılmıştır (Kuder & Richardson, 1937). Madde analizi yapıldıktan sonra iç tutarlılığı belirlemek için KR-20 yöntemi uygulanarak güvenirlik katsayısı hesaplanmıştır.

6. Testin uygulanması

Hatay İli Antakya İlçesinde Günyazı Ortaokulundan 8 5. sınıf öğrencisi ile soruların anlaşılabilirliğini kontrol etmek amacıyla 48 soruluk deneme testinin ön uygulaması 2022 Temmuz ayında yapılmıştır. Ön uygulama sonrası, 2022 Aralık ayında Hatay ili Antakya ilçesinde Beyhan Gencay Ortaokulu (n=69) ve İnönü Ortaokulunda (n=82) öğrenim gören 151 6. sınıf öğrencisine deneme testi uygulanmıştır. Birinci araştırmacı dersin öğretmeni ile gözlemci olarak katılım sağlamıştır. Deneme testi, okul sayısının birden fazla olması ve öğrenci sayısı dikkate alınarak kâğıt israfını engellemek ve değerlendirme kolaylığı için Google Formlar üzerinden dijital ortamda öğrencilere uygulanmıştır. Bunun yanı sıra sorularda kalem kullanma gereksinimi olmamasından dolayı dijital ortamda testin uygulanması öğrenciler için sorun yaratmamıştır. Nihai test ise 2023 Mayıs ayında Hatay ili Arsuz ilçesi Madenli Ortaokulunda 72 5. sınıf öğrencisine kâğıt üzerinde uygulanmıştır. Deneme testi için iki ders saati (iki ders arası teneffüsle birlikte toplam 90 dk.), nihai test için bir ders saati (40 dk.) süre verilmiştir.

7. Madde analizi

Madde analizi genellikle maddenin kolaylığı ve buna bağlı olarak güvenirliği, ayırt etme gücü ve buna bağlı olarak geçerliliği ve üçüncü olarak da çeldiricilerin işlerliği ile ilgili bulguları ortaya koymak için yapılır (Özçelik, 2010). Bu bulgular, bir maddenin nihai testte kalıp kalmayacağını, testte kalması halinde testin geçerlik ve güvenirliğine ne derece katkıda bulunabileceğinin ve bu maddenin nasıl geliştirilebileceğinin belirlenmesinde kullanılır.

Elde edilen verilerin madde analizleri MS Excel programında yapılmıştır. Madde analizi, öğrencilerin vermiş olduğu yanıtların ölçülmesi ve yorumlanmasıdır (DeVellis, 2022). Bu çalışmada alt ve üst gruplara göre madde güçlük indeksi ve madde ayıricılık indeksi hesaplanmıştır. Bu yöntemde öğrencilerin verdiği cevaplarda en başarılı olan %27'lik üst öğrenci grubu ile en başarısız olan %27'lik alt öğrenci grubu belirlenerek, toplam öğrenci grubunun %54'lük kısmı madde analizi için kullanılır (Livingstone, Reynolds & Wilson, 2006). Bu yöntemde göre madde güçlük indeksi (Formül 2) ve maddelerin ayırt edicilik indeksi (Formül 3) hesaplanmıştır.

Formül 2

Madde Güçlük Hesabı Formülü

$$P = \frac{(\text{Maddeyi üst grupta doğru cevaplayanların sayısı}) + (\text{Maddeyi alt grupta doğru cevaplayanların sayısı})}{(\text{Üst gruptaki öğrenci sayısı}) + (\text{Alt gruptaki öğrenci sayısı})}$$

Formül 3

Madde Ayırt Edicilik Hesabı Formülü

$$A = \frac{(\text{Maddeyi üst grupta doğru cevaplayanların sayısı}) - (\text{Maddeyi alt grupta doğru cevaplayanların sayısı})}{\text{Grupların herhangi birindeki öğrenci sayısı}}$$

Madde Yanlılığı

Maddelerin yanlılığını cinsiyet bakımından test edebilmek için her bir maddenin Değişen Madde Fonksiyonu (DMF) hesaplanmıştır. DMF; bir grupta cinsiyet, etnik köken gibi farklı kategorilerde bulunan bireylerin test maddelerindeki sorulara verdikleri cevapların kategorik değişkenlere göre analiz edilmesidir (Devine & Raju, 1982). Madde yanlılığını test etmek için maddelerin DMF düzeylerinin belirlenmesi sık kullanılan bir tekniktir (Doğan & Öğretmen, 2008).

DMF sonuçlarını çözümlemek için Lojistik Regresyon (LR) tekniği kullanılmıştır. DMF düzeylerinin belirlenmesinde LR tekniği önerilen bir yöntemdir (Camilli & Shepard, 1994). LR hesaplamalarında R² değeri DMF hakkında çıkarımlar elde edilmesini sağlar (Doğan & Öğretmen, 2008). LR analizinde DMF durumunu ölçmek için genellikle Nagelkerke indeksi

kullanılır (Erdem Keklik, 2012). Bu çalışmada Nagelkerke indeksine göre DMF düzeyi analiz edilmiştir. Zumbo ve Thomas (1997) tarafından önerilen R^2 ölçütleri Tablo 4'te verilmiştir.

Tablo 4

Lojistik Regresyon Tekniğinde DMF Düzeyi için R^2 Ölçütleri

DMF düzeyi	Ölçütler	Yorumlar
A	$R^2 < 0.13$	DMF yoktur.
B	$0.13 \leq R^2 < 0.26$	Orta düzeyde DMF vardır.
C	$R^2 \geq 0.26$	Yüksek düzeyde DMF vardır.

Araştırma Etiği

Bu araştırmanın planlanmasından uygulanmasına, verilerin toplanmasından verilerin analizine kadar olan tüm süreçte “Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi” kapsamında uyulması belirtilen tüm kurallara uyulmuştur. Yönergenin ikinci bölümü olan “Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler” başlığı altında belirtilen eylemlerden hiçbirisi gerçekleşmemiştir.

Bu çalışmanın yazım sürecinde bilimsel, etik ve alıntı kurallarına uyulmuş; toplanan veriler üzerinde herhangi bir tahrifat yapılmamış ve bu çalışma herhangi başka bir akademik yayın ortamına değerlendirme için gönderilmemiştir.

Bu araştırma Hatay Mustafa Kemal Üniversitesi Sosyal ve Beşerî Bilimler Araştırma ve Etik Komisyonu tarafından verilen izinle yürütülmüştür (01.02.2022 tarihli 26 sayılı karar). Hatay İl Milli Eğitim Müdürlüğünden veri toplama izni alınmıştır. Bunların yanı sıra, çalışmaya katılan öğrencilerin velilerine bilgilendirilmiş gönüllü olur/onam formunu gönderilmiş ve onayları imzalı olarak alınmıştır.

Etik kurul izin bilgileri

Etik değerlendirmeyi yapan kurul adı: Hatay Mustafa Kemal Üniversitesi Sosyal ve Beşerî Bilimler Araştırma ve Etik Komisyonu

Etik değerlendirme karar tarihi: 01.02.2022

Etik değerlendirme belgesi sayı numarası: 26

BULGULAR

Bu bölümde BTY dersi 5. sınıf PÇP ünitesinin programlama ile ilgili kazanımlarını BBTPU bağlamında yoklayan başarı testinin geçerlik ve güvenilirliğine yönelik analiz sonuçları yer almaktadır.

Kapsam Geçerliği

Uzmanların madde havuzundaki sorulara dönütleri Tablo 5'te sunulmuştur. KGO değerleri hesaplandıktan sonra KGİ değeri hesaplanmıştır. Ayre ve Scally (2014)'nin bildirdiği KGÖ değerlerine bakıldığında (Tablo 3), 14 uzmanın KGÖ değeri için anlamlılık düzeyi 0.57'dir. KGİ değeri KGÖ değerinden büyük ise ölçekte kalan maddelerin kapsam geçerliğine sahip olduğu sonucuna varılır (Lawshe, 1975). Bu çalışmada, KGİ değerinin KGÖ değerinden büyük olduğu görülmektedir. Sonuç olarak, deneme uygulamasında elde edilen değerlerden KGİ (0.92), KGÖ (0.57) değerinden büyük olduğundan deneme testinin (48 madde) kapsam geçerliği istatistiksel olarak anlamlı düzeydedir.

Tablo 5'te görüldüğü gibi uzmanlardan toplam 18 madde için “düzeltilmeli” görüşü gelmiş, gerekçelerine göre bu maddeler iyileştirilerek deneme testi oluşturulmuştur. Herhangi bir maddenin çıkarılması önerilmediği için deneme testi toplam 48 maddeden oluşturulmuştur.

Tablo 5*Uzman Görüşlerine Göre Deneme Formundaki Maddelerin Kapsam Geçerlik Oranları ve İndeksi*

Madde No.	Uygun	Düzeltilmeli	Çıkartılmalı	KGO	Madde	Uygun	Düzeltilmeli	Çıkartılmalı	KGO
M1	12	2	0	0.71	M25	13	1	0	0.85
M2	13	1	0	0.85	M26	14	0	0	1.0
M3	14	0	0	1.0	M27	14	0	0	1.0
M4	14	0	0	1.0	M28	14	0	0	1.0
M5	13	1	0	0.85	M29	14	0	0	1.0
M6	14	0	0	1.0	M30	14	0	0	1.0
M7	13	1	0	0.85	M31	14	0	0	1.0
M8	14	0	0	1.0	M32	14	0	0	1.0
M9	14	0	0	1.0	M33	14	0	0	1.0
M10	14	0	0	1.0	M34	14	0	0	1.0
M11	14	0	0	1.0	M35	13	1	0	0.85
M12	14	0	0	1.0	M36	14	0	0	1.0
M13	11	3	0	0.57	M37	14	0	0	1.0
M14	12	2	0	0.71	M38	14	0	0	1.0
M15	13	1	0	0.85	M39	14	0	0	1.0
M16	13	1	0	0.85	M40	14	0	0	1.0
M17	11	3	0	0.57	M41	14	0	0	1.0
M18	13	1	0	0.85	M42	13	1	0	0.85
M19	13	1	0	0.85	M43	14	0	0	1.0
M20	14	0	0	1.0	M44	14	0	0	1.0
M21	13	1	0	0.85	M45	14	0	0	1.0
M22	13	1	0	0.85	M46	14	0	0	1.0
M23	13	1	0	0.85	M47	14	0	0	1.0
M24	13	1	0	0.85	M48	14	0	0	1.0
Toplam Uzman Sayısı				14					
Kapsam Geçerlik Ölçütü (KGÖ)				0.57					
Kapsam Geçerlik İndeksi (KGI)				0.92					

Madde Analizi**Deneme testi madde analizi**

Deneme testi uygulandıktan sonra öğrenci yanıtları doğru ise 1, yanlış ise 0 olarak puanlanmıştır. Madde bazlı puanlamalardan her öğrencinin toplam puanı hesaplanmıştır. Toplam puanlar MS Excel yazılımında büyükten küçüğe doğru sıralanarak %27'lik üst grup ve %27'lik alt grup belirlenmiş, bu iki grup analize dahil edilmiştir. Bu grupların verileri üzerinde madde güçlük indeksi Formül 2; madde ayıricılık indeksi Formül 3 ile hesaplanmıştır. Deneme testi madde güçlük indeksi ve madde ayıricılık indeksi sonuçları Tablo 6'da sunulmuştur.

Deneme testindeki maddelerin ayırt ediciliği incelenmiştir. Deneme testinin madde ayıricılık indeksi ortalaması 0.52 olarak hesaplanmıştır. Tekin (2019) bir testte ortalama ayırt edicilik indeksinin 0.30'un üstünde olmasının yeterli olduğunu ifade etmiştir. Özçelik (2013) tarafından bildirilen ölçütlere göre deneme testinin ortalama madde ayıricılığının çok iyi olduğu tespit edilmiştir (Tablo 7). Maddelerin madde ayıricılık indeksi 0.29 ve 0.71 değerleri arasında değişmektedir (Tablo 6). Deneme testinin maddeleri, madde ayıricılık indekslerine göre

sınıflandırılmıştır (Tablo 8). Deneme testinde madde ayıricılık indeksi 0.40 değerine eşit veya üstünde olan 42 madde, 0.30 ve 0.40 değerleri arasında 5 madde, 0.20 ve 0.29 değerleri arasında bulunan 1 madde mevcuttur. Madde ayıricılık indeksi 0.29 değerinin altında olan madde bulunmamaktadır (Tablo 8).

Tablo 6*Deneme Testinin Madde Güçlük ve Madde Ayırt Edicilik İndeksi Sonuçları*

Madde	Madde Güçlük İndeksi	Madde Ayıricılık İndeksi	Madde	Madde Güçlük İndeksi	Madde Ayıricılık İndeksi
M1	0.70	0.46	M25	0.63	0.39
M2	0.51	0.63	M26	0.50	0.46
M3	0.52	0.51	M27	0.68	0.39
M4	0.54	0.49	M28	0.56	0.54
M5	0.55	0.61	M29	0.65	0.71
M6	0.57	0.56	M30	0.30	0.51
M7	0.39	0.68	M31	0.68	0.63
M8	0.56	0.59	M32	0.44	0.54
M9	0.70	0.46	M33	0.61	0.44
M10	0.59	0.59	M34	0.62	0.56
M11	0.55	0.66	M35	0.67	0.41
M12	0.52	0.71	M36	0.61	0.49
M13	0.76	0.49	M37	0.68	0.54
M14	0.74	0.51	M38	0.65	0.61
M15	0.78	0.44	M39	0.73	0.54
M16	0.70	0.61	M40	0.62	0.61
M17	0.71	0.34	M41	0.67	0.51
M18	0.74	0.51	M42	0.44	0.54
M19	0.80	0.39	M43	0.56	0.63
M20	0.77	0.41	M44	0.35	0.56
M21	0.60	0.61	M45	0.52	0.41
M22	0.59	0.49	M46	0.34	0.29
M23	0.71	0.49	M47	0.48	0.32
M24	0.60	0.61	M48	0.33	0.41
				Güçlük İndeksi	Ayıricılık İndeksi
Ortalama				0.59	0.52

Tablo 7’de görüldüğü gibi madde ayıricılık indeksi 0.20 ile 0.29 arasında olan maddeler testte zorunlu ise kullanılabilir niteliktedir. Ayıricılığı 0.30 ile 0.40 arasında olan maddeler iyi, ayıricılığı 0.40 değerinden daha yüksek olan maddeler ise çok iyi olarak değerlendirilir (Özçelik, 2013). Madde analizi sonucunda ayırt edicilik indeksi 0.40 ve daha büyük olan maddeler ile 0.29-0.39 değerleri arasında olan maddeler nihai teste seçilmiştir.

Tablo 7*Madde Ayırtıcılık İndeksi ve Madde Güçlük İndeksi Ölçütü*

Madde Ayırtıcılık İndeksi Ölçütü (Özçelik, 2013)	0.40 ve üstü	Çok iyi
	0.30 ile 0.40	İyi
	0.20 ile 0.29	Kullanılabilir
	0.19 ve altı	Revize edilmeli
Madde Güçlük İndeksi Ölçütü (Başol, 2015)	0.85 - 1	Çok Kolay (testten çıkarılmalı)
	0.61- 0.84	Kolay
	0.40 - 0.60	Orta güçlükte
	0.16 - 0.39	Zor
	0.15 - 0.00	Çok zor (testten çıkarılmalı)

Deneme testinin madde güçlük indeksi ortalaması 0.59 olarak hesaplanmıştır (Tablo 6). Başol (2015)'e göre deneme testinin ortalama güçlük düzeyinin orta olduğu söylenebilir (Tablo 7). Başarı testlerinde testin ortalama güçlüğü 0.50 civarında olması istenilen bir durumdur (Bayrakçeken, 2012; Turgut & Baykul, 2014; Yılmaz, 2012). Özçelik (2010) her bir madde için madde güçlüğü 0.50 dolaylarında toplanmak üzere 0.20 ile 0.80 arasında olmasının kabul edilebilir olduğunu ifade eder. Tablo 6'da görüldüğü gibi deneme testinin madde güçlük indeksi 0.30 ve 0.80 değerleri arasında değişmektedir. Bir teste kolay, orta ve zor maddelerin bulunması, madde güçlüğü indeksi ortalamasının 0.50 düzeyinde olması koşuluyla istenen bir durumdur (Hasançebi, Terzi & Küçük, 2020).

Tablo 8*Deneme Testinin Madde Ayırtıcılık İndeksi Kategorileri (Özçelik, 2013)*

Madde Ayırtıcılık İndeksi	Kategori	Madde Sayısı	Maddeler
0.40 ve üstü	Çok iyi	42	M1, M2, M3, M4, M5, M6, M7, M8, M9, M10, M11, M12, M13, M14, M15, M16, M18, M20, M21, M22, M23, M24, M26, M28, M29, M30, M31, M32, M33, M34, M35, M36, M37, M38, M39, M40, M41, M42, M43, M44, M45, M48
0.30-0.40	İyi	5	M17, M19, M25, M27, M47
0.20-0.29	Zorunlu ise kullanılmalı	1	M46

Tablo 9*Nihai Teste Seçilen Maddelerin Deneme Testi Sonuçlarının Analizi*

Soru Sıralaması	Deneme testinde madde no.	Madde Güçlük İndeksi	Madde Ayırıcılık İndeksi	Kazanımlar
1	M2	0.51	0.63	BT.5.5.1.6. Problemi çözmek için gerekli değişken, sabit ve işlemleri açıklar
2	M3	0.52	0.51	
3	M5	0.55	0.61	BT.5.5.1.7. Problem çözümünde kullanılabilecek operatörlere örnek verir
4	M7	0.39	0.68	
5	M11	0.55	0.66	BT.5.5.1.8. Problem çözümünde ifade ve eşitliklere örnek verir
6	M12	0.52	0.71	
7	M43	0.56	0.63	BT.5.5.1.10. Verilen bir problemin çözümünde operatörleri kullanır
8	M44	0.35	0.56	
9	M37	0.68	0.54	BT.5.5.1.11. Verilen bir problemde ifade ve eşitlikleri kullanarak çözüm üretir
10	M48	0.33	0.41	
11	M38	0.65	0.61	BT.5.5.1.17. Matematik ve bilgisayar bilimi arasındaki ilişkiyi tespit eder
12	M39	0.73	0.54	
13	M14	0.74	0.51	BT.5.5.2.4. Doğrusal mantık yapısını açıklar
14	M16	0.70	0.61	
15	M18	0.74	0.51	BT.5.5.2.5. Doğrusal mantık yapısını kullanan algoritmalar geliştirir
16	M20	0.77	0.41	
17	M21	0.60	0.61	BT.5.5.2.6. Karar yapısını ve işlevlerini açıklar
18	M22	0.59	0.49	
19	M28	0.56	0.54	BT.5.5.2.7. Karar yapıları içeren algoritmalar geliştirir
20	M40	0.62	0.61	
21	M24	0.60	0.61	BT.5.5.2.8. Döngü yapısını ve işlevlerini açıklar
22	M29	0.65	0.71	
23	M31	0.68	0.63	BT.5.5.2.9. Döngü yapısı içeren algoritmalar oluşturur
24	M34	0.62	0.56	
	Ortalama	0.59	0.58	

Denenmiş maddelerden nihai test oluşturulması sırasında yapılacak iş yoklanacak tüm davranışları içermesi koşulu ile mümkün olduğunca ayırıcı gücü en yüksek olan maddeleri seçmektir (Özçelik, 2013). Bu bağlamda her kazanım için iki madde seçilerek başarı testinin 24 maddelik son hali düzenlenmiştir. Her bir kazanım için madde ayırt ediciliği en yüksek olan iki madde alınmış ve diğer iki madde testten çıkarılmıştır (Tablo 9).

Deneme testinden elde edilen verilere göre nihai test için seçilen 24 maddenin deneme uygulamasındaki ayırt edicilik indeksi 0.41 ile 0.71; madde güçlük indeksi ise 0.33 ile 0.77 arasındadır. Ortama madde ayırıcılık indeksi 0.58 ve ortalama madde güçlük indeksi ise 0.59 olarak hesaplanmıştır (Tablo 9). Seçilen maddelerin ortalama madde güçlüğü deneme testinin tüm maddelerinin güçlük ortalamasıyla aynı kalırken madde ayırıcılık indeksi ortalaması deneme testindeki tüm maddelerin ayırıcılık ortalamasına göre 0.06 puan yüksektir. Bu artış, nihai forma seçilen maddelerin ayırıcılığının seçilmeyenlerden daha yüksek olmasından kaynaklanmaktadır.

Deneme Testinin Madde Yanlılığı Testi

Deneme testinin sonuçlarında her maddenin cinsiyete göre DMF gösterip göstermediğinin belirlenmesi için LR tekniği kullanılmıştır. Maddelerin LR tekniği sonucu elde edilen Nagelkerke R^2 değer farkları (ΔR^2) Tablo 10'da verilmiştir. Değerlerin $\Delta R^2=0.000$ ve $\Delta R^2=0.039$ arasında

değiştirdiği gözlemlenmiştir. Zumbo ve Thomas (1997) tarafından belirlenen ölçütlere göre (Tablo 4) deneme testindeki maddelerin DMF düzeyinin A seviyesinde olduğu, başka bir deyişle hiçbirinde DMF görülmediği sonucuna varılmıştır. Bu sonuç deneme testindeki maddelerin hiçbirinin kızlar ya da erkekler lehine anlamlı bir yanlılığı olmadığını ortaya koymaktadır.

Tablo 10

Deneme Testi için LR Yöntemi ile DMF Analizi Sonuçları

No	ΔR^2	No	ΔR^2	No	ΔR^2	No	ΔR^2
1	.003	13	.010	25	.039	37	.013
2	.011	14	.001	26	.007	38	.010
3	.001	15	.029	27	.001	39	.002
4	.002	16	.006	28	.025	40	.000
5	.005	17	.000	29	.000	41	.001
6	.002	18	.011	30	.001	42	.009
7	.013	19	.008	31	.013	43	.024
8	.004	20	.000	32	.000	44	.010
9	.002	21	.007	33	.002	45	.001
10	.001	22	.002	34	.001	46	.001
11	.001	23	.000	35	.000	47	.008
12	.001	24	.015	36	.005	48	.010

Nihai Test Madde Analizi

Deneme testinden seçilmiş olan 24 maddeden oluşan nihai test, Arsuz Madenli Ortaokulunda 72 5. sınıf öğrencisine ünite işlendikten hemen sonra uygulanmıştır. Nihai test Ek 1’de verilmiştir.

Tablo 11’de gösterilen nihai test uygulama sonuçlarına göre; madde güçlük indeksi aralığı 0.32 ile 0.75 değerleri arasındadır. Madde güçlük indeksi ortalaması 0.63 olarak hesaplanmıştır. Madde güçlük indeksinin ortalama değeri, konuyu yeni işlenmiş öğrenciler için, testin ortalama güçlük düzeyinin kolay olduğuna işaret etmektedir (Başol, 2015) (Tablo 11).

Nihai test uygulaması verilerinde madde ayırıcılık indeksi ise 0.30 ile 0.75 değerleri arasındadır (Tablo 11). Testin madde ayırıcılık indeksi ortalaması 0.48’dir. Ayırıcılık indeksi 0.20 ile 0.30 arasında olan maddeler zorunlu ise testte kullanılmalıdır. Ayırıcılığı 0.30 ile 0.40 arasında olan maddeler iyi, ayırıcılığı 0.40’tan daha yüksek olan maddeler ise çok iyi sayılabilir (Özçelik, 2013). Tablo 12’de maddeler ayırıcılık indekslerine göre sınıflandırılmıştır. Nihai test maddelerinin ayırıcılık indekslerinin iyi veya çok iyi, ortalamasının ise çok iyi olduğu sonucuna varılmıştır.

Tablo 11*Nihai Test Uygulamasının Madde Analizi*

Soru No.	Deneme Testindeki Madde No.	Madde Güçlük İndeksi	Madde Ayırıcılık İndeksi	Kazanımlar
1	M2	0.67	0.51	BT.5.5.1.6. Problemi çözmek için gerekli değişken, sabit ve işlemleri açıklar.
2	M3	0.32	0.39	
3	M5	0.50	0.62	BT.5.5.1.7. Problem çözümünde kullanılabilecek operatörlere örnek verir.
4	M7	0.50	0.63	
5	M11	0.60	0.75	BT.5.5.1.8. Problem çözümünde ifade ve eşitliklere örnek verir.
6	M12	0.63	0.53	
7	M43	0.74	0.52	BT.5.5.1.10. Verilen bir problemin çözümünde operatörleri kullanır.
8	M44	0.64	0.56	
9	M37	0.69	0.47	BT.5.5.1.11. Verilen bir problemde ifade ve eşitlikleri kullanarak çözüm üretir.
10	M48	0.75	0.36	
11	M38	0.74	0.35	BT.5.5.1.17. Matematik ve bilgisayar bilimi arasındaki ilişkiyi tespit eder.
12	M39	0.75	0.30	
13	M14	0.75	0.50	BT.5.5.2.4. Doğrusal mantık yapısını açıklar
14	M16	0.69	0.35	
15	M18	0.72	0.48	BT.5.5.2.5. Doğrusal mantık yapısını kullanan algoritmalar geliştirir.
16	M20	0.75	0.43	
17	M21	0.71	0.47	BT.5.5.2.6. Karar yapısını ve işlevlerini açıklar.
18	M22	0.65	0.52	
19	M28	0.58	0.53	BT.5.5.2.7. Karar yapıları içeren algoritmalar geliştirir.
20	M40	0.68	0.52	
21	M24	0.51	0.44	BT.5.5.2.8. Döngü yapısını ve işlevlerini açıklar.
22	M29	0.57	0.48	
23	M31	0.57	0.56	BT.5.5.2.9. Döngü yapısı içeren algoritmalar oluşturur.
24	M34	0.45	0.39	
	Ortalama	0.63	0.48	

Tablo 12*Nihai Testin Madde Ayırıcılık İndeksi Kategorileri (Özçelik, 2013)*

Madde Ayırıcılık İndeksi	Kategori	Madde Sayısı	Maddeler
0.40 ve üstü	Çok iyi	18	M2, M5, M7, M11, M12, M43, M44, M37, M14, M18, M20, M21, M22, M28, M40, M24, M29, M31
0.30-0.40	İyi	6	M3, M48, M38, M39, M16, M34

Nihai Testin Madde Yanlılığı Testi

Nihai testin uygulanması sonrası elde edilen veriler üzerinde her maddenin cinsiyete göre DMF gösterip göstermediğinin belirlenmesi için LR tekniği kullanılmıştır. Nihai testin sonuçlarından alınan maddelerin LR tekniği sonucu elde edilen Nagelkerke R^2 değer farkları (ΔR^2) Tablo 13'te görülmektedir. Değerlerin, $\Delta R^2=0.000$ ve $\Delta R^2=0.052$ arasında değiştiği gözlemlenmiştir. Zumbo ve Thomas (1997) tarafından belirlenen ölçütlere (Tablo 4) göre maddelerin DMF değer aralığı A

düzeyinde olup DMF görülen madde bulunmamaktadır. Bu sonuç, testin nihai formunda da maddelerin cinsiyet açısından yanlılık göstermediğine işaret etmektedir.

Tablo 13

Nihai test için LR Yöntemi ile DMF Analizi Sonuçları

No	ΔR^2	No	ΔR^2
1	.001	13	.002
2	.048	14	.001
3	.016	15	.001
4	.016	16	.002
5	.012	17	.006
6	.003	18	.003
7	.052	19	.003
8	.002	20	.023
9	.001	21	.009
10	.014	22	.002
11	.000	23	.001
12	.014	24	.000

KR-20 İç Tutarlık Analizi

Testte ölçülmek istenilen özelliğin kararlı bir biçimde ölçülmesine ölçme aracının güvenilirliği denir. Güvenirlik göstergesi değerleri 0 ile 1 arasındadır. Güvenirlik değerinin 1'e yaklaşması testin güvenilirliğinin yüksek olduğu anlamına gelmektedir (Özçelik, 2013). Testin güvenilirlik aralığı, maddeler farklı güçlük indekslerine sahip olduğu için, KR-20 yöntemiyle hesaplanmıştır (Kuder & Richardson, 1937). Deneme uygulaması için oluşturulan deneme testinin KR-20 puanı 0.93 olarak hesaplanmıştır. Nihai testin uygulandığı grupta kullanılan 24 maddelik nihai testin KR-20 puanı ise 0.86 olarak hesaplanmıştır. Güvenirlik katsayısı değerinin 0.80 ve 1.00 arasında olması testin yüksek derecede güvenilir olduğunu, 0.60 ve 0.80 arasında olması oldukça güvenilir olduğunu, 0.60 ve altında olması ise güvenilirliğin düşük olduğunu gösterir (Kalaycı, 2008). Bu durumda KR-20'nin 0.86 olması, nihai testin yüksek derecede güvenilir olduğunu gösterir.

TARTIŞMA ve SONUÇ

Yapılan çalışmada 5. sınıf PÇP ünitesinin programlama ile ilişkili kazanımlarını BBTPU bağlamında ölçmeye yönelik çoktan seçmeli maddelerden oluşan geçerli ve güvenilir bir başarı testi geliştirilmesi amaçlanmıştır. Test geliştirme sürecinde, testin amacının ve kapsamının belirlenmesi, madde havuzunun oluşturulması, ölçme biçiminin belirlenmesi, testin geçerliğinin sağlanması, oluşturulan testin uygulanması, testin güvenilirliğinin belirlenmesi ve madde analizlerinin tamamlanması basamakları izlenerek teste son hali verilmiştir (Özçelik, 2013).

5. sınıf PÇP ünitesinin programlama ile ilgili kazanımlarını ölçmek testin amacı olarak belirlenmiş, belirtke tablosu hazırlanmış ve madde havuzu oluşturulmuştur. Uzmanlardan alınan dönütlere göre başarı testinin kapsam geçerliği incelenmiştir. Başarı testinin kapsam geçerlik indeksi ($KGİ=0.92$), kapsam geçerlik ölçütünden ($KGÖ=0.57$) büyük olduğundan madde havuzundaki maddelerin (48 madde) kapsam geçerliliği istatistiksel olarak anlamlı düzeydedir. Çapraz ve Yeşilyurt (2018)'in çalışmasında da kapsam geçerliliği için aynı yöntem izlenmiştir. Çalışma kapsamında oluşturulan ve uzman görüşlerine göre iyileştirilen 48 maddelik deneme testi ön uygulama için 8 öğrenciye uygulanmıştır. Deneme uygulaması ise daha önceden BTY dersi alan 6. sınıf öğrencilerine uygulanmıştır. Her kazanım için madde ayırt edicilik indeksi en büyük olan iki madde teste dâhil edilerek başarı testinin 24 maddelik son hali düzenlenmiş ve nihai test olarak 72 5. sınıf öğrencisine uygulanmıştır.

Asım KANMAZ, Hamide ŞAHİNKAYASI

Developing an achievement test of programming with Blockly for middle school students

Deneme testi analizi sonrası elde edilen 24 maddelik testin ortalama madde güçlük indeksi 0.59 iken nihai uygulamada testin ortalama madde güçlük indeksinin 0.63'e yükseldiği görülmüştür. Nihai uygulama test maddeleri deneme testindeki madde güçlük indeksleriyle karşılaştırıldığında, iki test arasında 13 sorunun (3, 5, 6, 9, 12, 13, 14, 15, 16, 17, 18, 19, 21. maddeler) madde güçlük indeksi seviyesi Başol (2015) tarafından bildirilen ölçütlere göre değişmezken; 6 sorunun (1, 7, 8, 11, 20, 22. maddeler) orta veya zor seviyeden kolay seviyeye düştüğü, bir sorunun ise (4. madde) zor seviyeden orta seviyeye düştüğü görülmüştür. İki sorunun (2, 10. maddeler) orta seviyeden zor seviyeye çıktığı, 2 sorunun (23, 24. maddeler) güçlük indeksinin ise kolay seviyeden orta seviyeye çıktığı sonucuna varılmıştır. Maddelerin neredeyse üçte birinin nihai testte öğrencilere daha kolay gelmesi, nihai formun ünite kazanımların işlenmesinin hemen ardından 5. sınıflara uygulanmasından, öte yandan deneme testi uygulanmış olan 6. sınıf öğrencilerinin Blockly uygulaması çalışmalarını bir önceki sene yapmış olmasından dolayı konuyu tam olarak hatırlayamamalarından kaynaklanmış olabilir. Madde güçlüğü nihai testte artan dört madde ise deneme testinin almış olan 6. sınıfların nihai formun uygulandığı 5. sınıflara göre olgunluk düzeylerinin daha fazla olması, soruların 5. sınıf düzeyi için zor olması ile açıklanabilir.

Nihai testin ortalama ayırt edicilik indeksinin 0.48 olduğu görülmüştür. Bu sonuç, testin ayırt edicilik düzeyinin deneme testinde aynı maddelerin ayırt edicilik ortalamasına göre 0.10 puan düşük olsa da çok iyi düzeyde olduğunu gösterir (Bolat & Karamustafaoğlu, 2019; Hasançebi, Terzi & Küçük, 2020; Özçelik, 2013). Nihai testin ortalama madde güçlük indeksi ise 0.63 olarak hesaplanmıştır. Başarı testinde bu değer 0.50 civarı olması beklenmektedir (Tekin, 2019). Nihai testin ortalama madde güçlüğüne kolay olduğu sonucuna varılabilir (Başol, 2015). Bu sonuç nihai formun uygulandığı gruba PÇP ünitesi işlenir işlenmez testin uygulanmış olmasından kaynaklanmış olabilir.

Başarı testi nihai formunun KR-20 güvenirlik katsayısı 0.86 olarak hesaplanmıştır. Bu değerinin 0.70 değerinden büyük olması, nihai testin yüksek derecede güvenilir bir test olduğunu gösterir (Kehoe, 1995). İlgili alan yazınında yapılan çalışmalar incelendiğinde, geliştirilen testler için benzer sonuçlara ulaşanlar mevcuttur. Örneğin Çiçek Şentürk ve Selvi (2021) geliştirdikleri başarı testinde, testin son hali için ortalama madde güçlüğü 0.62; ortalama madde ayırt ediciliğini 0.47; KR-20 güvenirlik katsayısını 0.82 olarak hesaplamıştır. Ortaokul öğrencileri için diğer derslere yönelik hazırlanan başarı testlerinde de ortalama madde güçlük ve ayırtıcılık indekslerinin, alan yazını ölçütlerine uygun olarak yapılan bu çalışma ile benzer değerlerde olduğu sonucuna varılmıştır (Kargın & Gül, 2021; Saraç, 2018).

Yapılan bu çalışmada, 5. sınıf PÇP ünitesinin programlama ile ilgili kazanımlarını Blockly uygulaması bağlamında ölçmeye yönelik çoktan seçmeli maddelerin yer aldığı 24 maddelik güvenilir ve geçerli bir başarı testi hazırlanmıştır. Başarı testinin ortalama madde güçlüğü indeksinin 0.63 olması, ortalama ayırt edicilik indeksinin 0.48 olması, KR-20 değerinin 0.86 olması testin akademik çalışmalarda kullanılabileceğini göstermektedir. Bu başarı testi, 5. sınıf BTY dersi PÇP ünitesinin programlamayla ilgili kazanımlarının BBTPU üzerinden işlendiğinde kâğıt kalemle veya elektronik ortamda bu uygulama bağlamında ölçülmesini sağlayarak alan yazınına katkı sağlayacaktır.

Araştırmanın Sınırlılıkları

Araştırmanın sonuçları değerlendirilirken dikkate alınması gereken bazı sınırlılıkları vardır. Çalışmanın deneme testinin 151 6. sınıf öğrencisine uygulanması, nihai testin ise 72 5. sınıf öğrencisine uygulanması örneklem grubunun sınırlılıkları olarak değerlendirilebilir. Çalışmada testin deneme uygulaması PÇP ünitesini bir önceki yıl işlemiş olan 6. sınıflarda, nihai formun uygulaması ise ilgili üniteyi henüz işlemiş olan 5. sınıflarda yapılmıştır. Kahramanmaraş merkezli 6 Şubat depremleri nedeniyle nihai formun uygulaması başka bir ilçede ve üç şubeden toplam 72 öğrencinin katılımı ile gerçekleştirilmiştir. Deneme testinin çevrimiçi olarak Google Formlar üzerinden, nihai formun ise kâğıt kalemle uygulanmış olması çalışmanın bir diğer

sınırlılığıdır. Geliştirilen başarı testinin daha büyük ve farklı bölgelerdeki örneklemeler üzerinde uygulanması araştırmacılara önerilebilir.

Destek ve Teşekkür

Bu makale, birinci yazarın ikinci yazar danışmanlığında hazırlamakta olduğu “Bilgisayarsız Ortamda Blok Temelli Programlama Söz Dizim Yapısını Öğrenmeyi Destekleyen Öğretim Materyali Geliştirilmesi ve Etkililiğinin İncelenmesi” başlıklı doktora tezinin başarı testi geliştirme kısmından üretilmiştir.

Finansal Destek

Çalışmanın üretildiği doktora tezi Hatay Mustafa Kemal Üniversitesi BAP Koordinatörlüğü tarafından desteklenmiştir (Proje No: 22.D.006, Proje Adı: Bilgisayarsız Ortamda Blok Temelli Programlama Söz Dizim Yapısını Öğrenmeyi Destekleyen Öğretim Materyali Geliştirilmesi ve Etkililiğinin İncelenmesi, proje bitiş tarihi: 28.03.2025).

Araştırmacıların Katkı Oranı

Araştırmanın birinci yazarı madde havuzunun oluşturulması, okullarda uygulama sürecinin takip edilmesi, madde analizi ve kapsam geçerliği analizinin yapılması ve makale taslağının oluşturulmasında etkin rol almıştır. İkinci yazar ise araştırmanın planlaması, araştırma sürecinin yürütülmesi, uygulama ve etik kurul izin süreçleri, madde havuzunun incelenmesi ve iyileştirilmesi, madde analizleri ve güvenilirlik analizlerinin yorumlanması, sonuçların raporlanması ve makale yazımının iyileştirilmesi görevlerini üstlenmiştir.

Çatışma Beyanı

Araştırmanın yazarları olarak herhangi bir çıkar/çatışma beyanımız olmadığını ifade ederiz.

Ölçek İzni

Bu çalışma kapsamında geliştirilen “Blockly ile Programlama Başarı Testi” Ek. 1’de sunulmuştur. Araştırmacılar başarı testini ekte sunulan haliyle, metin içinde uygun atıf kullanarak ve kaynakçada kaynak olarak bu yayına yer vererek, ticari hedef gütmeyen bilimsel araştırmalar kapsamında ayrıca herhangi bir izin süreci gerçekleştirilmeden, kullanabilirler.

Yayın Etiği Beyanı

Bu araştırmanın planlanmasından, uygulanmasına, verilerin toplanmasından verilerin analizine kadar olan tüm süreçte “Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi” kapsamında uyulması belirtilen tüm kurallara uyulmuştur. Yönergenin ikinci bölümü olan “Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler” başlığı altında belirtilen eylemlerden hiçbirini gerçekleştirilmemiştir.

Bu çalışmanın yazım sürecinde bilimsel, etik ve alıntı kurallarına uyulmuş; toplanan veriler üzerinde herhangi bir tahrifat yapılmamış ve bu çalışma herhangi başka bir akademik yayın ortamına değerlendirme için gönderilmemiştir.

Katılımcılara çalışma hakkında bilgi verilmiş ve bilgilendirilmiş gönüllü olur/onam formu imzalatılmıştır.

Etik kurul izin bilgileri

Etik değerlendirmeyi yapan kurul adı: Hatay Mustafa Kemal Üniversitesi Sosyal ve Beşerî Bilimler Araştırma ve Etik Komisyonu

Etik değerlendirme karar tarihi: 01.02.2022

Etik değerlendirme belgesi sayı numarası: 26

KAYNAKÇA

- Açıkgöz, M., & Karşı, F. (2015). Alternatif ölçme-değerlendirme yaklaşımları kullanılarak iş ve enerji konusunda geliştirilen başarı testinin geçerlilik ve güvenilirlik analizi. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, 4(1), 1-25. <https://doi.org/10.17539/aej.71851>
- İpek Akbulut, H., & 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.
- Akdeniz, A. R., Öztürk, M. & Bakırcı, H. (2017). Bilgisayar destekli öğretim uygulamalarının sekizinci sınıf öğrencilerinin fen dersi akademik başarılarına ve bilginin kalıcılığına etkisi. *HAYEF: Journal of Education*, 14(2), 59-77. <http://dx.doi.org/10.26650/hayef.2017.14.2.0007>
- Alagöz, N. B., & Bilgiç, H. G. (2023). Erken yaşta programlama öğretiminde metafor kullanımı: Metaforlarla programlama öğreniyorum etkinlikleri. *Eğitim Teknolojisi Kuram ve Uygulama*, 13 (2), 388-414. <https://10.17943/etku.1229601>
- Aydın, S. (2019). *Algoritma Öğretiminde Scratch Kullanımının Ortaokul Öğrencileri Üzerindeki Başarıya Etkisi ve Öğrenci Görüşleri* [Yayımlanmamış Yüksek Lisans Tezi]. Uludağ Üniversitesi.
- Arslan, K., & Akçelik, M. (2019). Programlama eğitiminde Scratch'in kullanılması: Öğretmen adaylarının tutum ve algıları. *Ulusal Eğitim Akademisi Dergisi (UEAD)*, 3(1), 41-61. <https://doi.org/10.32960/uead.455502>
- Ayre, C., & Scally A. J. (2014). Critical values for Lawshe's content validity ratio: revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, 47 (1), 79-86. <https://10.1177/0748175613513808>
- Alisinanoğlu, F., & Şimşek, Ö. (2013). Okul öncesi dönemdeki çocukların yazmaya hazırlık becerilerini değerlendirme kontrol listesinin geçerlik ve güvenilirlik çalışması. *Kastamonu Eğitim Dergisi*, 21(3), 1163-1176.
- Aytekin, A., Sönmez Çakır, F., Yücel, Y. B., & Kulaöz, İ. (2018). Geleceğe yön veren kodlama bilimi ve kodlama öğrenmede kullanılabilecek bazı yöntemler. *Avrasya Sosyal ve Ekonomi Araştırmaları Dergisi*, 5(5), 24-41.
- Balcı, A. (2011). *Sosyal bilimlerde araştırma: Yöntem, teknik ve ilkeler* (9.baskı). Pegem Akademi.
- Bakırcı, F. (2019). *Blok tabanlı programlama aracının 6. sınıf öğrencilerinin programlama başarıları, algoritma geliştirme öz-yeterlilikleri ve güdülenmelerine etkisi* [Yayımlanmamış Yüksek Lisans Tezi]. Sakarya Üniversitesi.
- Başol, G. (2015). *Eğitimde Ölçme ve Değerlendirme*. Pegem Akademi Yayıncılık
- Bayrakçeken, S. (2012). Test Geliştirme. E. Karip (Ed.) İçinde, *Ölçme ve Değerlendirme* (5. Baskı) (ss. 294-324). Pegem Akademi.
- Best, J. W., & Kahn, J. V. (2006). *Research in Education*. Pearson Education India.
- Bolat, A., & Karamustafaoğlu, S. (2019). "Vücudumuzdaki Sistemler" ünitesi başarı testi geliştirme: Geçerlik ve güvenilirlik. *Gazi Eğitim Bilimleri Dergisi*, 5 (2), 131-159. <https://dx.doi.org/10.30855/gjes.2019.05.02.008>
- Camilli, G. & Shepard, L.A. (1994). *Methods for Identifying Biased Test Items*. Sage Publications, California.
- Can Şen, H., & Eryılmaz, A. (2011). Bir başarı testi geliştirme çalışması: basit elektrik devreleri başarı testi geçerlik ve güvenilirlik araştırması. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 8 (1), 1-39.
- Creswell, J. W. (2017). *Eğitim araştırmaları: Nicel ve nitel araştırmanın planlanması, yürütülmesi ve değerlendirilmesi*. H. Ekşi (Çev. Ed.). Edam Yayınevi.
- Çalık, M., & Ayas, A. (2003). Çözümlerde kavram başarı testi hazırlama ve uygulama. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 2(14), 1-17.
- Çapraz, C. & Yeşilyurt, S. (2018). Ölçek geliştirme çalışmalarında kullanılan kapsam geçerliği için bir yol haritası. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, 20(1), 251-264. <https://doi.org/10.17556/erziefd.297741>
- Çiçek Şentürk, Ö. & Selvi, M. (2021). Fen bilimleri dersi "insan ve çevre" ünitesi akademik başarı testi geliştirme: güvenilirlik ve geçerlik çalışması. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 41(2), 601-630. <https://doi.org/10.17152/gefad.940400>
- Demir, Ö., & Seferoğlu, S. S., (2019). Developing a Scratch-based coding achievement test. *Information and Learning Sciences*, 120, 383-406. <https://doi.org/10.1108/ILS-08-2018-0078>

- Demirer, V., & Sak, N. (2016). Programming education and new approaches around the world and in turkey / Dünyada ve Türkiye’de programlama eğitimi ve yeni yaklaşımlar. *Eğitimde Kuram ve Uygulama*, 12(3), 521-546.
- Demirtaşlı, N. (2017). Öğrenme, öğretim ve değerlendirme arasındaki ilişkiler. R. N. Demirtaşlı (Ed.) İçinde, *Eğitimde Ölçme ve Değerlendirme* (4. baskı) (ss. 1-24). Anı Yayıncılık.
- DeVellis, R. F. (2022). *Ölçek geliştirme: kuram ve uygulamalar*, (Scale development theory and applications), (Çev. Ed.) T. Totan, Nobel Yayınları.
- Devine, P. J., ve Raju N. S. (1982). Extent of overlap among four item bias methods. *Educational and Psychological Measurement*, 42, 1049–1066.
- Doğan, N., öğretmen, T. (2008). Değişen madde fonksiyonunu belirlemede mantel-haenszel, ki-kare ve lojistik regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim*, 33(148), 100-112.
- Erdem Keklik, D. (2012). *İki kategorili maddelerde tek biçimli değişen madde fonksiyonu belirleme tekniklerinin karşılaştırılması: bir simülasyon çalışması* [Yayımlanmamış Doktora Tezi]. Ankara Üniversitesi.
- Gönen, S., Kocakaya, S., & Kocakaya, F. (2011). Dinamik konusunda geçerliliği ve güvenilirliği sağlanmış bir başarı testi geliştirme çalışması. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 8(1), 40-57.
- Grover S., Jackiw, N., & Lundh, P. (2019) Concepts before coding: non programming interactives to advance learning of introductory programming concepts in middle school, *Computer Science Education*, 29 (2-3), 106-135, <https://doi.org/10.1080/08993408.2019.1568955>
- Gültekin, S. (2017). Testlerde kullanılacak madde türleri, hazırlama ilkeleri ve puanlaması. R. N. Demirtaşlı (Ed.) İçinde, *Eğitimde Ölçme ve Değerlendirme* (4.Baskı) (ss. 145-222). Anı Yayıncılık.
- Haladyna, T. M. (1997). *Writing test items to evaluate higher order thinking*. Allyn and Bacon: Needham Heights, MA.
- Haladyna T. M. (2011). *Handbook of test development*. Taylor and Francis Ltd.
- Hasançebi, B., Terzi, Y. & Küçük, Z. (2020). Madde güçlük indeksi ve madde ayırt edicilik indeksine dayalı çeldirici analizi. *Gümüşhane Üniversitesi Fen Bilimleri Dergisi*, 10 (1), 224-240. <https://10.17714/gumusfenbil.615465>
- Kargın, P. D., & Gül, Ş. (2021). Altıncı sınıf vücudumuzdaki sistemler ve sağlığı ünitesine yönelik bir başarı testi geliştirilmesi. *İhlara Eğitim Araştırmaları Dergisi*, 6(1), 1-26. <https://doi.org/10.47479/ihead.729412>
- Korkmaz, H., & Konukaldı, I. (2015). İlköğretim fen ve teknoloji eğitiminde disiplinlerarası tematik öğretim yaklaşımının öğrencilerin öğrenme ürünleri üzerine etkisi. *Buca Eğitim Fakültesi Dergisi*, 39, 1-22.
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2, 151-160.
- Kehoe, J. (1995). Basic item analysis for multiple-choice tests. *Practical Assessment, Research & Evaluation*, 4(10), 1-3. <https://doi.org/10.7275/07zg-h235>
- Kalaycı, Ş. (2008). *SPSS uygulamalı çok değişkenli istatistik teknikleri*. Asil Yayın Dağıtım.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel psychology*, 28(4), 563-575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Livingstone, R., Reynolds, C., & Wilson, V. (2006). *Measurement and Assessment in Education*. Prentice Hall.
- MEB (2017) “Bilişim teknolojileri ve yazılım dersi için hazırlanan materyaller öğretmen ile öğrencilerimiz kullanımına sunulmuştur”, <https://tegm.meb.gov.tr/www/bilisim-teknolojileri-ve-yazilim-dersi-icin-hazirlanan-materyaller-ogretmen-ile-ogrencilerimiz-kullanimina-sunulmustur/icerik/476>
- MEB (2018). *Bilişim teknolojileri ve yazılım dersi öğretim programı*, <https://mufredat.meb.gov.tr/ProgramDetay.aspx?PID=1383>
- MEB-TTKB, (2015). *Bilişim teknolojileri ve yazılım dersi programı*, Talim Terbiye Kurulu Başkanlığı.
- Piedade, J & Dorotea, N. (2023). Effects of Scratch-based activities on 4th-grade students’ computational thinking skills, *Informatics in Education*, 22(3), 499-523, <https://10.15388/infedu.2023.19>
- Saraç, H. (2018). Fen bilimleri dersi ‘maddenin değişimi’ ünitesi ile ilgili başarı testi geliştirme: geçerlik ve güvenilirlik çalışması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18(1), 416-445. <https://doi.org/10.17240/aibuefd.2018.-388815>

- Şahin, S. (2019). *Programlama öğretiminde ters yüz sınıf uygulamalarının ortaokul öğrencilerinin akademik başarısına etkisi* [Yayınlanmamış Yüksek Lisans Tezi]. İnönü Üniversitesi.
- Şenel, S. (2019). Programlama konusu “kazanımları” ile “sınıf içi ölçme süreçlerinin” bilişsel düzeylerinin karşılaştırılması, *International Journal of Computers in Education*, 1(2), 1-20.
- Oluk, A., Korkmaz Ö., & Oluk, H. (2018). Scratch’ın 5. sınıf öğrencilerinin algoritma geliştirme ve bilgi-işlemsel düşünme becerilerine etkisi. *Türk Bilgisayar ve Matematik Eğitimi Dergisi*, 9(1), 54-71. <https://10.16949/turkbilmat.399588>
- Özçelik, D. A. (2010). *Ölçme ve değerlendirme* (3. Baskı). Pegem Yayınları.
- Özçelik, D. A. (2013). *Test hazırlama kılavuzu* (5. Baskı). Pegem Yayınları.
- Özüdoğru, M., & Aksu, M. (2019). The development of an achievement test for the principles and methods of instruction course, *Trakya Eğitim Dergisi*, 9(4), 680-695. <https://doi.org/10.24315/tred.494903>
- Tekin, H. (2019). *Eğitimde ölçme ve değerlendirme* (27.Baskı). Yargı Yayınları.
- Turgut, M. F., & Baykul, Y. (2014). *Eğitimde ölçme ve değerlendirme* (6. baskı). Pegem Akademi.
- Yıldırım, A., & Şimşek, H. (2016). *Sosyal bilimlerde nitel araştırma yöntemleri*. Seçkin Yayıncılık.
- Yılmaz, H. (2012). *Eğitimde ölçme ve değerlendirme*. Çizgi Kitabevi.
- Zapata-Cáceres, M., Martín-Barroso, E., & Román-González, M. (2020). Computational thinking test for beginners: Design and content validation. *2020 IEEE Global Engineering Education Conference (EDUCON)*, ss. 1905–1914. IEEE.
- Zumbo, B. D., & Thomas, D. R. (1997). *A measure of effect size for a model-based approach for studying DIF*. Working Paper of the Edgeworth Laboratory for Quantitative Behavioral Science, University of Northern British Columbia: Prince George, B.C.

EXTENDED ABSTRACT

Introduction

Teaching the basics of programming skills at middle schools is one of the essential goals of Information Technologies and Software courses in Türkiye. Block-based programming tools such as Scratch, Blockly, Code.org, and Kodu Game Lab are used for coding in education in the framework of this course. Scratch is one of the block-based coding platforms where coding is done using the drag-and-drop method with a visual interface (Arslan & Akçelik, 2019). In addition to the Scratch program used in education, the Ministry of National Education (MEB) recommends Blockly, a block-based programming application developed by Google for 5th-graders, which can be used online and offline. This application allows teaching the basics of coding visually with the drag-and-drop method (Aytekin, Sönmez Çakır, Kulaöz & Yücel, 2018). This application started to be implemented in public and private schools in the 2017-2018 academic year within the scope of the cooperation protocol between the General Directorate of Basic Education and Google for the information technologies and software course (MEB, 2017). The course teachers can measure course achievements practically on the computer or on paper. Şenel (2019) states that the course teachers mostly use measurement tools containing multiple choice items as a measurement method.

There are some studies in the literature that develop and use block-based programming achievement tests (e.g. Oluk et al., 2018; Bakırcı, 2019; Şahin, 2019; Aydın, 2019; Demir & Seferoğlu, 2019; Zapata-Cáceres et al., 2020; Alagöz & Bilgiç, 2023). However, no comprehensive achievement test that measures the learning outcomes related to programming skills in the context of the Blockly block-based programming application developed for beginners in programming education has been found. This study aims to develop a valid and reliable academic achievement test for the Problem Solving and Programming unit of the 5th-grade Information Technologies and Software course. This study enables teachers and researchers to validly and reliably measure Blockly-related learning outcomes for 5th-grade students who have just started programming education and will contribute to the literature in these aspects.

Method

The study adopts a quantitative research design. As for test development stages, an approach in accordance with classical test theory was followed. Özçelik's (2013) achievement test development steps were applied. These steps determine the purpose and scope of the test, create an item pool, determine the measurement method, and ensure the validity and reliability and application of the test and item analysis.

An item pool of 48 questions was created based on the Blockly application environment, which measures the programming-related learning objectives of the unit in question, and expert opinions were applied for content validity. The items were corrected according to the feedback from experts. Content validity was ensured through expert opinion analysis method offered by Lawshe (1975). Criterion sampling was used as for the sampling method of the research. The trial test was created from an item pool prepared for 12 objectives and was applied to 151 6th grade students who had taken the Information Technologies and Software course in previous year. The final form of the test, consisting of 24 items, was administered to 72 5th grade students.

The trial application group of the research consists of a total of 151 (71 Girls, 80 Boys) 6th grade middle school students studying at Beyhan Gencay Middle School (n = 69) and Inonu Middle School (n = 82) in Antakya District of Hatay Province in 2021-2022 school year. The group of students who took the final test consists of 72 (34 boys, 38 girls) 5th grade students studying at Arsuz Madenli Middle School. It can be said that the gender ratios of the students are close. The sample of the research was determined using criterion sampling, which is one of the non-random purposive sampling methods.

The trial test of the scale developed by the authors, consisting of 48 items, was administered to students electronically via Google Forms. The 24-item final test, created based on the item analyses, was administered to the students as paper-and-pencil test.

Results

According to the item analyses results, the two items having the highest item discrimination index were selected for each learning objective to form the final test. The average item difficulty index of the final test, which consists of 24 items, was calculated as 0.63, the average item discrimination index was 0.48, and the reliability coefficient KR-20 value was calculated as 0.86. It can be concluded that the developed test is highly reliable and of moderate difficulty. Furthermore, the gender bias tests yielded no significant results, thus the final test has no gender bias.

Discussion and Conclusion

In this study, a reliable and valid 24-item achievement test, including multiple choice items, was prepared to measure the programming-related learning objectives of the 5th grade Problem Solving and Programming unit in the context of the Blockly application. This achievement test will contribute to the literature by demonstrating that the programming-related learning objectives of the 5th grade Information Technologies and Software course are measured when the Blockly block-based programming application is used as teaching environment.

Ek. 1. Blockly ile Programlama Başarı Testi

Yönerge: Bu test Bilişim Teknolojileri ve Yazılım dersi, Problem Çözme ve Programlama ünitesinin programlama ile ilgili kazanımlarını ölçmek için hazırlanmıştır. Soruları dikkatlice okuyup en doğru olan seçeneği işaretleyiniz.

Soru 1:

Aşağıdaki kod bloklarından hangisinde değişken kullanılmıştır?

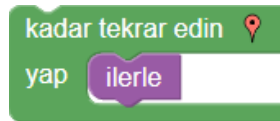
A)



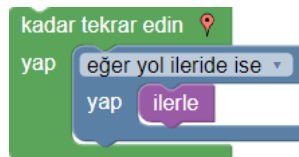
B)



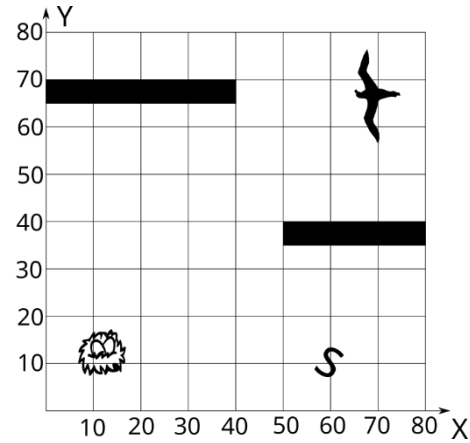
C)



D)



Soru 2:



Yukarıdaki görselde kuş solucanı alıp yuvaya götürecektir.

Bu süreçte aşağıdaki bilgilerden hangisi doğrudur?

- A) Kuşun konumu sabittir.
- B) Solucanın konumu sabittir.
- C) Yuvanın konumu sabittir.
- D) Engellerin konumu değişkendir.

Soru 3:

X = 80

Y = 20

Yukarıda verilen bilgilere göre aşağıdaki ifadelerden hangisi doğrudur?

- A) $X > 70$
- B) $Y < 20$
- C) $X < 80$
- D) $Y > 40$

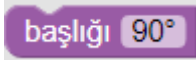
Soru 4:

Aşağıdakilerden hangisi operatör bloğu
örneğidir?

A)



B)



C)



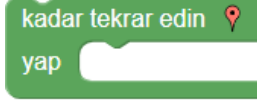
D)



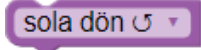
Soru 5:

Aşağıdakilerden hangisi bir şart ifadesidir?

A)



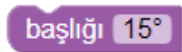
B)



C)



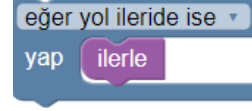
D)



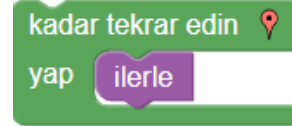
Soru 6:

Aşağıdaki kod bloklarından hangisi şart
ifadesi içerir?

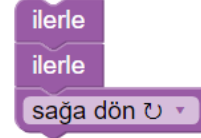
A)



B)



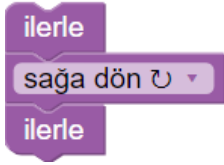
C)



D)



Soru 7:



Yukarıdaki kod bloğu, karakterin hangi hareketleri sırayla yapmasını sağlar?

- A) 1 kez ilerler, sağa döner
- B) 2 kez ilerler, sağa döner
- C) 1 kez ilerler, sağa döner, 1 kez ilerler
- D) Sağa döner, 2 kez ilerler

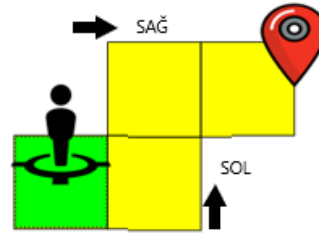
Soru 8:



Yukarıdaki görselde karakteri hedefe ulaştırabilmek için hangi kod bloğundan kaç adet kullanmamız gerekir?

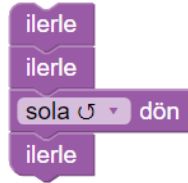
	BLOK	ADET
A)	ilerle	1
B)	sağa dön	2
C)	ilerle	3
D)	ilerle	5

Soru 9:



Yukarıdaki görselde karakteri hedefe ulaştırmak için hangi kod bloğu kullanılmalıdır?

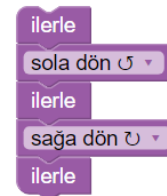
A)



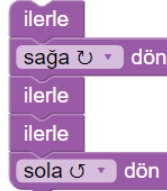
B)



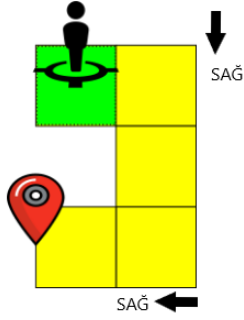
C)



D)



Soru 10:



Yukarıdaki görselde karakteri hedefe ulaştırmak için hangi kod bloğu kullanılmalıdır?

- A)
- B)
- C)
- D)

Soru 11:

Karar yapısı oluştururken aşağıdakilerden hangi kod bloğu kullanılır?

- A)
- B)
- C)
- D)

Soru 12:

Aşağıdaki kod bloklarından hangisi karar yapısı içerir?

- A)
- B)
- C)
- D)

Soru 13:

Döngü yapısının görevi aşağıdakilerden hangisidir?

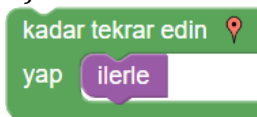
- A) Kod bloklarını tekrar ettirir.
- B) Karakterin rengini değiştirir.
- C) Karakterin hangi yöne gideceğini belirler.
- D) Şart ifadesi oluşturur.

Soru 14:

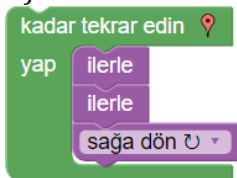


Aşağıdaki kod bloklarından hangisi yukarıda verilen görseldeki kod bloğu ile aynı görevi yapar?

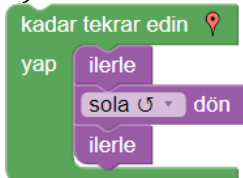
A)



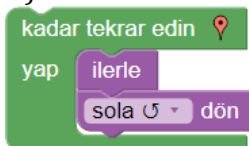
B)



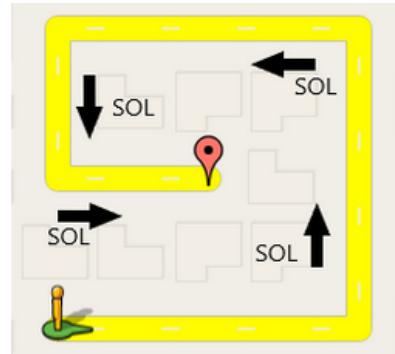
c)



D)



Soru 15:

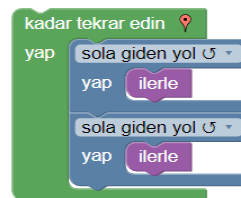


Yukarıdaki görselde karakteri hedefe ulaştırabilmek için hangi kod bloğu kullanılabilir?

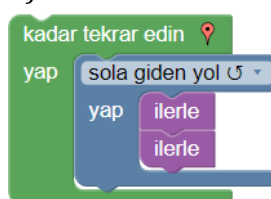
A)



B)



c)



D)



Soru 16:

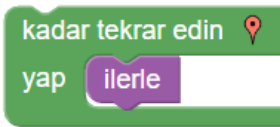


Yukarıdaki görselde karakteri **en az kod bloğuyla** kırmızı hedefe ulaştırmak için aşağıdakilerden hangisi en uygun kod bloğudur?

A)



B)



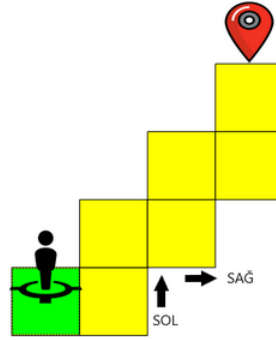
C)



D)



Soru 17:



Yukarıdaki görselde karakterin kırmızı hedefe ulaşması için aşağıdaki kod bloklarından hangisi kullanılmalıdır?

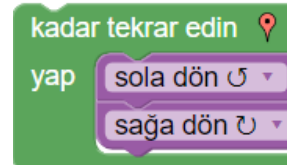
A)



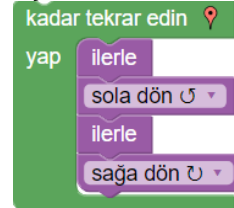
B)



C)



D)

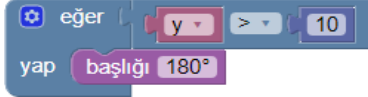


Soru 18:

“Eğer $Y < 10$ ise başlığı (yön derecesini)**180 derece yap”**

Yukarıdaki ifadenin kod bloğu
şeklinde yazılmış hâli
aşağıdakilerden hangisidir?

A)



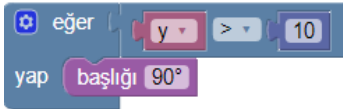
B)



C)

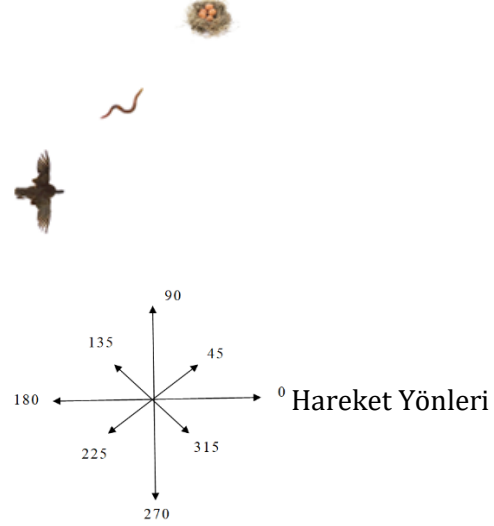


D)



Aşağıdaki sorularda kuşun yönünü belirlemek için **Hareket Yönlerini** kullanınız.

Soru 19:



Yukarıdaki görselde kuşun yuvaya solucanı
götürebilmesi için yönünü kaç derece
çevirmesi gerekir?

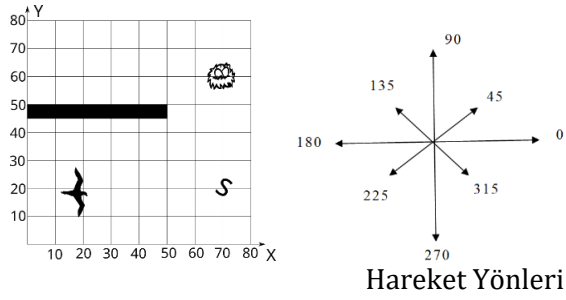
A) başlığı 0°

B) başlığı 45°

C) başlığı 90°

D) başlığı 180°

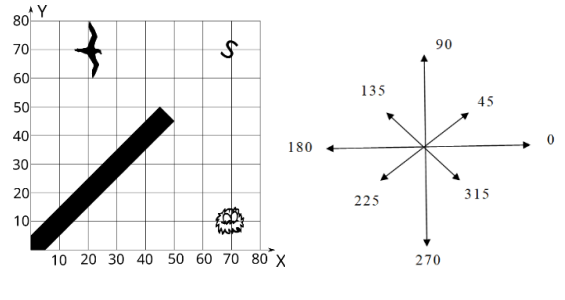
Soru 20:



Yukarıdaki görselde kuşun yuvaya solucanı götürmesi için yönünü sırasıyla kaç derece yapması gerekir?

- A) Solucanı almak için 0° derece yönünde, yuvaya uçmak için 90° derece yönünde.
- B) Solucanı almak için 45° derece yönünde, yuvaya uçmak için 180° derece yönünde.
- C) Solucanı almak için 90° derece yönünde, yuvaya uçmak için 90° derece yönünde.
- D) Solucanı almak için 135° derece yönünde, yuvaya uçmak için 90° derece yönünde.

Soru 21:

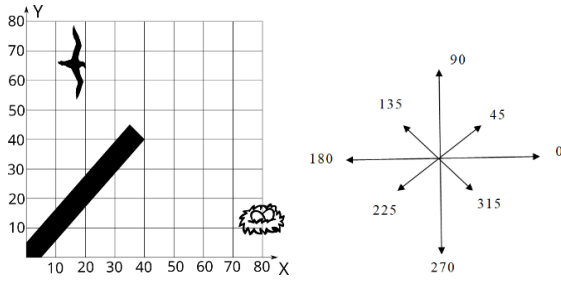


Yukarıdaki görselde kuşun solucanı olarak yuvaya doğru gitmesi için aşağıdaki karar yapısı bloğunda numaralı yerlere hangi kod blokları gelmelidir?

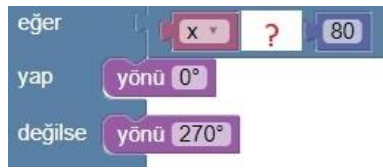


- A) 1 başlığı 0° 2 solucanı yok 3 başlığı 90°
- B) 1 başlığı 90° 2 solucanı yok 3 başlığı 270°
- C) 1 solucanı yok 2 başlığı 0° 3 başlığı 270°
- D) 1 solucanı yok 2 başlığı 90° 3 başlığı 0°

Soru 22:



Hareket Yönleri

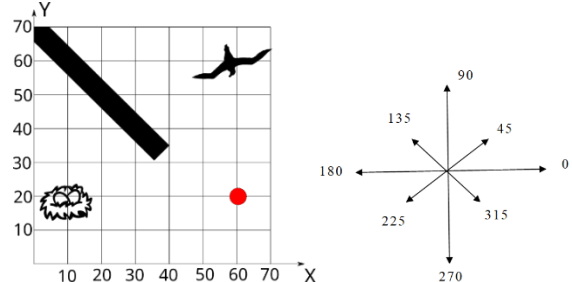


Yukarıdaki görselde kuşun yuvaya ulaşması için oluşturulan kod bloğunda soru işareti yerine hangi operatör gelmelidir?

(**X** : Kuşun **X eksen**i üzerindeki konumuna, **Y** : Kuşun **Y eksen**i üzerindeki konumuna karşılık gelen sayıyı ifade eder.)

- A) <
- B) >
- C) VE
- D) =

Soru 23:



Hareket Yönleri



Yukarıdaki görselde kuşun kırmızı noktaya ulaşması için oluşturulan kod bloğunda soru işareti yerine hangi şart ifadesi kullanılmalıdır?

(**X** : Kuşun **X eksen**i üzerindeki konumuna, **Y** : Kuşun **Y eksen**i üzerindeki konumuna karşılık gelen sayıyı ifade eder.)

A)



B)



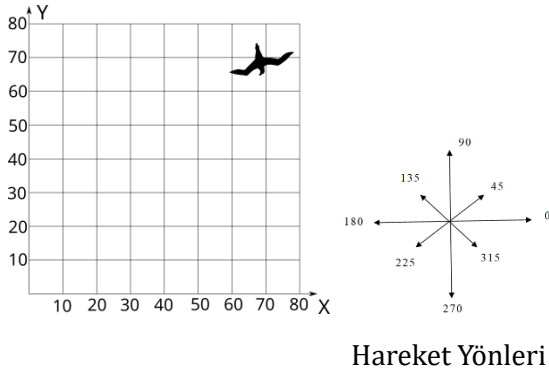
C)



D)



Soru 24:



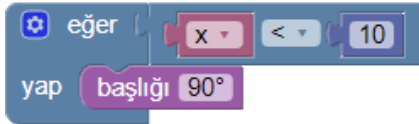
Cevaplar:

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

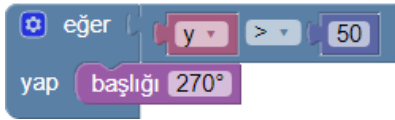
Yukarıdaki görselde kuşun **X=70** ve **Y=20** konumuna ulaşabilmesi için hangi kod bloğunun kullanılması gerekir?

(**X** : Kuşun **X eksen**i üzerindeki konumuna,
Y : Kuşun **Y eksen**i üzerindeki konumuna karşılık gelen sayıyı ifade eder.)

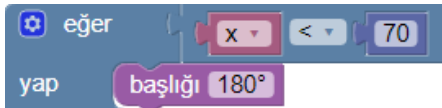
A)



B)



C)



D)

