

İLKOKUL 3. VE 4. SINIF ÖĞRENCİLERİNE YÖNELİK MATEMATİK BAŞARI TESTİ GELİŞTİRİLMESİ

Oğuz BALCI
Hicabi Çelebi Ortaokulu, Ordu
balcioguz52@gmail.com
ORCID ID: 0000-0003-0943-8116

Gökhan ÖZSOY
Ordu Üniversitesi Eğitim Fakültesi
gozsoy@odu.edu.tr
ORCID ID: 0000-0002-1250-624X

Araştırma Makalesi

Geliş Tarihi: 02.05.2021

Revize Tarihi: 01.06.2021

Kabul Tarihi: 19.06.2021

Atıf Bilgisi: Balcı, O. ve Özsoy, G. (2021). İlkokul 3. ve 4. sınıf öğrencilerine yönelik matematik başarı testi geliştirilmesi. *Sınıf Öğretmenliği Araştırmaları Dergisi (SÖAD)*, 1(1), 13-32.

ÖZ

Bu araştırmanın amacı ilkokul 3. ve 4. sınıf öğrencilerinin tüm kazanımları kapsayacak bir yaklaşımla matematik başarılarını belirlemeye yönelik, geçerli, güvenilir ve birbirine eşdeğer olan testler geliştirmektir. Bu doğrultuda 3. ve 4. sınıfta bulunan matematik kazanımları incelenerek aralarında kritik kazanım olanlar belirlenmiştir. Belirlenen bu kazanımlara yönelik en az üçer madde yazılmış ve bu maddeler uzmanların görüşlerine sunulmuştur. Uzmanlardan gelen dönütlere göre gerekli düzeltmeler yapılarak 3. ve 4. sınıfın deneme uygulamalarının her biri için 80 soru seçilmiştir. Testlerin deneme ve final uygulamaları, ilgili sınıf düzeyindeki kazanımların tamamını görmüş olan bir üst sınıftaki öğrencilerle Ordu ilinde yürütülmüştür. Nihai testlere madde seçerken maddelerin ayırt edicilik değerleri göz önünde bulundurularak testlerde belirlenen kritik kazanımların her birine yönelik madde bulunmasına dikkat edilmiştir. İlkokul 3. sınıf testlerine 28'er soru, 4. sınıf testlerine 22'şer soru seçilmiştir. İlkokul 3. sınıf testlerinin son uygulamaları 202 ilkokul 4. sınıf öğrencisi ile ilkokul 4. sınıf testlerinin final uygulamaları ise 221 ortaokul 5. sınıf öğrencisiyle yürütülmüştür. İlkokul 3. sınıf testlerinin KR-20 güvenilirlik katsayıları sırasıyla 0.834 ve 0.814; ortalama güçlük değerleri sırasıyla 0.627 ve 0.628 bulunmuştur. İlkokul 4. sınıf testlerinin KR-20 güvenilirlik katsayıları ise sırasıyla 0.841 ve 0.828; ortalama güçlük değerleri sırasıyla 0.610 ve 0.625 bulunmuştur. İlkokul 3. sınıf testleri arasındaki Spearman korelasyon katsayısı 0.880, 4. sınıf testleri arasındaki Spearman korelasyon katsayısı 0.829 olarak hesaplanmıştır. Ölçme araçlarının yapı geçerliğini incelemek amacıyla testler son uygulamasına katılan öğrencilerin yanında geliştirildikleri sınıf seviyesinin bir alt sınıfında öğrenim gören öğrencilere de uygulanmıştır. Araştırmanın sonucunda ilkokul 3. ve 4. sınıfa yönelik eşdeğer matematik başarı testleri oluşturulmuştur.

Anahtar Kelimeler: Matematik dersi, başarı testi geliştirme, 3. sınıf başarı testi, 4. sınıf başarı testi, eşdeğer testler.

MATHEMATICS ACHIEVEMENT TEST DEVELOPMENT FOR 3RD AND 4TH GRADE STUDENTS

ABSTRACT

The purpose of this research is developing reliable, valid and equivalent tests which are directed to determining the 3rd and 4th class students' math achievement. In this direction, the learning outcomes of maths lesson during 3rd and 4th classes have been analyzed and critical learning outcomes have been determined. For these learning outcomes, at least 3 matters have been written and those matters have been presented for the experts view. According to the feedbacks which have come from experts, necessary corrections have been made and 80 questions have been selected for 3rd and 4th class trial applications.

The trial and final applications of test have been carried out in Ordu province with upper class students who achieved all the learning outcomes. While choosing the matters for final tests, considering the distinguishing valuations of the matters, we have paid attention to the matters which must have been found out for each critical learning outcomes. 28 question have been selected for 3rd class tests and 22 questions have been selected for 4th class tests. The final practice of the 3rd classes tests have been carried out with 202 primary school students, the final practice of the 4th classes tests have been carried out with 221 secondary school students. The KR-20 reliability parameters of primary 3rd class tests have been found out as respectively 0.834 and 0.814; the average difficulty value have been found out as respectively 0.627 and 0.628. The KR-20 reliability parameters of primary 4th class tests have been found out as respectively 0.841 and 0.828; the average difficulty value have been found out as respectively 0.610 and 0.625. The coefficient of Spearman correlation among the primary 3rd class tests have been calculated as 0.880, the coefficient of Spearman

correlation among the primary 4th class tests have been calculated as 0.829. For analyzing the construct validity of assessment instruments, those tests have been performed low level students beside the students who joined the final performance. As a result of this research, equivalent maths tests for primary 3rd and 4th classes, have been formed.

Keywords: Math lesson, achievement test development, 3rd grade achievement test, 4th grade achievement test, equivalent tests.

Giriş

Matematik eğitimi ile öğrencilerin eğitim-öğretim süreci sonunda matematik öğretim programındaki hedef davranışları sergilemeleri yani istendik davranış değişikliklerini göstermesi beklenir. Değişikliğe uğramış ve yeni geliştirilmiş bu davranışlara eğitim sisteminin çıktıları adı verilir. Çıktı olarak adlandırılan davranışlar eğitim sisteminin ürünleridir (Baykul, 1992). Ortaya konan ürünler hedef davranışa ulaşıp ulaşılmadığı noktada öğretmene fikir verecektir. Bu durum ise eğitimsel çıktıların açık bir şekilde tanımlanmasını ve bireylerde meydana gelen öğrenmenin göstergelerinin neler olabileceğinin düşünülmesini gerektirmektedir (Fidan, 2013). Öğretmenler bu göstergeler neticesinde eğitim-öğretim faaliyetlerini nitelik, verimlilik ve etkililik bazında değerlendirme fırsatı bulacaktır. Bu fırsatı oluşturacak temel unsur da öğretmenin yapacağı ölçme işlemidir.

Ölçme, matematik eğitiminde de öğretim faaliyetlerinin önemli bir aşamasıdır. Matematik öğrenme sürecinin başında, devamında ve sonunda öğrencinin gerek gelişimsel gerekse de akademik başarı düzeyini belirlemek ve öğretimin istenilen oranda olup olmadığını görmek öğretmenler için büyük öneme sahiptir. Çünkü öğretmenlerin uyguladıkları öğretim etkinliklerinin başarısı ile öğrencinin akademik başarısı doğru orantılı olacaktır. Akademik başarı, çoğunlukla bilişsel beceri ve yeterliliği ölçmeye yaramakta ve derslerde gösterilen performansın betimlenmesine dayanmaktadır (Kaya, Bozaslan ve Genç, 2012). Matematik eğitiminde de betimleme işleminin genellikle ölçme işlemi ile yapıldığı düşünülmektedir.

Matematikte akademik başarıyı betimlerken ölçme araçlarından faydalanılması öğretmenlerin sayısal olarak daha objektif bilgiler elde etmesini sağlayacaktır. Turgut ve Baykul (2015) da ölçme işlemi için araç kullanımının önemine dikkat çekerek, araçların ölçülecek niteliği sayılarla ifade ettiğini ve böylece ölçmenin daha duyarlı bir şekilde yapıldığını ifade etmişlerdir. Ölçme aracının ölçmede sağladığı duyarlılığın yanında amaca hizmet ediyor olması da oldukça önemlidir. Bundan dolayı yapılacak çıkarımın doğasına göre ölçme amacının netleştirilmesi ve bu amaca uygun ölçme yöntemlerinin seçilmesi gerekmektedir (Cizek, 1997; Turgut ve Baykul, 2015).

Matematik dersinde öğrencinin öğrenme düzeyini ve öğretimin etkililiğini ölçebilmek için kullanılacak birçok ölçme aracı bulunmaktadır. Her ölçme aracının da kullanım amacına göre kendine has olumlu ve olumsuz yönleri vardır. Ancak öğretmenlerin, öğrencilerin öğrenme düzeylerini yani matematik başarılarını ölçmek için genellikle çoktan seçmeli başarı testlerini kullandıkları söylenebilir. Çakan (2004), öğretmenlerin ölçme aracı kullanımına ilişkin olarak yaptığı bir çalışmada, öğretmenlerin %56.2 oranla çoktan seçmeli maddeleri kullandıkları ve bunu %48 ile %46.1 oranla yazılı yoklama ve kısa cevaplı maddelerin izlediğini belirlemiştir.

Öğretmenlerin öğrencilerin matematik başarılarını çoğunlukla onların ölçme araçlarından aldıkları puanlara göre yorumladıkları düşünülmektedir. Öğretmenler tarafından en çok tercih edilen ölçme aracı çoktan seçmeli testler olduğundan bu testlerin standartlaştırılması oldukça önemlidir. Standart testler; yoğun deneysel çalışmaların ürünü olarak geliştirilen, belirli amaçlar doğrultusunda uzman kişiler tarafından hazırlanan (Koç, 1985), uygulama esasları, nasıl puanlanacağı ve puanların yorumlanmasında nasıl bir yol izleneceği sabit olan ölçme araçlarıdır (Cronbach, 1970; Shipley ve McAfee, 2009). Bu ölçme araçlarının doğru bir şekilde ve uzman kişilerce kullanımı eğitim-öğretimin etkililiği, rehberlik, seçme, yerleştirme gibi pek çok konuda alınan kararların yerindeliğini sağlamaktadır (Koç, 1985).

Günümüzde öğretmenlerin çoğu standart olmayan çoktan seçmeli testleri kullanmaktadır. Bu testler geçerli ve güvenilir olmadığı için öğretim sürecinin başından sonuna kadar öğretmeni yanlış yönlendirebilme ihtimaline sahiptir. Öğretmenler, standart olmayan çoktan seçmeli testlerden elde ettiği ölçme sonuçları ile öğretimi ilişkilendirecek ve buna göre bir sonraki aşamaya geçilip geçilemeyeceğine karar verecektir. Ancak gözden kaçan eksik öğrenmeler veya kavram yanlışları öğrencilerin sonraki öğrenmelerini zorlaştıracak gibi, matematik başarısını da düşürecektir. Yapılan çalışmalarda da öğrencilerin sahip olduğu yanlış anlamalar ve kavram yanlışları, onların sonraki öğrenmelerini etkilemiş ve değişime karşı direnç göstermelerine sebep olmuştur (Ayas ve Demirbaş, 1997; Hewson ve Hewson, 1983; Nakhleh, 1992; Pardo ve Partoles, 1995; Zoller, 1990). Bu nedenle öğretmenlerin kullanacakları çoktan seçmeli testlerin, hedef davranışları tam olarak ölçüp ölçmediği, güvenilir ya da geçerli olup olmadığı, kullanılabilirliği ve öğrenci öğrenmesinin hangi düzeyde olduğunu belirlemesi açısından standartlaşmış olmasını gerektirmektedir.

Değerlendirme, ölçme sonuçlarını belli bir ölçüte dayandırarak başarılı ya da başarısız gibi bir değer yargısına varılmasıdır (Eskiocak, 2004). Bu tanımdan da anlaşılacağı üzere ölçme ve değerlendirme öğretim sürecinin ayrılmaz bir ögesi olup (Heritage, 2007), gerek öğrencilerin öğrenme düzeylerini iyileştirmek ve geliştirmek gerekse de kazanımlara ulaşma düzeylerini tespit ederek öğrenciler hakkında karar vermek için kullanılır (Şimşek, Bars ve Zengin, 2017). Verilecek kararların gerçekçi olabilmesi için geçerli ve güvenilir ölçümlere ihtiyaç vardır (Atılğan, Kan ve Doğan, 2013). Geçerli ve güvenilir ölçümler de yine öğretmenlerin ölçmede kullanacakları standartlaşmış testlerle daha mümkün olabilecektir.

Bu çalışmanın amacı, matematik dersi öğretim programı temel alınarak ilkokul 3. ve 4. sınıf öğrencilerinin matematik dersindeki başarı durumlarını belirlemeye yönelik, geçerliği ve güvenilirliği sağlanmış birbirine eşdeğer matematik testleri geliştirmektir.

Çalışmanın amacına paralel olarak aşağıdaki sorulara yanıt aranmaktadır:

Geliştirilen testler;

- Kapsam geçerliğine sahip midir?
- Yapı geçerliğine sahip midir?
- Madde güçlük indeksleri kaçtır?
- Madde ayırtıcılık gücü indeksleri kaçtır?

Değişen zaman ve koşullardan dolayı öğretim programlarının değişmesi gerekmektedir (Kemertaş, 1999). Millî Eğitim Bakanlığı da 2018 yılında bireylerin değişime ve çağın beklentilerine ayak uydurabilmesi amacıyla matematik dersi öğretim programını güncellemiştir. Güncellenen program doğrultusunda 3. ve 4. sınıf öğrencilerine yönelik ilgili matematik programının tamamını hedef alan geçerliği ve güvenilirliği araştırılmış bir başarı testi geliştirme çalışmasına rastlanmamıştır.

Yapılan çalışmada aynı özellikleri ölçen, geçerliği ve güvenilirliği istatistiksel olarak hesaplanmış eşdeğer olan iki çoktan seçmeli test geliştirilmesi amaçlanmıştır. Bu nedenle geliştirilen testlerin çalışmalarında eşdeğer formdan oluşan testler kullanacak araştırmacılara daha güvenilir sonuçlar vereceği düşünülmektedir. Bu yönüyle çalışmanın alana katkı sağlaması beklenmektedir.

Bu çalışmada kullanılan maddelerin bazılarında öğretim araçları kullanılmıştır. Yapılan literatür taramasında öğretim materyalinin kullanıldığı çoktan seçmeli test maddelerine rastlanmamıştır. Çalışmanın bu noktada da diğer çalışmalardan farklı olduğu düşünülmektedir.

Yöntem

Araştırma Modeli

Bu çalışma, 3. ve 4. sınıfta öğrenim görmekte olan öğrencilerin matematik dersindeki başarı düzeylerinin belirlenmesine yönelik olarak kullanılabilir geçerliği ve güvenilirliği kanıtlanmış başarı testleri geliştirme çalışmasıdır. Testlerin geliştirilme sürecinde 3. sınıftan iki tane matematik başarı

Testi ve 4. sınıftan iki tane matematik başarı testi olmak üzere, aynı sınıf seviyesine yönelik birbirine eşdeğer iki test oluşturulması hedeflenmiştir.

Çalışma Grubu

Araştırmanın çalışma grubunu, geliştirilen testlerdeki kazanımların tamamını gören ve 2018-2019 eğitim-öğretim yılında 4. ve 5. sınıfta öğrenim görmekte olan öğrenciler oluşturmaktadır. Öğrencilerin buldukları okullar ve sınıflar elverişli örnekleme yoluyla Altınordu ilçesinden seçilmiştir. Geliştirilen 3. sınıf matematik başarı testlerinin son uygulaması, altı devlet ilkokulunda deneme uygulamasına katılmayan 202 ilkokul 4. sınıf öğrencisiyle; 4. sınıf matematik başarı testlerinin son uygulaması ise 4 devlet ortaokulunda deneme uygulamasına dâhil olmayan 221 ortaokul 5. sınıf öğrencisiyle yürütülmüştür.

Veri Toplama Araçları

Çalışmanın veri toplama araçları, 2018’de yenilenen Matematik Dersi Öğretim Programı’ndaki kazanımlara göre, öğrencilerin matematik başarısını ortaya çıkarmada kullanılması amaçlanan 3. ve 4. sınıf matematik başarı testleridir.

Araştırmacı tarafından geliştirilen matematik başarı testlerinin geliştirilme sürecinde aşağıdaki adımlara göre hareket edilmiştir (Turgut ve Baykul, 2015, s. 215-216):

- Testin (test puanlarının) hangi amaçla kullanılacağı belirlenmesi,
- Testle ölçülecek davranışların saptanması,
- Maddelerin yazılması,
- Maddelerin gözden geçirilmesi (redaksiyon),
- Deneme formunun hazırlanması,
- Deneme uygulamasının yapılması,
- Deneme uygulaması cevap kâğıtlarının puanlanması, madde analizi ve madde seçimi,
- Nihai testin oluşturulması ve istatistiklerinin kestirilmesi.

Verilerin Toplanması ve Analizi

Testin hangi amaçla kullanılacağı belirlenmesi

Bu çalışmada 3. ve 4. sınıf öğrencilerinin matematik başarısını ortaya çıkarmak amaçlandığından geliştirilecek testler öğrenci başarısını ölçmek için kullanılacaktır.

Testle ölçülecek davranışların belirlenmesi

Testte matematik dersi öğretim programında bulunan ilgili her bir kazanım için soru bulunması, testlerin uygulanmasını kolay ve ekonomik yapmayacaktır. Bundan dolayı konu alanının hedefleri doğrultusunda ilerleme ve gelişmeleri daha iyi yansıtan davranışlar seçilerek bunların ölçülmesi önerilmektedir (Özçelik, 2010). Çalışmada bu önerinin yanında Tablo 1’de verilen öğrenme alanlarındaki kazanım sayısının, toplam kazanım sayısındaki yüzdesi de gözetilerek, 3. ve 4. sınıf matematik öğretim programındaki kazanımlardan kritik kazanım olanları belirleme yoluna gidilmiştir.

Tablo 1

Matematik Dersi Öğretim Programındaki İlkokul 3. ve 4. Sınıf Kazanımlarının Öğrenme Alanlarına Göre Dağılımı

Öğrenme Alanları	3. Sınıf	4. Sınıf
Sayılar ve İşlemler	36	34

Geometri	10	12
Ölçme	23	21
Veri Toplama ve Değerlendirme	3	4
Toplam	72	71

Herhangi bir kavram onun ön şartı olan diğer kavramlar kazandırılmadan tam olarak verilemez (Altun, 2011). Bu çalışmada 3. ve 4. sınıf matematik öğretim programında bulunan öğrenme alanlarındaki kazanımlardan, aşamalılık ilkesine göre diğerini kapsayan kazanım kritik kazanım olarak belirlenmiştir. Kritik kazanımları belirleme sürecinde 1 konu alanı uzmanı, 2 sınıf öğretmeni ve 2 ortaokul matematik öğretmeni olmak üzere toplam 5 kişiden görüş alınmıştır. Matematik öğretim programındaki kazanımlar arasından belirlenen kritik kazanımların öğrenme alanlarına göre dağılımı 3. sınıf için Tablo 2’de, 4. sınıf için Tablo3’te gösterilmiştir.

Tablo 2

Matematik Öğretim Programında 3. Sınıf İçin Belirlenen Kritik Kazanımlar

Öğrenme Alanları	Kritik Kazanımlar
Sayılar ve İşlemler	M.3.1.1.3. Üç basamaklı doğal sayıların basamak adlarını, basamaklarındaki rakamların basamak değerlerini belirler. M.3.1.1.7. Aralarındaki fark sabit olan sayı örüntüsünü genişletir ve oluşturur. M.3.1.1.8. Tek ve çift doğal sayıları kavrar. M.3.1.2.3. İki sayının toplamını tahmin eder ve tahminini işlem sonucuyla karşılaştırır. M.3.1.1.10. 20’ye kadar olan Romen rakamlarını okur ve yazar. M.3.1.2.5. Bir toplama işleminde verilmeyen toplananı bulur. M.3.1.3.4. Doğal sayılarla toplama ve çıkarma işlemlerini gerektiren problemleri çözer. M.3.1.4.3. Biri çarpma işlemi olmak üzere iki işlem gerektiren problemleri çözer. M.3.1.5.3. Bölme işleminde bölünen, bölen, bölüm ve kalan arasındaki ilişkiyi fark eder. M.3.1.5.4. Biri bölme olacak şekilde iki işlem gerektiren problemleri çözer. M.3.1.6.4. Paydası 10 ve 100 olan kesirlerin birim kesirlerini gösterir. M.3.1.6.5. Bir çokluğun, belirtilen birim kesir kadarını belirler.
Geometri	M.3.2.1.2. Küp, kare prizma ve dikdörtgen prizmanın birbirleriyle benzer ve farklı yönlerini açıklar. M.3.2.1.3. Cetvel kullanarak kare, dikdörtgen ve üçgeni çizer; kare ve dikdörtgenin köşegenlerini belirler. M.3.2.2.2. Bir parçası verilen simetrik şekli dikey ya da yatay simetri doğrusuna göre tamamlar. M.3.2.3.1. Şekil modelleri kullanarak kaplama yapar, yaptığı kaplama örüntüsünü noktalı ya da kareli kâğıt üzerine çizer. M.3.2.4.2. Doğruyu, ışını ve açıyı tanıır. M.3.2.4.3. Doğru parçasını çizgi modelleri ile oluşturur; yatay, dikey ve eğik konumlu doğru parçası modellerine örnekler vererek çizimlerini yapar.
Ölçme	M.3.3.1.3. Cetvel kullanarak uzunluğu verilen bir doğru parçasını çizer. M.3.3.1.5. Metre ve santimetre birimlerinin kullanıldığı problemleri çözer. M.3.3.2.4. Şekillerin çevre uzunlukları ile ilgili problemleri çözer. M.3.3.3.1. Şekillerin alanını standart olmayan uygun malzeme ile kaplar ve ölçer. M.3.3.4.2. Paralarımızla ilgili problemleri çözer. M.3.3.5.4. Zaman ölçme birimlerinin kullanıldığı problemleri çözer. M.3.3.6.3. Kilogram ve gramla ilgili problemleri çözer. M.3.3.7.3. Litre ile ilgili problemleri çözer.
Veri İşleme	M.3.4.1.2. Grafiklerde verilen bilgileri kullanarak veya grafikler oluşturularak toplama ve çıkarma işlemleri gerektiren problemleri çözer. M.3.4.1.3. En çok üç veri grubuna ait basit tabloları okur, yorumlar ve tablodan elde ettiği veriyi düzenler.

Tablo 3
Matematik Öğretim Programında 4. Sınıf İçin Belirlenen Kritik Kazanımlar

Öğrenme Alanları	Kritik Kazanımlar
Sayılar ve İşlemler	<p>M.4.1.1.3. 4, 5 ve 6 basamaklı doğal sayıların bölüklerini ve basamaklarını, basamaklarındaki rakamların basamak değerlerini belirler ve çözümler.</p> <p>M.4.1.1.6. Belli bir kurala göre artan veya azalan sayı örüntüleri oluşturur ve kuralını açıklar.</p> <p>M.4.1.2.2. İki doğal sayının toplamını tahmin eder ve tahminini işlem sonucu ile karşılaştırır.</p> <p>M.4.1.3.4. Doğal sayılarla toplama ve çıkarma işlemini gerektiren problemleri çözer.</p> <p>M.4.1.4.6. Doğal sayılarla çarpma işlemini gerektiren problemleri çözer.</p> <p>M.4.1.5.6. Doğal sayılarla en az bir bölme işlemi gerektiren problemleri çözer.</p> <p>M.4.1.5.7. Aralarında eşitlik durumu olan iki matematiksel ifadeden birinde verilmeyen değeri belirler ve eşitliğin sağlandığını açıklar.</p> <p>M.4.1.5.8. Aralarında eşitlik durumu olmayan iki matematiksel ifadenin eşit olması için yapılması gereken işlemleri açıklar.</p> <p>M.4.1.6.3. Bir çokluğun belirtilen bir basit kesir kadarını belirler.</p> <p>M.4.1.6.4. Paydaları eşit olan en çok üç kesri karşılaştırır.</p> <p>M.4.1.7.2. Kesirlerle toplama ve çıkarma işlemlerini gerektiren problemleri çözer.</p>
Geometri	<p>M.4.2.1.2. Kare ve dikdörtgenin kenar özelliklerini belirler.</p> <p>M.4.2.1.5. İzometrik ya da kareli kâğıda eş küplerle çizilmiş olarak verilen modellere uygun basit yapılar oluşturur.</p> <p>M.4.2.2.2. Verilen şeklin doğruya göre simetriğini çizer.</p> <p>M.4.2.3.5. Standart açı ölçme araçları kullanarak ölçüsü verilen açığı oluşturur.</p>
Ölçme	<p>M.4.3.1.4. Uzunluk ölçme birimlerinin kullanıldığı en çok üç işlem gerektiren problemleri çözer.</p> <p>M.4.3.2.3. Şekillerin çevre uzunluklarını hesaplama ile ilgili problemleri çözer.</p> <p>M.4.3.3.2. Kare ve dikdörtgenin alanını toplama ve çarpma işlemleri ile ilişkilendirir.</p> <p>M.4.3.4.2. Zaman ölçme birimlerinin kullanıldığı problemleri çözer.</p> <p>M.4.3.5.5. Ton, kilogram, gram ve miligram ile ilgili problemleri çözer.</p> <p>M.4.3.6.5. Litre ve mililitre ile ilgili problemleri çözer.</p>
Veri İşleme	<p>M.4.4.1.4. Sütun grafiği, tablo ve diğer grafiklerle gösterilen bilgileri kullanarak günlük hayatla ilgili problemler çözer.</p>

Maddelerin Yazılması

Belirlenen kritik kazanımlar doğrultusunda hazırlanan belirtke tablosuna göre her bir kazanım için en az 2 ya da 3 soru hazırlanmıştır. Hazırlanan sorulardan bazıları şekil ve görseller ile desteklenmiş, en uygun soru şeklinin ne olduğunu anlayabilmek için farklı biçim ve içeriklerden oluşan sorular oluşturulmuştur

Maddelerin Gözden Geçirilmesi

Kritik kazanımlara yönelik hazırlanan 3. sınıf için 172, 4. sınıf için 129 sorunun ilgili kazanımı ölçüp ölçmediğine yönelik görüşler almak için konu alanı uzmanı ile öğretmenlere başvurulmuştur. Yine soruların bilimsel açıdan doğruluğu ve sınıf düzeyine uygunluğu için 3 sınıf öğretmeni, 2 ortaokul matematik öğretmeni ile matematik eğitimi alanında 1 uzmana danışılmıştır.

Gelen dönütlere göre sorularda düzeltmeler yapılarak, sorular dil bilgisi ve yazım hataları yönünden kontrol edilip değerlendirilmesi için bir Türkçe öğretmenine yazılı olarak verilmiştir.

Deneme Formunun Hazırlanması

Uzmanlardan ve öğretmenlerden gelen görüşler doğrultusunda, nihai testleri oluşturacak soruları belirlemek amacıyla, oluşturulacak testlerin kullanılabilirliği ve ekonomikliği de göz önünde bulundurularak her sınıf için pilot çalışmada kullanılacak soru sayısının 80 olmasına karar verilmiştir. Yine pilot çalışmada 4. ve 5. sınıf öğrencilerinin tek oturumda 80 soruyu çözmesinin sağlıklı olamayacağı düşünüldüğünden, sorular dört ayrı test şeklinde 40 dakikalık bir ders saatinde uygulanabileceği yönünde öğretmenler ve konu alanı uzmanı ile görüş birliğine varılmıştır. Testlerin oluşturulma esnasında sorular, ilk konudan son konuya doğru yerleştirilmiştir. Turgut ve Baykul (2015) da sorular cevaplanırken ortaya çıkabilecek zihinsel yorgunluğu engelleyebilmek için başarı testlerinde soruların konulara göre gruplandırılmasını önermektedir.

Deneme Uygulamasının Yapılması

Deneme uygulamaları ile maddelerin sayısal özelliklerini hesaplamak amaçlanmaktadır. İşlemeyen maddelerin testten çıkarılması veya yeniden düzenlenmesi, maddelerin güçlük düzeylerinin belirlenmesi, atlanan ve erişilemeyen maddelerin tespit edilmesi, çeldiricilerin ne derece işlediğinin belirlenmesi, madde ayırt ediciliğinin hesaplanması, teste alınacak madde sayısı ile test için ideal sürenin belirlenmesi buradaki asıl amaçtır (Tekindal, 2011).

Deneme uygulamalarında önce ilgili okul yönetimi ve sınıf öğretmeniyle bir araya gelinerek çalışmanın içeriği açıklanmıştır. Testler sınıflara uygulanmadan önce de sınavı uygulayacak sınıf öğretmenleriyle bir toplantı yapılarak, onlara uygulamanın ayrıntıları hakkında bir bilgilendirme yapılmıştır. Toplantının sonunda 3. ve 4. sınıf öğrencilerinin yapılacak deneme uygulamalarında kullanacakları öğretim araçları yeterli düzeyde olacak şekilde ilgili sınıfın öğretmenlerine öğrencilere dağıtmaları için verilmiştir.

Testler 2018-2019 Eğitim-öğretim yılının 1. döneminde Ordu ili Altınordu ilçesinde uygulanmıştır. Deneme formlarının oluşturulması sırasında kararlaştırıldığı gibi her bir testin 40 dakikada yapılmasına ve aynı gün içerisinde en fazla iki testin uygulanacak olmasına dikkat edilmiştir.

Deneme Uygulaması Cevap Kâğıtlarının Puanlanması, Madde Analizi ve Madde Seçimi

Matematik başarı testlerinin deneme uygulamalarından elde edilen veriler TESTAN programı ile analiz edilmiştir. Madde ayırt edicilik indeksi, maddenin ilgili davranışa sahip olanla olmayanı ne ölçüde ayırdığını gösterir (Turgut ve Baykul, 2015). Crocker ve Algina'ya (1986) göre teste konulacak maddeler için Tablo 4'teki değerler genel olarak ölçüt kabul edilir.

Tablo 4

Madde Ayırt Edicilik İndeksi Tablosu

Madde Ayırtıcılık Gücü	Madde Değerlendirmesi
0.19 ve daha küçük olan maddeler	Teste konulmaz
0.20-0.29 arasında olan maddeler	Teste düzeltilerek konulabilir
0.30 ve daha büyük olan maddeler	Teste aynen konulabilir

Yapılan analiz sonuçlarına göre Tablo 4'teki bilgiler dikkate alınarak madde ayırt edicilik indeksi 0.20'den küçük olan maddeler elenmiştir. Testlerin nihai formuna madde ayırt edicilik indeksi 0.20 - 0.29 değer aralığında olan maddelerin düzenlenerek, 0.30'dan büyük olan maddelerin ise olduğu gibi konulmasına karar verilmiştir.

Deneme uygulamalarından elde edilen analiz sonuçlarına göre 3. sınıf için kritik kazanımları ölçmeye yönelik olarak hazırlanan maddelerden bazılarının ayırt edicilik indeksi 0.20'nin altında bulunmuştur. Bu nedenle bu kritik kazanımları nihai testlerde temsil edecek yeterli miktarda eşdeğer

madde elde edilememiştir. Bunun üzerine bu kritik kazanımlara yönelik 25 sorudan oluşan yeni bir test hazırlanarak bu testin deneme uygulaması 4. sınıfta öğrenim gören 101 öğrenci ile yapılmıştır. Deneme uygulamasından elde edilen veriler TESTAN programı ile analiz edilmiş ve analiz sonuçlarına göre 3. sınıfın nihai testlerinde kullanılmak üzere aynı kazanımı ölçen en az 2 eşdeğer madde elde edilmiştir. Aynı şekilde 4. sınıf için yapılan analiz sonuçlarına göre de belirlenen kritik kazanımlardan her birini ölçmeye yönelik en az 2 eşdeğer maddenin ek bir uygulamaya gerek kalmadan elde edildiği görülmüştür.

İlkokul 3. ve 4. sınıflar için eşdeğer testler oluşturulurken aynı kazanımı ölçen ikiden çok maddenin olduğu durumlarda, uzman görüşü alınarak bu maddelerden biçim ve özellik olarak birbirine en yakın olan 2 tanesi seçilerek testlere yerleştirilmiştir.

Nihai Testin Oluşturulması ve İstatistiklerinin Kestirilmesi

Birbirine eşdeğer olarak oluşturulması amaçlanan ve son şekilleri verilen nihai testlerde, 3. sınıf düzeyinde 28'er madde, 4. sınıf düzeyinde ise 22'şer madde bulunmaktadır. Yapılan deneme uygulamaları sonucunda nihai testlere seçilen maddelerin ayırt edicilik indekslerine ait değer aralığı Tablo 5'de verilmiştir.

Tablo 5
Nihai Testlere Seçilen Maddelerin Ayırt Edicilik İndeksleri

Sınıf Düzeyi	Test 1	Test 2
3. sınıf	0.296 – 0.730	0.222 – 0.762
4. sınıf	0.333 – 0.768	0.222 – 0.818

Madde güçlüğü, maddeyi doğru cevaplayanları yüzde olarak ifade etmekte ve maddenin zorluğu hakkında fikir vermektedir. Yapılan deneme uygulamaları sonucunda nihai testlere seçilen maddelerin güçlük indekslerine ait değer aralığı Tablo 6'da verilmiştir.

Tablo 6
Nihai Testlere Seçilen Maddelerin Güçlük İndeksleri

Sınıf Düzeyi	Test 1	Test 2
3. sınıf	0.151 – 0.817	0.185 – 0.889
4. sınıf	0.500 – 0.815	0.400 – 0.852

Bulgular

Matematik Başarı Testlerinin Güvenirliğinin Belirlenmesi

Geliştirilen 3. ve 4. sınıf matematik başarı testlerinin güvenilirliklerinin hesaplanmasında Kuder-Richardson 20 (KR-20) yöntemi kullanılmıştır. Nihai testlerin asıl uygulamalar sonucunda elde edilen analiz sonuçlarına göre KR-20 güvenilirlik katsayıları 3. ve 4. Sınıf testleri için Tablo 7'de verilmiştir.

Tablo 7
Nihai Testlere Seçilen Maddelerin KR-20 Güvenirlik Katsayıları

Sınıflar	Test-1	Test-2
3. Sınıf	0.834	0.814
4. Sınıf	0.841	0.828

Tablo 7'deki değerler göz önünde bulundurulduğunda 3. ve 4. sınıflar için geliştirilen her bir başarı testinin güvenilirliğinin yüksek olduğu yorumu yapılabilir.

Matematik Başarı Testlerinde Yer Alan Maddelerin Ayırt Edicilik İndekslerinin Belirlenmesi

İlkokul 3. sınıf için geliştirilen matematik başarı testlerinin nihai formları altı devlet ilkokulunda 4. sınıfta öğrenim gören 202 öğrenciye uygulanmış ve bu uygulamalar sonucunda elde edilen veriler üzerinden madde analizleri yapılmıştır. Analiz sonuçları Tablo 8’de verilmiştir.

Tablo 8

3. Sınıf Matematik Başarı Testlerindeki Maddelerin Ayırt Edicilik İndeksleri

MADDE NO	AYIRT EDİCİLİK İNDEKSİ	
	TEST-1	TEST-2
1	0.400	0.727
2	0.455	0.418
3	0.382	0.709
4	0.364	0.364
5	0.509	0.455
6	0.345	0.636
7	0.655	0.636
8	0.691	0.727
9	0.545	0.564
10	0.527	0.600
11	0.218*	0.302
12	0.691	0.491
13	0.545	0.564
14	0.436	0.365
15	0.364	0.618
16	0.600	0.433
17	0.618	0.455
18	0.618	0.345
19	0.436	0.618
20	0.655	0.618
21	0.655	0.455
22	0.564	0.273*
23	0.636	0.527
24	0.400	0.673
25	0.636	0.600
26	0.491	0.527
27	0.455	0.491
28	0.364	0.600
Ortalama	0.509	0.528

*Uzman görüşü alınarak yeniden düzenlenen maddeler.

Analiz sonuçlarına göre maddelerin ayırt edicilik indeksleri 0.20 – 0.30 aralığında olan iki maddenin uzman görüşü alınıp yeniden düzenlenerek teste konulmasına karar verilmiştir.

İlkokul 4. sınıf için geliştirilen matematik başarı testlerinin nihai formları dört devlet ortaokulunda 5. sınıfta öğrenim gören 221 öğrenciye uygulanmış ve bu uygulamalar sonucunda elde edilen veriler üzerinden madde analizleri yapılmıştır. Analiz sonuçları Tablo 9’da verilmiştir.

Tablo 9

4. Sınıf Matematik Başarı Testlerindeki Maddelerin Ayırt Edicilik İndeksleri

MADDE NO	AYIRT EDİCİLİK İNDEKSİ	
	TEST-1	TEST-2
1	0.467	0.617
2	0.433	0.467
3	0.633	0.550
4	0.583	0.533
5	0.667	0.800

6	0.833	0.657
7	0.517	0.767
8	0.650	0.650
9	0.467	0.583
10	0.667	0.667
11	0.783	0.733
12	0.683	0.467
13	0.567	0.567
14	0.267*	0.383
15	0.633	0.533
16	0.717	0.500
17	0.633	0.783
18	0.567	0.350
19	0.683	0.833
20	0.783	0.633
21	0.683	0.667
22	0.733	0.450
Ortalama	0.620	0.599

*Uzman görüşü alınarak yeniden düzenlenen maddeler.

Analiz sonuçlarına göre maddelerin ayırt edicilik indeksleri 0.20 – 0.30 aralığında olan bir maddenin uzman görüşü alınıp yeniden düzenlenerek teste konulmasına karar verilmiştir.

Matematik Başarı Testlerinde Yer Alan Maddelerin Güçlük İndekslerinin Belirlenmesi

Nihai testler oluşturulurken madde seçiminde madde ayırt edicilik indeksleri ana unsur olarak dikkate alınmıştır. Eşdeğer testler oluşturulurken maddelerin testlere dağılımında madde güçlük indekslerine bakılarak testlerin ortalama güçlüklerinin yakın olmasına özen gösterilmiştir. Oluşturulan nihai testlerle asıl uygulamalar yapılmış ve elde edilen veriler sonucunda tekrar madde analizleri yapılarak 3. sınıf matematik başarı testlerindeki maddelerin güçlük indeksleri Tablo 10’da verilmiştir.

Tablo 10

3. Sınıf Matematik Başarı Testlerindeki Maddelerin Güçlük İndeksleri

MADDE NO	GÜÇLÜK İNDEKSİ	
	TEST-1	TEST-2
1	0.782	0.618
2	0.773	0.773
3	0.700	0.591
4	0.727	0.782
5	0.727	0.718
6	0.809	0.609
7	0.618	0.627
12	0.655	0.755
13	0.655	0.664
14	0.727	0.718
15	0.800	0.618
16	0.664	0.382
17	0.382	0.736
18	0.582	0.718
19	0.764	0.527

20	0.673	0.636
21	0.418	0.500
22	0.409	0.845
23	0.627	0.700
24	0.527	0.573
25	0.645	0.682
26	0.391	0.591
27	0.664	0.536
28	0.727	0.664
Ortalama	0.627	0.628

Tablo 10'daki analiz sonuçlarına göre 3. sınıf matematik başarı testlerinde orta güçlükteki maddelerin çoğunlukta olduğu yorumu yapılabilir. Testin geçerliği artırılmak isteniyorsa, orta güçlükteki maddelerden oluşturulması uygun düşer (Gömleksiz ve Erkan, 2010).

İlkokul 4. sınıf matematik başarı testlerinde yer alan maddelerin güçlük indeksleri Tablo 11'de verilmiştir.

Tablo 11

4. Sınıf Matematik Başarı Testlerindeki Maddelerin Güçlük İndeksleri

MADDE NO	GÜÇLÜK İNDEKSİ	
	TEST-1	TEST-2
1	0.700	0.542
2	0.750	0.617
3	0.633	0.692
4	0.692	0.717
5	0.533	0.567
6	0.533	0.558
7	0.642	0.617
8	0.625	0.642
9	0.500	0.525
10	0.583	0.583
11	0.542	0.533
12	0.558	0.733
13	0.650	0.683
14	0.817	0.775
15	0.633	0.633
16	0.608	0.600
17	0.533	0.508
18	0.600	0.542
19	0.608	0.583
20	0.542	0.667
21	0.642	0.667
22	0.500	0.775
Ortalama	0.610	0.625

Tablo 11'deki analiz sonuçlarına göre İlkokul 3. sınıf matematik başarı testlerinde olduğu gibi 4. sınıf matematik başarı testlerinde de orta güçlükteki maddelerin çoğunlukta olduğu görülmektedir. Orta güçlükteki maddelerin ağırlıkta kullanıldığı testler, bilenlerle bilmeyenler arasındaki değişkenliği en iyi tanımlamaya hizmet edebilir. (Gömleksiz vd., 2010).

Matematik Başarı Testlerinin Yapı Geçerliğinin Belirlenmesi

Matematik başarı testlerinin yapı geçerliği belirlenirken testler, geliştirildikleri sınıf düzeyindeki kazanımları görmüş final uygulamasına katılan öğrencilerin yanında, bu kazanımları görmemiş olan bir alt sınıftaki öğrencilere de uygulanmıştır. Uygulama Ordu ilinde 2. dönemin sonuna doğru gerçekleştirilmiştir. Uygulama sonucunda elde edilen veriler normal dağılım göstermediğinden dolayı analizlerde parametrik olmayan Mann Whitney-U testi kullanılmıştır.

İlkokul 3. sınıf için geliştirilen matematik başarı testlerinin yapı geçerliğini incelemek amacıyla, testler final uygulamasına katılan 202 ilkokul 4. sınıf öğrencisinin yanında 120 ilkokul 2. sınıf öğrencisine uygulanmıştır.

İlkokul 3. sınıf için geliştirilen 1. Matematik Başarı Testinin yapı geçerliğini belirlemeye yönelik yapılan analiz sonuçları Tablo 12’de verilmiştir.

Tablo 12

3. Sınıf İçin Geliştirilen 1. Matematik Başarı Testinin Yapı Geçerliğini Belirlemeye Yönelik yapılan Mann Whitney-U Testi Sonucu

Ölçek	Gruplar	N	Sıra Ortalama	Sıra Toplamları	U	Z	p
3. Sınıf 1. Matematik Testi	2. Sınıf	120	84.75	41833.00	2910.00	-11.417	0.000
	4. Sınıf	202	207.09	41833.00			
	Toplam	322					

Tablo 12 incelendiğinde grupların matematik başarı testinden aldıkları puanlar arasında anlamlı düzeyde bir farkın olduğu tespit edilmiştir ($p<0.05$). Analiz sonuçlarına göre geliştirilen 3. sınıf 1. Matematik Başarı Testinin yapı geçerliğinin sağlandığı söylenebilir.

Tablo 13

3. Sınıf İçin Geliştirilen 2. Matematik Başarı Testinin Yapı Geçerliğini Belirlemeye Yönelik yapılan Mann Whitney-U Testi Sonucu

Ölçek	Gruplar	N	Sıra Ortalama	Sıra Toplamları	U	Z	p
3. Sınıf 2. Matematik Testi	2. Sınıf	120	82.88	9945.50	2685.50	-11.698	0.000
	4. Sınıf	202	208.21	42057.50			
	Toplam	322					

Tablo 13 incelendiğinde grupların matematik başarı testinden aldıkları puanlar arasında anlamlı bir fark bulunmuştur ($p<0.05$). Analiz sonuçlarına göre geliştirilen 3. sınıf 2. Matematik Başarı Testinin de yapı geçerliğinin sağlandığı söylenebilir.

İlkokul 4. sınıf için geliştirilen matematik başarı testlerinin yapı geçerliğini incelemek amacıyla, testler final uygulamasına katılan 221 ortaokul 5. sınıf öğrencisinin yanında, 125 ilkokul 3. sınıf öğrencisine uygulanmıştır.

İlkokul 4. sınıf için geliştirilen 1. Matematik Başarı Testinin yapı geçerliğini belirlemeye yönelik analiz sonuçları Tablo 14’te verilmiştir.

Tablo 14

4. Sınıf İçin Geliştirilen 1. Matematik Başarı Testinin Yapı Geçerliğini Belirlemeye Yönelik yapılan Mann Whitney-U Testi Sonucu

Ölçek	Gruplar	N	Sıra Ortalama	Sıra Toplamları	U	Z	p
4. Sınıf 1. Matematik Testi	3. Sınıf	125	119.87	14984.00	7109.00	-7.512	0.000
	5. Sınıf	221	203.83	45047.00			
	Toplam	346					

Tablo 14 incelendiğinde grupların matematik başarı testinden aldıkları puanlar arasında anlamlı düzeyde bir farkın olduğu tespit edilmiştir ($p<0.05$). Analiz sonuçlarına göre geliştirilen 4. sınıf 1. Matematik Başarı Testinin yapı geçerliğinin sağlandığı söylenebilir.

İlkokul 4. sınıf için geliştirilen 2. Matematik Başarı Testinin yapı geçerliğini belirlemeye yönelik analiz sonuçları Tablo 15’te verilmiştir.

Tablo 15

4. Sınıf İçin Geliştirilen 2. Matematik Başarı Testinin Yapı Geçerliğini Belirlemeye Yönelik yapılan Mann Whitney-U Testi Sonucu

Ölçek	Gruplar	N	Sıra Ortalama	Sıra Toplamları	U	Z	p
4. Sınıf 2. Matematik Testi	3. Sınıf	125	116.01	14501.00	6626.00	-8.055	0.000
	5. Sınıf	221	206.02	45530.00			
	Toplam	346					

Tablo 15'te görüldüğü üzere grupların matematik başarı testinden aldıkları puanlar arasında anlamlı bir fark bulunmuştur ($p < 0.05$). Analiz sonuçlarına göre geliştirilen 4. sınıf 2. Matematik Başarı Testinin de yapı geçerliğinin sağlandığı söylenebilir.

Matematik Başarı Testlerinin Eşdeğer Olup Olmadığının Belirlenmesi

İlkokul 3. ve 4. sınıf için geliştirilen matematik başarı testlerinin final uygulamalarından elde edilen veriler normal dağılım göstermemiştir. Her iki sınıf seviyesinde yapılan final uygulamalarından elde edilen veriler normal dağılım göstermediğinden geliştirilen nihai testlerin aralarındaki ilişki Spearman korelasyon katsayısı hesaplanarak incelenmiş ve elde edilen analiz sonuçları Tablo 16'da sunulmuştur.

Tablo 16

Nihai Testlerin Aralarındaki Korelasyon Katsayıları

Testler	Korelasyon Katsayısı	
	3. Sınıf Test-2	4. Sınıf Test-2
3. Sınıf Test-1	0.880	-
4. Sınıf Test-1	-	0.829

Tablo 16'ya bakıldığında 3. sınıf matematik başarı testleri arasında yüksek düzeyde bir ilişki olduğu görüldüğü için bu testlerin eşdeğer olduğu söylenebilir. Benzer şekilde 4. sınıf matematik başarı testleri arasında da yüksek düzeyde bir ilişki olduğu ifade edilebileceği için bu testlerin de eşdeğer olduğu söylenebilir. Korelasyon katsayısının 0.70 – 1.00 aralığında değer alması ilişkinin yüksek olduğunu göstermektedir (Başol, 2016).

Sonuç, Tartışma ve Öneriler

Araştırma sonucunda geçerliği ve güvenilirliği sağlanmış birbirine eş değer olan 3. sınıf için iki, 4. sınıf için iki olmak üzere, toplam dört matematik başarı testi geliştirilmiştir. Geliştirilen İlkokul 3.sınıf matematik başarı testlerinde 28'er soru bulunmakta olup testler arasındaki Spearman korelasyon katsayısı 0.880 olarak hesaplanmıştır. Testlerin KR-20 güvenilirlik katsayıları sırasıyla 0,834 ve 0.814; ortalama güçlük değerleri ise sırasıyla 0.627 ile 0.628 olarak bulunmuştur. Oluşturulan İlkokul 4.sınıf matematik başarı testlerinde ise 22'şer soru bulunmakta olup testler arasındaki Spearman korelasyon katsayısı 0.829 olarak hesaplanmıştır. Testlerin KR-20 güvenilirlik katsayıları sırasıyla 0,841 ve 0.828; ortalama güçlük değerleri ise sırasıyla 0.610 ile 0.625 olarak bulunmuştur.

Başarı testi geliştirilmesine yönelik literatürde birçok çalışma bulunmaktadır. Bu çalışmalar incelendiğinde matematik öğretim programlarındaki kazanımlar doğrultusunda geliştirilen testlerin olduğu görülmektedir (Başer, 1996; Duru, 2007; Üçüncü, 2010; Fidan, 2013; Erdoğan, 2015; Şahin ve Keşan, 2017). Yapılan bu çalışmada da yöntem olarak diğer başarı testi geliştirme çalışmalarıyla benzer test geliştirme aşamaları kullanılmıştır.

Çalışmadan elde edilen bulgulara dayanarak yapılacak araştırmalara ışık tutması ve yol göstermesi bakımından aşağıdakiler önerilebilir:

1. Geliştirilen matematik başarı testleri, ilgili sınıf düzeyindeki akademik çalışmalarda öğrencilerin başarılarının belirlenmesinde kullanılabilir.
2. Testler birbirine eşdeğer iki formdan oluştuğundan kullanılması öngörülen akademik çalışmalarda araştırmacılara ön test ve son test için seçenek oluşturabilir.
3. Öğretmenler, matematik öğretim programındaki kazanımlar doğrultusunda geliştirilen testleri eğitim-öğretim yılının sonunda öğrencilerinin başarı durumlarını belirlemek amacıyla kullanabilirler.
4. Öğretmenler, testleri eğitim-öğretim yılının başında öğrencilerin matematik dersindeki hazırbulunuşluk düzeylerinin belirlenmesinde kullanarak bir görüşe varabilirler.
5. Geliştirilen matematik testlerine ait madde analizleri uygulama yapılan deneme grubu ile sınırlıdır. Bu nedenle testler daha farklı örneklem grupları üzerinde yeniden uygulanarak testlerin madde analizleri yapılabilir.
6. Testler 3. ve 4. sınıfta öğrenim gören öğrencilerin matematik başarısını belirlemek amacıyla geliştirilmiştir. Farklı sınıf seviyelerinde matematik başarı testleri geliştirilebilir.,
7. Uygulamada ekonomiklik ve kullanılabilirlik olması istendiğinden belirlenen kritik kazanımlar dışındaki kazanımlar elenmiştir. Bu nedenle bu çalışmada elenmiş olan kazanımların da dâhil edildiği geniş çaplı bir başarı testi geliştirilebilir.
8. Çalışmada öğrencilerin matematik başarısını ölçebilmek amacıyla çoktan seçmeli testler geliştirilmiştir. Öğrencilerin matematik başarılarını ölçebilmek için daha farklı ölçme araçları geliştirilebilir.
9. İlgili sınıf seviyeleri için geliştirilen testlerin bir üst sınıfta dönemin başında uygulanması, öğretmenlere öğrencilerin hazırbulunuşluk düzeyleri hakkında fikir verebilir.
10. Testlerde öğretim araçlarının kullanılmasına yönelik maddeler yer almaktadır. Bundan sonraki çalışmalarda geliştirilecek başarı testlerinde de öğretim araçlarının kullanımına yönelik maddelere yer verilebilir.

Yazarların Katkı Oranı

Çalışmaya 1. yazar %70 oranında, 2. yazar %30 oranında katkı sağlamıştır.

Çıkar Çatışması

Bu çalışmada çıkar çatışması ortaya çıkarabilecek herhangi bir durum yoktur.

Kaynaklar

- Altun, M. (2011). *Matematik Öğretimi* (16. Baskı). Bursa: Aktüel Alfa Akademi Yayıncılık.
- Atılğan, H., Kan, A., ve Doğan, N. (2013). *Eğitimde ölçme ve değerlendirme*. Anı Yayıncılık.
- Ayas, A., ve Demirbaş, A. (1997). Turkish secondary students' conception of introductory chemistry concepts. *Journal of Chemical Education*, 74(5), 518-521.
- Başer, N. (1996). *Ders Geçme ve Kredi Sisteminde Lise Öğrencileri İçin Bir Matematik Başarı Testi Tasarımı ve Uygulanabilirliğinin Araştırılması*. Doktora tezi. Dokuz Eylül Üniversitesi/Sosyal Bilimler Enstitüsü, İzmir.

- Başol, G. (2016). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem Akademi Yayıncılık.
- Baykul, Y. (1992). Eğitim sisteminde değerlendirme. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 7(7).
- Cizek, G. J. (1997). Learning, achievement and assessment: constructs at a crossroads. In G. D. Phye (Ed.). *Handbook of Classroom Assessment. Learning, Achievement and Adjustment*. California: Academic Press.
- Crocker, L. ve Algina, J. (1986). *Introduction to Classical & Modern Test Theory*. New York: Holt, Rinehalt and Winston.
- Cronbach, L.J. (1970). *Essentials of psychological testing*. New York: Harper and Row.
- Çakan, M. (2004). Öğretmenlerin ölçme-değerlendirme uygulamaları ve yeterlik düzeyleri: İlk ve ortaöğretim. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 37(2), 99-114.
- Dane, A., Kudu, M. ve Balkı, N. (2009). Lise Öğrencilerinin Algılarına Göre Matematik Başarısını Olumsuz Yönde Etkileyen Faktörler. *Erzincan Üniversitesi, Fen Bilimleri Enstitüsü Dergisi*, 2(1), 7-35.
- Dede, Y. ve Dursun, Ş. (2004). Öğrencilerin matematik başarısını etkileyen faktörler: Matematik öğretmenlerinin görüşleri bakımından. *Gazi Üniversitesi Eğitim Fakültesi Dergisi*, 24(2), 217-233.
- Duru, E.D. (2007). *Oluşturmacı yaklaşıma göre tasarlanmış bir ders paketinin uygulanmasıyla ulaşılan başarının geliştirilen başarı testi ile değerlendirilmesi*. Yüksek Lisans Tezi. Dokuz Eylül Üniversitesi/Eğitim Bilimleri Enstitüsü, İzmir.
- Ebel, R. L. (1965). *Measuring Educational Achievement*. New Jersey: Prentice-Hall Education Series.
- Erdoğan, F. (2015). *İşbirlikli öğrenme yönteminin ilkokul 4. sınıf matematik dersinde öğrencilerin akademik başarılarına ve üst bilişsel farkındalıklarına etkileri*. Yüksek Lisans Tezi. Dokuz Eylül Üniversitesi/Eğitim Bilimleri Enstitüsü, İzmir.
- Eskiocak, S., Gökmen, S. S., Erbaş, H., Çakır, E. ve Gülen, Ş. (2004). Trakya Üniversitesi Tıp Fakültesinde Son 5 Yılda Yapılan Biyokimya Sınav Sorularının Analizi. *Turk J Biochem*, 29: 273-276.
- Fidan, E. (2013). *İlkokul öğrencileri için matematik dersi sayılar öğrenme alanında başarı testi geliştirilmesi*. Yüksek Lisans Tezi. Ankara Üniversitesi/Eğitim Bilimleri Enstitüsü, Ankara.
- Gömlüksiz, M. ve Erkan, S. (2010). *Eğitimde ölçme ve değerlendirme* (2. Baskı). Ankara: Nobel Yayın Dağıtım.
- Heritage, M. (2007). What do teachers need to know and do?. *Phi Delta Kappan*, 89(2), 140 – 145.
- Hewson, M.G. ve Hewson, P.W. (1983). Effect of instruction using students' prior knowledge and conceptual change strategies on science learning. *Journal of Research in Science Teaching*, 20(8), 731-743.
- İlhan, M., Çetin, B. ve Kinay, İ. (2015). Standart Testlere Yönelik İnanç Ölçeği'nin (STYİÖ) Türkçe Uyarlaması: Geçerlik ve Güvenirlik Çalışması. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 12(29).

- Kaya, A., Bozaslan, H. ve Genç, G. (2012). Üniversite öğrencilerinin anne-baba tutumlarının problem çözme becerilerine, sosyal kaygı düzeylerine ve akademik başarılarına etkisi. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, 18(1), 208-225.
- Kemertaş, İ. (1999). *Uygulamalı genel öğretim yöntemleri: "Öğretimde planlama ve değerlendirme"* İstanbul: Birsen Yayınevi.
- Koç, N. (1985). Standart başarı testlerinin, bir eğitim sisteminde verilen çeşitli kararlardaki yeri ve önemi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 17(1-2). Ankara: 159-172.
- Millî Eğitim Bakanlığı (2018). Matematik dersi öğretim programı (ilkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. Sınıflar).
- Nakhleh, M.B. (1992), Why some students don't learn chemistry. *Journal of Chemical Education*, 69(3), 191-196.
- Özçelik, D. A. (2010). *Test Hazırlama Kılavuzu*. Ankara: Pegem Akademi.
- Özçelik, D.A. (1998). *Ölçme ve Değerlendirme*. Ankara: ÖSYM Yayınları.
- Pardo, J.Q. ve Partoles, J.J.S. (1995). Students and teachers misapplication of Le Chatelier's principle: implications for the teaching of chemical equilibrium. *Journal of Research in Science Teaching*, 32(9), 939-957.
- Sezgin, M. (2007). *Öğrencilerin matematik başarısına etki eden faktörler (10.sınıf örneği)*. Yüksek Lisans Tezi. Beykent Üniversitesi/Sosyal Bilimler Enstitüsü, İstanbul.
- Shipley, K. ve McAfee, J. (2009). *Assessment in speech-language pathology: A resource manual*. Clifton Park, NY: Delmar Cengage Learning.
- Şahin, Z., ve Keşan, C. (2017). Bir başarı testi geliştirme çalışması: Beşinci sınıf öğrencilerinin geometri kazanımlarını ölçmeye yönelik bir başarı testi geçerlik ve güvenilirliğin araştırılması. *International Journal of New Trends in Arts, Sports & Science Education (IJTASE)*, 6(3), 47-57.
- Şimşek, Ö., Bars, M. ve Zengin, Y. (2017). Matematik Öğretiminin Ölçme ve Değerlendirme Sürecinde Bilgi ve İletişim Teknolojilerinin Kullanımı. *Uluslararası Eğitim Programları ve Öğretim Çalışmaları Dergisi*, 7(13), 189-207.
- Tachie, S. A. ve Chrishe, R. (2013). High failure rate in mathematics examinations in rural senior secondary schools in Mthatha District, Eastern Cape: Learners' attributions. *Stud Tribes Tribals*, 11(1), 67-73.
- Tekindal, S. (Ed.) (2011). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem A Yayıncılık.
- Turgut, M. F. ve Baykul, Y. (2015). *Eğitimde ölçme ve değerlendirme (Yedinci Baskı)*. Pegem A yayıncılık, Ankara.
- Turgut, M. F. ve Baykul, Y. (2015). *Eğitimde ölçme ve değerlendirme (Yedinci Baskı)*. Pegem A yayıncılık, Ankara.
- Zoller, U. (1990). Students' misunderstandings and misconceptions in college freshman chemistry (General and Organic). *Journal of Research in Science Teaching*, 27(10), 1053-1065.

Extended Abstract

Introduction

Assessment is an important stage of teaching activities in mathematics education. Because, using measurement tools while describing academic achievement in Mathematics lesson will enable teachers to obtain more objective information numerically. Thanks to this information, teachers will be able to have an idea about the effectiveness of education and make decisions for the development of their students. Valid and reliable measurements are needed for the decisions to be taken to be realistic (Atılğan, Kan, & Doğan, 2013). Valid and reliable measurements will be more possible with standardized tests that teachers will use in measurement.

Method

The aim of the study is to develop valid, reliable and equivalent tests to determine the achievement level of primary school 3rd and 4th grade students in Math classes. Learning outcomes of Math curricula in the relevant grade were taken as basis in the development process of the tests. In this respect, math learning acquisitions of primary school 3rd and 4th grade classes were examined in order to determine a critical learning outcome as it would be neither easy nor economical to apply the test if it involved questions for each of these learning outcomes. A total of 301 questions including 172 for the primary school 3rd grade and 129 for 4th grade were prepared for the critical outcomes identified and presented to expert opinion. Necessary corrections were made on the questions based on the feedbacks received from the experts and 80 questions were selected for each of the trial applications of primary school 3rd and 4th grades.

Trial and final applications of the tests were carried out in Ordu province in 2018-2019 academic year with upper class students who have already studied all of the learning outcomes of the relevant grade. Distinctiveness values of items were considered while selecting the items for the final tests based on the analysis results obtained from trial applications and attention was given to include items related to each of the critical learning outcomes identified. However, difficulty values of the items were also considered when developing equivalent tests and items were appointed to the tests in such a way that test difficulty levels would be similar. 28 items were included in the tests of the primary school 3rd grade while 22 items were included in the tests of the 4th grade which were planned to be equivalent.

Final forms of Maths achievement tests developed for the primary school 3rd grade were applied on 202 4th grade students in six different public primary schools and the final forms of Maths achievement tests developed for the 4th grade were applied on 221 5th grade students in four different public secondary schools. Two test items used in the test for the primary school 3rd grade and one test item used in the test for the primary school 4th grade were found to have a distinctiveness between 0.20-0.30, which were then modified in line with expert opinions and included back in the tests. It was observed that primary school 3rd grade tests and 4th grade tests mainly involved items of moderate difficulty level. K-20 method was used in the calculation of reliability coefficients of the tests developed. Reliability coefficients of the primary school 3rd grade tests were calculated as 0.834 and 0.814, respectively and mean difficulty levels as 0.627 and 0.628, respectively. Reliability coefficients of the primary school 4th grade tests were calculated as 0.841 and 0.828, respectively and mean difficulty levels as 0.610 and 0.625, respectively. As the data obtained from final applications carried out for both grades did not present a normal distribution, the correlation among the final tests developed was examined by calculating the Spearman correlation coefficient. It was found that the Spearman correlation coefficient among the tests of the 3rd grade was 0.880 whereas the Spearman correlation coefficient among the tests of the 4th grade was 0.829.

Findings

These tests were also applied to students studying at one grade lower than the grade they had been developed for in order to examine the construct validity of the achievement tests. As the data obtained at the end of the application did not present a normal distribution, Mann Whitney-U test was used in analyses. According to the results of the analysis, it was seen that construct validity was obtained in each test.

Conclusion, Discussion and Recommendations

As a result of all these applications, a total of four valid, reliable and equivalent Maths achievement tests consisting of two tests for the 3rd grade and two tests for the 4th grade were obtained.

The tests were developed to determine the mathematics achievement of students in 3rd and 4th grade. Mathematics achievement tests can be developed at different grade levels. In addition, tests can be developed for different courses with a similar method.

Ekler

Geliştirilen başarı testlerinin tamamına Oğuz Balcı'nın "İLKOKUL 3. VE 4. SINIF ÖĞRENCİLERİNE YÖNELİK MATEMATİK BAŞARI TESTİ GELİŞTİRİLMESİ" başlıklı Yüksek Lisans Tezi'nden ulaşılabilir.

3. SINIF İÇİN GELİŞTİRİLEN 1. MATEMATİK BAŞARI TESTİ İÇİN ÖRNEK MADDELER

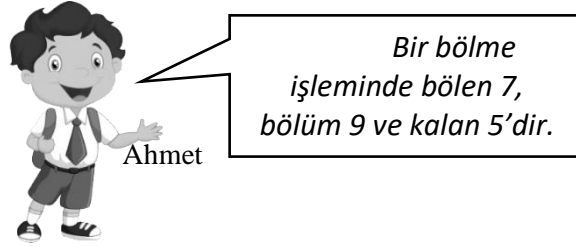
Örnek Madde 1:

"Beş yüz beş"

Yukarıda okunuşu verilen sayı ile ilgili aşağıdaki ifadelerden hangisi doğrudur?

- A) Onlar basamağında 5 rakamı vardır.
- B) Yüzler basamağındaki rakamın basamak değeri 500'dir.
- C) Beş basamaklı bir sayıdır.

Örnek Madde 2:



Ahmet'in ifade ettiği bölme işleminde bölünen sayı kaçtır?

- A) 63
- B) 68
- C) 72

Örnek Madde 3:

Her iki günde bir 6 litre içme suyu tüketen bir aile 10 günde kaç litre içme suyu tüketmektedir?

- A) 15
- B) 30
- C) 60

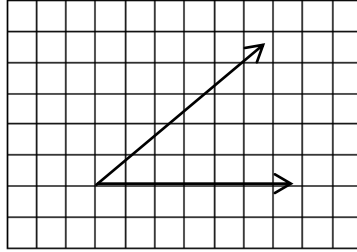
4. SINIF İÇİN GELİŞTİRİLEN 1. MATEMATİK BAŞARI TESTİ İÇİN ÖRNEK MADDELER

Örnek Madde 1:

8 katı 2048 olan sayının 44 fazlası kaçtır?

- A) 256 B) 100 C) 344 D) 300

Örnek Madde 2:



Yukarıda kareli kâğıtta verilen açının ölçüsü kaç derecedir?
(Bu soruda açıölçer kullanınız.)

- A) 40 B) 50 C) 60 D) 70

Örnek Madde 3:

Oğulcan doğduğunda kilosu 2 kg 600 g'dı. Oğulcan'ın kilosu 1.ay 600 g ve 2.ay 800 g artmıştır.

Buna göre 2. ayın sonunda Oğulcan'ın kilosu aşağıdakilerden hangisine eşittir?

- A) 3 kg B) 3 kg 200 g C) 3 kg 600 g D) 4 kg