

ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
İLKÖĞRETİM ANABİLİM DALI
MATEMATİK EĞİTİMİ PROGRAMI

İLKOKUL ÖĞRENCİLERİ İÇİN MATEMATİK DERSİ SAYILAR
ÖĞRENME ALANINDA BAŞARI TESTİ GELİŞTİRİLMESİ

YÜKSEK LİSANS TEZİ

Esra Fidan

Ankara

Temmuz, 2013

ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
İLKÖĞRETİM ANABİLİM DALI
MATEMATİK EĞİTİMİ PROGRAMI

İLKOKUL ÖĞRENCİLERİ İÇİN MATEMATİK DERSİ SAYILAR
ÖĞRENME ALANINDA BAŞARI TESTİ GELİŞTİRİLMESİ

YÜKSEK LİSANS TEZİ

Esra Fidan

Danışman: Prof. Dr. Sinan Olkun

Ankara

Temmuz, 2013

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼'ne,

Bu alıřmaj¼rimiz tarafından İlköđretim Anabilim DalındaY¼KSEK LİSANS
TEZİ olarak kabul edilmiřtir.

Başkan

Prof. Dr. Sinan OLKUN (Danıřman)

¼ye

Yrd. Do. Dr. Didem AKY¼Z

¼ye

Yrd. Do. Dr. Kaan Z¼lfikar DENİZ

Onay

Yukarıdaki imzaların, adı geen öđretim ¼yelerine ait olduđunu onaylarım.

.../.../2013

Prof. Dr. İsmail G¼VEN

Enstit¼ M¼d¼r¼

ÖNSÖZ

Eğitim-öğretim sürecinin etkililiğinin, öğrencilerin hedeflenen kazanımları edinme düzeylerinin, dolayısıyla başarının ölçülmesi için geçerlik, güvenirlik gibi psikometrik özellikler bakımından yeterli, sınıf düzeyine uygun ölçme araçlarına ihtiyaç vardır. İyi hazırlanmış ölçme araçları kullanılarak güvenilir ve kullanışlı bilgiler elde edilmesi, süreçte alınacak eğitimsel kararlara dayanak oluşturma açısından oldukça önemlidir.

Hemen her alanda olduğu gibi matematik alanında da öğrencilerin başarılarının doğru ve güvenilir bir şekilde ölçülmesi oldukça önemlidir. Öğrencilerin matematik dersindeki ilerlemelerini izleyerek öğretim sürecini yönlendirmek ve öğrencilerin bir eğitim-öğretim döneminin sonunda eriştikleri düzeylerin belirlenmesi için geçerli ve güvenilir araçlar kullanılması uygun olacaktır.

Bu çalışmada matematik alanında öğrenci başarısının belirlenmesi yönünde var olan ihtiyaçlardan hareketle ilkokul öğrencilerinin matematik dersi sayılar öğrenme alanındaki başarılarının belirlenmesinde kullanılabilecek testler geliştirilmiştir. Testlerin geçerlik, güvenirlik gibi psikometrik özellikleri irdelenmiş; bu testlerle eğitim-öğretim sürecindeki eksikliklerin saptanarak gerekli geliştirme ve iyileştirmelerin yapılmasına ve matematik eğitimi alanında yapılan başka çalışmalara katkı sağlanacağı düşünülmüştür.

Tez sürecinde; değerli fikirleri ve yapıcı eleştirileri ile yol gösteren danışmanım Prof. Dr. Sinan Olkun'a, her ihtiyaç duyduğumda yardım eden Hacettepe Üniversitesi Ölçme ve Değerlendirme Bölümü Araştırma Görevlisi Sakine Göçer Şahin'e, başta Ayşe Babacan olmak üzere emeği geçen sınıf öğretmenlerine, ayrıca manevi desteğini her an hissettiğim aileme sonsuz teşekkürlerimi sunarım.

Esra Fidan

ÖZET

İLKOKUL ÖĞRENCİLERİ İÇİN MATEMATİK DERSİ SAYILAR ÖĞRENME ALANINDA BAŞARI TESTİ GELİŞTİRİLMESİ

Esra Fidan

Yüksek Lisans, İlköğretim Anabilim Dalı

Danışmanı: Prof. Dr. Sinan Olkun

Mayıs 2013, xiv + 112 sayfa

Bu çalışmanın amacı, ilkokul öğrencilerinin matematik dersi sayılar öğrenme alanındaki başarı düzeylerini belirlemeye yönelik geçerli ve güvenilir başarı testleri geliştirilmesidir. Bu amaçla, öncelikle 1-4. sınıf matematik dersi öğretim programında bulunan kazanımlar incelenerek ölçülecek kritik kazanımlar belirlenmiş, daha sonra kritik kazanımlar için sorular yazılarak hazırlanan deneme formları ön uygulama ve uzman görüşleri ile düzenlenmiş ve her sınıf düzeyindeki testin yaklaşık 300 öğrenci ile pilot uygulaması yapılmıştır. Araştırmaya, İstanbul ilinde farklı sosyoekonomik düzeylerdeki 3 ilçede bulunan 6 ilköğretim okulundan toplam 1351 öğrenci katılmıştır. Testlerin pilot uygulamalarından elde edilen veriler üzerinden geçerlik, güvenirlik ve madde analizleri yapılmıştır. Testlerin kapsam geçerliklerinin sağlanmasında hem kritik kazanımların belirlenmesi, hem de soruların yazılması aşamalarında uzman görüşlerine başvurulmuş; ayrıca belirtke tabloları hazırlanmıştır. Yapı geçerliği analizinde, testlerin uygulandığı sınıf düzeyleri arasındaki puan ortalamaları farklılığı için Mann Whitney-U, ilçeler arasındaki puan ortalamaları farklılığı için ise Kruskal Wallis veya ANOVA istatistikleri uygulanmıştır. Ölçüt dayanaklı geçerlik analizi için, öğrencilerin öğretmenlerinden alınan görüşler ile testlerden aldıkları puanlar arasındaki korelasyon hesaplanmıştır. Testlerin güvenirliği KR-20 güvenirlik katsayısı ile belirlenmiştir. Testlerin son formuna seçilecek maddelerin belirlenmesi için maddelere ait güçlük ve ayırtıcılık indeksleri

hesaplanmış, madde istatistikleri değerlendirilerek güçlük ve ayıricılık indeksleri açısından testlere konulmaya uygun olan nitelikli maddelerin seçilmesiyle testler son haline getirilmiştir. Bulgular, testlerin sınıf düzeyleri ve ilçeler arasında beklenen yöndeki farkları ortaya çıkardığını, öğretmen görüşleri ile öğrencilerin test puanları arasındaki korelasyon katsayılarının 1-4. sınıf testleri için sırasıyla 0,79, 0,75, 0,82 ve 0,76 olduğunu göstermektedir. KR-20 güvenirlik katsayıları 1-4. sınıf testleri için sırasıyla 0,80, 0,92, 0,93 ve 0,95 bulunmuştur. Testlerin hesaplanan ortalama güçlük değerleri 1-4. sınıflar için sırasıyla 0,79, 0,74, 0,69 ve 0,59 bulunmuştur. Maddelere ait güçlük ve ayıricılık indeksleri değerlendirilerek mümkün olduğunca orta güçlükte ve yüksek ayıricılığa sahip maddelerin seçilmesiyle testler son haline getirilmiştir. Testlerin son formları 1-4. sınıflar için sırasıyla 13, 15, 16 ve 24 maddeden oluşmaktadır. Sonuç olarak ilkokul 1-4. sınıf matematik dersi sayılar öğrenme alanındaki kazanımlar çerçevesinde yeterli düzeyde geçerli ve güvenilir başarı testleri elde edilmiştir.

Anahtar Kelimeler: Matematik Eğitimi, Test Geliştirme, Başarı Testi

SUMMARY

DEVELOPMENT OF ACHIEVEMENT TESTS IN THE NUMBER DOMAIN OF MATHEMATICS COURSE FOR ELEMENTARY SCHOOL STUDENTS

Esra Fidan

M. S., Department of Primary Education

Thesis Advisor: Prof. Sinan Olkun

May 2013, xiv + 112 pages

The purpose of this study was to develop achievement tests in the number domain of mathematics course for elementary school students. With this purpose, first, by examining the objectives in the 1-4th classes mathematics curricula, critical objectives to be measured have been identified. Then, by composing items for these critical objectives, trial forms of the tests have been prepared and edited by pre-application results and views of experts. Pilot applications of the tests for each grade have been conducted with about 300 students. In total, 1351 students from 6 schools in 3 districts with different socioeconomic status in İstanbul have been participated in the study. On the data gathered in pilot applications of the tests, validity, reliability and item statistics have been conducted. To provide content validity of the tests, views of experts have been asked both in the processes of identification of critical gains and composition of items. Moreover, tables of specifications have been prepared for each test. In the analysis of construct validity, Mann Whitney-U test has been applied for the difference between class levels and Kruskal Wallis or ANOVA tests has been applied for the difference between different districts. For the analysis criterion referenced validity, correlations have been calculated between teachers' evaluation points and test scores of students. Reliability of the tests has been

determined by KR-20 reliability coefficient. Calculating item difficulty and discrimination power indexes and by evaluating these indexes, those items which are qualified enough to be put in the final forms of the tests have been selected. Findings have shown that tests can reveal the expected differences between class levels and different districts and that correlation coefficients between teachers' evaluation points and test scores of students for the 1-4th class tests are 0,79, 0,75, 0,82 and 0,76, respectively. KR-20 reliability coefficients for the 1-4th class tests have found as 0,80, 0,92, 0,93 and 0,95, respectively. Mean difficulty of the 1-4th class tests have been calculated as 0,79, 0,74, 0,69 and 0,59, respectively. By evaluating difficulty and discrimination power indexes and selecting those items with medium difficulty and high discrimination power, tests have been put in the final forms. Final forms of the 1-4th class tests have consisted of 13, 15, 16 and 24 items, respectively. In conclusion, sufficiently valid and reliable achievement tests have been constructed in the framework of the objectives in the number domain of mathematics curricula for 1-4th classes.

Key Words: Mathematics education, test development, achievement test

İÇİNDEKİLER

JÜRİ ÜYELERİNİN İMZA SAYFASI.....	iv
ÖNSÖZ	v
ÖZET	vi
SUMMARY.....	viii
İÇİNDEKİLER	x
ÇİZELGELER LİSTESİ	xii
EKLER LİSTESİ.....	xiv
BÖLÜM 1. GİRİŞ	1
Problem	1
Amaç	9
Önem.....	11
Sayıtlılar.....	12
Sınırlılıklar.....	12
BÖLÜM 2. İLGİLİ ARAŞTIRMALAR	13
Yurt İçinde Yapılan Araştırmalar	13
Yurt Dışında Yapılan Araştırmalar	18
BÖLÜM 3. YÖNTEM.....	22
Araştırma Modeli.....	22
Evren ve Örneklem	22
Veri Toplama Araçları	24
Birinci Sınıf Testi Kritik Kazanımlarının Belirlenmesi ve Uzman Görüşleri ile Netleştirilmesi	27
Birinci Sınıf Testi Kritik Kazanımları İçin Soruların Yazılması ve Görüşleri ile Düzenlenmesi	30
İkinci Sınıf Testi Kritik Kazanımlarının Belirlenmesi ve Uzman Görüşleri ile Netleştirilmesi	32

İkinci Sınıf Testi Kritik Kazanımları İçin Soruların Yazılması ve Görüşleri ile Düzenlenmesi	35
Üçüncü Sınıf Testi Kritik Kazanımlarının Belirlenmesi ve Uzman Görüşleri ile Netleştirilmesi	38
Üçüncü Sınıf Testi Kritik Kazanımları İçin Soruların Yazılması ve Görüşleri ile Düzenlenmesi	41
Dördüncü Sınıf Testi Kritik Kazanımlarının Belirlenmesi ve Uzman Görüşleri ile Netleştirilmesi	44
Dördüncü Sınıf Testi Kritik Kazanımları İçin Soruların Yazılması ve Görüşleri ile Düzenlenmesi	48
Verilerin Toplanması	51
Verilerin Analizi	52
BÖLÜM 4. BULGULAR VE YORUMLAR	56
4. 1. Birinci Sınıf Başarı Testinin Psikometrik Özelliklerine İlişkin Bulgular ve Yorum	56
4. 2. İkinci Sınıf Başarı Testinin Psikometrik Özelliklerine İlişkin Bulgular ve Yorum	62
4. 3. Üçüncü Sınıf Başarı Testinin Psikometrik Özelliklerine İlişkin Bulgular ve Yorum	68
4. 4. Dördüncü Sınıf Başarı Testinin Psikometrik Özelliklerine İlişkin Bulgular ve Yorum	74
BÖLÜM 5. TARTIŞMA VE SONUÇ	82
BÖLÜM 6. ÖNERİLER	93
KAYNAKÇA	94
EKLER	98

ÇİZELGELER LİSTESİ

Çizelge 1. İlköğretim 1-4. Sınıflar Matematik Dersi Kazanımların Öğrenme Alanlarına Dağılımı.....	9
Çizelge 2. Öğrencilerin Farklı İlçe, Okul Ve Sınıf Düzeylerine Dağılımı	23
Çizelge 3. Testlere Göre Katılımcıların Yaş ve Cinsiyet Dağılımları	24
Çizelge 4.1. Sınıf Testinden Alınan Puanların Sınıf Düzeylerine Göre Dağılımı İçin Yapılan Normallik Testi Sonuçları	57
Çizelge 5.1. Sınıf Testi Puanlarının Sınıf Düzeylerine Göre Farklılıkları İçin Uygulanan t-Test Sonuçları	57
Çizelge 6.1. sınıf Testinden Alınan Puanların İlçelere Göre Dağılımı İçin Yapılan Normallik Testi Sonuçları	58
Çizelge 7.1. Sınıf Testi Puanlarının İlçelere Göre Kruskal Wallis H-Testi Sonuçları.....	59
Çizelge 8. İlçelerin 1. Sınıf Test Puan Ortalama Farklılıkları İçin Uygulanan Mann Whitney-U Testi Sonuçları	59
Çizelge 9. Uygulanan Sınıf Düzeylerine Göre 1. Sınıf Testi Deneme Formuna Ait Madde Analizi Sonuçları	61
Çizelge 10. 2. Sınıf Testinden Alınan Puanların Sınıf Düzeylerine Göre Dağılımı İçin Yapılan Normallik Testi Sonuçları	63
Çizelge 11. 2. Sınıf Testi Puanlarının Sınıf Düzeylerine Göre Farklılıkları İçin Uygulanan Mann Whitney-U Testi Sonuçları	63
Çizelge 12.2. Sınıf Testinden Alınan Puanlarının Dağılımı İçin Yapılan Normallik Testi Sonuçları	64
Çizelge 13.2. Sınıf Testi Puanlarının İlçelere Göre Kruskal Wallis H-Testi Sonuçları.....	65
Çizelge 14. İlçelerin 2. Sınıf Test Puan Ortalama Farklılıkları İçin Uygulanan Mann Whitney-U Testi Sonuçları	65
Çizelge 15. Uygulanan Sınıf Düzeylerine Göre 2. Sınıf Testi Deneme Formuna Ait Madde Analizi Sonuçları	66
Çizelge 16.3. Sınıf Testinden Alınan Puanların Sınıf Düzeylerine Göre Dağılımı İçin Yapılan Normallik Testi Sonuçları	69

Çizelge 17.3. Sınıf Testi Puanlarının Sınıf Düzeylerine Göre Farklılıkları İçin Uygulanan Mann Whitney-U Testi Sonuçları	69
Çizelge 18.3. Sınıf Testinden Alınan Puanlarının Dağılımı İçin Yapılan Normallik Testi Sonuçları	70
Çizelge 19.3. Sınıf Testi Puanlarının İlçelere Göre Kruskal Wallis H-Testi Sonuçları.....	71
Çizelge 20. İlçelerin 3. Sınıf Test Puan Ortalama Farklılıkları İçin Uygulanan Mann Whitney-U Testi Sonuçları	71
Çizelge 21. Uygulanan Sınıf Düzeylerine Göre 3. Sınıf Testi Deneme Formuna Ait Madde Analizi Sonuçları.....	73
Çizelge 22. 4. Sınıf Testinden Alınan Puanların Sınıf Düzeylerine Göre Dağılımı İçin Yapılan Normallik Testi Sonuçları	75
Çizelge 23.4. Sınıf Testi Puanlarının Sınıf Düzeylerine Göre Farklılıkları İçin Uygulanan t-Test Sonuçları	76
Çizelge 24. 4. Sınıf Testinden Alınan Puanlarının Dağılımı İçin Yapılan Normallik Testi Sonuçları	76
Çizelge 25. 4. sınıf test puanlarının ilçelere göre tek yönlü varyans analizi (ANOVA) sonuçları	77
Çizelge 26. 4. Sınıf Test Puan Ortalama Farklılıkları İçin Uygulanan Scheffe Testi Sonuçları.....	78
Çizelge 27. Uygulanan Sınıf Düzeylerine Göre 4. Sınıf Testi Deneme Formuna Ait Madde Analizi Sonuçları.....	79
Çizelge 28. Testlerin Asıl Formlarındaki Maddelerin Güçlük Kategorilerine Göre Dağılımı.....	87

EKLER LİSTESİ

Ek-1.1. Sınıf Matematik Dersi Sayılar Öğrenme Alanı Başarı Testi Belirtke Tablosu	98
Ek-2.2. Sınıf Matematik Dersi Sayılar Öğrenme Alanı Başarı Testi Belirtke Tablosu	101
Ek-3. 3. Sınıf Matematik Dersi Sayılar Öğrenme Alanı Başarı Testi Belirtke Tablosu	104
Ek-4. 4. Sınıf Matematik Dersi Sayılar Öğrenme Alanı Başarı Testi Belirtke Tablosu	107
Ek-5. Ek-5. 1-4. Sınıf Matematik Dersi Sayılar Öğrenme Alanı Başarı Testlerine Ait Örnek Sorular.....	110

BÖLÜM 1

GİRİŞ

Bu bölümde araştırma problemi açıklanmış; amaç, önem, sayılılar, sınırlılıklar ve tanımlara yer verilmiştir.

Problem

Eğitim; sosyal bir olgu olan insanı merkeze aldığından bilişim, teknoloji, psikoloji, sosyoloji gibi insanı ve insan yaşamını konu alan pek çok alanda meydana gelen gelişmelerden doğrudan etkilenen bir alandır. Oldukça hızlı gelişim gösteren bu alanlarda yapılan çalışmalar sonucu ortaya çıkarılan teknolojik ürün, bilgi veya teorilerin eğitim alanına entegre edilerek eğitim sistemi içerisindeki dinamiklerinsürekli yenilenip, geliştirilmesi çağın gerekliliklerine uyum sağlayabilmek için bir zorunluluk halini almıştır.

Toplumların gelişme ve ilerlemesini sağlama misyonunu üstlenen eğitim, bunu sağlayabilmek için toplumu oluşturan bireylerde olumlu yönde değişimler meydana getirmeyi hedeflemektedir. En yaygın tanımı ile eğitim “bireyin davranışlarında kendi yaşantısı yoluyla ve kasıtlı olarak, kalıcı izli istendik davranış değişikliği meydana getirme süreci” (Ertürk, 1972, s. 9) olarak bilinmektedir. Bu tanımda da vurgulanan istendik davranış değişikliği ifadesi toplumdan topluma ve zamana göre farklılaşmaktadır. Toplumların bireylerden bekledikleri davranış ve nitelikler zamanla değişim göstermektedir. Bununla birlikte, bireylerin bu davranış ve nitelikleri edinme yöntemleri, başka bir deyişle davranış değişikliği oluşturma süreci olarak nitelendirilen öğrenme kavramında da değişimler kaçınılmaz olmaktadır.

Son yıllarda özellikle bilişsel psikoloji alanındaki gelişimlerin de etkisiyle öğrenmenin doğası ve nasıl gerçekleştiği hakkında benimsenen teorilerdedeğişimler gözlenmekte, bunlara paralel olarak öğrenme

kavramının tanımı da birtakım değişikliklere uğramaktadır. Önceki dönemlerde sıklıkla vurgulanan davranış değişikliği kavramının yerini son yıllarda bireylerin bilişsel yapılarındaki değişimlerin aldığı dikkat çekmektedir.

Öğrenme kavramı söz konusu olduğunda akla gelen önemli bir kavram başarıdır. Cizek (1997) öğrenme ile başarı kavramlarının zaman zaman birbirinin yerine kullanılsa da aslında ayrı kavramlar olduğunu iddia ederek öğrenmenin deneyimler yoluyla bilişsel yapıların nispeten kalıcı bir şekilde yeniden düzenlenmesi olduğunu, buna karşın başarının daha çok performansla ve davranışın gösterilmesiyle anlaşıldığını vurgulamaktadır. Ayrıca başarının şans, kopya gibi faktörlerden kaynaklanabileceği belirtilerek başarı için öğrenmenin her zaman bir koşul olmadığına, bu yüzden puanlarla ifade edilen başarının öğrenme hakkında yanıltıcı bir gösterge olabileceğine dikkat çekilmektedir. Ancak, başarıyı zaman zaman etkileyen bu faktörlerin var olması öğrenmenin başarı için ön koşul olmadığını iddia etmek için yeterli değildir. Bir alanda performans gösterebilmek, davranış sergileyebilmek için o alan ile ilgili birtakım bilgi ve becerilerin kazanılmış olması gerekir. Eskiden kabul edilen davranışçı öğrenme teorileri düşünme, akıl yürütme ve problem çözmenin karmaşık süreçlerini açıklamada yetersiz kaldıklarından öğrencilerin bilgi ve becerileri sorma ve sorgulama aracı olarak kullanmalarını önemseyen yeni öğrenme teorilerine ihtiyaç duyulduğuna vurgu yapılmaktadır. Ancak, her ne kadar son yıllarda öğrenme teorilerinde ve tanımında davranış değişiminden çok bilişsel yapılardaki değişime önem verilse de bireylerin bilişsel yapılarındaki değişimin göstergesi olarak yine edinilen bir kazanımın gözlenebilecek şekilde ortaya konması söz konusudur. Piaget, Bruner gibi bilişsel öğrenme kuramcılarına göre öğrenme bireyin karşılaştığı yeni durumları zihninde önceden var olan şemalarla karşılaştırıp oluşan dengesizlik durumunu düzenlemek için bilişsel yapılarını yenilemesi şeklinde gerçekleşmektedir (Şirin, 2008). Bilişsel yapılardaki bu yenilenmenin, dolayısıyla öğrenmenin gerçekleştiğine delil olabilecek göstergelerin tamamen bireyin zihninde var olan olgular olarak kabul edilmesinin yerine eğitim sürecindeki yaşantılarda problem çözme gibi üst düzey bilişsel etkinlik gerektiren durumlar yoluyla sergilenebilir olarak görülmesi önerilebilir.

Eğitim alanında herhangi bir yeniliğe gidilirken ya da kararlar alınırken eğitimin temel amacının ne olduğu üzerine düşünülmesi ve bu amacın göz önünde tutulması gerekir. Eğitimin asıl amacı sorgulandığında eğitim sürecinin sonuçları, başka bir ifade ile eğitimsel çıktılar akla gelmektedir. Bu durum da eğitimsel çıktılarının açık bir şekilde tanımlanmasını ve bireylerde meydana gelen öğrenmenin göstergelerinin neler olabileceğinin düşünülmesini gerektirmektedir.

Bireylerin bir alan ile ilgili kazanımları edindiklerine, bilişsel yapılarında değişiklikler meydana geldiğine, dolayısıyla öğrenmenin gerçekleştiğine delil olarak kabul edilebilecek göstergelerin yoklanması oldukça önemlidir. Ancak, sonuçların kullanım amacına uygun olarak ölçmenin ne amaçla yapılacağına açıkça belirlenmesi gerekmektedir. Eğitimde ölçme, eğitimsel kararları etkileyecek, öğrenci gelişimi ve öğretim sürecinin işleyişi hakkında bilgi sağlayacak birtakım çıkarımlar amaçlanarak yapılmaktadır. Yapılacak çıkarımın doğasına göre ölçme amacının netleştirilmesi ve bu amaca uygun ölçme yöntemleri seçilmesi gerekmektedir (Cizek, 1997; Turgut ve Baykul, 2012). Bu durumda, örneğin öğrenci gelişiminin ve öğretimin etkililiğinin takip edilmesi amacıyla ünite sonu testleri gibi biçimlendirici testler, öğrencilerin mevcut durumda öğrenme sürecinin hangi düzeyinde bulunduğu belirlenmesi amacıyla da düzey belirleme testleri yapılabilmektedir.

Ölçme kavramı Türkiye’de 1950’li yıllarda program geliştirme, rehberlik gibi alanlarda kullanılmaya başlanmıştır (Koç, 1993). Bu kavrama ait farklı kaynaklarda çeşitli tanımlarına rastlamak mümkündür. Bunlardan birine göre ölçme “Öğrencilerin güçlü ve zayıf yönlerini keşfetme ve rapor etme; öğretimi planlama ve geliştirme; ilerlemeyi değerlendirme ve öğrenciler hakkında kararlar alma amaçlarına yönelik planlı bilgi toplama ve sentezleme süreci” olarak ifade edilmektedir (Cizek, 1997, p. 10). Başka bir tanımda ise “Birey ya da nesnelerin niteliklerinin uygun araçlar kullanılarak gözlenip, gözlem sonuçlarının sembollerle ifade edilmesi” şeklinde tarif edilmiştir. Özetle ölçme için belli bir amaç için bireylerin bir özelliğe sahip olma düzeylerinin belirlenmesi tanımı yapılabilir. Cizek (1997) ölçme tanımında süreçten birincil düzeyde etkilenen ve sonuçlarından yararlanan olarak öğrencinin vurgulanması gerektiğine dikkat çekerek araştırma, izlenebilirlik gibi

amaçlarla kullanımlarının ikincil düzeyde olacağını belirtmektedir. Eğitimin ve öğrenme sürecinin merkezinde bulunan ve birincil derecede önemli elemanı olan öğrencinin ne kadar ilerleme gösterdiğinin, öğrenme eksiklerinin ve sunulan öğretim hizmetinin etkililiğinin ölçülmesi gibi çabaların tümü de yine öğrenci yararına olmaktadır.

Eğitimsel çıktıların nasıl ölçüleceğine karar vermeden önce bu çıktıların ölçülebilir olup olmadığı düşünülmelidir. Eğitim alanında karşılaşılan hemen her özellik ölçülebilir olarak görülmektedir (Ebel, 1965; Turgut ve Baykul, 2012). Ebel (1965)ölçme için gerekli olanın doğrulanabilir bir az-çok ilişkisi olduğunu belirterek prensip olarak gözlenebilir bir fark oluşturan tüm eğitimsel çıktıların ölçülebileceğini iddia etmektedir. Eğitim sürecinde bireylere sunulan öğrenme-öğretme etkinlikleri yardımıyla sahip oldukları bilgi, beceri ve tutumlarında değişiklik ve ilerlemeler oluşturmak amaçlandığı düşünüldüğünde bu sürecin başında ve sonunda bulunulan düzeyler arasındaki farkın ölçme için temel oluşturacağı söylenebilir. Bir özelliğe sahip olma açısından yapılan bir dizi az-çok karşılaştırmaları sonucunda bu özelliği ölçen bir ölçek oluşturulması ve bu ölçeği oluşturan adımlara ya da aralıklara sayılar atanarak bu ölçekle nicel ölçümlerin yapılması mümkün hale gelebilir (Ebel, 1965). Başarı gibi bir kavramın da bireylerde bulunma derecesini belirtmede az-çok gibi karşılaştırmalar sayısal bir değer içermediğinden ölçme gibi algılanmayabilmekte, başka bir ifade ile yapılan ölçümlerin sayısallaştırılması problemi ile karşılaşılmaktadır. Ebel (1965) bir özelliğin sayısallaştırılması sürecinin yaygın olarak kullanılan ifade ile o özelliğin işlevsel tanımı olduğunu belirtmektedir. Turgut ve Baykul (2012) da ölçme sürecinin her zaman ölçülecek özelliğin tanımlanmasıyla başladığını vurgulamaktadır. Ölçülecek bir özelliğin işlevsel tanımının yapılması o özelliğin bireylerde bulunma durumunu gözlemlemeye ve ölçmeye yarayacak farklı ve kendine özgü belli bir ölçme metodu geliştirilmesi olarak ifade edilebilir. Eğitimde ölçülecek bir yapı olarak başarı için başarı testlerinden alınan puanlar işlevsel tanım olarak kabul edilebilir. Turgut ve Baykul (2012) da ölçme işlemi için araç kullanımı gerektiğine dikkat çekerek araçların ölçülecek niteliğin sayılarla ifade edilmesini ve böylece daha duyarlı bir şekilde yapılmasını sağladığını belirtmektedirler.

Başarı düzeyinin belirlenmesinde kullanılan ölçme araçları arasında standart testler önemli bir yere sahiptir. Standart başarı testlerinin doğru bir şekilde ve uzman kişilerce kullanımı eğitim-öğretimin etkililiği, rehberlik, seçme, yerleştirme gibi pek çok konuda alınan kararların yerindeliğini sağlamaktadır (Koç, 1985). Standart testler geçerlik, güvenirlik gibi psikometrik özellikler bakımından yeterliliği kanıtlanmış ve temsil edici örneklemelere göre normları belirlenmiş ölçme araçları olduklarından sonuçlarının diğer testlere göre çoğunlukla daha güvenli bilgi sağladığı kabul görmektedir. Bu testler bireylerin belli bir norm grubu içerisindeki göreceli konumlarının karşılaştırılması veya sınıf düzeylerine göre yıldan yıla kaydedilen ilerlemenin ortaya çıkarılması amaçlarıyla kullanılabilmektedir (Anastasi, 1988). Öğrencilerin eğitim sürecinde ne kadar ilerleme gösterdiklerinin izlenmesi oldukça önemlidir. Bu ilerlemelerin izlenmesinde farklı öğretmenlerin kendi ölçütlerine göre yaptıkları değerlendirmeler farklı olacağından standart başarı testlerindeki norm değerleri bu değerlendirmelere standart bir referans noktası sağlama açısından yararlı olmaktadır (Ansley, 1997). Öğrencileri akranlarıyla karşılaştırmanın yanında öğrencilerin bağlantılı yetenek testlerinin de yardımıyla sahip oldukları öğrenme güçlüklerini saptamada da yardımcı olmaktadır. Wechsler Bireysel Başarı Testi ve Kaufman Eğitimsel Başarı Testi gibi standart başarı testlerinin ilerlemeyi takip etme, öğrenci hatalarının analizi ile bireysel eğitim planları oluşturma, öğretimi planlama, değerlendirme, öğrenci yerleştirme ve bilimsel araştırma gibi amaçlarla kullanılabilmektedir (Lichtenberger ve Smith, 2005).

Ancak standart testlerin eleştiri aldığı birtakım noktalar da bulunmaktadır. Bu testler eğitim programlarında bulunan kazanımların çok az bir bölümünü kapsadıkları, maddelerin çoktan seçmeli olmasından dolayı üst düzey bilişsel kazanımları ölçemedikleri ve öğretmenlerin bu testlerin kapsamına yönelik öğretim yapmalarına neden oldukları için eleştirilmektedir (Ansley, 1997). Ayrıca bu testlerin revize edilmeden uzun yıllar boyunca dünya çapında çok yaygın ve aşırı bir şekilde kullanılmaları da eleştirilebilir. Standart testlerin hazırlanması ve uygulanması kadar sonuçlarının yorumlanması da oldukça önemli bir konudur. Ansley (1997) bu

sonuçların bireylere uygulanan başka ölçme sonuçlarından ve başarıyı etkileyen diğer faktörlerden bağımsız olarak değerlendirildiğinde yanlış yorumlara yol açabildiğini belirterek standart başarı testlerinin öğrencilerin güçlü veya zayıf olma durumlarını ve varsa özel eğitime ihtiyaçlarını belirlemek için sadece bir araç olarak kullanılması gerektiğini belirtmektedir. Öğrencilerin belli norm değerlerine göre yüksek ya da düşük başarı gösterdiklerini ortaya çıkararak karşılaştırma yapmak amacıyla uygulanan standart testlerin sonuçlarının yorumlanmasında dikkatli olunması ve aşırı genellemelerden kaçınılması gerekmektedir.

Eğitimde başarıyı ölçmek için kullanılan testlerin büyük çoğunluğu standart olmayan öğretmen yapımı testlerdir. Sanchez ve Ice (2005) standart testlerin öğrencilerin ne kadar öğrendiğini karşılaştırma amaçlı ölçerken, öğretmen yapımı testlerin sadece öğrencilerin belli koşullar altında ne kadar performans gösterebildiklerini ölçmeyi değil, aynı zamanda öğrenmeyi artırmayı ve öğretimi yönlendirmeyi de amaçladığını belirtmektedir. Öğretim sürecinde bu testlerin belirli sıklıkla uygulanması hem öğretmenlere hem de öğrencilere öğrenmenin ne kadar gerçekleştiği hakkında dönütler verme açısından oldukça yararlıdır. Bunun yanında öğrenme sürecinde sonuçlardan haberdar olmak öğrencilerde motivasyonu artırıcı bir etken olabilir. Öğretmenler tarafından hazırlanan bu testler standart testlerden farklı olarak hem işlemsel becerileri hem de kavramsal bilgileri ölçmektedir (Sanchez ve Ice, 2005). Bu testlerin kavramsal bilgileri de ölçebiliyor olması standart testlerden farklı olarak kullanılacak ölçme aracı ya da soru türü konusunda daha çok seçeneğe sahip olmasından kaynaklanabilir. Küçük gruplara uygulanan sınıf testlerinde açık uçlu sorular daha çok tercih edilse de testin kapsamının artırılması için çoktan seçmeli maddelere yer verilmesi önerilmektedir (Anastasi, 1988). Ancak, çoktan seçmeli test maddelerinin ve cevap seçeneklerinin oluşturulması test geliştirme konusunda bir miktar eğitim ve deneyim gerektirdiğinden ve zaman aldığından tercih edilmediği gözlenmektedir. Testlerin oluşturulduğu madde türlerinin hazırlanma sürecinin güçlüğü, kapsamı gibi yönlerden birbirine göre üstünlükleri bulunmaktadır. Çoktan seçmeli ve kısa cevaplı maddelerden oluşan testlerin psikometrik özelliklerinin karşılaştırıldığı çalışmalarda kısa cevaplı

maddelerden oluşan testlerin madde ayırıcılık gücü indeksleri çoktan seçmeli testlerinkinden daha yüksek bulunmuştur (Akdeniz, 2003; Aydın, 1993). Bu durumun çoktan seçmeli maddelerin cevaplanmasında test ile ölçülmek istenen özelliğe sahip olmayan bireylerin de doğru seçeneğe ulaşabilmelerinde etkili olan şans faktöründen kaynaklandığı düşünülebilir.

Test geliştirme çalışması planlı ilerlemesi gereken bir süreçtir. Anastasi (1988) testler hazırlarken iyi bir test hazırlama planına uyulmaması durumunda test maddelerinin konulara ya da bilişsel düzeylere dağılımında dengesizlikler gibi problemler ortaya çıkacağı konusunda uyarmaktadır. Test geliştirme sürecinin başında testin amacının ve kapsamının netleştirilmesi gerekir (Özçelik, 2010). Geliştirilecek olan testin öğretim sürecinde biçimlendirici değerlendirme ya da belli bir sürecin sonunda öğrenci düzeyini belirleme amaçlarından hangisi için yapıldığının belirlenmesi gerekir. Testin amacı belirlendikten sonra kapsamının belirlenmesine geçilmelidir. Ölçülecek özellik ile ilgili gözlenebilir davranışların tanımlandığı bir hedefler kümesi oluşturulmalı ve bu hedefler kümesi ölçme sonucunda atfedilen etiketle uyumlu olmalıdır (Kane, 2006). Geliştirilecek testin amacına uygun olarak testin kapsamı eğitimsel çıktılar ya da öğrenmenin gerçekleştiğine delil sayılabilecek davranışlar halinde açık bir şekilde belirlenmelidir.

Başarı testleri ile ölçülmeye çalışılan özellikler eğitim-öğretim sürecinde kazanılması hedeflenen bilgi, beceri ve tutumlar olmalıdır. Eğitimde hedefler, öğretimi yönlendirmenin yanında yapılacak ölçmelere de kapsanacak konular ve önem verilecek noktalar hakkında rehberlik etme işlevine sahiptir (Demirel, 2010; Ebel, 1965). Öğretim programlarında her sınıf düzeyi için öğrencilerin edinmeleri hedeflenen kazanımlar belirlenmiştir. Dolayısıyla bu kazanımlarla ifade edilen bilgi ve beceriler ölçme için temel oluşturmaktadır. Test geliştiren kişi hem eğitim sürecinin bütünüyle hem de testin özel olarak geliştirildiği ders ya da konuyla ilgili kazanımları göz önüne almalıdır (Ebel, 1965). Böylece hazırlanacak testin sınıf düzeyine uygunluğu ve kapsamlılığı için dayanak sağlanmış olur. Ebel (1965) özellikle standart başarı testi veya geniş ölçekli testlerin hazırlanmasında farklı materyaller kullanan farklı öğretmenler tarafından eğitilen öğrencilerin tümü için uygun tek bir testin hazırlanması konusunda güçlükler olduğuna dikkat çekmektedir.

Bu noktada tüm öğrenciler için ortak olan öğretim programındaki kazanımlardan yararlanılması hem test edilecek bilgi ve becerilerin belirlenmesi hem de farklılıkların etkisinin azaltılması için uygun bir çözüm olabilir. Test geliştirme konusunda hem yurt içinde hem de yurt dışında yapılan pek çok çalışmada (Aydın Gürler, 2011; Deering, 1988; Gönen, Kocakaya ve Kocakaya, 2011; Jitendra, Parker ve Kameenui, 1997; Üçgül Öcal, 2011; Üçüncü, 2010) testlerle ölçülecek özelliklerin belirlenmesinde uygulanmakta olan öğretim programındaki kazanımların temel alındığı görülmektedir.

Matematik alanında başarı belirlemeye yönelik çalışmaların yurt dışındaki örnekleri arasında Amerika Birleşik Devletleri'nde ülke çapında öğrencilerin başarı düzeylerini belirli aralıklarla ölçmek için yapılan National Assessment of Educational Progress (NAEP) bulunmaktadır. Ayrıca, 1995'ten beri dört yılda bir uluslararası düzeyde uygulanarak öğrencilerin matematik ve fendeki başarı düzeylerini karşılaştırmalı olarak rapor eden Trends in International Mathematics and Science Study (TIMSS) de başarı ölçme konusunda oldukça kabul gören bir uygulamadır. Bu iki ölçme uygulamasının ortak yanı, katılımcı okullarda uygulanmakta olan öğretim programları temel alınarak öğrencilerin öğretim programlarındaki kazanımları edinme düzeylerini ölçmeleridir. TIMSS 2011'de ölçme için kullanılan çerçeve incelendiğinde matematik testinde sayı, ölçme, geometri, cebir, veri ve olasılık gibi konu alanlarının kapsandığı görülmektedir(Mullis, Martin, Ruddock, O'Sullivan ve Preuschoff, 2009). Bu konu alanları Türkiye'de uygulanan matematik öğretim programındaki öğrenme alanları ile oldukça benzerdir.

Matematik dersi için öğretim programında yer alan kazanımlar temelinde öğrencilerde matematik alanında öğrenmenin göstergesi olarak kabul edilecek davranışların belirlemesi ve bunların yoklanmasıyla öğrenci başarısının ölçülmesinin gerekliliği açıktır. Bu amaçla, uygun test geliştirme adımları izlenerek ihtiyaç duyulan ölçme aracının geliştirilmesi gerekmektedir.

İlköğretim matematik dersi öğretim programı incelendiğinde programın öğrenme alanlarından oluştuğu görülmektedir. İlköğretimin 1-4. sınıflarında kazanımlar; sayılar, geometri, ölçme ve veri olmak üzere dört öğrenme alanına ayrılmıştır. Bu sınıf düzeylerindeki kazanımların öğrenme alanlarına dağılımı Çizelge 1’de gösterilmektedir:

Çizelge1

İlköğretim 1-4. Sınıflar Matematik Dersi Kazanımlarının Öğrenme Alanlarına Dağılımı

Sınıflar Öğrenme Alanları	1. Sınıf	2. Sınıf	3. Sınıf	4. Sınıf
Sayılar	27 (%61)	33 (%56)	28 (%41)	40 (%48)
Geometri	7 (%16)	10 (%17)	19 (%27)	17 (%20)
Ölçme	9 (%21)	13 (%22)	18 (%26)	24 (%28)
Veri	1 (%2)	3 (%5)	4 (%6)	3 (%4)
Toplam	44 (%100)	59 (%100)	69 (%100)	84 (%100)

Çizelge1 incelendiğinde belirtilen sınıfların tümünde kazanımların oldukça büyük bölümünün sayılar öğrenme alanında bulunduğu görülmektedir. Sayılar öğrenme alanı, genel olarak öğrencilerde sayma ve sayı kavramları ile sayılarla işlem yapma becerilerine yönelik kazanımlar içermektedir.

Amaç

Bu çalışmanın amacı ilkokul öğrencilerinin matematik dersi sayılar öğrenme alanındaki başarı düzeylerini belirlemeye yönelik geçerli ve güvenilir başarı testleri geliştirmektir.

Araştırmanın amacına uygun olarak aşağıdaki sorulara cevap aranmıştır:

1. Birinci sınıf başarı testi,

- a) Kapsam geçerliğine sahip midir?
- b) Yapı geçerliğine sahip midir?
- c) Ölçüt dayanaklı geçerliğe sahip midir?
- d) Madde güçlük indeksleri kaçtır?
- e) Madde ayırıcılık gücü indeksleri kaçtır?

2. İkinci sınıf başarı testi,

- a) Kapsam geçerliğine sahip midir?
- b) Yapı geçerliğine sahip midir?
- c) Ölçüt dayanaklı geçerliğe sahip midir?
- d) Madde güçlük indeksleri kaçtır?
- e) Madde ayırıcılık gücü indeksleri kaçtır?

3. Üçüncü sınıf başarı testi,

- a) Kapsam geçerliğine sahip midir?
- b) Yapı geçerliğine sahip midir?
- c) Ölçüt dayanaklı geçerliğe sahip midir?
- d) Madde güçlük indeksleri kaçtır?
- e) Madde ayırıcılık gücü indeksleri kaçtır?

4. Dördüncü sınıf başarı testi,

- a) Kapsam geçerliğine sahip midir?
- b) Yapı geçerliğine sahip midir?
- c) Ölçüt dayanaklı geçerliğe sahip midir?
- d) Madde güçlük indeksleri kaçtır?
- e) Madde ayırıcılık gücü indeksleri kaçtır?

Önem

Eğitim-öğretim sürecinde öğrenciler için belirlenen hedeflere ulaşıp ulaşılmadığının belirlenmesi için ölçme-değerlendirme uygulamaları vazgeçilmez derecede önemlidir (Koç, 1987). Ölçme-değerlendirmenin yapılabilmesi ise ölçülecek özelliklere uygun araç kullanımını gerektirmektedir. Eğitimde başarının ölçülmesi için geçerlik, güvenirlik gibi psikometrik özellikler bakımından yeterli, sınıf düzeyine uygun ölçme araçlarına ihtiyaç vardır. Öğrenciler hakkında ölçme sonuçlarının verdiği bilgi çok değişken olmasına rağmen iyi hazırlanmış ölçme araçları ile kesin, güvenilir ve kullanışlı bilgi elde edilmesi mümkündür (Cizek, 1997; Ebel, 1965). Bu bilgiler, alınacak eğitimsel kararlara dayanak oluşturmaktadır. İyi hazırlanmış başarı testleri ile öğrenci başarısının periyodik olarak ölçülmesi tüm öğrenciler için öğrenmeyi artırıcı ve kolaylaştırıcı bir etkiye sahiptir (Anastasi, 1988). Eğitim süreci boyunca ne kadar öğrenildiğinin ve başarı düzeyinin bilinmesi öğrenciler için de motive edici olabilir.

Eğitim-öğretim sürecinde öğrencilerin yıldan yıla kaydettikleri gelişmeleri daha objektif bir şekilde belirleyebilmek ve özel bir müdahaleye ihtiyaç duyan öğrencileri saptayabilmek için de sınıf düzeylerine uygun başarı testleri gerekmektedir. Ayrıca matematik eğitimi alanında farklı öğretim yöntemlerinin öğrenci başarısına etkisi, bazı beceri veya tutum düzeylerinin öğrenci başarısı ile ilişkisi, farklı sosyoekonomik düzeylerdeki bölgeler ve okullar arasındaki başarı farkları gibi konuları inceleyen pek çok deneysel çalışmada öğrenci başarısı değişkeni için öğrencilerin karne veya yazılı sınav notları değerlendirilmektedir. Karne notları ve yazılı sınavlardan alınan puanlar ise öğrencilerin başarılarını göstermede kimi zaman yeterli olmamakta veya farklı okullarda yapılan uygulamalarda objektif ve standart bir ölçme anlamına gelmemektedir. Bu yüzden öğrencilerde gerçekleşen öğrenmenin nitelik ve niceliği ile başarı düzeyinin belirlenmesinde kullanılacak objektif, geçerli ve güvenilir ölçme araçları gerekmektedir.

Tüm derslerde olduğu gibi matematik alanında da öğrencilerin başarılarının doğru ve güvenilir bir şekilde ölçülmesi oldukça önemlidir. Öğrencilerin matematik dersindeki ilerlemelerini izlemek ve öğretim sürecini yönlendirmek amacıyla öğretmenler tarafından yazılı yoklama sınavı, ünite sonu testi gibi ölçme araçları hazırlanıp kullanılmaktadır. Ancak bu araçlar kısa vadede öğretmen ve öğrencilere dönüt sağlama açısından yeterli görülse de öğrencilerin bir eğitim-öğretim döneminin sonunda eriştikleri düzeylerin belirlenmesi için geçerli ve güvenilir araçlar kullanılması daha uygun olacaktır. Matematik eğitimi alanında yapılan araştırmalarda da sıklıkla öğrencilerin matematik başarı düzeylerini belirleyen ölçme araçlarına ihtiyaç duyulmaktadır.

Öğrenci başarısının belirlenmesi yönünde var olan ihtiyaçları karşılamak için standart başarı testleri en güvenilir ölçme araçları olarak kabul edilmektedir. Ancak, Türkiye’de matematik alanında böyle testler bulunmamaktadır. Bu yüzden matematik başarısını ölçmek için nitelikli başarı testleri geliştirilmesi oldukça gerekli görülmektedir.

Sayıtlılar

Öğrencilerin veri toplama aracındaki soruları dikkatle cevapladıkları kabul edilmektedir.

Sınırlılıklar

Çalışmada, araştırmanın amacına uygun olarak 1-4. sınıflar için test geliştirme çalışması yapılmış, testlerin geçerlik ve güvenirlik analizleri test geliştirme aşamasında elde edilen verilerle yapılmıştır. Testler için veri toplama işlemi kasım ayında yapılmıştır. Testlerdeki maddeler kısa cevaplı soru türünden olup, maddelerin cevaplayıcıların özgün bilgi veya ürün içeren cevaplar üretmelerine olanak sunacak bir yapıya sahip olmamaları ölçülecek becerilerin düzeyi açısından bir sınırlılık olarak görülebilir. Geliştirilen testler, ilkökul 1-4. sınıflar matematik dersi sayılar öğrenme alanı kazanımları ile sınırlıdır.

BÖLÜM 2

İLGİLİ ARAŞTIRMALAR

Test geliştirme konusunda yurtiçinde ve yurtdışında daha önce yapılmış, öncelikle bu çalışmanın amacına benzer şekilde matematik alanında ve kazanımlar çerçevesinde başarı belirlemeye yönelik test geliştirme çalışmaları araştırılmıştır. Bunun yanında farklı alanlarda yapılmış başarı testi ve ölçek geliştirme çalışmaları incelenmiştir. Bu bölümde, yöntem bakımından başarı testi geliştirme ile benzerlik gösteren test ve ölçek geliştirme çalışmalarında izlenen süreçlere yer verilmiştir.

2. 1. Yurt İçinde Yapılan Araştırmalar

Matematik alanında yapılan başarı testi geliştirme çalışmaları arasında bu çalışmanın amaç ve kapsamı ile benzerlik gösteren çalışmalara rastlanmıştır (Başer, 1996; Üçgül Öcal, 2011; Üçüncü, 2010). Farklı sınıf düzeylerine yönelik matematik başarı testlerinin geliştirildiği bu çalışmalarda oldukça benzer bir test geliştirme yönteminin izlendiği dikkat çekmiştir.

Üçgül Öcal (2011) tarafından yapılan 6. sınıf matematik dersi sayılar öğrenme alanında hazırbulunuşluk testi geliştirme çalışmasında 6. sınıf matematik dersi sayılar öğrenme alanı kazanımları için ön öğrenme sayılan, bir önceki sınıflara ait kritik kazanımlar, uzman görüşleri ile belirlenmiş; bunlar için yazılan maddelerle 29'ar maddeden oluşan iki eşdeğer deneme formu oluşturulmuştur. Testin deneme uygulaması 162 beşinci sınıf öğrencisi ile yapılmış, elde edilen veriler üzerinde yapılan analizler sonucunda aynı kazanım için yazılan iki maddeden birinin seçilmesiyle 30 maddelik asıl form

oluşturulmuştur. Testin geliştirilmesi aşamasından sonra geçerlik ve güvenilirlik çalışması için ayrıca bir uygulama yapılmış, bundan sonra yapılan analizlerde maddelerin ayırıcılık gücü indeksleri 0,49 –0,92 arasında, güçlük indeksleri ise 0,24 –0,75 arasında bulunmuştur. Ölçme aracının yapı geçerliğini incelemek için test 5. sınıf öğrencilerinin yanında 4. ve 6. sınıf öğrencilerine de uygulanmış, üç grubun test puanları arasındaki anlamlı fark yapı geçerliğinin olduğu şeklinde yorumlanmıştır. Testin KR-20 güvenilirlik katsayısı 0,92 olarak hesaplanmıştır. Sonuçta elde edilen bulgular aracın geçerli ve güvenilir olduğunu göstermiştir.

Üçüncü (2010) tarafından geliştirilen Doğal Sayılarla Çarpma İşlemi Başarı Testi'nin geliştirilmesi için öncelikle 2–5. sınıf doğal sayılarla çarpma işlemi kazanımları incelenerek belirtke tabloları hazırlanmış, ölçülecek her kazanıma yönelik en fazla 3 çoktan seçmeli madde yazılmıştır. Testlerin kapsam geçerliği için program geliştirme, ölçme değerlendirme ve matematik eğitimi uzmanlarından görüşler alınmış, öneriler doğrultusunda gerekli düzeltmeler yapılarak 2–5. sınıfta bulunan toplam 317 öğrenci ile deneme uygulaması yapılmıştır. Madde ve test istatistikleri ITEMAN programında analiz edilmiştir. Ayırıcılık gücü 0,30'dan büyük olan maddelerin seçilmesiyle son haline ulaştırılan testlerin KR-20 güvenilirlik katsayıları 0,72-0,88 arasında değişmektedir.

Başer (1996) tarafından 9. sınıf öğrencileri için matematik başarı testinin geliştirilmesinde öncelikle uzman görüşleri ile her birinde 75 çoktan seçmeli soru bulunan iki test oluşturulmuştur. Birinci test 449, 2. test 453 öğrenciye uygulanmış, ancak madde analizi rastgele seçilen 370 sonuç üzerinden yapılmıştır. Madde analizi sonuçlarına göre ayırıcılık gücü dikkate alınarak madde seçimi yapılmış, ayırıcılık gücü 0,20'den büyük olan 75 madde ile test son formuna ulaştırılmıştır. Testte, çoğunluğu uygulama olmak üzere bilişsel alanın tüm basamaklarında soru bulundurulmuştur. Testin kapsam geçerliği uzman öğretmen görüşleri değerlendirilerek sağlanmış, ölçüt dayanaklı geçerlik için test puanları ile okul matematik başarıları arasındaki korelasyon 0,64 bulunmuştur. Güvenirlik için test-tekrar test ve KR-20 yöntemleri uygulanmış, sonuçlar sırasıyla $r=0,84$ ve $r = 0,95$

bulunmuştur. Bu sonuçlara göre güvenilirliğin yüksek olduğu ve maddelerin birbiri ile yüksek derecede uyumlu olduğu şeklinde yorumlanmıştır.

Yukarıdaki üç çalışmada (Başer, 1996; Üçgül Öcal, 2011; Üçüncü, 2010) öncelikle testte ölçülecek kazanımların belirlendiği, kazanımlar için 2 ya da 3 soru yazılarak oluşturulan deneme formunun uygulanması sonucu elde edilen veriler üzerinden yapılan madde analizlerine göre aynı kazanımı ölçmeye yönelik maddelerden ayırıcılık değerleri 0,30'dan ya da 0,20'den yüksek olanın seçilmesiyle testlerin son haline getirildiği görülmüştür.

Başarı testi geliştirme ile ilgili literatürde matematik alanından farklı olarak fizik, kimya, genetik gibi alanlarda yapılmış çalışmalara (Aydın Gürler, 2011; Gönen, vd., 2011; Kılıç ve Sağlam, 2009; Tosun ve Taşkesenligil, 2011) rastlanmıştır, bunlarda da matematik alanında yapılanlardakine benzer yöntemlerin izlendiği görülmüştür.

Fizik alanında yapılan bir test geliştirme çalışmasında Gönen, Kocakaya ve Kocakaya (2011) tarafından yapılan fizik dersi dinamik konusunda başarı testi geliştirme çalışması 10. sınıf fizik dersi öğretim programındaki hedef ve davranışlar dikkate alınarak dinamik konusu ile ilgili belirtke tablosu hazırlanmış, bu tabloya göre 60 soru yazılmıştır. Bu sorular Bloom taksonomisine göre sınıflandırılmış, uzman görüşleri ile de deneme formuna konulacak madde sayısı 39 olarak belirlenmiştir. Madde analizi sonucunda, kapsam geçerliğini bozmayacak şekilde ayırt edicilik indeksleri 0,25'in altında olan 14 madde elenerek başarı testi son haline getirilmiştir. Testin ortalama güclüğü 0,36 olarak hesaplanmış ve bu değer testin değişik öğrenci seviyelerine hitap edebilecek yeterlikte ve güçlükte olduğu şeklinde yorumlanmıştır. Testin güvenirlik katsayısı 0,81 olarak bulunmuş ve bu değer 0,70'ten yüksek olduğundan güvenilirliğin tatmin edici olduğu sonucuna varılmıştır.

Kılıç ve Sağlam (2009) ortaöğretim öğrencilerinin genetiğin temel kavramlarını edinimlerini belirlemeye yönelik iki aşamalı soru maddelerinden oluşan testin geliştirilmesinde; içeriğin belirlenmesi, öğrencilerin yanlış anlamaları hakkında bilgi edinilmesi ve maddelerin yazılması olmak üzere 3

bölümden oluşan bir yöntem izlemişlerdir. İçeriğin belirlendiği ilk bölümde bilgi önermeleri ve kavram haritasıyla konunun kapsamı belirlenmiştir. İkinci bölümde öğrencilerin yanlış anlamalarını belirlemek için ilgili araştırmalar incelenmiş, öğrencilerle görüşmeler yapılmış, açık uçlu ve çoktan seçmeli sorular içeren bir test uygulanmıştır. Üçüncü bölümde, iki aşamalı çoktan seçmeli test maddeleri yazılmıştır. Maddelerin 231 öğrenciye uygulanmasıyla elde edilen verilerle madde güçlük indeksi, ayırıcılık gücü indeksi, çeldiricilerin işlerliği, testin güvenirlik ve geçerliği belirlenmiştir. Kapsam geçerliği için uzman görüşleri alınmış ve kavram haritası hazırlanmış, yapı geçerliği için farklı tür liseler arasındaki puan farklılıkları değerlendirilmiştir. Testin güvenirliği KR-20 güvenirlik katsayısı ile hesaplanmıştır. Öğrencilerin test maddelerinin hem birinci aşamasına hem de her iki aşamasına birden verdikleri cevaplar değerlendirilmiştir. Yanlış cevap kombinasyonları incelenmiş ve öğrencilerin genetik kavramları ile ilgili yanlış anlamaları belirlenmiştir. Sonuçta madde ayırıcılık güçleri 0,34 ile 0,74 arasına, güçlük indeksleri 0,25 ile 0,67 arasında değişen ve ortalama güçlüğü 0,43 olan 14 maddelik bir test oluşturulmuştur.

Tosun ve Taşkesenligil'in (2011) üniversite düzeyindeki öğrencilerin çözümler ve fiziksel özellikleri konusunda geliştirdikleri başarı testinde çoktan seçmeli, açık uçlu, kısa cevaplı ve doğru-yanlış sorularına yer verilmiştir. Testte ölçülecek özelliklerin belirlenmesi amacıyla Genel Kimya dersini veren öğretim elemanlarıyla, konu kapsamında hangi özelliklere değindiklerini ve öğrencilerin anlamakta zorlandıkları konular hakkında görüşmeler yapılmış, ayrıca öğrencilere konu ile ilgili sorular yöneltilerek kavram yanılgıları belirlenmiştir. Soru yazmaya başlamadan önce ölçülecek davranışlar, konular ve Bloom'un revize edilmiş bilişsel öğrenme düzeylerinin yer aldığı belirtke tablosu hazırlanmış, buna uygun 32 maddeden oluşan bir test hazırlanmıştır. Testin 78 öğrenci ile yapılan pilot çalışması sonucuna göre ayırıcılıkları düşük veya öğrenciler tarafından anlaşılmayan maddeler testten çıkarılmış veya düzenlenmiştir. Toplamda 31 maddeden oluşan deneme formundaki maddelerin 23'ü çoktan seçmeli, 1'i 6 boşluk doldurma içeren kısa cevaplı, 1'i doğru-yanlış, 6'sı ise açık uçlu soru tipinde oluşturulmuştur. Testin deneme uygulaması 160 üniversite öğrencisi ile yapılmış, madde analizlerine göre

ayırıcılık gücü indeksleri 0,20'nin altında olan maddeler testten çıkarılmıştır. Testin güvenilirliği KR-20 güvenilirlik formülü ile hesaplanmıştır.

Öğrencilerin öğretim programındaki kazanımları edinme düzeylerinin incelenmesi yönüyle bu çalışmanın amacı ile benzerlik gösteren bir araştırmada (Aydın Gürler, 2011) ilköğretim 6. Sınıf Fen ve Teknoloji programı Maddenin Tanecikli Yapısı ve Özellikleri ünitesinde başarı testi geliştirilmiştir. Öncelikle kazanımlara uygun 81 madde hazırlanmış ve bu üniteyi önceki yıl görmüş olan 7. Sınıf öğrencilerine uygulanmıştır. Öğrencilerin cevaplarına göre her bir maddenin güçlük ve ayırıcılık gücü indeksi hesaplanmıştır. Madde güçlük indeksi ve ayırıcılık gücü düşük olan 13 maddenin testten çıkarılmasıyla soru sayısı 68'e düşmüştür. Elde edilen 68 soru, 6. sınıf öğrencilerine ünite işlenmeden ön test ve ünite işlendikten sonra son test şeklinde 7 hafta ara ile uygulanmıştır. Bulgulara göre bir kazanım hariç, diğer kazanımlara ait ön ve son testteki doğru sayıları anlamlı bir farklılık göstermektedir.

Sanlı Richards(2004) tarafından 8. Sınıf Akademik Yetenek Testi'nin geliştirilmesinde de benzer bir test geliştirme yaklaşımı izlenmiş, öncelikle test puanlarının kullanım amacı, daha sonra testle ölçülecek özellikler belirlenmiş ve bu özellikleri ölçmeye yönelik test maddeleri yazılmıştır. Test için yazılan maddeler konu alanı öğretmenlerinin ve test geliştirme konusunda uzman olan kişilerin incelemesine sunulmuştur. Alınan görüşlere göre uygun düzeltmeler yapıldıktan sonra, maddelerin 299 öğrenci üzerinde deneme uygulaması yapılmıştır. Sonuçlar analiz edilerek maddelerin güçlük ve ayırıcılık değerlerine göre istenen özelliklere sahip maddelerin seçilmesiyle testin asıl formu oluşturulmuştur.

Yapılan literatür taramasında amacı başarı testi geliştirme olmayan, ancak kullanılan yöntem açısından başarı testi geliştirme çalışmaları ile benzerlik gösteren bir ölçek geliştirme çalışmasına rastlanmıştır. Seber (2001) tarafından Çoklu Zekâ Alanlarında Kendini Değerlendirme Ölçeği'nin geliştirilmesinde farklı sosyoekonomik düzeylerdeki bölgelerden toplam 380 öğrenci ve 13 öğretmene araştırmacı tarafından kaynak tarama yolu ile derlenerek geliştirilen ve uzman görüşleri doğrultusunda şekillenen ölçeğin

deneme formu uygulanmıştır. Ölçeğin kapsam geçerliği için on iki uzmandan görüş alınmış, yapı geçerliği için faktör analizi yapılmış, yordama geçerliği için öğrenci puanlarıyla, öğrencilerin öğretmenleri tarafından yapılan değerlendirmeler arasındaki ilişki hesaplanmıştır. Güvenirlik için, test-tekrar test ve iç tutarlılık yöntemleri kullanılmıştır. Bulgulara göre, sekiz boyuttan ve her bir boyutta sekizer madde olmak üzere toplam 64 maddeden oluşan, ilköğretim beşinci sınıf öğrencilerinin çoklu zekâ alanlarında kendilerini değerlendirebilecekleri geçerli ve güvenilir bir ölçeğin geliştirildiği sonucuna varılmıştır.

2. 2. Yurt Dışında Yapılan Araştırmalar

Matematik alanında başarı belirlemeye yönelik yurtdışında yapılan test geliştirme çalışmalarında genellikle bu çalışmanın amacına benzer şekilde okullarda uygulanmakta olan öğretim programı temelinde başarı ölçmeyi amaçlayan çalışmalara rastlanmıştır (Deering, 1988; Gucken, 1986; Jitendra, vd., 1997).

Gucken'in (1986) lise öğrencileri için yüzde kavramı ile ilgili başarı testi geliştirme çalışmasında öncelikle 7, 8 ve 9. sınıflar için kullanılmakta olan ders kitapları incelenerek yüzde kavramı ile ilgili olan 25 konu belirlenmiş, bu konular temelinde 30 kazanım oluşturulmuştur. Kazanımlar için toplamda 242 çoktan seçmeli madde hazırlanıp matematik öğretmenleri ile matematik ve program alanından akademisyenlerden oluşan 17 kişilik bir panele sunularak hem kazanımların hem de soru maddelerinin uygunlukları konusunda görüşler alınmıştır. Maddelerin çeldiricileri, sorular 100 öğrenciye açık uçlu şekilde sorularak en sık görülen yanlış cevaplardan oluşturulmuştur. Uzman görüşleri sonucu uygun görülen 220 madde iki form halinde olacak şekilde ayrılmış, geçerlik, güvenilirlik ve madde analizleri için 548 adet 10. sınıf öğrencisi ile deneme uygulaması yapılmıştır. Madde güçlük indeksinin 0,20 ile 0,80 arasında, ayırıcılık indeksinin 0,20'den büyük olması ve her kazanım için 2 madde bulunması ölçütleri göz önünde bulundurularak madde seçimleri yapılmıştır. Testin güvenilirliği test-tekrar test, paralel formlar ve KR-20 uygulamaları ile test edilmiş ve yüksek güvenilirlik katsayıları elde edilmiştir.

Jitendra, Parker ve Kameenui (1997) tarafından matematik başarısını öğretim programı temelinde ölçmeye yönelik testler geliştirmenin amaçlandığı çalışmada 3, 4 ve 5. sınıf temel matematik dersi öğretmen ve öğrenci kitapları incelenmiş, öğretim programındaki kazanımlar yerine program içeriğini daha iyi yansıttığı düşünülen öğrenme etkinlikleri ölçme için belirlenmiştir. Test maddeleri, kitaplardaki problem türlerine uygun olarak oluşturulmuştur. Hazırlanan testlerin her sınıf düzeyinde 240 öğrenci ile güz ve bahar dönemlerinde deneme uygulamaları yapılmıştır. Testlerin son formuna seçilecek maddeler; güçlük indekslerinin 0,15 ile 0,85 arasında olması, test toplam güçlüğü'nün %30'dan aşağı düşmemesi, 0.90'ın üzerinde güvenirlik katsayısı gibi ölçütlere göre belirlenmiştir. Sonuçta güvenirliği 0,80'in üzerinde, farklı zamanlardaki uygulamalar arasındaki farkı ortaya çıkaran testler elde edilmiştir.

Amerika'da Missouri eyaletinde 2-10. sınıflar düzeyinde İngilizce, sosyal bilimler, matematik ve fen alanları için geliştirilmiş bir başarı testi olan Missouri Mastery and Achievement Test (MMAT)'in amacı öğretim programında bulunan temel beceriler temelinde öğrencilerin kazanmaları önemli görülen becerilerin edinimini ölçmek olarak belirlenmiştir (Deering, 1988). Bu amaçla öncelikle 3, 6, 8 ve 10. sınıflar düzeyinde testler geliştirilmiş, bu testler uygulandıkları bölgelerde ölçme açısından oldukça büyük beğeni aldıktan sonra aynı yöntemle 2, 4, 5, 7 ve 9. sınıf düzeyleri için de testler geliştirilmiştir. Testlerin geliştirilme sürecinde öğretmenler ölçülecek becerilerin belirlenmesi, test maddelerinin yazılması ve seçilmesinde görev almışlardır. Belirlenen 400 beceri için akademisyenlerden oluşan Kapsam Belirleme Komitesi (The Test Content Specification Committee) test kapsamında ölçülecek davranışları belirlemişlerdir. Bu davranışlar ayrıca öğretmenlerden oluşan bir çalıştayda değerlendirilerek test maddeleri için taslaklar oluşturulmuştur. Sonuçta testlerde ölçülecek 350 davranış belirlenmiştir. Devam eden çalıştaylarda akademisyenlerin danışmanlığında öğretmenler tarafından her davranış için ortalama 14 olmak üzere toplamda 4800 test maddesi yazılmıştır. Maddeler tanınmış akademisyenler tarafından ölçme teknikleri açısından incelenmiş, testlerin deneme uygulaması temsil edici bir örneklem seçilerek 350 okulda 36000 öğrenci ile mart ve nisan

aylarında gerçekleştirilmiştir. Elde edilen verilere göre test geliştirme alanından uzmanların ve öğretmenlerin görüşleri ile her davranış için 7 madde seçilerek testler son haline getirilmiştir.

Matematik alanında başarı testi geliştirme çalışmaları arasında test maddelerinin içeriği ve bilişsel düzeyi açısından farklı bir çalışmaya rastlanmıştır. Paek, Holland ve Suppes (1999) tarafından yapılan çalışmada matematikte üstün yetenekli 9-13 yaş arası çocuklar için öğretim programı temelli bilgileri hatırlama yerine zeka bulmacaları gibi zorlayıcı ve matematiksel becerileri kullanmayı gerektiren maddelerden oluşan yapılan Stanford Education Program for Gifted Youth Mathematics Aptitude Test (SEMAT) geliştirilmiştir. Çoğunluğu daha önce Şili’de geliştirilmiş olan bir testten modifiye edilen 12’si aritmetik, 11’i geometri ve 9’u mantık konusundan olmak üzere toplam 32 madde üstün yetenekliler eğitim programına devam eden 248 öğrenciye uygulanmıştır. Öğrencilerin SEMAT puanlarının standart başarı testlerinden aldıkları puanlarla ilişkisi, puanların öğrencilerin üstün yetenekliler programdaki başarılarını yordama durumları ve öğrenciler arasındaki ayırt edicilik gücü incelenmiştir. Test maddelerinin güçlük ve yaşa uygunlukları programdaki öğrencilerin öğretmenlerinin görüşleri ile belirlenmiştir. Testin güvenilirliği için hesaplanan KR-20 katsayısı 0,79 bulunmuştur. Sonuç olarak öğrencilerin SEMAT puanları ile standart başarı testlerinden aldıkları puanlar arasında yüksek bir ilişki bulunmamış, ancak puanların öğrencilerin devam ettikleri programdaki başarılarını iyi bir şekilde yordadığı görülmüştür. Ayrıca maddelerin ayırtıcılık değerleri yüksek bulunmuştur. Hazırlanan test üstün yetenekli çocuklar eğitim programında bazı matematik derslerine öğrenci seçmede bir ölçüt olarak kullanılmaktadır.

Paek, Holland ve Suppes’in (1999) çalışmalarında diğer başarı testi geliştirme çalışmalarından farklı olarak maddelerin hazırlanıp denenerek nitelikli maddelerin seçiminin yapılmadığı, ancak önceden başka bir testte var olan soruların uyarlanarak geçerlik-güvenirlik çalışmasının yapıldığı dikkat çekmektedir.

Hem yurtiçinde hem yurtdışında yapılan başarı testi geliştirme çalışmaları incelendiğinde genel hatlarıyla büyük benzerlikler gösteren

yöntemlerin izlendiği dikkat çekmektedir. Başarı testi geliştirmede ortak denebilecek bir yöntemle öncelikle testte ölçülecek kazanımların belirlendiği, her kazanım için 2 ya da 3 madde yazılarak oluşturulan deneme formunun uygulanması sonucu yapılan madde analizlerine göre aynı kazanımı ölçmeye yönelik maddelerden güçlük ve ayıricılık değerleri uygun olan madde ya da maddelerin seçilmesiyle testlerin son haline getirildiği görülmüştür. Geliştirilen testlerin güvenirliklerinin belirlenmesinde büyük çoğunlukla KR-20 formülü kullanılmıştır.

BÖLÜM 3

YÖNTEM

Bu bölümde araştırmanın modeli, evren ve örnekleme, veri toplama araçları, verilerin toplanması ve analizi açıklanmıştır. Veri toplama araçları kısmında her sınıf düzeyindeki testin uygulama formu için madde hazırlama aşamaları ayrıntılı bir şekilde açıklanmıştır.

Araştırma Modeli

Bu çalışma, ilkokul öğrencilerinin matematik dersi sayılar öğrenme alanındaki başarı düzeylerinin belirlenmesinde kullanılacak geçerli ve güvenilir başarı testlerigeliştirme çalışmasıdır.

Evren ve Örneklem

Araştırmada geliştirilen testlerin deneme uygulamaları testlerde kapsanan kazanımların tamamını görmüş öğrenciler üzerinde yapılmış; her bir sınıf düzeyi için geliştirilen test eğitim-öğretim yılının 1. döneminde bir üst sınıfta bulunan öğrencilere uygulanmıştır. Bunun yanında, testlerin kazanımları görmüş ve görmemiş öğrencileri ayırabilme durumlarını analiz edebilmek için her test, orta sosyoekonomik düzeydeki okullardan birinde bulunan geliştirildiği sınıf düzeyinde bir sınıfta da uygulanmıştır. Böylece 1-4. sınıflar düzeyinde geliştirilecek testler için araştırmanın evreni İstanbul ilindeki ilköğretim okullarının 1–5. sınıflarına devam eden öğrenciler olarak belirlenmiştir.

Araştırmanın örnekleminin belirlenmesinde öncelikle evren içerisindeki ilçelerden alt, orta ve üst sosyoekonomik düzeylerde birer ilçenin, daha sonra bu ilçelerdeki ilköğretim okullarından ikişer okulun, son olarak bu okullardan ikişer sınıfın seçilmesi şeklinde bir yol izlenmiştir. Farklı sosyoekonomik

düzeylerden birer ilçenin seçiminde uygun örnekleme yöntemi kullanılmış ve alt, orta ve üst sosyoekonomik düzeylerden sırasıyla Sultanbeyli, Tuzla ve Kadıköy ilçeleri seçilmiştir. Bu ilçelerden seçilecek okulların belirlenmesinde öncelikle bu ilçelerde 2012-2013 eğitim-öğretim yılında 1-5. sınıfların bulunduğu ilköğretim okulları belirlenmiş, daha sonra bu ilköğretim okulları arasından rastgele ikişer okul seçilmiştir. Son olarak belirlenen okulların her birinden hedeflenen sayıda öğrenciye ulaşılabilmesi için her sınıf düzeyindeki şubeler arasından öğrenci sayılarına göre bir ya da ikisi rastgele seçilmiştir. Çizelge 2’de okulların seçildiği ilçe ve öğrenci sayılarına ait bilgiler gösterilmektedir:

Çizelge 2

Öğrencilerin Farklı İlçe, Okul ve Sınıf Düzeylerine Göre Dağılımı

İlçeler	Okullar	Öğrenci sayıları				
		1. Sınıf	2.sınıf	3. sınıf	4. sınıf	5. sınıf
Sultanbeyli	Okul 1		49	50	50	51
	Okul 2		47	49	48	59
Tuzla	Okul 3	26	87	92	78	49
	Okul 4		60	48	60	62
Kadıköy	Okul 5		47	54	54	48
	Okul 6		55	49	40	39
Toplam		26	345	342	330	308
Genel Toplam		1351				

Çizelge 2’de görüldüğü gibi testlerin deneme uygulamaları 3 farklı ilçeden yaklaşık olarak eşit sayıda öğrencinin katılımıyla her sınıf düzeyinde en az 300 öğrenci ile gerçekleştirilmiştir. Her bir sınıf düzeyindeki test için katılımcı öğrencilerin yaş ve cinsiyetlerine ait bilgiler Çizelge 3’te gösterilmektedir:

Çizelge 3

Testlere Göre Katılımcıların Yaş ve Cinsiyet Dağılımları

Testler	Yaş			Cinsiyet	
	Ortalama	Minimum	Maksimum	Kız	Erkek
1. Sınıf Testi	7,8	6	9	170 (%50)	170 (%50)
2. Sınıf Testi	8,8	7	10	157(%47,9)	171(%52,1)
3. Sınıf Testi	9,9	8	12	170(%51,7)	159(%48,3)
4. Sınıf Testi	10,7	9	12	173(%51,6)	162(%48,4)

Çizelge 3'te görüldüğü gibi her sınıf düzeyindeki testin deneme uygulamasına katılan öğrencilerin cinsiyetlere dağılımı yaklaşık olarak eşittir.

Veri Toplama Araçları

Veri toplama araçları, araştırmanın amacına uygun olarak geliştirilecek başarı testlerinde ölçülmesi hedeflenen 1–4. sınıf matematik öğretim programı sayılar öğrenme alanında belirlenen kritik kazanımların her biri için 2 ya da 3'er soru yazılmasıyla oluşturulan testlerdir. Test geliştirmede aşağıdaki adımların izlenmesi önerilmektedir:

1. Testin hangi amaçla kullanılacağını belirlenmesi
2. Testle ölçülecek davranışların saptanması
3. Maddelerin yazılması
4. Maddelerin gözden geçirilmesi
5. Deneme formunun hazırlanması
6. Deneme uygulamasının yapılması
7. Deneme uygulaması cevap kağıtlarının puanlanması, madde analizi ve madde seçimi
8. Nihai testin oluşturulması ve istatistiklerinin kestirilmesi”(Turgut ve Baykul, 2012, pp. 215-216)

Eğitim alanında ölçme ve değerlendirme; öğretim programının değerlendirilmesi, öğrenme eksikliklerinin belirlenmesi, yapılan öğretimin değerlendirilmesi, öğrenci başarısının belirlenmesi gibi çeşitli amaçlar ile yapılabilmektedir (Turgut ve Baykul, 2012). Bu çalışmada geliştirilmesi amaçlanan testlerin amacı öğrenci başarısını ölçmektir. Geliştirilecek testlerin amacı belirlendikten sonra testlerin hazırlanmasında genel olarak izlenen adımlar aşağıda sıralanmış, daha sonra her bir test için bu adımlara uyularak deneme formunun hazırlanması ayrıntılı olarak açıklanmıştır.

1. Ölçülecek kritik kazanımların belirlenmesi

Öğrenme sürecinin sonunda erişilen düzeyi belirlemeye yönelik hazırlanan başarı testlerinde öğretim programında yer alan kazanımların her biri için soru bulundurulması testte öğrencilerin cevaplayabileceğinden çok fazla sayıda soru olmasına, bu da testin kullanılabilirliğinin azalmasına neden olabilir. Bu durumda konu alanının hedefleri doğrultusunda ilerleme ve gelişmeleri daha iyi yansıtan davranışlar seçilerek bunların ölçülmesi önerilmektedir (Özçelik, 2010). Bu araştırmada 1-4. sınıf matematik öğretim programı sayılar öğrenme alanı kazanımlarından içerik yönünden kapsamlı olanlar kritik kazanım olarak belirlenmiştir. Bunun yanında gerektirdiği bilgi ya da beceri yönünden birbirini gerektiren kazanımlar ayrı ayrı kritik kazanımlar olarak alınmıştır. Böylece testlerdeki soru sayısı fazla artırılmadan tüm kazanımların gerektirdiği bilgi ve becerilerin kapsanmasına çalışılmıştır.

2. Uzman görüşü ile kritik kazanımların netleştirilmesi

Araştırmacı tarafından kritik kazanım olarak belirlenen kazanımların gerçekten alana ait kritik kazanımlar olduğunun netleştirilmesi için uzman görüşü alınmıştır. Bir alanda başarı testi geliştiren kişi testte yoklayacağı bilgi ve becerileri yalnızca kendisinin değil alandaki tecrübeli öğretmenlerin de edinilmeye değer bulup bulmadığını sorgulamalıdır (Ebel, 1965). Görüş ve önerileri alınan uzmanlar matematik eğitimi alanından bir akademisyen ve ilgili sınıflara matematik öğreten dört sınıf öğretmenidir.

3. Kritik kazanımlar için soruların yazılması

Uzman görüşleri de değerlendirilerek netleştirilen kritik kazanımların her biri için 2 ya da 3 soru hazırlanmıştır. Belirlenen bir kazanımı ölçmede

hangi tür ve içerikte bir sorunun daha iyi olduğunun belirlenmesi için bu sorular birbirinden olabildiğince farklı biçim ve içeriklerde oluşturulmaya çalışılmıştır. Örneğin; toplama işlemi becerisi içeren bir kazanım için yan yana veya alt alta yazılmış toplama işlemi şeklinde farklı sorular yazılmıştır.

4. Yazılan sorular için uzman görüşü alınması

Belirlenen kritik kazanımlar için yazılan soruların uygunluğunun kontrol edilmesi amacıyla alan ile ilgili akademisyen ve öğretmenlerin görüşleri alınmıştır. Soruların hem belirlenen kazanımı ölçebilme, hem de teknik açıdan uygun olma durumları irdelenmiştir. Soruların teknik açıdan incelenmesinde dil ve anlatımının açıklığı ve matematik açısından kazanıma ve sınıf düzeyine uygunluğu için matematik eğitimi alanından bir akademisyen ve belirtilen sınıf düzeylerinde matematik öğretimi yapan üç sınıf öğretmeninden; ölçme ilke ve teknikleri açısından uygunluğu için ise bir ölçme ve değerlendirme uzmanından görüşler alınmıştır. Görüşlerin alınmasında her soru için değerlendirme ve önerilerin belirtilmesi için aşağıdaki form kullanılmıştır:

	Katılıyorum	Katılmıyorum	<u>Varsa düzeltme</u>
Soru belirtilen sınıf düzeyine uygundur.			<u>önerileriniz:</u>
Soru kazanıma uygundur.			
Sorunun anlatımı anlaşılırdır.			

5. Testlerin oluşturulması

Test maddeleri için uzmanlardan alınan görüş ve öneriler değerlendirilip bunlara göre gerekli değişiklik ve düzeltmelerin yapılmasından sonra maddeler sayfa düzenine konulmuştur. Turgut ve Baykul (2012) cevaplama oluşabilecek zihinsel yorgunluğu önlemek için başarı testlerinde maddelerin konulara göre gruplandırılmasını önermektedirler. Ayrıca, maddelerin kolaydan zora doğru sıralanmasının cevaplayıcıların soruları istekle cevaplamalarını sağlayacağı belirtilmektedir. Testlerde her kazanım için yazılan 2 ya da 3 madde ard arda sıralanmıştır. Öğretim programındaki kazanımlar aşamalı şekilde ilerlediği için kazanımların

programdaki sıralamasının aynı zamanda güçlük sıralaması olduğu düşünülerek maddelerin kolaydan zora sıralama oluşturmaları sağlanmıştır.

6. Ön uygulama yapılması

Hazırlanan testmaddeler ile ilgili herhangi bir sorun ya da eksiklik olup olmadığının kontrol edilmesinin testlerin büyük gruplara uygulanmadan önce olası aksaklıkların fark edilip önlem alınması için yararlı olacağı düşünülerek her sınıf düzeyinden 20-30 öğrenci ile ön uygulamalar yapılmıştır. Ön uygulamada karşılaşılan eksik ya da sorunlu noktalarda düzeltmeler yapılarak testler deneme uygulaması için hazır hale getirilmiştir.

Test geliştirme sürecinin adımlarına uygun olarak testlerin deneme formlarının hazırlanmasında kritik kazanımların belirlenmesi ve uzman görüşleri ile netleştirilmesi, kritik kazanımlar için soruların yazılması ve uzman görüşleri ile ön uygulama sonuçlarında göre düzenlenmesi şeklinde izlenen yol aşağıda her sınıf düzeyindeki test için ayrı ayrı ele alınmıştır.

Birinci Sınıf Testi Kritik Kazanımlarının Belirlenmesi ve Uzman Görüşleri ile Netleştirilmesi

Kritik kazanım sayılarının belirlenmesinde uzmanların bu sınıf düzeyindeki öğrenciler için bir başarı testinde bulunacak ortalama soru sayısı hakkındaki görüşleri dikkate alınmış, sınıf öğretmenlerince 1. sınıf düzeyinde bir başarı testindeki soru sayısının ortalama 10 olması uygun görülmüştür.

Doğal sayılar alt öğrenme alanındaki kazanımlardan “100 içinde ileriye doğru birer ve onar ritmik sayar” ve “20 içinde geriye birer sayar” ritmik sayma becerisinin önemi dolayısıyla kritik kazanımlar olarak belirlenmiştir.

“Miktarı 20’ye kadar olan bir grup nesneyi, onluk ve birliklerine ayırarak gösterir, bu nesnelere karşılık gelen sayıyı rakamlarla yazar ve okur” nesne sayısı açısından “Nesne sayısı 10’dan az olan bir topluluktaki nesnelerin sayısını belirler ve bu sayıyı rakamla yazar” kazanımını kapsadığından kritik kazanım olarak seçilmiştir. Kazanım ayrıca rakamları yazma becerisi içerdiğinden “Rakamları okur ve yazar” kazanımını da içine almaktadır. Ancak, kısa cevaplı soru türüne uygun olarak “Miktarı 20’ye kadar

olan bir grup nesneye karşılık gelen sayıyı rakamlarla yazar” şekliyle alınmıştır.

“20’ye kadar olan bir sayıya karşılık gelen çokluğu belirler” kazanımındaki sayıya karşılık gelen bir ‘çokluk’ belirleme işi sayı ile nesneler birebir eşleme ve karşılaştırma becerileri gerektirdiğinden “Miktarları 20’den az nesnelerden oluşan iki gruptaki nesneleri bire bir eşler, grupların nesne sayılarını karşılaştırır” kazanımındaki becerileri kapsamaktadır. Bu da “Nesne gruplarını azlık ve çokluklarına göre karşılaştırır” kazanımını kapsadığından, kazanımların ilk kazanım içerisinde değerlendirilmesi uygun görülmüştür.

“Bir ‘çokluktan’ belirtilen sayı kadarını ayırır” kazanımı kısa cevaplı soru ile ölçülemeyeceği düşünülmüş kritik kazanımlar arasına alınmamıştır.

“Sıra bildiren ilk yirmi sayıyı kullanır” kazanımı sayıların sahip olduğu sıra belirtme işlevini içerdiğinden kritik kazanım olarak alınmıştır.

Doğal sayılarla toplama işlemi alt öğrenme alanındaki “Toplamanın bir araya getirme, ekleme ve çoğaltma anlamlarını fark eder”, “Toplama işleminde sıfırın etkisini nedenleriyle açıklar”, “Toplama işleminde, toplananların yerleri değiştirildiğinde toplamın değişmediğini gösterir” ve “Toplamları 20’ye kadar olan iki doğal sayıyı zihinden toplar” kazanımlarındaki açıklama, zihinden gibi becerilerin kısa cevaplı soru ile ölçülemeyeceği düşünülmüş kritik kazanımlara dahil edilmemiştir.

“Toplamları 20’ye kadar olan iki doğal sayının toplamını bulur, matematik cümlesini yazar ve modelle gösterir” kazanımı kritik kazanımlara seçilmiş, ancak kısa cevaplı soru türüne uygun olan “Toplamları 20’ye kadar olan iki doğal sayının toplamını bulur” şekliyle alınmıştır.

“Toplamları 20’yi geçmeyen iki doğal sayının toplandığı bir işlemde, toplam ile toplananlardan biri verildiğinde verilmeyen toplananı bulur” kazanımının öncelikle verilmeyen toplananı bulmak için çıkarma işlemi yapılacağından çıkarma işlemi alt öğrenme alanındaki ilgili kazanımda kapsandığı düşünülmüştür. Ancak burada çıkarma işlemi açıkça sunulmadığından çıkarma işlemi yapmanın gerekliliğini fark etme becerisini içerdiği düşünülmüş kritik kazanım olarak seçilmiştir.

“20’ye kadar olan doğal sayıları, iki doğal sayının toplamı biçiminde yazar” kazanımı ile sayılar arasındaki ilişkileri anlama açısından paralel olan “Toplamları 10 veya 20 olan sayı ikililerini belirler” kazanımı, 10 ve 20 sayılarını köprü olarak kullanarak zihinden toplama ve tahmin becerileri için önemli görüldüğünden ölçülecek kritik kazanımlara dahil edilmiştir.

“Doğal sayılarla toplama işlemini gerektiren problemleri çözer ve kurar” kazanımındaki problem çözme becerisi önemli görülmüş, ancak problem kurmanın kısa cevaplı soru ile ölçülemeyeceği düşünüldüğünden “Doğal sayılarla toplama işlemini gerektiren problemleri çözer” şekliyle alınmıştır.

Çıkarma işlemi alt öğrenme alanında “Çıkarmanın ayırma, azaltma ve eksiltme anlamlarını fark eder” ve “20’ye kadar olan iki doğal sayının farkını zihinden bulur” kazanımları fark etme ve zihinden işlem becerilerinin kısa cevaplı soru ile ölçülemeyeceği düşünülerek kritik kazanımlara alınmamıştır.

“20’ye kadar olan iki doğal sayının farkını bulur, matematik cümlesini yazar ve modellerle gösterir” kazanımı “Bir doğal sayıdan aynı doğal sayı çıkarıldığında “sıfır” elde edildiğini gösterir” kazanımını içerdiğinden seçilmiş, ancak “20’ye kadar olan iki doğal sayının farkını bulur” şekliyle alınmıştır.

“Bir çıkarma işleminde verilmeyen eksilen veya çıkanı bulur” kazanımı verilmeyen eksileni ve çıkanı bulmak için ters işlemlerini yapma gerekliliğini fark etme becerisi içerdiğinden kritik kazanımlara dahil edilmiştir.

“Doğal sayılarla çıkarma işlemini gerektiren problemleri çözer ve kurar” kazanımındaki problem çözme becerisi önemli görülmüş, ancak problem kurma kısmının kısa cevaplı soru ile ölçülemeyeceği düşünüldükten “Doğal sayılarla çıkarma işlemini gerektiren problemleri çözer” şekliyle alınmıştır.

“Uygun şekil veya nesneleri iki eş parçaya böler ve yarımı belirtir” kazanımının gerektirdiği beceriyi kapsayan “Yarım ve bütün arasındaki ilişkiyi açıklar” kazanımı kritik kazanım olarak belirlenmiştir.

Birinci sınıf matematik öğretim programı sayılar öğrenme alanındaki kazanımlar uzman görüşleri de dikkate alınarak yukarıdaki gibi irdelenmiş ve 1. sınıf testi için 13 adet kritik kazanım belirlenmiştir.

Birinci Sınıf Testi Kritik Kazanımları İçin Soruların Yazılması ve Uzman Görüşleri ile Düzenlenmesi

“100 içinde ileriye doğru birer ve onar ritmik sayar” kazanımı için biri birer, diğeri onar ritmik sayma şeklinde ilerleyen sayı sıralamalarında verilmeyen sayıların yazılmasının istendiği iki boşluk doldurma sorusu yazılmıştır. Birer ritmik sayma sorusunda öğretim programındaki birer ritmik saymanın 1’den başlatılması uyarısı dikkate alınmıştır.

Sorular için olumlu uzman görüşleri alınmış, yalnızca boşluk doldurma için “hangi sayılar yazılmalıdır?” ifadesi ölçme ve değerlendirme uzmanının görüşü doğrultusunda “uygun sayıları yazınız” olarak değiştirilmiştir.

“20 içinde geriye birer sayar” kazanımı için 20’den geriye doğru sıralanmış sayılardan verilmeyen iki tanesinin yazılmasının istendiği boşluk doldurma türünde bir soru yazılmıştır.

“Miktarı 20’ye kadar olan bir grup nesneye karşılık gelen sayıyı rakamlarla yazar” kazanımı için birinde 16 ayıcık, diğesinde 12 balon resminin verildiği “kaç tane var” türündeiki soru yazılmıştır. Bu sorular için alınan uzman görüşleri olumlu bulunduğundan herhangi bir değişiklik yapılmamıştır.

“20’ye kadar olan bir sayıya karşılık gelen çokluğu belirler” kazanımı için boş bir kutucuk verilerek cevaplayıcılardan içine 14 tane yuvarlak çizimleri istenmiş, soru için alınan uzman görüşleri olumlu bulunmuştur.

“Sıra bildiren ilk yirmi sayıyı kullanır” kazanımı için sorulardan birinde resimde bir koşu yarışmasındaki koşucular altlarında isimleri ile gösterilerek bunlardan adı verilen birinin kaçınıcı olduğu, diğesinde ise önünde 12 kişi olan kişinin baştan kaçınıcı olduğu sorulmuştur. Sorular için olumlu uzman görüşleri alınmış, yalnızca koşucuların koşma yönlerinin açık şekilde anlatılması önerildiğinden soru uygun şekilde düzenlenmiştir.

“Toplamları 20’ye kadar olan iki doğal sayının toplamını bulur” kazanımı için biri alt alta, diğeri yan yana toplama işlemi şeklinde yazılmış sırasıyla iki ve üç sayının toplanmasını içeren iki soru yazılmıştır.

“Toplamları 20’yi geçmeyen iki doğal sayının toplandığı bir işlemde, toplam ile toplananlardan biri verildiğinde verilmeyen toplananı bulur” kazanımı için biri toplananlarından biri verilmemiş alt alta toplama işlemi, diğeri problem olan iki soru yazılmış, sorular için olumlu görüşler alınmıştır.

“Toplamları 10 veya 20 olan sayı ikililerini belirler” kazanımı için birinde toplamları 10, diğeri 20 olan iki sayının yazılmasının istendiği, dolayısıyla birden çok doğru cevabı olan iki soru yazılmış, olumlu görüşler alınmıştır.

“Doğal sayılarla toplama işlemini gerektiren problemleri çözer ve kurar” için biri iki diğeri üç sayının toplanmasını gerektiren iki soru yazılmıştır. Ölçme uzmanının sorulardan birindeki “hediye ederlerse” ifadesinin olasılık durumu içerdiğini belirtmesi ile ifade “verdiler” olarak değiştirilmiştir.

“20’ye kadar olan iki doğal sayının farkını bulur” kazanımı için üç soru yazılmıştır. Aynı kazanım içerisinde kapsandığı düşünülen “Bir doğal sayıdan aynı doğal sayı çıkarıldığında ‘sıfır’ elde edildiğini gösterir” kazanımının da ölçülmesi için sorulardan ikisi sayıların alt alta ve yan yana verildiği bir doğal sayıdan aynı doğal sayının çıkarılmasını gerektirecek şekilde seçilmiştir.

“Bir çıkarma işleminde verilmeyen eksilen veya çıkanı bulur” kazanımı için birincisinde verilmeyen eksilenin, ikincisinde verilmeyen çıkanın istendiği çıkarma işlemleri içeren iki soru yazılmıştır. Sorularda verilemeyen eksilen veya çıkan için konulan noktalı yere uygun sayının yazılması istenmiştir.

“Doğal sayılarla çıkarma işlemini gerektiren problemleri çözer” kazanımı için birincisinde iki basamaklı bir sayıdan bir basamaklı bir sayının çıkarılmasını, ikincisinde ise verilmeyen çıkanın bulunması için çıkarma işlemi gerektiren iki problem yazılmıştır:

Soru 1:Canan’ın akvaryumunda 14 balık vardı. Balıklardan 3 tanesi ölürse akvaryumda kaç balık kalır?

Soru 2:Arda’nın 15 lirası vardı. Kantinden alışveriş yaptıktan sonra geriye 12 lirası kaldığına göre kantinde kaç lirasını harcamıştır?

İkinci sorudaki “ölürse” ifadesinin 1. sınıf düzeyindeki öğrenciler için olumsuz olabileceği yönündeki uzman görüşü doğrultusunda, ifade “Canan, balıklarından 3 tanesini arkadaşına verdiğinde” şeklinde değiştirilmiştir.

“Yarım ve bütün arasındaki ilişkiyi açıkla” kazanımı için biri bütünden yarıma, diğeri yarımdan bütüne dönüşüm içeren, üçüncüsü ise yarım olarak verilen bir daire şeklinin bütününe çizilmesini isteyen üç soru oluşturulmuştur. Sorular uzmanlar tarafından uygun bulunmuştur.

Uygulama için hazırlanan 1. sınıf testi toplamda 26 adet kısa cevaplı sorudan oluşturulmuştur.

İkinci Sınıf Testi Kritik Kazanımlarının Belirlenmesi ve Uzman Görüşleri ile Netleştirilmesi

Kritik kazanımların sayılarının belirlenmesinde uzmanlardan bu sınıf düzeyindeki öğrenciler için bir başarı testinde bulunacak ortalama soru sayısı için alınan görüşler dikkate alınmıştır. Sınıf öğretmenlerine göre 2. sınıf düzeyinde bir başarı testinde bulunacak soru sayısı 10-15 arasında olmalıdır.

“Deste ve düzineyi örneklerle açıkla” kazanımı görüşü alınan sınıf öğretmenleri tarafından yalnızca deste ve düzinedeki nesne sayısı bilgisini içerdiği ve bu sınıf düzeyinde kritik bir kazanım olmadığı belirtildiğinden testte ölçülecek kazanımlara alınmamıştır.

“100’den küçük doğal sayıların basamaklarını adlandırır, basamaklardaki rakamların basamak değerlerini belirtir” kazanımının, basamakları adlandırmada onluk ve birlik bilgisi gerektirdiğinden “Nesne sayısı 100’den az olan bir çokluğu, onluk ve birlik gruplara ayırarak bunlara karşılık gelen sayıyı yazar ve okur” kazanımını kapsadığı düşünülmüş ve uzmanlar tarafından önemli görüldüğünden kritik kazanımlara alınmıştır.

“Sayı örüntüleri oluşturur” kazanımının ritmik artan sayılardan oluşan örüntüler oluşturma becerisini de içerdiğinden “100 içinde ikişer ve beşer, 40 içinde dörder, 30 içinde üçer ileriye ve geriye doğru sayar” kazanımını kapsamakta olduğu düşünülmüş ve kritik kazanım olarak seçilmiştir.

“100’den küçük en çok dört doğal sayıyı büyükten küçüğe veya küçükten büyüğe doğru sıralar” kazanımı karşılaştırma yapılacak sayı adedi

bakımından “100’den küçük iki doğal sayıyı karşılaştırarak aralarındaki ilişkiyi belirtir” kazanımını kapsadığından kritik kazanımlara dahil edilmiştir.

“Sıra bildiren sayıları sözlü ve yazılı olarak kullanır” ve “İki basamaklı doğal sayıların hangi onluğa daha yakın olduğunu belirler” uzmanlardan alınan görüşlere göre 2. sınıf için kritik kazanımlar olarak görülmemiştir.

Doğal sayılarla toplama işlemi alt öğrenme alanındaki kazanımlar incelendiğinde öncelikle eldeli toplama işleminin eldesiz toplama işlemini kapsadığı düşünülmüş, ancak eldeli toplama işlemi yapmadaki olası bir hatanın sadece eldenin akılda tutulup onlar basamağındaki sayıya eklenmesinden kaynaklanma durumunda bir karışıklık olacağı düşüncesi ile bu iki becerinin ayrı ayrı ölçülmesine karar verilmiştir. Sonuçta “Toplamları 100’e kadar olan doğal sayıların eldesiz toplama işlemini yapar” ve “Eldeli toplama işlemini yapar” kazanımları ayrı kritik kazanımlar olarak alınmıştır.

“İki doğal sayının toplandığı işlemde verilmeyen toplananı belirler” verilmeyen toplananı belirlemek için çıkarma işlemi yapmanın gerekliliğini fark etme becerisi gerektirdiğinden kritik kazanım olarak seçilmiştir.

“Toplamları 100’u geçmeyen, 10 ve 10’un katı olan doğal sayıların toplamını zihinden bulur”, “Toplamları 50’yi geçmeyen iki doğal sayıyı zihinden toplar” ve “Toplamı 100’e kadar olan iki doğal sayının toplamını tahmin eder ve tahminini işlem sonucuyla karşılaştırır” kazanımları zihinden işlem ve tahmin becerilerinin kısa cevaplı soru ile ölçülemeyeceği düşünüldükten kritik kazanımlara dahil edilmemiştir.

“Doğal sayılarla toplama işlemini gerektiren problemleri çözer ve kurar” kazanımındaki problem çözme becerisi önemli görülmüş, ancak problem kurmanın kısa cevaplı soru ile ölçülemeyeceği düşünüldüğünden “Doğal sayılarla toplama işlemini gerektiren problemleri çözer” şekliyle alınmıştır.

“100’den küçük ve onluk bozmayı gerektirmeyen iki doğal sayının farkını bulur” öncelikle onluk bozmayı gerektiren çıkarma işlemi ile ilgili kazanım içerisinde kapsandığı düşünüldükten kritik kazanımlara alınmamıştır. Ancak onluk bozma gerektirmeyen çıkarma işlemi yapıp onluk bozma gereken durumda yapamayan öğrenciler ile ilgili uzmanların gözlemlerine dayanılarak hem bu kazanım hem de “Onluk bozmayı gerektiren iki doğal sayının farkını bulur” kazanımı kritik kazanımlara dahil edilmiştir.

“100’den küçük ve 10’un katı olan iki doğal sayının farkını zihinden bulur” ve “100’e kadar olan doğal sayılarla yapılan çıkarma işleminin sonucunu tahmin eder, tahminini işlem sonucuyla karşılaştırır” kazanımları zihinden işlem yapma ve tahmin etme becerilerinin kısa cevaplı soru ile ölçülemeyeceği düşünülerek kritik kazanımlara dahil edilmemiştir.

“Doğal sayılarla yapılan bir çıkarma işleminde verilmeyen eksileni veya çıkanı belirler” kazanımı verilmeyen eksileni bulmak için toplama işlemi veya verilmeyen çıkanı bulmak için çıkarma işlemi yapmanın gerekliliğini fark etme becerisi içerdiğinden kritik kazanım olarak belirlenmiştir.

“Doğal sayılarla toplama ve çıkarma işlemlerini gerektiren problemleri çözer ve kurar” çıkarma ile toplama işlemlerinin birlikte olduğu iki adımlı problemleri kapsadığından kritik kazanım olarak belirlenmiş, ancak problem kurmanın kısa cevaplı soru ile ölçülemeyeceği düşünüldüğünden “Doğal sayılarla çıkarma işlemini gerektiren problemleri çözer” şekliyle alınmıştır.

“İki sayıyı çarpmanın ne anlama geldiğini modellerle açıklar” kazanımı modellerle açıklama becerisinin kısa cevaplı soru ile ölçülemeyeceği düşünülerek kritik kazanım olarak alınmamıştır.

“Toplamları 20’ye kadar ve toplananları aynı olan toplama işlemlerini, çarpma işlemine dönüştürerek çarpma işlemi yapar” kazanımının kritik kazanımlar arasına alınmasına öğretmenler tarafından gerek görülmemiştir.

“10’a kadar olan doğal sayıları 2, 3, 4 ve 5 sayılarıyla çarpar” kazanımı ilk kez 2. sınıf öğretim programında bulunan çarpma işlemi alt öğrenme alanına ait temel işlemsel becerileri içerdiğinden kritik kazanımlara seçilmiştir.

“Çarpma işleminde “1” ve “0”ın etkisini açıklar”, “Çarpma işleminde çarpanların yerleri değiştirildiğinde çarpımın değişmeyeceğini gösterir” ve “Çarpımı 100’u geçmeyen ve bir çarpanı 10 olan çarpma işlemlerini zihinden yapar” kazanımları açıklama, gösterim ve zihinden işlem becerilerinin kısa cevaplı soru ile ölçülemeyeceği düşünülerek kritik kazanımlara alınmamıştır.

“Biri çarpma işlemi olmak üzere en çok iki işlem gerektiren problemleri çözer ve kurar” kazanımı biri çarpma işlemi olan iki adımda çözülen problemleri kapsadığından kritik görülüş, ancak problem kurmanın kısa cevaplı soru ile ölçülemeyeceği düşünüldüğünden “Biri çarpma işlemi olmak üzere en çok iki işlem gerektiren problemleri çözer” şekliyle alınmıştır.

“Kalansız olarak gruplandırılabilen en çok 20 nesneyi; birerli, ikişerli, üçerli, dörderli ve beşerli gruplandırarak grup sayısını belirtir” kazanımında bölme işleminin tekrarlı sayma anlamı, “En çok 20 nesneyi kalansız olarak 2, 3, 4 ve 5 gruba eşit olarak paylaştırarak her gruptaki nesne sayısını belirtir” kazanımında ise eşit paylaşım anlamı vardır. Bu kazanımlar bölme işleminde sırasıyla verilmeyen bölen ve bölümün bulunması ile ilgili olduğundan yapılacak işlem açısından aralarında önemli bir ayrım görülmemiştir. Uzman görüşleri ile de desteklenen ikinci kazanım kritik kazanım olarak seçilmiştir.

“Eksileni 20’yi geçmeyen ve çıkanları aynı olan ardışık çıkarma işlemini, bölme işlemine dönüştürerek bölme işlemini yapar” kazanımının kritik kazanımlara alınmasına öğretmenler tarafından gerek görülmemiştir.

“Bütün, yarım ve çeyrek arasındaki ilişkiyi açıklar” kazanımı ilk kez 2. sınıf öğretim programında karşılaşılan çeyrek kavramını içerdiğinden kritik kazanım olarak alınmıştır.

İkinci sınıf matematik öğretim programı sayılar öğrenme alanındaki kazanımlar uzman görüşleri de dikkate alınarak yukarıdaki gibi analiz edilmiş ve 2. sınıf testi için 15 adet kritik kazanım belirlenmiştir.

İkinci Sınıf Testi Kritik Kazanımları için Soruların Yazılması ve Uzman Görüşleri ile Düzenlenmesi

“100’den küçük doğal sayıların basamaklarını adlandırır, basamaklardaki rakamların basamak değerlerini belirtir” kazanımı için basamak adı ve basamak değerlerinin istendiği iki soru yazılmıştır.

“Sayı örüntüleri oluşturur” kazanımı için yazılan sorulardan ikisi bir sayı örüntüsünde verilmeyen sayının bulunmasını diğeri ise bir problem içerisinde tablo şeklinde verilmiş bir örüntüdeki verilmeyen sayının bulunmasını istemektedir. Sorularda kullanılan sayılar “100 içinde ikişer ve beşer, 40 içinde dörder, 30 içinde üçer ileriye ve geriye doğru sayar” kazanımını da kapsayacak şekilde seçilmiştir. Alınan uzman görüşlerine uyularak “hangi sayı yazılmalıdır?” yerine “uygun sayıyı yazınız” getirilmiştir.

“100’den küçük en çok dört doğal sayıyı büyükten küçüğe veya küçükten büyüğe doğru sıralar” kazanımı için 100’den küçük üç ve dört sayının sıralanmasının istendiği iki soru yazılmıştır. Sorulardan birinde üç

kardeşin yaşları verilerek büyük, ortanca ve küçüğün belirlenmesi istenmiş, diğer soruda dört sayının küçükten büyüğe doğru sıralanması istenmiştir.

“Toplamları 100’e kadar olan doğal sayıların eldesiz toplama işlemini yapar” kazanımı için toplanan sayısı birinde iki, diğerinde öğretim programındaki gibi üç olan ve eldesi olmayan iki soru yazılmıştır.

“Eldeli toplama işlemini yapar” kazanımı için toplanan sayısı birinde iki, diğerinde üç olan eldeli toplama işlemleri içeren iki soru yazılmıştır.

“İki doğal sayının toplandığı işlemde verilmeyen toplananı belirler” kazanımı için biri alt alta, diğeri yan yana yazılmış toplama işleminde verilmeyen toplananın istendiği iki soru yazılmıştır. Sorular için olumlu görüşler alınmış, yalnızca verilmeyen toplanan yerine konan kutucuğun 2. sınıf düzeyinde bir sayının yerine kullanımının anlatımı güçleştirebileceği düşünülerek soru kutucuk yerine noktalı yer olacak şekilde düzenlenmiştir.

“Doğal sayılarla toplama işlemini gerektiren problemleri çözer” kazanımı için aşağıdaki gibi birincisinde iki, ikincisinde üç doğal sayının toplanmasını gerektiren iki soru yazılmıştır:

Soru 1:Emel’in kumbarasında 46 lira vardı. Kumbaraya 14 lira daha atarsa kumbarada kaç lira birikmiş olur?

Soru 2: Elif’in 23, Pelin’in 18 ve Okan’ın 21 hikaye kitabı vardır. Üçü kitaplarını bir araya getirdiğinde kaç kitap olur?

İlk sorudaki “vardı” kelimesinin sorudaki zaman uyumunu bozduğuna dair uzman görüşe göre bu kelime “var” olarak düzeltilmiştir. İkinci soru “kitaplarını bir araya getirme” durumu yerine günlük hayatta karşılaşılabilecek bir durum olarak “kitaplarını kütüphaneye getirme” şeklinde düzenlenmiştir.

“100’den küçük ve onluk bozmayı gerektirmeyen iki doğal sayının farkını bulur” kazanımı için biri alt alta diğeri yan yana verilmiş iki doğal sayı ile çıkarma işlemi içeren iki soru yazılmıştır.

“Onluk bozmayı gerektiren iki doğal sayının farkını bulur” kazanımı için yazılan iki sorudan birinde birler basamağı 0’dan farklı, diğerinde 0 olan sayıların kullanıldığı onluk bozmayı gerektiren çıkarma işlemleri verilmiştir.

“Doğal sayılarla yapılan bir çıkarma işleminde verilmeyen eksileni veya çıkanı belirler” kazanımı için birinde verilmeyen eksilen, ikincisinde verilmeyen çıkanın istendiği çıkarma işlemleri; üçüncüsünde ise verilmeyen eksilenin çıkarma işlemde değil problem durumu içerisinde sorulduğu üç soru yazılmıştır. Sorulardan birinde verilmeyen çıkan yerine konan kutucuğun bu sınıf düzeyinde bir sayının yerine kullanılmasının anlatımını güçleştirebileceği düşünülerek soru kutucuk yerine noktalı yer olacak şekilde düzenlenmiştir.

“Doğal sayılarla toplama ve çıkarma işlemlerini gerektiren problemleri çözer” kazanımı için yazılan üç sorudan ilkinde onluk bozmayı gerektirmeyen çıkarma işlemi, ikincisinde verilmeyen eksilenin istendiği çıkarma işlemi, üçüncüsünde ise iki işlem gerektiren problemler kullanılmıştır:

Soru 1: Didem’in 75 lirası vardır. 22 lirasını harcadıktan sonra kaç lirası kalır?

Soru 2: Bir markette 45 tane ekmek vardı. Ekmeklerden bir miktarı satıldıktan sonra geriye 23 ekmek kaldı. Kaç ekmek satılmıştır?

Soru 3: Ömer’in 18 bilyesi var. Bilyelerinden 5’ini Emre’ye, 6’sını Can’a verirse kaç bilyesi kalır?

“10’a kadar olan doğal sayıları 2, 3, 4 ve 5 sayılarıyla çarpar” kazanımı için kazanımda belirtilen sayılarla çarpma işlemi içeren üç soru yazılmıştır.

“Biri çarpma işlemi olmak üzere en çok iki işlem gerektiren problemleri çözer” kazanımı için ilki bir çarpma, ikincisi bir çarpma ve bir çıkarma işlemi gerektiren iki problem yazılmış, olumlu uzman görüşleri alınmıştır.

“En çok 20 nesneyi kalansız olarak 2, 3, 4 ve 5 gruba eşit olarak paylaştırarak her gruptaki nesne sayısını belirtir” kazanımı için aşağıdaki iki soru yazılmıştır:

Soru 1: Elif kumbarasına her gün 2 lira atarak 10 lirayı kaç günde biriktirir?

Soru 2: Bir anne 20 lirayı 4 çocuğuna eşit olarak paylaştırırsa her çocuğa kaç lira düşer?

Bir öğretmenin 1. soruda gruptaki nesne sayısı yerine oluşan grup sayısının sorulduğu uyarısı üzerine soru “Elif 5 günde 10 lira biriktirmiştir. Buna göre 1 günde kaç lira biriktirmiştir?” şeklinde düzenlenmiştir.

“Bütün, yarım ve çeyrek arasındaki ilişkiyi açıkla” kazanımı için bütün, yarım ve çeyrek arasında dönüşümler içeren üç soru yazılmıştır. Sorular için olumlu görüşler alınmış, yalnızca “bütün” kelimesi, anlatımda oluşabilecek belirsizliği önlemek amacıyla “1 bütün” şeklinde değiştirilmiştir.

Uygulama için hazırlanan 2. sınıf testi toplamda 35 adet kısa cevaplı sorudan oluşturulmuştur.

Üçüncü Sınıf Testi Kritik Kazanımlarının Belirlenmesi ve Uzman Görüşleri ile Netleştirilmesi

Kritik kazanımların sayılarının belirlenmesinde uzmanlardan bu sınıf düzeyindeki öğrenciler için bir başarı testinde bulunacak ortalama soru sayısı için alınan görüşler dikkate alınmıştır. Sınıf öğretmenlerine göre 3. sınıf düzeyinde bir başarı testinde bulunacak soru sayısı ortalama 15 olmalıdır.

“Üç basamaklı doğal sayıları okur ve yazar” kazanımı, öğretim programında ilk kez bu sınıf düzeyinde karşılaşılan üç basamaklı sayıların okunmasını ve okunuşu verilen sayıların doğru şekilde yazılmasını içerdiğinden kritik kazanım olarak seçilmiştir.

“Üç basamaklı doğal sayıların basamak adlarını, basamaklarındaki rakamların basamak değerlerini belirtir” kazanımı basamak değerlerinin önemi dolayısıyla kritik kazanımlara dahil edilmiştir.

“En çok üç basamaklı doğal sayıları en yakın onluğa yuvarlar” uzmanlar tarafından bu sınıf düzeyinde kritik kazanım olarak görülmemiştir.

“1000’den küçük en çok beş doğal sayıyı, büyükten küçüğe veya küçükten büyüğe doğru sembol kullanarak sıralar” kazanımı büyüklük-küçüklük karşılaştırması yapılacak sayı adedi bakımından “1000’den küçük iki doğal sayıyı karşılaştırır ve aralarındaki ilişkiyi sembol kullanarak belirtir” kazanımını kapsadığı düşünülerek kritik kazanım olarak alınmıştır.

“Bir örüntüdeki ilişkiyi belirler ve örüntüyü genişletir” kazanımı ritmik artan sayılardan oluşan örüntüler oluşturma becerisini de içerdiğinden “100 içinde altışar, yedişer, sekizer ve dokuzar ileriye doğru sayar” kazanımını kapsamakta olduğu düşünülmüş ve kritik kazanım olarak seçilmiştir.

“Tek ve çift doğal sayıları belirtir” kazanımının içerdiği tek ve çift kavramları ileriki sınıf düzeylerinde karşılaşılabilecek bölünebilme gibi konulara

temel oluřturmaları aısından uzmanlar tarafından da nemli kavramlar olarak grldğnden bu kazanım kritik kazanımlar arasına alınmıřtır.

“20’ye kadar olan Romen rakamlarını okur ve yazar” kazanımı grřne bařvurulan sınıf ğretmenleri tarafından kritik kazanım olarak dřnlmediğinden kritik kazanımlara dahil edilmemiřtir.

Topamları en ok  basamaklı olan doğal sayılarla eldesiz ve eldeli toplama iřlemine yapar” kazanımıtoplamları  basamaklı olan eldeli ve eldesiz toplama iřlemleri nemli grlerek kritik kazanımlara alınmıřtır.

“İki doğal sayının toplamını tahmin eder ve tahminini iřlem sonucuyla karřılařtırır” ve “Topamları 100’ gemeyen en ok iki doğal sayıyı zihinden toplar” kazanımlarıtahmin ve zihinden iřlem becerilerinin kısa cevaplı soru ile llemeyeceğ dřnldğnden kritik kazanımlara alınmamıřtır.

“Doğal sayılarla toplama iřlemine gerektiren problemleri zer ve kurar” kazanımındaki problem zme becerisi nemli grlmř, ancak problem kurmanın kısa cevaplı soru ile llemeyeceğ dřnlerek “Doğal sayılarla toplama iřlemine gerektiren problemleri zer” řekliyle alınmıřtır.

“En ok  basamaklı doğal sayılarla ıkarma iřlemi yapar” kazanımı ıkarma iřlemi yapma becerisinin nemi dolayısıyla kritik kazanım olarak seilmiřtir.

“İki basamaklı doğal sayılarla zihinden ıkarma iřlemine yapar” ve“10’un katı olan  basamaklı doğal sayılardan, 10’un katı olan en ok  basamaklı doğal sayıları zihinden ıkarır” kazanımlarının ierdiğ zihinden iřlem becerisinin kısa cevaplı soru ile llemeyeceğ dřnlerek bu kazanımlar kritik kazanımlara dahil edilmemiřtir.

“Doğal sayılarla toplama ve ıkarma iřlemlerini gerektiren problemleri zer ve kurar” kazanımı ıkarma iřlemine problem zmede kullanma becerisi ierdiğinden nemli grlmř, ancak problem kurmanın kısa cevaplı soru ile llemeyeceğ dřnlerek “Doğal sayılarla toplama ve ıkarma iřlemlerini gerektiren problemleri zer” řekliyle alınmıřtır.

“Çarpım tablosunu oluşturur” kazanımı,3. sınıf düzeyinde 2. sınıftan farklı olarak çarpım tablosunun tamamını içerdiğinden seçilmiştir.

“Eldeli çarpma işlemini yapar, eldenin ne anlama geldiğini açıklar” kazanımı çarpma işleminde eldeyi akılda tutma becerisini de içerdiğinden kritik kazanım olarak seçilmiş, ancak kısa cevaplı soru ile ölçülebilmesi açısından uygun olan “Eldeli çarpma işlemini yapar” şekliyle alınmıştır. Bu kazanım ayrıca “Çarpımları 1000’den küçük olacak şekilde en çok üç basamaklı iki doğal sayıyla çarpma işlemi yapar” kazanımını kapsamaktadır.

“En çok iki basamaklı doğal sayıları 10 ile; bir basamaklı doğal sayıları 100 ile kısa yoldan çarpar” kazanımı kısa yoldan işlem yapma becerisi kısa cevaplı soru ile ölçülemeyeceğinden kritik kazanım olarak alınmamıştır.

“Doğal sayılarla çarpma işlemini gerektiren problemleri çözer ve kurar”kazanımı problem çözme becerisi içerdiğinden önemli görülmüş,ancak problem kurmanın kısa cevaplı soru ile ölçülemeyeceği düşünülerek “Doğal sayılarla çarpma işlemini gerektiren problemleri çözer” şekliyle alınmıştır.

“İki basamaklı doğal sayıları bir basamaklı doğal sayılara böler” kazanımı, 2. sınıf düzeyindenbölünen sayı bakımından kapsamlı olduğu için 3. sınıf düzeyinde kritik kazanımlar arasında görülmüştür.

“Biri bölme olacak şekilde iki işlem gerektiren problemleri çözer ve kurar” problem çözme becerisi içerdiğinden önemli görülmüş, ancak problem kurmanın kısa cevaplı soru ile ölçülemeyeceği düşünülerek “Biri bölme olacak şekilde iki işlem gerektiren problemleri çözer” şekliyle alınmıştır.

“Bir bütünü eş parçalara ayırarak eş parçalardan her birinin kesrin birimi olduğunu belirtir” kazanımındaki kesir birimi ile ilgili kavramsal bilginin kısa cevaplı soru ile ölçülemeyeceği düşünülerek kritik kazanımlar arasına alınmamıştır.

“Payı paydasından küçük ve paydası en çok iki basamaklı doğal sayı olan kesirler elde eder” kazanımı performans testi ile ölçülebilecek beceri içerdiğinden kritik kazanımlar arasına alınmamıştır.

“Paydası en çok iki basamaklı doğal sayı olan en çok üç kesri karşılaştırır ve sıralar” kazanımı, karşılaştırma yapmak için gerekli olan pay ve payda kavramları bilgisini de içerdiğinden kritik kazanımlara alınmıştır.

“Bir çokluğun belirtilen kesrin birimi kadarını belirler” kazanımı kesirlere ait temel bir kavram olan kesir birimi kavramını içerdiğinden kritik bir kazanım olarak görülmüştür.

Üçüncü sınıf matematik öğretim programı sayılar öğrenme alanındaki kazanımlar uzman görüşleri de dikkate alınarak yukarıdaki gibi analiz edilmiş ve 3. sınıf testi için 16 adet kritik kazanım belirlenmiştir.

Üçüncü Sınıf Testi Kritik Kazanımları için Soruların Yazılması ve Uzman Görüşleri ile Düzenlenmesi

“Üç basamaklı doğal sayıları okur ve yazar” kazanımı için okunuşu verilen bir sayının yazılışı, rakamlarla verilen bir sayının okunuşu ve onluk taban blokları ile verilen üç basamaklı bir sayının yazılışının istendiği üç soru yazılmış, sorular için olumlu görüşler alındığından değişiklik yapılmamıştır.

“Üç basamaklı doğal sayıların basamak adlarını, basamaklarındaki rakamların basamak değerlerini belirtir” kazanımı için üç soru yazılmıştır:

Soru 1: 570 sayısındaki 5 rakamı hangi basamaktadır?

Soru 2: 475 sayısındaki 4 rakamının basamak değeri kaçtır?

Soru 3: Ezgi üç basamaklı abc sayısını yazarken yanlışlıkla a rakamını 9 yerine 8 olarak yazmıştır. Buna göre sayı gerçek değerinden ne kadar eksik yazılmıştır?

İkinci sorudaki basamak değerinin birinci sorudaki basamak adından farkına dikkat çekmek amacıyla “basamak değeri” ifadesinin altı çizilmiştir. Üçüncü sorudaki harfli ifadenin 3. sınıfta anlamayı güçleştireceği yönündeki uzman görüşüne göre soru “Ezgi üç basamaklı bir sayıyı yazarken en baştaki rakamı yanlışlıkla 9 yerine 8 olarak yazmıştır” şeklinde düzenlenmiştir.

“1000’den küçük en çok beş doğal sayıyı, büyükten küçüğe veya küçükten büyüğe doğru sembol kullanarak sıralar” kazanımı için birinde dört,

diğerinde beş sayının sıralamasının istendiğı iki soru yazılmış, sorular için olumlu uzman görüşleri alınmıştır.

“Bir örüntüdeki ilişkiyi belirler ve örüntüyü genişletir” kazanımı yazılan sorulardan ikisindesayı örüntüleri verilip örüntüde boş bırakılan yere gelmesi gereken sayı, diğerinde ise bir problem içerisinde tablo ile verilen sayı örüntüsündeki verilmeyen sayının bulunması istenmiştir. Örüntülerdeki sayılar“100 içinde altışar, yedişer, sekizer ve dokuzar ileriye doğru sayar” kazanımını kapsar şekilde seçilmiş, sorular için olumlu görüşler alınmıştır.

“Tek ve çift doğal sayıları belirtir” kazanımı için sayıların basamak sayısı gibi kavramlar karıştırılmadan yalnızca tek ve çift olma durumlarını ölçebilmek için en küçük tek ve çift sayıların sorulduğu iki soru yazılmış, sorular içinolumlu görüşler alınmıştır.

“Toplamları en çok üç basamaklı olan doğal sayılarla eldesiz ve eldeli toplama işlemini yapar” kazanımı için iki soru yazılmış, bir önceki sınıfta toplanan sayısı en fazla üç olan toplama işlemi yapıldığı düşünülerek, sorulardan biri toplanan sayısı dört olacak şekilde düzenlenmiştir.

“Doğal sayılarla toplama işlemini gerektiren problemleri çözer” için biri tek diğeri iki toplama işlemi gerektiren problemler yazılmıştır.

“En çok üç basamaklı doğal sayılarla çıkarma işlemi yapar” kazanımı için üç soru yazılmıştır. Bu sorulardan ilki onluk bozmayı gerektirmeyecek, ikincisi 0'dan bir sayının çıkarılması için onluk bozmayı gerektirecek, üçüncüsü ise birler basamağındaki çıkarma işlemi için onlar basamağındaki bir onluk alındıktan sonra onlar basamağı 0 kalacak şekilde düzenlenmiştir. Sorular için alınan uzman görüşleri olumlu bulunmuştur.

“Doğal sayılarla toplama ve çıkarma işlemlerini gerektiren problemleri çözer” kazanımı için biri bir çıkarma, diğeri bir çıkarma ve bir toplama işlemi gerektiren iki soru yazılmış, sorular için olumlu uzman görüşleri alınmıştır:

“Çarpım tablosunu oluşturur” kazanımı için bir önceki sınıftaki “10'a kadar olan sayıları iki, üç, dört ve beş ile çarpar” kazanımının dışında kalan,

çarpım tablosunda 5'in üzerindeki sayıların çarpımların sorulduğu üç soru yazılmış, sorular için uzman görüşleri olumlu bulunmuştur.

“Eldeli çarpma işlemini yapar” kazanımı için basamaklarındaki rakamların çarpımı elde gerektiren bir, iki veya üç basamaklı sayıların çarpımını içeren iki soru yazılmıştır. Sorularda bu sınıfın sayı sınırlılıkları içerisinde kalınarak çarpımları üç basamaklı olan sayılar kullanılmıştır.

“Doğal sayılarla çarpma işlemini gerektiren problemleri çözer” kazanımı için biri iki basamaklı iki sayının çarpımını, diğerleri öğretim programında örnek gösterilen çarpma ile birlikte toplama işlemini gerektiren ve sonucu üç basamaklı sayı olan iki problem yazılmış, problemler için olumlu uzman görüşleri alınmıştır.

“İki basamaklı doğal sayıları bir basamaklı doğal sayılara böler” kazanımı için belirtilen basamak sayılarına uygun sayılar kullanılarak iki soru yazılmış, sorular için olumlu uzman görüşleri alınmıştır.

“Biri bölme olacak şekilde iki işlem gerektiren problemleri çözer” kazanımı için bölme işleminin yanında toplama veya çıkarma işlemlerinden birini gerektiren iki problem yazılmış, olumlu görüşler alınmıştır.

“Paydası en çok iki basamaklı doğal sayı olan en çok üç kesri karşılaştırır ve sıralar” kazanımı için payları ya da paydaları eşit iki ya da üç kesrin karşılaştırılmasının istendiği aşağıdaki üç soru yazılmıştır:

Soru 1: $\frac{5}{10}$ $\frac{3}{10}$ Yandaki kesirlerin arasına <, > ve = işaretlerinden hangisi gelmelidir?

Soru 2: $\frac{1}{5}$, $\frac{1}{3}$ ve $\frac{1}{6}$ kesirlerini tabloya belirtilen sıraya göre yerleştiriniz.

<u>Büyük</u>	<u>Ortanca</u>	<u>Küçük</u>

Soru 3: Bir pastanın $\frac{1}{12}$ 'i Merve'ye, $\frac{3}{12}$ 'ü Emre'ye ve $\frac{4}{12}$ 'ü Caner'e verildiğinde en çok ve en az pastayı kim yemiştir?

En çok:
En az:

İlk iki soru, payları veya paydaları eşit kesirleri karşılaştırma becerilerini ölçmek için yazılan soruların karşılaştırılacak kesirlerin adedi açısından birbirine benzer olması gerektiği düşünülerek biri paydaları, diğeri payları eşit iki kesrin karşılaştırılmasını içerecek şekilde düzenlenmiştir.

“Bir çokluğun belirtilen kesrin birimi kadarını belirler” kazanımı için yazılan sorularda 3. sınıf bölme işlemi sayı sınırlılıkları düşünülerek iki basamaklı sayılar kullanılmıştır. Ön uygulamada sorulardan birindeki $90 \div \frac{1}{5}$ 'ini bulma işleminin öğrenciler tarafından anlaşılmadığı görüldüğünden soru, sayı ile birlikte çokluk kullanılarak “90 liranın $\frac{1}{5}$ 'i” şeklinde yeniden yazılmıştır.

Uygulama için hazırlanan 3. sınıf testi toplamda 38 adet kısa cevaplı sorudan oluşturulmuştur.

Dördüncü Sınıf Testi Kritik Kazanımlarının Belirlenmesi ve Uzman Görüşleri ile Netleştirilmesi

Kritik kazanım sayılarının belirlenmesinde uzmanların bu sınıf düzeyindeki öğrenciler için bir başarı testinde bulunacak ortalama soru sayısı hakkındaki görüşleri dikkate alınmış, sınıf öğretmenlerince 4. sınıf düzeyinde bir başarı testindeki soru sayısının 20-25 arasında olması uygun görülmüştür.

“4, 5 ve 6 basamaklı doğal sayıları okur ve yazar” kazanımı, öğretim programında ilk kez bu sınıf düzeyinde karşılaşılan 4, 5 ve 6 basamaklı sayıların doğru bir şekilde okunmasını ve okunuşu verilen sayıların doğru şekilde yazılmasını içerdiğinden kritik kazanım olarak seçilmiştir.

“4, 5 ve 6 basamaklı doğal sayıların bölüklerini ve basamaklarını, basamaklarındaki rakamların basamak değerlerini belirtir” basamak değeri kavramının öneminden kritik kazanımlara alınmıştır. Bu kazanım ayrıca “4, 5 ve 6 basamaklı doğal sayıları çözümler” kazanımını kapsamaktadır.

“Doğal sayıları en yakın onluğa veya yüzlüğe yuvarlar” kazanımı alınan uzman görüşü doğrultusunda kritik kazanımlara dahil edilmiştir.

“Bir örüntüyü sayılarla ilişkilendirir ve eksik olan bölümü tamamlar” örüntüdeki kuralı belirleme becerisi içerdiğinden önemli görülmüştür.

“En çok altı basamaklı doğal sayıları sıralar” kazanımı sayıları karşılaştırma becerisinin öneminden dolayı kritik kazanım olarak seçilmiştir.

“En çok dört basamaklı doğal sayılarla toplama işlemini yapar” kazanımı 3. sınıftaki ilgili kazanımdan toplanan sayıların basamak sayısı yönüyle farklı olduğundan 4. sınıfta kritik bir kazanım olarak görülmüştür.

“Toplamı en çok dört basamaklı olan iki doğal sayının toplamını tahmin eder ve tahminini işlem sonucu ile karşılaştırır”, “Toplamları en çok dört basamaklı olacak şekilde en çok dört basamaklı doğal sayıları, 100’un katlarıyla zihinden toplar”, “En çok üç basamaklı iki doğal sayının farkını tahmin eder, tahminini işlem sonucu ile karşılaştırır” ve “Üç basamaklı doğal sayılardan 100’un katı olan doğal sayıları zihinden çıkarır” kazanımları tahmin ve zihinden işlem becerilerinin kısa cevaplı soru ile ölçülemeyeceği düşünülmüş ve kritik kazanım olarak seçilmemiştir.

“Doğal sayılarla toplama işlemini gerektiren problemleri çözer ve kurar” ve “Doğal sayılarla çıkarma işlemini gerektiren problemleri çözer ve kurar” kazanımları problem çözme becerisinden dolayı seçilmiş, ancak problem kurmanın kısa cevaplı soru ile ölçülemeyeceği düşünülmüş ve “Doğal sayılarla toplama/çıkarma işlemini gerektiren problemleri çözer” şeklinde alınmıştır.

“En çok dört basamaklı doğal sayılarla çıkarma işlemini yapar” dört basamaklı sayılarla yapılan işlemlerin öneminden dolayı seçilmiştir.

“Çarpımı en çok beş basamaklı doğal sayı olacak şekilde iki doğal sayıyla çarpma işlemini yapar” kazanımı çarpımın basamak sayısı açısından bu sınıf düzeyinde yoklanması gereken kritik bir kazanım olarak görülmüştür.

“Üç doğal sayı ile yapılan çarpma işleminde sayıların birbirleriyle çarpılma sırasının değişmesinin, sonucu değiştirmediğini gösterir” kısa cevaplı soru ile ölçülemeyeceği düşünülmüş ve kritik kazanımlara alınmamıştır.

“En çok üç basamaklı doğal sayıları 10, 100 ve 1000’in en çok dokuz katı olan doğal sayılarla kısa yoldan çarpar”, “En çok üç basamaklı doğal sayıları 10, 100 ve 1000 ile zihinden çarpar”, “En çok iki basamaklı doğal sayıları 5, 25 ve 50 ile kısa yoldan çarpar” ve “En çok iki basamaklı iki doğal

sayının çarpımını tahmin eder ve tahminini işlem sonucu ile karşılaştırır” kazanımları; kısa yoldan veya zihinden işlem yapma ve tahmin becerileri kısa cevaplı soru ile ölçülemeyeceğinden kritik kazanımlara alınmamışlardır.

“Doğal sayılarla çarpma işlemini gerektiren problemleri çözer ve kurar” kazanımındaki problem çözme becerisi önemli görülmüş, ancak problem kurmanın kısa cevaplı soru ile ölçülemeyeceği düşünüldüğünden “Doğal sayılarla çarpma işlemini gerektiren problemleri çözer” şekliyle alınmıştır.

“Üç basamaklı doğal sayıları en çok iki basamaklı doğal sayılara böler” kazanımı bölünen ve bölenin basamak sayısı yönüyle bir önceki sınıftan farklı olduğundan kritik kazanım olarak seçilmiştir.

“Bölme işleminde bölümün basamak sayısını işlem yapmadan belirler”, “Son üç basamağı sıfır olan en çok beş basamaklı doğal sayıları 10, 100 ve 1000’e kısa yoldan böler” ve “Bir bölme işleminin sonucunu tahmin eder ve tahminini işlem sonucu ile karşılaştırır” kazanımlarının içerdiği kısa yoldan işlem yapma ve tahmin becerilerinin kısa cevaplı soru ile ölçülemeyeceği düşünüldüğünden kritik kazanımlara dahil edilmemiştir.

“İki adımlı işlemleri yapar” kazanımı işlem önceliği için temel oluşturacağı düşünüldükten kritik bir kazanım olarak görülmüştür.

“Doğal sayılarla bölme işlemini gerektiren problemleri çözer ve kurar” kazanımı problem çözme becerisi içerdiğinden kritik kazanım olarak seçilmiş, problem kurma kısa cevaplı soru ile ölçülemeyeceğinden “Doğal sayılarla bölme işlemini gerektiren problemleri çözer” şekliyle alınmıştır.

“Payı ve paydası en çok iki basamaklı doğal sayı olan kesirleri, kesrin birimlerinden elde ederek isimlendirir” kazanımı basit, bileşik ve tam sayılı kesirleri ayırt etme becerileri içerdiğinden kritik kazanım olarak alınmıştır.

“Payı ve paydası en çok iki basamaklı olan kesirleri sayı doğrusunda gösterir” kesirlerin sayı doğrusundaki gerçek yerini gösterme becerisinin kısa cevaplı soru ile ölçülemeyeceği düşünüldükten kritik kazanımlara alınmamıştır.

“Eşit paydalı en çok dört kesri, büyükten küçüğe veya küçükten büyüğe doğru sıralar” kazanımı “Kesirleri karşılaştırır” kazanımını kapsadığından kritik kazanım olarak alınmıştır.

“Payları eşit, paydaları birbirinden farklı en çok dört kesri, büyükten küçüğe veya küçükten büyüğe doğru sıralar” kazanımı paydaları farklı olan kesirlerde açıkça anlaşılmayan kesir birimlerini de belirlemeyi gerektirerek karşılaştırma yapma becerisi gerektirdiğinden kritik kazanımlara alınmıştır.

“Bir çokluğun belirtilen bir basit kesir kadarını belirler” kazanımı bir önceki sınıftaki kesrin birimi kadarını belirleme becerisinin daha ileri bir durumu olarak görüldüğünden kritik bir kazanım olarak alınmıştır.

“Paydaları eşit kesirlerle toplama işlemi yapar” ve kesirlerle çıkarma işlemi alt öğrenme alanındaki “Paydaları eşit kesirlerle çıkarma işlemi yapar” kazanımları ilk kez bu sınıf düzeyinde karşılaşılan kesirlerle toplama ve çıkarma işlemlerini içerdiğinden kritik kazanımlar olarak seçilmiştir.

“Kesirlerle toplama ve çıkarma işlemlerini gerektiren problemleri çözer ve kurar” kazanımındaki kesirlerle verilmiş çokluklar içeren problemleri çözme becerisi önemli görülmüş, ancak problem kurmanın kısa cevaplı soru ile ölçülemeyeceği düşünülerek “Kesirlerle toplama ve çıkarma işlemlerini gerektiren problemleri çözer” şekliyle alınmıştır.

“Bir bütün 10 ve 100 eş parçaya bölündüğünde, ortaya çıkan kesrin birimlerinin ondalık kesir olduğunu belirtir” kazanımı kesir birimi ile ilgili kavramsal bilgi kısa cevaplı soru ile ölçülemeyeceğinden seçilmemiştir.

“Ondalık kesirleri virgöl kullanarak yazar” kazanımı verilen kesirleri virgöl kullanarak yazmayı içerdiğinden kritik kazanım olarak seçilmiştir.

“Ondalık kesirlerin tam kısmını, kesir kısmını ve basamak adlarını belirtir” kazanımı ondalık kesirleri tam ve kesir kısımlarına ayırma ve basamak kavramı içerdiğinden önemli görülerek kritik kazanımlara alınmıştır.

“İki ondalık kesri karşılaştırarak aralarındaki ilişkiyi büyük, küçük veya eşit sembolüyle gösterir” ondalık kesirleri tam ve ondalık kısımlarına dikkat ederek karşılaştırma gerektirdiğinden kritik kazanımlara alınmıştır.

Dördüncü sınıf matematik öğretim programı sayılar öğrenme alanındaki kazanımlar uzman görüşleri de dikkate alınarak yukarıdaki gibi irdelenmiş ve 1. sınıf testi için 24 adet kritik kazanım belirlenmiştir.

Dördüncü Sınıf Testi Kritik Kazanımları için Soruların Yazılması ve Uzman Görüşleri ile Düzenlenmesi

“4, 5 ve 6 basamaklı doğal sayıları okur ve yazar” kazanımı okunuşu verilen bir sayının yazılışının, rakamlarla verilen bir sayının okunuşunun ve onluk taban blokları ile verilen dört basamaklı bir sayının yazılışının istendiği üç soru yazılmış, sorular için olumlu uzman görüşleri alınmıştır.

“4, 5 ve 6 basamaklı doğal sayıların bölüklerini ve basamaklarını, basamaklarındaki rakamların basamak değerlerini belirtir” kazanımı için 4. sınıf düzeyinde 4, 5 ve 6 basamaklı sayılarla ilk kez karşılaşıldığından, basamak değeri ve basamak adı sorulacak rakamların binler, on binler ve yüzbinler basamaklarında olduğu aşağıdaki üç soru yazılmıştır. Sayılardaki rakamların basamak adı ve değerinin sorulduğu sorular olumlu bulunmuş, “Pinar beş basamaklı abcdef sayısını yazarken yanlışlıkla a rakamını 7 yerine 6 olarak yazmıştır. Pinar sayıyı gerçek değerinden ne kadar eksik yazmış olur?” sorusundaki harfli ifadenin 4. sınıfta anlamayı güçleştireceği görüşü ile soru “Pinar beş basamaklı bir sayıyı yazarken en baştaki rakamı yanlışlıkla 7 yerine 6 olarak yazmıştır” şeklinde düzenlenmiştir.

“Doğal sayıları en yakın onluğa veya yüzlüğe yuvarlar” kazanımı için verilen dört basamaklı sayıların en yakın onluğa ya da en yakın yüzlüğe yuvarlanmasının istendiği iki soru yazılmış, olumlu görüşler alınmıştır.

“Bir örüntüyü sayılarla ilişkilendirir ve eksik olan bölümü tamamlar” kazanımı için yazılan ilk iki soruda şekillerle verilmiş örüntülerin bir sonraki adımındaki şekil sayısının bulunması istenmiş, örüntülerin içerdiği sayısal ilişkilerin sınıf düzeyine uygun olmasına dikkat edilmiştir. Programda sayılarla oluşturulmuş örüntülere de yer verildiğinden üçüncü soru 7’şer artan sayı örüntüsü içerecek şekilde düzenlenmiştir.

“En çok altı basamaklı doğal sayıları sıralar” kazanımı için altı basamaklı üç sayının sıralaması veprogramdaki gibi verilen rakamları kullanarak dört basamaklı en büyük sayının yazılması istenmiştir.

“En çok dört basamaklı doğal sayılarla toplama işlemini yapar” kazanımı için 4. sınıf düzeyinde 3. sınıftakinden farklı olarak sonuçları dört basamaklı olacak şekilde toplama işlemi içeren üç soru yazılmış, sorular için olumlu görüşler alındığından herhangi bir değişiklik yapılmamıştır.

“Doğal sayılarla toplama işlemini gerektiren problemleri çözer” kazanımı için biri üç ve dört basamaklı iki sayının toplanmasını, diğeri iki toplama işlemi gerektiren iki soru yazılmıştır. Ön uygulamada öğrencilerin sorulardan birinde istenen toplam öğrenci sayısı yerine işlemlerin yalnızca ilk adımını yaptıkları gözlemlendiğinden “öğrenci sayısı” ifadesinin altı çizilmiştir.

“En çok dört basamaklı doğal sayılarla çıkarma işlemini yapar” kazanımı için öğretim programında belirtildiği gibi ara basamaklarında 0 olan sayılar da içerecek şekilde iki soru yazılmış, olumlu görüşler alınmıştır.

“Doğal sayılarla çıkarma işlemini gerektiren problemleri çözer” kazanımı için dört basamaklı sayılarla çıkarma işlemi içeren iki problem yazılmış, olumlu uzman görüşleri alındığından değişiklik yapılmamıştır.

“Çarpımı en çok beş basamaklı doğal sayı olacak şekilde iki doğal sayıyla çarpma işlemini yapar” kazanımı için çarpımı beş basamaklı olan iki ya da üç doğal sayı ile çarpma işlemleri yazılmış, olumlu görüşler alınmıştır.

“Doğal sayılarla çarpma işlemini gerektiren problemleri çözer” kazanımı için biri iki doğal sayının çarpımını, diğeri çarpma ile birlikte toplama işlemi gerektiren iki problem yazılmış, olumlu uzman görüşleri alınmıştır.

“Üç basamaklı doğal sayıları en çok iki basamaklı doğal sayılara böler” kazanımı için üç basamaklı doğal sayıların bir ya da iki basamaklı sayılara bölünmesini içeren iki soru yazılmış, sorular için olumlu görüşler alınmıştır.

“İki adımlı işlemleri yapar” kazanımı için toplama, çıkarma, çarpma ve bölme işlemlerinden ikisi ile oluşturulmuş iki adımlı işlemler içeren iki soru

yazılmıştır. Sorularda, öğretim programındaki uyarı dikkate alınarak işlem öncelikleri parantez kullanılarak belirtilmiştir.

“Doğal sayılarla bölme işlemini gerektiren problemleri çözer” kazanımı için bu sınıf düzeyindeki bölme işleminin sayı sınırlılıkları içerisinde üç basamaklı sayıların bir veya iki basamaklı sayılara bölünmesini gerektiren üç problem yazılmış, bunlardan biri hız problemlerinin bu sınıf düzeyine uygun olmadığı yönündeki uzman görüşü dikkate alınarak testten çıkarılmıştır.

“Payı ve paydası en çok iki basamaklı doğal sayı olan kesirleri, kesrin birimlerinden elde ederek isimlendirir” kazanımı için şekil ile verilen basit, bileşik ve tam sayılı kesirlerin isimlendirilmesinin istendiği üç soru yazılmıştır. Bileşik ve tam sayılı kesirlerin sorulduğu sorularda birbirinin yerine yazılabileceği düşünülerek istenen kesir adı açıkça yazılmıştır. Sorular için alınan uzman görüşleri olumlu bulunmuştur.

“Eşit paydalı en çok dört kesri, büyükten küçüğe veya küçükten büyüğe doğru sıralar” kazanımı için yazılan ilk iki soruda paydaları eşit kesirlerin, üçüncüsünde ise somut nesnelerle ilişkilendirilmiş üç kesrin karşılaştırılması istenmiştir. Sorularda, öğretim programında kesirler için belirtilen sayı sınırlılıkları içinde kalınarak pay ve paydadaki sayılar en çok iki basamaklı şekilde seçilmiştir. Sorular için olumlu görüşler alınmıştır.

“Payları eşit, paydaları birbirinden farklı en çok dört kesri, büyükten küçüğe veya küçükten büyüğe doğru sıralar” kazanımı için yazılan iki soruda payları eşit kesirlerin, üçüncüsünde ise somut nesnelerle ilişkilendirilmiş üç kesrin karşılaştırılması istenmiştir. Bu sınıfın sayı sınırlılıkları içinde kalınarak pay ve paydadaki sayılar en çok iki basamaklı seçilmiştir.

“Bir çokluğun belirtilen bir basit kesir kadarını belirler” kazanımı için verilen çokluğun bir basamaklı sayılarla oluşturulmuş basit kesir kadarının bulunmasının istendiği iki soru yazılmış, olumlu uzman görüşleri alınmıştır.

“Paydaları eşit kesirlerle toplama işlemi yapar” kazanımı için yazılan iki soruda, öğretim programında belirtildiği gibi basit ve bileşik kesirlerin toplanmasında pay ve paydanın en fazla iki basamaklı olmasına, tam sayılı kesirlerin yalnızca tam sayılı kesirlerle toplanmasına ve tam sayılı kesirlerin bir basamaklı sayılarla oluşturulmasına dikkat edilmiştir.

“Paydaları eşit kesirlerle çıkarma işlemi yapar” kazanımı için de kesirlerle toplama işlemi ile ilgili kazanımda dikkat edilen noktalara uyularak iki çıkarma işlemi sorusu yazılmış, sorular için olumlu görüşler alınmıştır.

“Kesirlerle toplama ve çıkarma işlemlerini gerektiren problemleri çözer” kazanımı için yazılan iki sorudan birinde iki kesir ile toplama, ikincisinde kesirlerle hem toplama hem çıkarma işlemi gerektiren problemler sorulmuş, sorular için alınan uzman görüşleri olumlu bulunarak değişiklik yapılmamıştır.

“Ondalık kesirleri virgöl kullanarak yazar” kazanımı için yazılan sorularda paydası 10 ve 100 olarak verilen ya da şekil ile gösterilen kesirlerin virgöl kullanılarak yazılması istenmiştir. Kesirler, programdaki gibi kesir kısmı en fazla iki basamaklı olacak şekilde yazılmış; olumlu görüşler alınmıştır.

“Ondalık kesirlerin tam kısmını, kesir kısmını ve basamak adlarını belirtir” kazanımı için verilen ondalık kesirlerin tam veya kesir kısmındaki sayıların basamak adlarının sorulduğu üç soru yazılmıştır. Sorularda, bu sınıftaki ondalık kesirlerin sayı sınırlılıkları içerisinde kalınarak kesir kısmı en fazla iki basamaklı olan sayılar kullanılmış, olumlu görüşler alınmıştır.

“İki ondalık kesri karşılaştırarak aralarındaki ilişkiyi büyük, küçük veya eşit sembolüyle gösterir” kazanımı için tam kısımları aynı, ondalık kısımları farklı iki ondalık kesir arasına $<$, $>$, $=$ işaretlerinden uygun olanın yazılmasının istendiği iki soru yazılmış, sorular için olumlu görüşler alındığından değişiklik yapılmamıştır.

Uygulama için hazırlanan 4. sınıf testi toplamda 58 adet kısa cevaplı sorudan oluşturulmuştur.

Verilerin Toplanması

Araştırmanın verileri, İstanbul İl Milli Eğitim Müdürlüğü’nün araştırma uygulama izni ile belirlenen okullarda testler uygulanarak toplanmıştır. Hazırlanan testlerin deneme formları her bir öğrenciye verilerek öğrencilerden testlerdeki sorularını sınıf ortamında belirlenen süre içerisinde cevaplamaları istenmiştir. Öğrencilerin birbirlerinden etkilenmemeleri ve yardımlaşmamaları sağlanarak tüm soruları cevaplamaları istenmiştir. Turgut ve Baykul (2012) bir sınavda cevaplayıcıların soruları dikkatle ve hızlı bir şekilde cevaplamaya motive edilmelerinin ve aynı zamanda sınav süresinin

cevaplayıcıların hemen hepsinin tüm soruları cevaplamaları için yeterli uzunlukta olmasının ölçmede güvenilirliği sağlamak için gerekli önlemler olduğunu belirtmektedirler. Verilerin toplanmasında bu hususlara uygun olarak öğrenciler dikkatle ve olabildiğince hızlı bir şekilde cevaplamaları konusunda uyarılmış; ayrıca ölçme araçları daha önceden sınıf öğretmenlerinden alınmış olan görüşlere ve ön uygulamalardaki deneyimlere uygun sürelerde uygulanmıştır. Buna göre yalnızca 4. sınıf testi 10 dakika ara verilerek iki ders saatinde, diğer testler birer ders saatinde uygulanmıştır.

Araştırmada uygulamalar 2012 yılının kasım ayında gerçekleştirilmiştir.

Verilerin Analizi

Veri toplama araçlarının deneme uygulamalarıyla elde edilen verilerin değerlendirilmesiyle her bir sınıf düzeyindeki test için katılımcıların testlerdeki maddelere verdikleri doğru cevaplar 1, yanlış ve boş bırakılan cevaplar 0 şeklinde puanlanmıştır. Testlerin geçerlik ve güvenirlik analizleri bu uygulamalar ile elde edilen veriler üzerinden yapılmıştır. Analizlerde SPSS ve ITEMAN programları kullanılmıştır.

Testlerin geçerliği için kapsam ve yapı geçerliği ile ölçüt dayanaklı geçerlik analizleri yapılmıştır.

Testlerin kapsam geçerliğinin sağlanması için hem testlerle ölçülecek kritik kazanımların belirlenmesi, hem de test maddelerinin yazılması aşamalarında uzman görüşleri alınmıştır. Testlerin kapsam geçerliği analizinde her bir test ile ölçülen kazanımları ve bu kazanımların Bloom taksonomisindeki bilişsel düzeylerini gösteren belirtke tabloları hazırlanmıştır.

Testlerde yapı geçerliğini incelemek için faktör analizi, iç tutarlılık analizi, başka test sonuçları ile korelasyon ve hipotez testi teknikleri uygulanmaktadır (Anastasi, 1988; Büyüköztürk, 2007). Yapı geçerliği hakkında bilgi sağlamada başvuru hipotez testi uygulamasında aralarında fark olması beklenen grupların test puanları arasındaki farkın anlamlılığı yapı geçerliğinin olduğu şeklinde yorumlanmaktadır (Büyüköztürk, 2007; Şencan, 2005). Bu çalışmada her test geliştirildiği sınıf düzeyinin bir üst sınıfında

uygulanmasının yanında orta sosyoekonomik düzeydeki bir okuldageliştirildiği sınıf düzeyinde bir sınıfta uygulanmış ve bu grupların testlerden aldıkları puanlar arasındaki farkın anlamlılığına bakılmıştır. Bunun için kullanılacak analiz yöntemini belirlemek için öncelikle verilerin bağımsız örneklemeler t-test varsayımlarını karşılama durumları irdelenmiştir. Bağımsız örneklemeler t-testin uygulanması için karşılaştırılacak gruplardaki puan dağılımlarının normal olması gerekmektedir (Büyüköztürk, Çokluk ve Köklü, 2011). Araştırmada testlerin geliştirildiği sınıf düzeyinde ve bir üst sınıftaki dağılımların normal olduğu durumlarda bağımsız örneklemeler t-test analizi, normal dağılımın olmadığı durumlarda ise parametrik olmayan Mann Whitney-U testi uygulanmıştır. Sonuçların anlamlılığına göre testlerin yapı geçerliği değerlendirilmiştir.

Araştırmada, ayrıca, farklı sosyoekonomik düzeylerdeki ilçelerden alınan öğrencilerin test puan ortalamaları arasındaki farkların anlamlılığına bakılmış, sonuçlar yapı geçerliği açısından değerlendirilmiştir. Yapılan literatür taramasında sosyoekonomik düzey ile hem genel akademik başarı hem de matematik başarısı arasında orta düzeyde veya güçlü bir pozitif ilişki olduğunu ortaya çıkaran çalışmalara (Dursun, 2004; Şirin, 2005; Perry ve McConney, 2010) rastlandığından bu ilçelere göre katılımcıların puan ortalamaları arasında anlamlı düzeyde fark olması beklenmiştir. Üst, orta ve alt sosyoekonomik düzeylerdeki ilçelerin sırasıyla Kadıköy, Tuzla ve Sultanbeyli olduğu göz önünde bulundurularak puan ortalamalarının bu sıralamaya uygun şekilde üst sosyoekonomik düzeydeki ilçeler lehine farklı olması beklenmiştir. Testlerin uygulandığı 3 ilçe arasındaki puan ortalamalarının farklılığını test etmek için verilerin tek yönlü varyans analizi (ANOVA) varsayımlarını karşılama durumlarına göre ANOVA veya parametrik olmayan testlerden Kruskal Wallis H testi kullanılmıştır.

Bir ölçme aracının ölçüt dayanaklı geçerliği, önceden yeterli düzeyde geçerli ve güvenilir olduğu ortaya konmuş olan başka bir ölçme aracı ile korelasyonuna göre değerlendirilebilir (Turgut ve Baykul, 2012). Özçelik (2010) geliştirilen bir testin sonuçları ile aynı özelliği ölçtüğü bilinen başka bir ölçmenin sonuçlarının uyumluluğunu testin uyum geçerliğine sahip olduğu şeklinde ifade etmektedir. Araştırmada testlerin uygulandığı sınıfların

öğretmenlerinden öğrencileri matematik dersindeki algılayış ve başarıları açısından 1 ile 5 arasında bir puan vererek değerlendirmeleri istenmiştir. Testlerin ölçüt dayanaklı geçerlik analizi için bu puanlarla öğrencilerin testlerden aldıkları toplam puanlar arasındaki korelasyonlar hesaplanmıştır. Başer (1996) tarafından yapılan başarı testi geliştirme çalışmasında da geliştirilen testin ölçüt dayanaklı geçerliği testin uygulandığı öğrenci grubunun matematik ders notları ile testten aldıkları puanlar arasındaki korelasyon hesaplanarak değerlendirilmiştir.

Ölçme araçlarının uygulama ve puanlamalarıyla ilgili özelliklerine göre farklı güvenirlik analizi yöntemleri gerekmede ve bu analizlerin sonuçlarına göre testlere ait tutarlılık, kararlılık gibi farklı yorumlar yapılabilmektedir (Turgut ve Baykul, 2012). Bu araştırmada testler için hazırlanan maddelerin denenmesi tek bir uygulama ile yapılmıştır. Turgut ve Baykul'a (2012) göre bir ölçme aracının güvenirliğini tek uygulama ile belirlemenin bir yöntemi olarak ölçme aracındaki maddelerin birbiri ile tutarlılığını ortaya çıkaran Kuder-Richardson-20 (KR-20) testi uygulanabilir. Başarı testlerinin geliştirildiği pek çok çalışmada güvenirliğin KR-20 güvenirlik katsayısı hesaplanarak rapor edildiği görülmektedir. Testlerin uygulanması sonucu elde edilen veriler yardımıyla güvenirlik analizi için testlerdeki maddelerin birbiri ile tutarlılığı hakkında kestirim yapmayı sağlayan KR-20 güvenirlik formülü kullanılmıştır. KR-20 güvenirlik katsayıları hem formülle, hem de ITEMAN programında hesaplanmıştır.

Testlerin asıl formları için seçilecek nitelikli maddelerin belirlenmesi amacıyla maddelere ait güçlük ve ayırtıcılık gücü indeksleri ITEMAN programında hesaplanmıştır.

Bir testteki maddelerin güçlük indeksi her bir madde için doğru cevap veren öğrenci sayısının tüm öğrenci sayısına oranlanmasıyla elde edilmektedir (Özçelik, 2010; Turgut ve Baykul, 2012). Dolayısıyla 0 ile 1 aralığında değerler alacak olan güçlük indeksinin 0 değerine yakın bulunması maddenin zor olduğu, 1 değerine yakın bulunması ise kolay bir madde olduğu şeklinde yorumlanmıştır. Araştırmada testlerin uygulanması sonucu

elde edilen veriler bir madde istatistik programında analiz edilerek maddelere ait güçlük indeksleri ortaya çıkarılmıştır.

Bir test maddesinin ayıricılık gücü ölçülen davranışa sahip olan ve olmayan cevaplayıcıları ayırma gücü olarak tanımlanmaktadır (Özçelik, 2010; Turgut ve Baykul, 2012). Araştırmada testlerin uygulanması sonucu elde edilen veriler ITEMAN programında analiz edilerek maddelere ait ayıricılık indeksleri çift serili korelasyon katsayısı ile hesaplanmıştır. Böylece her maddenin testten alınan toplam puan ile korelasyonu maddenin yoklanan kazanımı edinen ve edinmeyen öğrencileri birbirinden ayırabilme gücünü ortaya çıkarmıştır.

BÖLÜM 4

BULGULAR VE YORUMLAR

Bu bölümde testlerin deneme formlarının uygulanması sonucu elde edilen veriler üzerinde araştırma sorularına yönelik yapılan analizler sonucu ortaya çıkan bulgular ve bu bulgulara göre yapılan yorumlar her test için ayrı ayrı ele alınmıştır:

4. 1. Birinci Sınıf Başarı Testinin Psikometrik Özelliklerine İlişkin Bulgular ve Yorumlar

Araştırmanın amacına uygun olarak geliştirilen birinci sınıf başarı testinin kapsam geçerliği, yapı geçerliği, ölçüt dayanaklı geçerlik, güvenirlik, madde güçlük indeksi ve madde ayıricılık indekslerine ait analiz sonuçları elde edilmiştir.

Testin kapsam geçerliği, test ile ölçülecek kazanımların belirlenmesinde 1. sınıf matematik dersi öğretim programı (Milli Eğitim Bakanlığı [MEB], 2009) sayılar öğrenme alanındaki kazanımların tamamı incelenerek kritik olarak görülenlerin seçilmesi ile sağlanmıştır. Bu kazanımların seçilmesinde araştırmacı tarafından belirlenen kritik kazanımlar için matematik eğitimi alanından bir akademisyen ve 1-4. sınıflara matematik öğreten dört sınıf öğretmeninden görüş ve öneriler alınarak testin kapsam geçerliği artırılmıştır. Ayrıca teste dahil edilen kazanımları, bu bunların Bloom taksonomisine göre bilişsel düzeylerini, bu kazanımı ölçmek üzere seçilmiş maddenin güçlük ve ayıricılık değerlerini gösteren belirtke tablosu hazırlanmıştır. Belirtke tablosu EK-1’de görülmektedir. Hazırlanan belirtke tablosuna göre testte ölçülmesi hedeflenen her bir kritik kazanım için bir madde bulunması 1. sınıf sayılar öğrenme alanındaki kritik kazanımların tamamının testte kapsandığını göstermektedir.

Testin yapı geçerliğine gösterge olarak testin geliştirildiği sınıf düzeyinde ve bir üst sınıftaki uygulaması sonucu elde edilen puan ortalamaları farklılığını test etmek için öncelikle sınıf düzeylerine göre puan dağılımlarının normalliği incelenmiştir. Bunun için yapılan Kolmogorov-Smirnov ve Shapiro-Wilk test sonuçları Çizelge 4'te görülmektedir:

Çizelge 4

1. Sınıf Testinden Alınan Puanların Sınıf Düzeylerine Göre Dağılımı İçin Yapılan Normallik Testi Sonuçları

Sınıflar	Kolmogorov-Smirnov			Shapiro-Wilk		
	İstatistik	Sd	p	İstatistik	Sd	p
1. Sınıf	0,145	26	.169	0,970	26	.630
2. Sınıf	0,142	56	.006	0,936	56	.005

Çizelge 4 incelendiğinde normallik testlerine göre puan dağılımlarının 1. sınıf düzeyinde normal olduğu ($p > .05$), 2. sınıf düzeyinde normal olmadığı ($p < .05$) görülmektedir. Ancak, 2. sınıf puan testi puan dağılımının çarpıklık ve basıklık katsayıları incelenmiş; çarpıklık ve basıklık katsayıları sırasıyla $-0,461$ ve $-0,902$ bulunmuştur. Çarpıklık katsayılarının ± 1 aralığında kalmasının dağılımların normalden önemli bir sapma göstermediği (Büyüköztürk, vd., 2011) göz önüne alınarak 2. sınıf testi puan dağılımı normal kabul edilmiştir. Birinci ve 2. sınıftaki puan ortalamaları arasındaki farkı ortaya çıkarmak için bağımsız örneklem t-test uygulanmıştır. Sonuçlar Çizelge 5'te gösterilmektedir:

Çizelge 5

1. Sınıf Testi Puanlarının Sınıf Düzeylerine Göre Farklılıkları İçin Uygulanan t-Test Sonuçları

Sınıflar	N	X	SS	Sd	T	p
1.Sınıf	26	9,38	4,52	80	-11,59	.000
2.Sınıf	56	20,30	3,69			

Çizelge 5'e göre 1. sınıf testinin uygulandığı 2. sınıftaki öğrencilerin puan ortalaması ($Ort_2=20,30$, $SS=4,52$) 1. sınıftakilerden ($Ort_1=9,38$, $SS=3,69$) istatistiksel olarak anlamlı düzeyde farklıdır ($t(80) = -11,59$, $p=.000$). Bu bulgu, 1. sınıf başarı testinin içerdiği konuları bir önceki yıl görmüş olan ve henüz görmemiş olan öğrencileri ayırt edebildiğini göstermektedir.

Testin yapı geçerliği hakkında bilgi sağlaması için ayrıca testten alınan puanlarda farklı sosyoekonomik düzeylerdeki ilçeler arasında fark olup olmadığına bakılmıştır. Bunun için öncelikle tek yönlü varyans analizi varsayımlarından bağımlı değişken ölçeğinin en az eşit aralık düzeyinde olması, bağımsız değişkenin kategorik olması, dağılımların normalliği ve varyansların homojenliği varsayımlarını karşılama durumları incelenmiştir (Büyüköztürk, vd., 2011). Kullanılan test eşit aralık düzeyindedir. Bağımsız değişken olan ilçeler değişkeni kategorik olduğundan bağımsız değişkenin kategorik olması varsayımı sağlanmaktadır. İlçelere göre elde edilen puan dağılımlarının normalliğini belirlemek için yapılan Kolmogorov-Simirnov ve Shapiro-Wilk test sonuçları Çizelge 6'da görülmektedir:

Çizelge 6

1. sınıf Testinden Alınan Puanların İlçelere Göre Dağılımı İçin Yapılan Normallik Testi Sonuçları

İlçeler	Kolmogorov-Simirnov			Shapiro-Wilk		
	İstatistik	Sd	P	İstatistik	Sd	p
Sultanbeyli	0,176	96	.000	0,918	96	.000
Tuzla	0,154	116	.000	0,939	116	.000
Kadıköy	0,179	102	.000	0,871	102	.000

Çizelge6 incelendiğinde ilçelere göre 1. sınıf testi puan dağılımlarının normal olmadığı görülmektedir ($p < .05$). Testten alınan puanların varyansların homojenliği varsayımını karşılama durumunu belirlemek için uygulanan Levene testi sonuçları ise 1. sınıf testi puanlarında ilçelere göre varyansların homojen olmadığını göstermiştir ($L = 17,914$; $p < .05$). Bu bulgulara göre 1. sınıf test puanları tek yönlü varyans analizi (ANOVA)

varsayımlarını karşılamadığından test puan ortalamalarının ilçelere göre anlamlı farklılık gösterip göstermediğinin belirlenmesi için parametrik olmayan testlerden Kruskal-Wallis H testi kullanılmıştır. Bulgular Çizelge 7’de gösterilmektedir:

Çizelge 7

1. Sınıf Testi Puanlarının İlçelere Göre Kruskal Wallis H-Testi Sonuçları

İlçe	N	Sıra Ort.	Sd	χ^2	p
Sultanbeyli	96	113,72	2	54,331	.000*
Tuzla	116	149,94			
Kadıköy	102	207,30			

* p < .05

Çizelge 7’de görüldüğü gibi 1. sınıf testi puanları ilçelere göre anlamlı bir farklılık göstermektedir ($\chi^2 = 54,331$, $SD=2$, $p= .000$). Bu farklılığın hangi ilçeler arasında olduğunun belirlenmesi için ilçelerin ikili kombinasyonları arasında Mann Whitney-U testleri uygulanmıştır. Sonuçlar Çizelge 8’de gösterilmektedir:

Çizelge 8

İlçelerin 1. Sınıf Test Puan Ortalama Farklılıkları İçin Uygulanan Mann Whitney-U Testi Sonuçları

İlçeler	N	Sıralar ortalaması	Sıralar toplamı	Mann Whitney-U	Z	p
Sultanbeyli	96	91,65	8798,00	4142,000	-3,224	.001*
Tuzla	116	118,79	13780,00			
Sultanbeyli	96	70,57	6775,00	2119,000	-6,923	.000*
Kadıköy	102	126,73	12926,00			
Tuzla	116	89,65	10399,00	3613,000	-4,985	.000*
Kadıköy	102	132,08	13472,00			

* p < .05

Çizelge8’e göre tüm ilçelerin 1. sınıf testi puan ortalamaları arasındaki fark anlamlı bulunmuştur. Araştırmada Sultanbeyli, Tuzla ve Kadıköy ilçeleri sırasıyla alt, orta ve üst sosyoekonomik düzeylerdeki ilçeler olarak alınmıştır.

Puan ortalamalarının Sultanbeyli ve Tuzla arasında Tuzla lehine, Sultanbeyli ve Kadıköy arasında Kadıköy lehine, Tuzla ve Kadıköy arasında Kadıköy lehine anlamlı düzeyde farklı olduğu görülmektedir. Bu bulgular 1. sınıf testinin farklı sosyoekonomik düzeylerdeki ilçeler arasında beklendik yöndeki farklılığı ortaya çıkarabildiğinden yapı geçerliğinin olduğunu göstermektedir.

Testin ölçüt dayanaklı geçerliğinin ortaya konması için testin uygulandığı sınıfların öğretmenlerinden öğrencileri matematik dersindeki algılayış ve başarıları açısından 1 ile 5 arasında bir puan vererek değerlendirmeleri istenmiş; bu puanlarla öğrencilerin testten aldıkları toplam puanlar arasındaki korelasyon hesaplanmıştır. Bu korelasyon değeri 1. sınıf testi için 0,79 bulunmuştur. Bu bulgu, test ile aynı özelliği ölçmeye yönelik bir başka ölçme sonucu olan öğretmen görüşü ile ilişkisinin yüksek olduğunu göstermektedir.

Testin güvenirlik analizi için hesaplanan KR-20 iç tutarlılık güvenirlik katsayısı 0,80 bulunmuştur. Bu katsayı 1. sınıf testini oluşturan maddelerin birbirleri ile ve testin tamamı ile tutarlılığının, dolayısıyla testin iç tutarlılık anlamında güvenirliğinin yüksek olduğunu göstermektedir.

Birinci sınıf testinin bir üst sınıf olan 2. sınıfta ve kendi sınıf düzeyinde bir sınıfta uygulanması sonucu elde edilen verilerle hesaplanan madde güçlük ve ayıricılık indeksleri sınıf düzeyleri arasındaki farkın anlaşılır olması için Çizelge 9'da birlikte verilmiştir. Testte kapsanan tüm kazanımları görmüş olan bir üst sınıftaki öğrencilerden elde edilen veriler testin geçerlik, güvenirlik ve madde seçimi analizlerinde esas alınmıştır. Bu verilere göre ortaya çıkan madde güçlük indekslerinin 0,40 ile 0,98 arasında değerler aldığı görülmektedir. Bu testteki 26 maddeden, 13'ü testin asıl formuna konulmak üzere seçilmiştir. Seçilen maddelerin güçlük indeksleri 0,40 – 0,96 aralığındadır. Testte yer alan 2. maddenin ayıricılık indeksinin oldukça düşük ($r_{j2} = 0,03$) bulunmuş, bunun dışındaki maddelerin ayıricılık indekslerinin 0,35 ile 0,91 arasında değerler aldığı görülmektedir. Testin asıl formu için seçilen maddelerin ayıricılık gücü indeksleri de 0,35 ile 0,91 aralığındadır.

Çizelge 9

Uygulanan Sınıf Düzeylerine Göre 1. Sınıf Testi Deneme Formuna Ait Madde Analizi Sonuçları

Madde No	Güçlük İndeksi		Ayırıcılık İndeksi	
	2.Sınıf	1.Sınıf	2.Sınıf	1.Sınıf
1*	0,95	0,92	0,42	0,94
2	0,98	0,81	0,03	0,88
3*	0,95	0,73	0,54	0,86
4*	0,96	0,77	0,40	0,92
5	0,98	0,96	0,46	0,87
6*	0,92	0,77	0,45	0,92
7	0,84	0,58	0,46	0,06
8*	0,59	0,38	0,35	0,62
9	0,83	0,38	0,56	0,92
10*	0,86	0,27	0,63	0,57
11	0,83	0,15	0,71	0,63
12*	0,73	0,08	0,69	0,78
13	0,74	0,15	0,87	0,74
14*	0,78	0,08	0,91	0,72
15	0,85	0,19	0,67	0,75
16*	0,84	0,38	0,89	0,94
17	0,69	0,23	0,50	0,66
18*	0,80	0,12	0,72	1,00
19	0,83	0,08	0,63	1,00
20	0,67	0,00	0,57	0,00
21*	0,40	0,08	0,52	0,25
22	0,84	0,12	0,69	0,83
23*	0,76	0,15	0,78	0,63
24	0,63	0,65	0,71	0,60
25	0,78	0,38	0,47	0,40
26*	0,55	0,00	0,61	0,00

* Testin asıl formuna konulmak üzere seçilen maddeler

Çizelge 9 incelendiğinde maddelerin büyük çoğunluğu için sınıf düzeylerine göre güçlük indeksleri arasında oldukça büyük fark olduğu; testin kendi sınıf düzeyindeki güçlük indekslerinin üst sınıftakilerden düşük değerler aldığı görülmektedir. Testin ilk birkaç maddesi için güçlük düzeylerinin birbirine yakın bulunmasının, 1. sınıfta da bu maddelerin uygulama tarihine kadar kapsanmış konulardan olmalarından veya bu maddelerin gerektirdiği bilgi ve becerilerin okul öncesinde kazanılmış olmasından kaynaklandığı düşünülmüştür.

Birinci sınıf testinin deneme formundan alınabilecek maksimum puan 26'dır. Testin kendi sınıf düzeyinde uygulanması sonucu ortaya çıkan puan ortalaması 9,38, bir üst sınıfta uygulanması sonucu ortaya çıkan puan ortalaması 20,56 bulunmuştur. Buna göre, ortalama puanların testten alınabilecek maksimum puana oranlanmasıyla elde edilen ortalama güçlük değerleri testin kendi sınıf düzeyinde 0,36, bir üst sınıfta ise 0,79 olarak hesaplanmıştır. Buna göre 1. sınıf testinin kendi sınıf düzeyinde zor, üst sınıf düzeyinde ise kolay olduğu söylenebilir. Testin ortalama ayırıcılık değeri ise testin kendi sınıf düzeyinde 0,73, bir üst sınıfta 0,59 olarak hesaplanmıştır. Kendi sınıf düzeyinde daha yüksek olmakla birlikte ortalama ayırıcılık değerinin yüksek olduğu söylenebilir.

4. 2. İkinci Sınıf Başarı Testinin Psikometrik Özelliklerine İlişkin Bulgular ve Yorumlar

İkinci sınıf başarı testinin kapsam geçerliği, yapı geçerliği, ölçüt dayanaklı geçerlik, güvenirlik, madde güçlük indeksi ve madde ayırıcılık indekslerine ait analiz sonuçları elde edilmiştir.

Testin kapsam geçerliği, ölçülecek kazanımların belirlenmesinde 2. sınıf matematik dersi öğretim programı sayılar öğrenme alanındaki kazanımların tamamı incelenerek birkaç kazanımı kapsadığı düşünülen kritik kazanımların seçilmesi ile sağlanmıştır. Bu kazanımların seçilmesinde araştırmacı tarafından belirlenen kritik kazanımlar için matematik eğitimi alanından bir akademisyen ve 1-4. sınıflara matematik öğreten dört sınıf öğretmeninden görüş ve öneriler alınarak testin kapsam geçerliği artırılmıştır. Testte ölçülen kazanımları ve bu kazanımların Bloom taksonomisindeki

bilişsel düzeyleri belirtke tablosu hazırlanmıştır. Belirtke tablosu EK-2’de görülmektedir. Belirtke tablosuna göre testte ölçülmesi hedeflenen her bir kritik kazanım için bir maddenin bulunduğu göz önüne alındığında testin kapsam geçerliğine sahip olduğu söylenebilir.

Testin yapı geçerliğine gösterge olarak testin geliştirildiği sınıf düzeyinde ve bir üst sınıftaki uygulaması sonucu elde edilen puan ortalamaları farklılığını test etmek için öncelikle sınıf düzeylerine göre puan dağılımlarının normalliği incelenmiştir. Bunun için yapılan Kolmogorov-Simirnov ve Shapiro-Wilk test sonuçları Çizelge 10’da görülmektedir:

Çizelge 10

2. Sınıf Testinden Alınan Puanların Sınıf Düzeylerine Göre Dağılımı İçin Yapılan Normallik Testi Sonuçları

Sınıflar	Kolmogorov-Simirnov			Shapiro-Wilk		
	İstatistik	Sd	p	İstatistik	Sd	p
2. Sınıf	0,133	31	0,172	0,952	31	0,179
3. Sınıf	0,119	47	0,091	0,960	47	0,106

Çizelge 10’a göre normallik testleri puan dağılımlarının normal olduğunu ($p>0,05$) göstermektedir. Ancak, 3. sınıf puan dağılımının çarpıklık katsayısı -1,202 bulunmuş, bu nedenle dağılımın normal olmadığı sonucuna varılmıştır. İkinci sınıf testinin uygulandığı 2. ve 3. sınıfların puanları arasındaki farkın test edilmesi için uygulanan Mann Whitney-U testi sonuçları Çizelge 11’de görülmektedir:

Çizelge 11

2. Sınıf Testi Puanlarının Sınıf Düzeylerine Göre Farklılıkları İçin Uygulanan Mann Whitney-U Testi Sonuçları

Sınıf Düzeyi	N	Sıralar ortalaması	Sıralar toplamı	Mann Whitney-U	Z	p
2	31	27,97	867,00	371,000	-4,656	0,000*
3	47	47,11	2214,00			

* $p < 0,05$

Çizelge 11’de görüldüğü gibi 2. sınıf testinin uygulandığı sınıflar arasında puanların sıra ortalamaları 3. sınıf lehine anlamlı düzeyde farklılık göstermektedir ($U = 371,000$, $p = 0.00$, $z = -4,656$). Bu bulgu, 2. sınıf testinin kapsadığı kazanımları görmüş ve görmemiş olan öğrencileri ayırabildiğini göstermektedir.

Testin yapı geçerliği hakkında bilgi sağlaması için ayrıca ikinci sınıf testinden alınan puanlarda farklı sosyoekonomik düzeylerdeki ilçeler arasında fark olup olmadığına bakılmıştır. Bunun için öncelikle tek yönlü varyans analizi varsayımlarından bağımlı değişken ölçeğinin en az eşit aralık düzeyinde olması, bağımsız değişkenin kategorik olması, dağılımların normalliği ve varyansların homojenliği varsayımlarını karşılama durumları incelenmiştir (Büyüköztürk, vd., 2011). Kullanılan test eşit aralık düzeyindedir. Bağımsız değişken olan ilçeler değişkeni kategorik olduğundan bağımsız değişkenin kategorik olması varsayımı sağlanmaktadır. İlçelere göre 2. sınıf testinden elde edilen puan dağılımının normalliğini belirlemek için yapılan Kolmogorov-Smirnov ve Shapiro-Wilk test sonuçları Çizelge 12’de gösterilmektedir:

Çizelge 12

2. Sınıf Testinden Alınan Puanlarının Dağılımı İçin Yapılan Normallik Testi Sonuçları

İlçeler	Kolmogorov-Smirnov			Shapiro-Wilk		
	İstatistik	Sd	P	İstatistik	Sd	p
Sultanbeyli	0,105	99	.009	0,920	99	.000
Tuzla	0,150	95	.000	0,922	95	.000
Kadıköy	0,179	103	.000	0,841	103	.000

Çizelge 12 incelendiğinde 2. sınıf testinden alınan puanların dağılımının normal olmadığı görülmektedir ($p < .05$). Testten alınan puanların varyansların homojenliği varsayımını karşılama durumunu belirlemek için Levene testi uygulanmıştır. İkinci sınıf testi puanlarında ilçelere göre varyansların homojen olmadığı görülmüştür ($L = 4,967$; $p < .05$). Bu bulgulara göre 2. sınıf test puanları tek yönlü varyans analizi (ANOVA) varsayımlarını

karşılamadığından puan ortalamalarının ilçelere göre anlamlı farklılık gösterip göstermediğinin belirlenmesi için parametrik olmayan testlerden Kruskal-Wallis H testi kullanılmıştır. Bulgular Çizelge 13'te gösterilmektedir:

Çizelge 13

2. Sınıf Testi Puanlarının İlçelere Göre Kruskal Wallis H-Testi Sonuçları

İlçe	N	Sıra Ort.	Sd	χ^2	P
Sultanbeyli	99	108,14	2	39,546	.000*
Tuzla	95	154,12			
Kadıköy	103	183,55			

* $p < .05$

Çizelge13'te görüldüğü gibi 2. sınıf testi puanları ilçelere göre anlamlı bir farklılık göstermektedir ($\chi^2 = 39,546$, $SD=2$, $p=.000$). Farklılığın hangi ilçeler arasında olduğu belirlemek için ilçelerin ikili kombinasyonları arasında Mann Whitney-U testleri uygulanmıştır. Sonuçlar Çizelge14'tedir:

Çizelge14

İlçelerin 2. Sınıf Test Puan Ortalama Farklılıkları İçin Uygulanan Mann Whitney-U Testi Sonuçları

İlçeler	N	Sıralar ortalaması	Sıralar toplamı	Mann Whitney-U	Z	p
Sultanbeyli	99	83,01	8218,00	3268,000	-3,674	.000*
Tuzla	95	112,60	10697,00			
Sultanbeyli	99	75,13	7437,00	2487,000	-6,297	.000*
Kadıköy	103	126,85	13065,00			
Tuzla	95	89,52	8504,50	3944,500	-2,359	.018*
Kadıköy	103	108,70	11196,50			

* $p < .05$

Çizelge14'e göre tüm ilçelerin 2. sınıf test puan ortalamaları arasındaki fark anlamlı bulunmuştur.Araştırmada Sultanbeyli, Tuzla ve Kadıköy ilçeleri sırasıyla alt, orta ve üst sosyoekonomik düzeylerdeki ilçeler olarak alınmıştır.Bulgular puan ortalamalarının Sultanbeyli ve Tuzla arasında Tuzla lehine, Sultanbeyli ve Kadıköy arasında Kadıköy lehine, Tuzla ve Kadıköy

arasında Kadıköy lehine anlamlı düzeyde farklı olduğunu göstermektedir. İkinci sınıf testinin farklı sosyoekonomik düzeylerdeki ilçeler arasında beklenen yöndeki farkı ortaya çıkarabilmesi yapı geçerliğinin olduğunu göstermektedir.

Testin ölçüt dayanaklı geçerliğinin ortaya konması için testin uygulandığı sınıfların öğretmenlerinin öğrencilere verdikleri puanlarla öğrencilerin testten aldıkları toplam puanlar arasındaki korelasyon 0,75 bulunmuştur. Bu bulgu, test ile aynı özelliği ölçmeye yönelik bir başka ölçme sonucu olan öğretmen görüşü ile ilişkisinin yüksek olduğunu göstermektedir.

Testin güvenirlik analizi için hesaplanan KR-20 iç tutarlılık güvenirlik katsayısı 0,92 bulunmuştur. Bu katsayı 2. sınıf test maddelerinin hem birbirleri ile hem de testin bütünüyle tutarlı olduğunu, dolayısıyla testin iç tutarlılık anlamında güvenirliğinin oldukça yüksek olduğunu göstermektedir.

İkinci sınıf testinin bir üst sınıf olan 3. sınıfta ve kendi sınıf düzeyinde uygulanması sonucu elde edilen verilerle maddelere ait güçlük ve ayıricılık indeksleri hesaplanmıştır. Bulgular Çizelge13'te gösterilmektedir. Çizelge 15 incelendiğinde maddelerin güçlük indekslerinin 0,37 ile 0,95 arasında değerler aldığı görülmektedir. Bu testteki 35 maddeden, 15'i testin asıl formuna konulmak üzere seçilmiştir. Seçilen maddelerin güçlük indeksleri 0,45 –0,95 aralığındadır. Testte yer alan maddelerin ayıricılık indekslerinin 0,53 ile 0,94 arasında değerler aldığı görülmektedir. Testin asıl formu için seçilen maddelerin ayıricılık gücü indeksleri de 0,65 ile 0,94 aralığındadır.

Çizelge15

Uygulanan Sınıf Düzeylerine Göre 2. Sınıf Testi Deneme Formuna Ait Madde Analizi Sonuçları

Madde No	Güçlük İndeksi		Ayıricılık İndeksi	
	3.Sınıf	2.Sınıf	3.Sınıf	2.Sınıf
1	0,94	0,74	0,94	0,83
2	0,89	0,74	0,87	0,75
5*	0,77	0,48	0,76	0,72
3*	0,72	0,39	0,78	0,81
4	0,65	0,39	0,63	0,54

6	0,94	0,52	0,91	0,87
7*	0,82	0,68	0,74	0,54
8	0,94	0,77	0,53	0,86
9*	0,95	0,81	0,80	0,66
10	0,91	0,68	0,65	1,00
11*	0,73	0,39	0,65	0,65
12*	0,53	0,13	0,77	0,46
13	0,75	0,45	0,79	0,94
14*	0,87	0,65	0,83	0,88
15	0,90	0,68	0,83	0,91
16	0,74	0,61	0,55	0,40
17*	0,78	0,48	0,68	0,44
18	0,72	0,48	0,76	0,82
19*	0,67	0,29	0,85	0,87
20*	0,71	0,39	0,78	0,49
21	0,74	0,58	0,78	0,88
22	0,70	0,32	0,63	0,67
23	0,87	0,65	0,81	0,41
24	0,81	0,45	0,79	0,43
25*	0,63	0,55	0,82	0,93
26	0,87	0,68	0,94	0,58
27	0,79	0,55	0,69	0,60
28*	0,84	0,61	0,94	0,52
29*	0,55	0,23	0,77	0,90
30	0,37	0,06	0,77	-0,11
31*	0,51	0,19	0,77	0,69
32	0,55	0,16	0,68	0,72
33	0,68	0,19	0,74	0,32
34*	0,45	0,13	0,67	0,78
35	0,46	0,19	0,64	0,89

* Testin asıl formuna konulmak üzere seçilen maddeler

Çizelge15 incelendiğinde maddelerin büyük çoğunluğu için sınıf düzeylerine göre güçlük indeksleri arasında oldukça büyük fark olduğu; testin kendi sınıf düzeyindeki güçlük indekslerinin üst sınıftakilerden düşük değerler aldığı görülmektedir.

İkinci sınıf testinin deneme formundan alınabilecek maksimum puan 35'tir. Testin kendi sınıf düzeyinde uygulanması sonucu ortaya çıkan puan ortalaması 16,29, bir üst sınıfta uygulanması sonucu ortaya çıkan puan ortalaması 25,64 bulunmuştur. Ortalama puanların testten alınabilecek maksimum puana oranlanmasıyla elde edilen ortalama güçlük değerleri testin kendi sınıf düzeyinde 0,47, bir üst sınıfta ise 0,74 olarak hesaplanmıştır. Testin ortalama ayıricılık değeri ise testin kendi sınıf düzeyinde 0,68, bir üst sınıfta 0,76 olarak hesaplanmıştır.

4. 3. Üçüncü Sınıf Başarı Testinin Psikometrik Özelliklerine İlişkin Bulgular ve Yorumlar

Üçüncü sınıf başarı testinin kapsam geçerliği, yapı geçerliği, ölçüt dayanaklı geçerlik, güvenirlik, madde güçlük indeksi ve madde ayıricılık indekslerine ait analiz sonuçları elde edilmiştir.

Testin kapsam geçerliği, ölçülecek kazanımların belirlenmesinde 3. sınıf matematik dersi öğretim programı sayılar öğrenme alanındaki kazanımların tamamı incelenerek birkaç kazanımı kapsadığı düşünülen kritik kazanımların seçilmesi ile sağlanmıştır. Bu kazanımların seçilmesinde araştırmacı tarafından belirlenen kritik kazanımlar için matematik eğitimi alanından bir akademisyen ve 1-4. sınıflara matematik öğreten dört sınıf öğretmeninden görüş ve öneriler alınarak testin kapsam geçerliği artırılmıştır. Testte ölçülen kazanımları ve bu kazanımların Bloom taksonomisindeki bilişsel düzeyleri belirtke tablosu hazırlanmıştır. Belirtke tablosu EK-3'te görülmektedir. Testte ölçülmesi hedeflenen her bir kritik kazanım için bir maddenin bulunduğu göz önüne alındığında testin kapsam geçerliğine sahip olduğu söylenebilir.

Testin yapı geçerliğine gösterge olarak testin geliştirildiği sınıf düzeyinde ve bir üst sınıftaki uygulaması sonucu elde edilen puan ortalamaları farklılığını test etmek için öncelikle sınıf düzeylerine göre puan

dağılımlarının normalliği incelenmiştir. Bunun için yapılan Kolmogorov-Simirnov ve Shapiro-Wilk test sonuçları Çizelge 16'da görülmektedir:

Çizelge 16

3. Sınıf Testinden Alınan Puanların Sınıf Düzeylerine Göre Dağılımı İçin Yapılan Normallik Testi Sonuçları

Sınıflar	Kolmogorov-Simirnov			Shapiro-Wilk		
	İstatistik	Sd	p	İstatistik	Sd	p
3. Sınıf	0,105	26	.200	0,955	26	.299
4. Sınıf	0,119	51	.069	0,977	51	.422

Çizelge 16'ya göre normallik testleri puan dağılımlarının normal olduğunu ($p>.05$) göstermektedir. Ancak, testin 3. sınıftaki puan dağılımının çarpıklık katsayısı -1,157 bulunmuş, bu nedenle dağılımın normal olmadığı sonucuna varılmıştır. Üçüncü sınıf testinin uygulandığı 3 ve 4. sınıfların puanları arasındaki farkın test edilmesi için uygulanan Mann-Whitney U testi sonuçları Çizelge 17'de görülmektedir:

Çizelge 17

3. Sınıf Testi Puanlarının Sınıf Düzeylerine Göre Farklılıkları İçin Uygulanan Mann Whitney-U Testi Sonuçları

Sınıf Düzeyi	N	Sıralar ortalaması	Sıralar toplamı	Mann Whitney-U	Z	P
3	26	32,73	851,00	500,000	-1,758	.079
4	51	42,20	2152,00			

Çizelge 17, 3. sınıf testinin uygulandığı 3. ve 4. sınıf düzeylerinin sıra ortalamaları arasında 4. sınıf lehine fark olduğu, ancak bu farkın istatistiksel olarak anlamlı olmadığını göstermektedir ($U = 500,00$, $p = 0.79$, $z = -1,758$).

Testin yapı geçerliği hakkında bilgi sağlaması için ayrıca testten alınan puanlarda farklı sosyoekonomik düzeylerdeki ilçeler arasında fark olup

olmadığına bakılmıştır. Bunun için öncelikle varyans analizi varsayımlarından bağımlı değişken ölçeğinin en az eşit aralık düzeyinde olması, bağımsız değişkenin kategorik olması, dağılımların normalliği ve varyansların homojenliği varsayımlarını karşılama durumları incelenmiştir (Büyüköztürk, vd., 2011). Kullanılan test eşit aralık düzeyindedir. Bağımsız değişken olan ilçeler değişkeni kategorik olduğundan bağımsız değişkenin kategorik olması varsayımı sağlanmaktadır. Testin puan dağılımlarının normalliğini belirlemek için yapılan Kolmogorov-Smirnov ve Shapiro-Wilk test sonuçları Çizelge 18’de gösterilmektedir:

Çizelge 18

3. Sınıf Testinden Alınan Puanlarının Dağılımı İçin Yapılan Normallik Testi Sonuçları

İlçeler	Kolmogorov-Smirnov			Shapiro-Wilk		
	İstatistik	Sd	P	İstatistik	Sd	p
Sultanbeyli	0,091	98	.043	0,957	98	.003
Tuzla	0,108	111	.003	0,946	111	.000
Kadıköy	0,201	94	.000	0,849	94	.000

Çizelge 18 incelendiğinde 3. sınıf testinden alınan puanların dağılımının normal olmadığı görülmektedir ($p < .05$). Ancak 3. sınıf testi puan dağılımının çarpıklık ve basıklık katsayıları incelenmiş; çarpıklık ve basıklık katsayıları sırasıyla $-0,753$ ve $-0,287$ bulunmuştur. Çarpıklık katsayılarının ± 1 aralığında kalmasının dağılımların normalden önemli bir sapma göstermediği (Büyüköztürk, vd., 2011) göz önüne alınarak 3. sınıf testi puan dağılımı normal kabul edilmiştir.

Testten alınan puanların varyansların homojenliği varsayımını karşılama durumunu belirlemek için Levene testi uygulanmıştır. Üçüncü sınıf testi puanlarında ilçelere göre varyansların homojen olmadığı görülmüştür ($L = 8,181$; $p < .05$).

Üçüncü sınıf test puanları tek yönlü varyans analizi (ANOVA) varsayımlarını karşılamadığından puan ortalamalarının ilçelere göre farklılık gösterip göstermediğinin belirlenmesi için parametrik olmayan testlerden Kruskal Wallis H testi kullanılmış,bulgular Çizelge 19'da gösterilmiştir:

Çizelge 19

3. Sınıf Testi Puanlarının İlçelere Göre Kruskal Wallis H-Testi Sonuçları

İlçe	N	Sıra Ort.	Sd	χ^2	p
Sultanbeyli	98	123,91	2	18,599	.000*
Tuzla	111	154,64			
Kadıköy	94	178,16			

* $p < .05$

Çizelge19'da görüldüğü gibi 3. sınıf testi puan ortalamaları ilçelere göre anlamlı bir farklılık göstermektedir ($\chi^2 = 18,599$, $SD=2$, $p=.000$). Bu farklılığın hangi ilçeler arasında olduğunun belirlenmesi için ilçelerin ikili kombinasyonları arasında Mann Whitney-U-testleri uygulanmıştır. Sonuçlar Çizelge20'de gösterilmektedir:

Çizelge20

İlçelerin 3. Sınıf Test Puan Ortalama Farklılıkları İçin Uygulanan Mann Whitney-U Testi Sonuçları

İlçeler	N	Sıralar ortalaması	Sıralar toplamı	Mann Whitney-U	Z	p
Sultanbeyli	98	93,45	9158,50	4307,500	-2,596	.009*
Tuzla	111	115,19	12786,50			
Sultanbeyli	98	79,95	7835,50	2487,000	-4,218	.000*
Kadıköy	94	113,75	10692,50			
Tuzla	111	95,45	10595,00	4379,000	-1,983	.047*
Kadıköy	94	111,91	10520,00			

* $p < .05$

Çizelge 20'ye göre tüm ilçelerin 3. sınıf test puan ortalamaları arasındaki fark anlamlı bulunmuştur. Araştırmada Sultanbeyli, Tuzla ve Kadıköy ilçeleri sırasıyla alt, orta ve üst sosyoekonomik düzeylerdeki ilçeler olarak alınmıştır. Puan ortalamalarının Sultanbeyli ve Tuzla arasında Tuzla lehine, Sultanbeyli ve Kadıköy arasında Kadıköy lehine, Tuzla ve Kadıköy arasında Kadıköy lehine anlamlı düzeyde farklı olduğu görülmektedir. Bu bulgular 3. sınıf testinin farklı sosyoekonomik düzeylerdeki ilçeler arasında beklendik yöndeki farklılığı ortaya çıkarabildiğinden yapı geçerliğinin olduğunu göstermektedir.

Testin ölçüt dayanaklı geçerliğinin ortaya konması için testin uygulandığı sınıfların öğretmenlerinin öğrencilere verdikleri puanlarla öğrencilerin testten aldıkları toplam puanlar arasındaki korelasyon hesaplanmıştır. Üçüncü sınıf testi için 0,82 olarak bulunan korelasyon değeri, öğrencilerin testten aldıkları puanlarla öğretmenlerinin onlar için yaptıkları değerlendirmeler arasında güçlü bir ilişki olduğunu göstermektedir.

Testin güvenirlik analizi hesaplanan KR-20 iç tutarlılık güvenirlik katsayısı 0,93 bulunmuştur. Bu katsayı 3. sınıf test maddelerinin birbirleri ile ve testin bütünüyle tutarlı olduğunu; testin iç tutarlılık anlamında güvenirliğinin oldukça yüksek olduğunu göstermektedir.

Üçüncü sınıf testinin bir üst sınıf olan 4. sınıfta ve kendi sınıf düzeyinde uygulanması sonucu elde edilen verilerle maddelere ait güçlük ve ayıricılık indeksleri hesaplanmıştır. Bulgular Çizelge21'de gösterilmektedir. Maddelerin güçlük indekslerinin 0,30 ile 0,97 arasında değerler aldığı görülmektedir. Bu testteki 38 maddeden, 16'sı testin asıl formuna konulmak üzere seçilmiştir. Seçilen maddelerin güçlük indeksleri 0,42 –0,84 aralığındadır. Testte yer alan 35 ve 36. maddelerin ayıricılık indeksi oldukça düşük (0,19 ve 0,18) bulunmuş, bunun dışındaki maddelerin ayıricılık indekslerinin 0,35 ile 0,94 arasında değerler aldığı görülmüştür. Testin asıl formu için seçilen maddelerin ayıricılık gücü indeksleri de 0,35 ile 0,94 aralığındadır.

Çizelge21

Uygulanan Sınıf Düzeylerine Göre 3. Sınıf Testi Deneme Formuna Ait Madde Analizi Sonuçları

Madde No	Güçlük İndeksi		Ayırıcılık İndeksi	
	4.Sınıf	3.Sınıf	4.Sınıf	3.Sınıf
1	0,94	0,96	0,47	0,57
2	0,97	0,92	0,82	0,73
3*	0,71	0,77	0,66	0,84
4	0,85	0,85	0,77	0,19
5*	0,75	0,65	0,80	0,96
6	0,41	0,39	0,63	0,89
7	0,87	0,88	0,84	0,66
8*	0,84	0,69	0,80	0,81
9	0,87	0,81	0,88	0,96
10	0,86	0,65	0,76	1,00
11*	0,80	0,65	0,78	0,97
12*	0,46	0,31	0,54	0,89
13	0,72	0,54	0,54	0,95
14	0,91	0,92	0,47	-0,03
15*	0,74	0,77	0,67	0,73
16	0,82	0,77	0,66	0,40
17*	0,48	0,31	0,85	0,84
18	0,91	0,92	0,44	0,80
19	0,71	0,58	0,85	0,92
20*	0,68	0,38	0,83	0,92
21	0,70	0,65	0,78	0,85
22*	0,42	0,23	0,94	0,94
23	0,84	0,69	0,83	0,56
24	0,80	0,54	0,68	0,73
25*	0,79	0,42	0,73	0,66
26	0,75	0,31	0,83	0,74
27*	0,58	0,00	0,73	0,00
28*	0,53	0,00	0,78	0,00

29	0,33	0,00	0,76	0,00
30	0,63	0,46	0,86	0,81
31*	0,64	0,15	0,89	0,76
32	0,63	0,58	0,85	1,00
33*	0,54	0,08	0,90	0,56
34*	0,57	0,31	0,35	0,54
35	0,30	0,04	0,19	0,63
36	0,63	0,27	0,18	0,19
37	0,37	0,04	0,86	0,63
38*	0,49	0,00	0,86	0,00

* Testin asıl formuna konulmak üzere seçilen maddeler

Çizelge21 incelendiğinde 3. sınıf test maddelerinin büyük çoğunluğu için sınıf düzeylerine göre güçlük indeksleri arasında oldukça büyük fark olduğu; testin kendi sınıf düzeyindeki güçlük indekslerinin üst sınıftakilerden düşük değerler aldığı görülmektedir. Testin ilk birkaç maddesi için güçlük düzeylerinin birbirine yakın bulunmasının, 3. sınıfta da bu maddelerin uygulama tarihine kadar kapsanmış konulardan olmalarından kaynaklandığı düşünülmüştür.

Üçüncü sınıf testinin deneme formundan alınabilecek maksimum puan 38'dir Testin kendi sınıf düzeyinde uygulanması sonucu ortaya çıkan puan ortalaması 18,50, bir üst sınıfta uygulanması sonucu ortaya çıkan puan ortalaması 26,01 bulunmuştur. Ortalama puanların testten alınabilecek maksimum puana oranlanmasıyla elde edilen ortalama güçlük değerleri testin kendi sınıf düzeyinde 0,49, bir üst sınıfta ise 0,68 bulunmuştur. Testin ortalama ayırıcılık değeri ise testin kendi sınıf düzeyinde 0,72, bir üst sınıfta 0,71 olarak hesaplanmıştır.

4. 4. Dördüncü Sınıf Başarı Testinin Psikometrik Özelliklerine İlişkin Bulgular ve Yorumlar

Dördüncü sınıf başarı testinin kapsam geçerliği, yapı geçerliği, ölçüt dayanaklı geçerlik, güvenirlik, madde güçlük indeksi ve madde ayırıcılık indekslerine ait analiz sonuçları elde edilmiştir.

Testin kapsam geçerliği, ölçülecek kazanımların belirlenmesinde 4. sınıf matematik dersi öğretim programı sayılar öğrenme alanındaki kazanımların tamamı incelenerek kritik kazanımların seçilmesi ile sağlanmıştır. Bu kazanımların seçilmesinde araştırmacı tarafından belirlenen kritik kazanımlar için matematik eğitimi alanından bir akademisyen ve 1-4. sınıflara matematik öğreten dört sınıf öğretmeninden görüş ve öneriler alınarak testin kapsam geçerliği artırılmıştır. Testte ölçülen kazanımları ve bu kazanımların Bloom taksonomisindeki bilişsel düzeyleri için hazırlanan belirtke tablosu EK-4'te görülmektedir. Belirtke tablosuna göre testte ölçülmesi hedeflenen her bir kritik kazanım için bir madde bulunması 4. sınıf sayılar öğrenme alanındaki kritik kazanımların tümünün kapsandığını göstermektedir.

Testin yapı geçerliğine gösterge olarak testin geliştirildiği sınıf düzeyinde ve bir üst sınıftaki uygulaması sonucu elde edilen puan ortalamaları farklılığını test etmek için öncelikle sınıf düzeylerine göre puan dağılımlarının normalliği incelenmiştir. Bunun için yapılan Kolmogorov-Smirnov ve Shapiro-Wilk test sonuçları Çizelge 22'de görülmektedir:

Çizelge 22

4. Sınıf Testinden Alınan Puanların Sınıf Düzeylerine Göre Dağılımı İçin Yapılan Normallik Testi Sonuçları

Sınıflar	Kolmogorov-Smirnov			Shapiro-Wilk		
	İstatistik	Sd	p	İstatistik	Sd	p
4. Sınıf	0,103	27	.200	0,969	27	.583
5. Sınıf	0,083	49	.200	0,973	49	.311

Çizelge 22 incelendiğinde normallik testlerine göre puan dağılımlarının her iki sınıf düzeyinde de normal olduğu ($p > .05$) görülmektedir. Ayrıca, 4. sınıf için çarpıklık ve basıklık katsayısı sırasıyla -0,200; -0,701 ve 5. sınıf için de sırasıyla 0,116; -0,689 bulunduğundan dağılımların normalden sapma göstermediği sonucuna varılmıştır. Dördüncü sınıf test puan ortalamalarının sınıf düzeylerine göre farklı olup olmadığını incelemek için bağımsız örneklem t-test uygulanmıştır. Sonuçlar Çizelge 23'te gösterilmektedir:

Çizelge 23

4. Sınıf Testi Puanlarının Sınıf Düzeylerine Göre Farklılıkları İçin Uygulanan t-Test Sonuçları

Sınıflar	N	X	SS	Sd	t	p
4.Sınıf	27	22,11	9,50	74	-2,37	.020
5.Sınıf	49	26,82	7,53			

Çizelge 23 incelendiğinde 4. sınıf testinin uygulandığı 5. sınıftaki öğrencilerin puan ortalaması ($Ort_5 = 26,82$, $SS=7,53$) 4. sınıftakilerden ($Ort_4 = 22,11$, $SS=9,50$) istatistiksel olarak anlamlı düzeyde farklıdır ($t(74) = -2,37$, $p=.020$). Bu bulgu, 4. sınıf testinin kapsadığı kazanımları görmüş ve görmemiş öğrencileri ayırabildiğinden yapı geçerliğine dair bilgi vermektedir.

Farklı sosyoekonomik düzeylerdeki ilçeler arasında 4. sınıf testinden alınan puan ortalamalarının farklı olup olmadığını belirlemek için öncelikle varyans analizi varsayımlarından bağımlı değişken ölçeğinin en az eşit aralık düzeyinde olması, bağımsız değişkenin kategorik olması, dağılımların normalliği ve varyansların homojenliği varsayımlarını karşılama durumları incelenmiştir (Büyüköztürk, vd., 2011). Kullanılan test eşit aralık düzeyindedir. Bağımsız değişken olan ilçeler değişkeni kategorik olduğundan bağımsız değişkenin kategorik olması varsayımı sağlanmaktadır. Testin puan dağılımlarının normalliğini belirlemek için yapılan Kolmogorov-Smirnov ve Shapiro-Wilk test sonuçları Çizelge 24'te gösterilmektedir:

Çizelge24

4. Sınıf Testinden Alınan Puanlarının Dağılımı İçin Yapılan Normallik Testi Sonuçları

İlçeler	Kolmogorov-Smirnov			Shapiro-Wilk		
	İstatistik	Sd	P	İstatistik	Sd	p
Sultanbeyli	0,073	110	.195	0,976	110	.046
Tuzla	0,105	111	.004	0,959	111	.002
Kadıköy	0,108	87	.015	0,954	87	.003

Çizelge 24 incelendiğinde Kolmogorov-Smirnov ve Shapiro-Wilk test sonuçlarına göre 4. sınıf testinden alınan puanların dağılımının normal olmadığı görülmektedir ($p < .05$). Ancak ilçelere göre puan dağılımlarının çarpıklık ve basıklık katsayıları incelenmiş; bu katsayılar Sultanbeyli için sırasıyla $-0,101$ ve $-0,849$, Tuzla için $-0,322$ ve $-0,898$, Kadıköy için $-0,548$ ve $-0,523$ bulunmuştur. Çarpıklık katsayılarının ± 1 aralığında kalmasının dağılımların normalden önemli bir sapma göstermediği (Büyüköztürk, vd., 2011) göz önüne alınarak ilçelere göre 4. sınıf testi puan dağılımları normal kabul edilmiştir.

Testten alınan puanların varyansların homojenliği varsayımını karşılama durumunu belirlemek için Levene testi uygulanmıştır. Dördüncü sınıf testi puanlarında ilçelere göre varyansların homojen olduğu görülmüştür ($L = 0,223$; $p > .05$, $p = .80$).

Dördüncü sınıf test puanları tek yönlü varyans analizi (ANOVA) varsayımlarını karşıladığından test puanlarının ilçelere göre anlamlı farklılık gösterip göstermediğinin belirlenmesi için tek yönlü varyans analizi (ANOVA) kullanılmıştır. Bulgular Çizelge 25'te gösterilmektedir:

Çizelge 25

4. sınıf test puanlarının ilçelere göre tek yönlü varyans analizi (ANOVA) sonuçları

Varyans kaynağı	S.D.	Katsayılar Toplamı	Katsayılar Ortalaması	F	p
Gruplar Arası	2	2094,465	1047,232	6,011	.003*
Grup İçi	305	53138,169	174,224		
Toplam	307	55232,633			

* $p < .05$

Çizelge 25'teki bulgulara göre 4. sınıf test puanı ortalamalarına göre ilçeler arasındaki fark istatistiksel olarak anlamlı bulunmuştur ($F = 6,011$; $p < .05$). Ortalamalardaki farklılığın kaynağını belirlemek amacıyla uygun post-

hoc tekniklerinden Scheffe çoklu karşılaştırma testi uygulanmıştır. Sonuçlar Çizelge 26'da gösterilmektedir:

Çizelge 26

4. Sınıf Test Puan Ortalama Farklılıkları İçin Uygulanan Scheffe Testi Sonuçları

İlçe (i)	İlçe (J)	$X_i - X_j$	Sh	p
Sultanbeyli	Tuzla	-4,426	1,776	.046*
Sultanbeyli	Kadıköy	-6,250	1,894	.005*
Tuzla	Kadıköy	-1,824	1,890	.628

* < .05

Çizelge26incelendiğinde ilçeler arasındaki anlamlı puan ortalaması farklılığının Sultanbeyli ile Tuzla ve Sultanbeyli ile Kadıköy arasında olduğu görülmektedir. Araştırmada Sultanbeyli, Tuzla ve Kadıköy ilçeleri sırasıyla alt, orta ve üst sosyoekonomik düzeylerdeki ilçeler olarak alınmıştır. Bulgular, puan ortalamalarının Sultanbeyli ile Tuzla arasında Tuzla lehine ve Sultanbeyli ile Kadıköy arasında Kadıköy lehine farklı olduğunu göstermektedir. Tuzla ile Kadıköy ilçeleri arasındaki fark Kadıköy lehine olmakla birlikte istatistiksel olarak anlamlı bulunmamıştır. Bu bulgular 4. sınıf testinin farklı sosyoekonomik düzeylerdeki ilçeler arasında beklendi yöndeki farklılığı ortaya çıkarabildiğinden yapı geçerliğinin olduğunu göstermektedir.

Testin ölçüt dayanaklı geçerliğinin ortaya konması için testin uygulandığı sınıfların öğretmenlerinden öğrencileri matematik dersindeki algılayış ve başarıları açısından 1 ile 5 arasında bir puan vererek değerlendirmeleri istenmiş; bu puanlarla öğrencilerin testten aldıkları toplam puanlar arasındaki korelasyon hesaplanmıştır. Bu korelasyon değerleri 4. sınıf testi için 0,76 bulunmuştur. Bu bulgu, test puanları ile öğretmen değerlendirmeleri arasında güçlü bir ilişkinin olduğunu göstermektedir.

Testin güvenirlik analizi için hesaplanan KR-20 iç tutarlılık güvenirlik katsayısı 0,95 bulunmuştur. Bu katsayı, 4. sınıf test maddelerinin birbirleri ile

ve testin bütünüyle tutarlılığını, dolayısıyla testin iç tutarlılık anlamında güvenilirliğinin oldukça yüksek olduğunu göstermektedir.

Dördüncü sınıf testinin bir üst sınıf olan 5. sınıfta ve kendi sınıf düzeyinde uygulanması sonucu elde edilen verilerle maddelere ait güçlük ve ayıricılık indeksleri hesaplanmıştır. Bulgular Çizelge27’de gösterilmektedir. Maddelerin güçlük indekslerinin 0,16 ile 0,89 arasında değerler aldığı görülmektedir. Bu testteki 58 maddeden, 24’ü testin asıl formuna konulmak üzere seçilmiştir. Seçilen maddelerin güçlük indeksleri 0,40 –0,88 aralığındadır. Testte yer alan 37 ve 38. maddelerin ayıricılık indeksi düşük (0,15 ve 0,22) bulunmuş, bunun dışındaki maddelerin ayıricılık indekslerinin 0,33 ile 0,87 arasında değerler aldığı görülmüştür. Testin asıl formu için seçilen maddelerin ayıricılık gücü indeksleri de 0,33 ile 0,87 aralığındadır.

Çizelge27

Uygulanan Sınıf Düzeylerine Göre 4. Sınıf Testi Deneme Formuna Ait Madde Analizi Sonuçları

Madde No	Güçlük İndeksi		Ayıricılık İndeksi	
	5.Sınıf	4.Sınıf	5.Sınıf	4.Sınıf
1	0,86	0,74	0,35	0,55
2	0,69	0,70	0,66	0,91
3*	0,53	0,56	0,73	0,88
4*	0,57	0,67	0,69	0,80
5	0,46	0,70	0,56	1,00
6	0,28	0,19	0,75	0,88
7	0,58	0,89	0,57	0,95
8*	0,44	0,63	0,73	0,80
9	0,88	0,85	0,48	0,71
10	0,89	0,85	0,49	0,83
11*	0,88	0,89	0,54	0,80
12*	0,69	0,67	0,59	0,61
13	0,87	0,89	0,64	0,59
14	0,85	0,85	0,43	0,39
15*	0,78	0,89	0,57	-0,08

16	0,81	0,89	0,34	0,85
17	0,82	0,81	0,59	0,75
18*	0,40	0,37	0,68	0,50
19	0,75	0,78	0,62	0,38
20*	0,48	0,41	0,68	0,85
21*	0,60	0,52	0,79	0,88
22	0,57	0,44	0,74	0,74
23	0,45	0,37	0,72	0,59
24*	0,49	0,41	0,68	0,76
25	0,75	0,67	0,81	0,89
26*	0,47	0,30	0,64	0,63
27	0,24	0,11	0,61	0,79
28	0,60	0,41	0,85	0,68
29*	0,61	0,41	0,86	0,94
30	0,47	0,22	0,69	0,61
31*	0,54	0,33	0,74	0,90
32*	0,48	0,19	0,81	0,75
33	0,56	0,41	0,85	1,00
34	0,77	0,19	0,69	0,29
35	0,29	0,15	0,75	0,94
36*	0,74	0,07	0,76	0,70
37	0,82	0,63	0,15	0,33
38	0,76	0,59	0,22	0,39
39*	0,81	0,52	0,33	0,36
40	0,39	0,07	0,70	0,65
41	0,38	0,15	0,74	0,25
42*	0,38	0,15	0,77	0,42
43*	0,45	0,11	0,87	0,85
44	0,41	0,15	0,86	0,78
45*	0,67	0,00	0,77	0,00
46	0,67	0,00	0,71	0,00
47	0,75	0,04	0,77	-0,01
48*	0,69	0,00	0,77	0,00

49*	0,67	0,07	0,75	0,50
50	0,27	0,00	0,79	0,00
51	0,62	0,00	0,82	0,00
52*	0,60	0,00	0,87	0,00
53	0,57	0,00	0,84	0,00
54*	0,56	0,04	0,79	,024
55	0,19	0,00	0,59	0,00
56	0,16	0,00	0,63	0,00
57*	0,78	0,11	0,49	0,60
58	0,32	0,07	0,43	0,16

* Testin asıl formuna konulmak üzere seçilen maddeler

Çizelge27 incelendiğinde maddelerin çoğunluğu için testin kendi sınıf düzeyindeki güçlük indekslerinin üst sınıftakilerden düşük değerler aldığı görülmektedir. Testin ilk birkaç maddesi için güçlük düzeylerinin birbirine yakın bulunmasının, 4. sınıfta da bu maddelerin uygulama tarihine kadar kapsanmış konulardan olmalarından olmasından kaynaklandığı düşünülmüştür.

Dördüncü sınıf testinin deneme formundan alınabilecek maksimum puan 58'dir. Testin kendi sınıf düzeyinde uygulanması sonucu ortaya çıkan puan ortalaması 22,11, bir üst sınıfta uygulanması sonucu ortaya çıkan puan ortalaması 34,09 bulunmuştur. Ortalama puanların testten alınabilecek maksimum puana oranlanmasıyla elde edilen ortalama güçlük değerleri testin kendi sınıf düzeyinde 0,38, bir üst sınıfta ise 0,59 bulunmuştur. Testin ortalama ayırıcılık değeri ise testin kendi sınıf düzeyinde 0,65, bir üst sınıfta 0,66 olarak hesaplanmıştır. Ortalama güçlük düzeyine göre testin kendi sınıf düzeyinde zor, bir üst sınıf düzeyine orta güçlükte olduğu, ayırıcılık değerlerinin yüksek olduğu görülmektedir.

BÖLÜM 5

TARTIŞMA VE SONUÇ

Bir testte ölçülecek davranışlar testin amacına göre belirlenmekte, program değerlendirme amaçlı testlerde öğretim programındaki tüm kazanımların, başarı belirleme amaçlı testlerde ise belirlenen kritik kazanımların ölçülmesi gerekmektedir (Turgut ve Baykul, 2012). Bu çalışmada geliştirilecek olan testler ile ilkokul düzeyinde öğrencilerin matematik dersi sayılar öğrenme alanındaki başarılarının belirlenmesi amaçlandığından her sınıf düzeyinde sayılar öğrenme alanındaki tüm kazanımlar arasından belirlenen kritik kazanımlar seçilerek testlerle ölçülecek özellikler olarak belirlenmiştir.

Kapsam geçerliği, bir testin belirlenen bilgi, beceri ve davranışlar kümesini temsil etme derecesini göstermektedir (Anastasi, 1988; Büyüköztürk, 2007; Turgut ve Baykul, 2012). Kapsam geçerliğinin belirlenmesi için başvurulacak yöntemlerden biri belirtke tablosu hazırlayarak testte ölçülecek davranışları ve bunların gerektirdiği zihinsel süreçleri ortaya koymaktır (Turgut ve Baykul, 2012). Bu çalışmada öncelikle her bir testte ölçülmesi hedeflenen kazanımları ve bu kazanımların Bloom taksonomisine göre bilişsel düzeylerini gösteren belirtke tabloları hazırlanmış, daha sonra belirtke tablolarına uygun olacak şekilde test maddeleri hazırlanmıştır. Sonuç olarak geliştirilen testlerde her bir kritik kazanımı ölçmeye yönelik bir madde bulundurulmuştur. Böylelikle, testlerle ölçülmesi hedeflenen tüm kazanımların kapsanması sağlanmıştır.

Bir testin kapsam geçerliğini sağlamakta kullanılan yöntemlerden biri uzman kişilerin görüşüne başvurmaktır (Üçgöl Öcal, 2011). Test geliştirme konusunda yapılan çalışmalarda testlerle ölçülecek özelliklerin belirlenmesinde sıklıkla alan uzmanlarından görüş alma yöntemine

başvurulduğu görülmüştür (Deering, 1988; Sanlı Richards, 2004; Seber, 2001; Tosun ve Taşkesenligil, 2011; Üçgöl Öcal, 2011). Bu çalışmada da kapsam geçerliğinin sağlanabilmesi için testlerin ölçülecek kazanımlar kümesini yeterince kapsadığına dayanak oluşturmak için hem kritik kazanımların belirlenmesi, hem de bu kazanımlar için test maddelerinin yazılması aşamalarında uzman görüşlerine başvurulmuştur. Bu aşamalarda Ankara Üniversitesi İlköğretim Bölümü, Boğaziçi Üniversitesi İlköğretim Bölümü ve Hacettepe Üniversitesi Ölçme ve Değerlendirme Anabilim Dalı'ndan birer öğretim elemanı ve ilköğretim okullarında testlerin geliştirildiği sınıf düzeylerinde öğretim yapan sınıf öğretmenlerinden görüş ve öneriler alınmıştır.

Kritik kazanımların belirlenmesinde akademisyenlerden ve ilkokul düzeyinde öğretim yapan sınıf öğretmenlerinden alınan görüşler kazanımların kapsadığı kavramsal bilgilerin ve işlemsel becerilerin önemi konusunda yol gösterici olmuştur. Maddelerin yazılması aşamasında ise, öğretmenlerin önerileri gerçek sınıf ortamlarındaki uygulamalara ve deneyimlere dayandığından oldukça yardımcı olmuştur. Bunun yanında, özellikle maddelerin ölçülmek istenen kazanıma ve sınıf düzeylerine uygunluğu konusunda sınıf öğretmenlerinin görüşleri çok yararlı bulunmuştur. Örneğin; 1. sınıf düzeyinde “Bir bütünün içinde kaç yarım vardır?” şeklinde yazılan bir madde için sınıf öğretmenleri bütün ve yarım kavramlarının somut nesnelerle ilişkilendirilmesinin bu sınıf düzeyindeki öğrenciler için zihinde canlandırarak anlamayı kolaylaştıracağı şeklindeki önerileri üzerine “Bir bütün ekmeğin içinde kaç yarım ekmek vardır?” olarak düzenlenmiştir. “Saatte 75 km sabit hızla yol alan bir araç 900 km’lik bir yolu kaç saatte alır?” şeklinde yazılmış bir madde için sınıf öğretmenleri tarafından hız kavramının sınıf düzeyine uygun olmadığına belirtilmesi üzerine bu madde deneme formundan çıkarılmıştır. Ölçme uzmanından ise maddelerin ölçme ve değerlendirme açısından uygunluğu konusunda önemli uyarı ve öneriler alınmış, maddeler bunlara göre düzenlenmiştir. Örneğin; “Aşağıdaki örüntüde boş bırakılan yere hangi sayı gelmelidir?” yerine “Aşağıdaki örüntüde boş bırakılan yere uygun sayıyı yazınız” ifadesinin öğrencileri cevabı söylemek yerine yazmaya yönlendirmek için daha uygun olacağı önerisi dikkate

alınmıştır. Tüm bu görüş ve öneriler maddelerde gerekli düzenlemelerin yapılarak testlerin deneme uygulamaları için hazırlanmasında oldukça yararlı olmuştur.

Kapsam geçerliği, bir testin ölçeceği bilgi ve becerileri örnekleme derecesinin yanında test sonuçlarının ilgisiz değişkenlerden bağımsız olup olmadığını da göstermektedir (Anastasi, 1988). Başka bir deyişle, bir testin neyi ölçmeyi amaçlıyorsa yalnızca onu ölçme derecesidir. Bu araştırmada geliştirilen testlerdeki her bir madde, yalnızca o madde ile ölçülmesi amaçlanan kazanıma yönelik yazılmış, böylelikle her kazanımın ilgisiz beceri veya kazanımlardan olabildiğince arınmış şekilde ölçülmesi sağlanmıştır.

Farklı kaynaklarda yapı geçerliğinin bir göstergesi olarak iç tutarlılığın değerlendirilebileceği belirtilmektedir (Anastasi, 1988; Büyüköztürk, 2007). Bir testin homojenlik derecesi yapı geçerliği hakkında bilgi sağladığından iç tutarlılık korelasyonlarının yüksek bulunması yapı geçerliğinin yüksek olduğunu göstermektedir. Araştırmada testlerin iç tutarlık düzeyleri KR-20 güvenirlik katsayısı hesaplanarak test edilmiş ve bu katsayılar 1-4. sınıf testleri için sırasıyla 0,80, 0,92, 0,93 ve 0,95 bulunmuştur. Güvenirlik katsayısının genel olarak 0,70'ten yüksek olması test puanlarının iç tutarlılık anlamındaki güvenirliği için uygun görülmektedir(Büyüköztürk, 2007). Hesaplanan bu değerlere göre testlerin iç tutarlılıklarının yüksek olduğu, dolayısıyla yapı geçerliğine sahip oldukları söylenebilir. Bu çalışma ile amaç ve yöntem açısından oldukça benzer olan Üçgül Öcal'ın (2011) başarı testi geliştirme çalışmasında da iç tutarlılık için hesaplanan KR-20 güvenirlik katsayısı 0,92 bulunmuş, bu değer testin yapı geçerliğinin olduğu şeklinde yorumlanmıştır.

Belirli bir özellik açısından aralarında fark olması beklenen iki grup arasındaki farkı ortaya çıkaran bir testin de ölçtüğü özellik ya da yapının tek boyutlu olduğu hakkında bilgi verdiği belirtilmektedir (Anastasi, 1988). Araştırmada testler geliştirildikleri sınıf düzeyindeki ve bir üst sınıftaki öğrencilere uygulanmış; bir üst sınıftaki öğrenciler tüm kazanımları bir önceki yıl görmüş olduklarından bu iki sınıf düzeyi arasında puan ortalamaları açısından anlamlı derecede fark olması beklenmiştir. Ayrıca, testlerin

deneme uygulaması farklı sosyoekonomik düzeylerdeki ilçelerden yaklaşık olarak eşit sayıda öğrenci üzerinde yapılmış, test puanlarının üst sosyoekonomik düzeylerdeki ilçeler lehine farklı olması beklenmiştir. Test geliştirme ile ilgili literatürde rastlanan çalışmalarda geliştirilen başarı testlerinin farklı sınıf düzeylerinde (Üçgöl Öcal, 2011) veya farklı bölgelerdeki okullarda (Kılıç ve Sağlam, 2009; Seber, 2001) uygulanarak bu düzeyler arasındaki farkın yapı geçerliği anlamında değerlendirildiği görülmektedir. Araştırmanın bulguları 3. sınıf testi hariç, testlerin tümü için sınıf düzeyleri arasındaki puan ortalaması farkının anlamlı olduğunu göstermiştir. Üçüncü sınıf testi için de istatistiksel olarak anlamlı olmamasına rağmen beklenen şekilde üst sınıf lehine puan farkı bulunmuştur. Üçgöl Öcal (2011) tarafından yapılan 6. sınıf matematik dersi hazırbulunuşluk testi geliştirme çalışmasında da benzer şekilde testin deneme formu 4, 5 ve 6. sınıf öğrencilerine uygulanmış, bu üç sınıf düzeyi arasındaki farkın anlamlı bulunması testin yapı geçerliğinin olduğu şeklinde yorumlanmıştır. Bulgulara göre, yalnızca 4. sınıf testi için orta ve üst sosyoekonomik düzeydeki ilçeler hariç diğer tüm testler literatürdeki sosyoekonomik düzey arttıkça başarının arttığı yönündeki bulguları (Dursun, 2004; Şirin, 2005; Perry ve McConney, 2010) doğrular şekilde ilçeler arasında beklenen yönde farklılığı ortaya çıkarmıştır. Dolayısıyla testlerin, ölçülmesi planlanan yapı olan başarı açısından geçerli olduğu sonucuna varılmıştır.

Testlerin ölçüt dayanaklı geçerliklerinin belirlenmesi için testlerin deneme uygulamalarına katılan öğrenciler için öğretmenlerinden alınan değerlendirme puanları ile testlerden aldıkları toplam puanlar arasındaki korelasyonlar 1-4. sınıf testleri içinsırasıyla 0,79, 0,75, 0,82 ve 0,76 bulunmuştur. Bu değerler, aynı özelliği ölçmeye yönelik farklı iki ölçme sonucu olan test puanları ile öğretmen değerlendirme puanları arasında güçlü bir ilişkinin olduğunu, dolayısıyla testlerin ölçüt dayanaklı geçerliğe sahip olduklarını göstermiştir. Başer (1996) tarafından yapılan 9. sınıf öğrencileri için matematik başarı testi geliştirme çalışmasında da ölçüt dayanaklı geçerlik için test puanları ile okul matematik başarısı arasındaki korelasyon 0,64 bulunmuştur. Bu korelasyon değeri ile karşılaştırıldığında

araştırmada geliştirilen 1-4. sınıf testlerinin ölçüt dayanaklı geçerliklerinin oldukça yüksek olduğu görülmektedir.

Geliştirilmesi amaçlanan testlerin güvenirliklerini belirlemek için KR-20 güvenirlik katsayısı hesaplanmıştır. Hesaplanan KR-20 güvenirlik katsayıları 1-4. sınıf testleri için sırasıyla 0,80, 0,92, 0,93 ve 0,95 bulunmuştur. Güvenirlik katsayısının genel olarak 0,70'ten yüksek olması test puanlarının iç tutarlılık anlamındaki güvenirliği için uygun görülmektedir(Büyüköztürk, 2007). Dolayısıyla hesaplanan bu değerlere göre testlerin güvenirliklerinin yüksek olduğu söylenebilir. Turgut ve Baykul'a (2012) göre KR-20 güvenirlik katsayısının yüksek bulunması testin güvenirliğinin yüksek olmasının yanında ölçülen özelliğin tek boyutlu olduğunun da bir göstergesidir. Literatürde bu çalışmanın amacı ile benzerlik gösteren başarı testi geliştirme çalışmalarında da testlerin güvenirliğini belirlemede KR-20 güvenirlik katsayısının hesaplandığı, 0,72 ile 0,95 arasında değişen değerlerde KR-20 güvenirlik katsayıları rapor edildiği görülmüştür(Başer, 1996; Gönen, vd., 2011; Paek, vd., 1999; Üçgül Öcal, 2011; Üçüncü, 2010). Bu çalışmada geliştirilen testler için hesaplanan KR-20 güvenirlik katsayılarının ilgili çalışmalarda hesaplanan katsayılar ile paralel olmakla birlikte oldukça yüksek değerlerde olduğu sonucuna varılmıştır.

Test geliştirme sürecinde, ölçülecek bilgi ve beceriler için hazırlanan maddelerin deneme uygulamasından sonra elde edilen veriler üzerinden bazı analizler yapılır. Madde analizi olarak adlandırılan bu analizler; güçlük ve ayıricılık gibi madde istatistiklerinin hesaplanması, bunlara göre doğrudan teste konulabilecek maddelerin seçilmesi, düzeltilmesi gereken maddelerin belirlenmesi ve test için uygun olmayan maddelerin ayıklanması amaçlarıyla yapılır(Turgut ve Baykul, 2012). Bu çalışmada her test için ayrı ayrı madde analizleri uygulanmış; maddelere ait güçlük ve ayıricılık gücü indeksleri hesaplanmıştır. Daha sonra bu indeksler literatürde belirtilen ölçütlere göre değerlendirilmiştir.

Madde güçlüğü maddeyi doğru cevaplayan katılımcıların sayısının tüm katılımcıların sayısına oranlanmasıyla elde edilen güçlük indeksi değeriyle ifade edilmektedir (Özçelik, 2010; Turgut ve Baykul, 2012). Güçlük indeksinin

0 değerine yakın bulunması maddenin zor olduğu, 1 değerine yakın bulunması ise kolay bir madde olduğu şeklinde yorumlanır. Genel olarak madde güçlük indeksi 0,81 ve üzerinde olan maddeler oldukça kolay, 0,61 – 0,80 arasında olanlar kolay, 0,41 – 0,60 arasında olanlar orta güçlükte, 0,21 – 0,40 arasında olanlar zor ve 0,20 ve altında olanlar oldukça zor madde olarak değerlendirilmektedir (Üçgül Öcal, 2011). Araştırma bulgularına göre madde güçlük indeksleri bu ölçütlere göre değerlendirilmiş ve testlerin asıl formlarına konulan maddelerin güçlük kategorilerine göre Çizelge28'deki gibi dağıldıkları sonucuna varılmıştır.

Çizelge28

Testlerin Asıl Formlarındaki Maddelerin Güçlük Kategorilerine Göre Dağılımı

Test \ Güçlük	1. Sınıf Testi	2. Sınıf Testi	3. Sınıf Testi	4. Sınıf Testi
Oldukça Kolay	6	4	1	2
Kolay	4	7	7	8
Orta Güçlükte	2	4	8	12
Zor	1	0	0	2
Toplam	13	15	16	24

Çizelge 28 incelendiğinde 3 ve 4. sınıf testlerinde bulunan maddelerin %50'sinin orta güçlükte, 1 ve 2. sınıf testlerinde ise maddelerin çoğunlukla kolay olduğu görülmektedir. Testlerin hesaplanan ortalama güçlük değerleri 1-4. sınıflar için sırasıyla 0,79, 0,74, 0,69 ve 0,59 bulunmuştur. Bu değerlere bakıldığında özellikle birinci sınıf testinin oldukça kolay olduğu sonucuna varılabilir. Bu durumun, testin deneme uygulamasının bir üst sınıftaki öğrencilere yapılmasından kaynaklanabileceği gibi, daha olası bir şekilde 1. sınıf matematik öğretim programının sayılar öğrenme alanında bulunan bazı kazanımların gerektirdiği becerilerin okul öncesinde kazanılmış beceriler olmasından kaynaklanabileceği düşünülmüştür. Okul öncesi dönemdeki çocukların 1. sınıf programının özellikle doğal sayılar alt öğrenme alanında bulunan ritmik sayma, nesne grubuna karşılık gelen sayıyı bulma gibi

becerileri hem aile hem de okul öncesi eğitim yaşantıları yoluyla büyük ölçüde kazanmış oldukları araştırmacının önceki gözlemlerine dayalı olarak bilinmektedir. Dolayısıyla 1. sınıf testinin kolay olması kazanımların bu özelliğinden kaynaklandığı şeklinde yorumlanmıştır.

Testlerin tümü için geçerli olmak üzere dikkat çeken bir bulgu olarak öğretim programına göre eğitim-öğretim yılı sonunda işlenen konulara yönelik maddelerin doğru cevaplanma yüzdelerinin düşük olduğu belirtilmeye değer görülmüştür. Öğretim programındaki kazanımlar için zaman sınırlılıklarından dolayı yılsonuna gelen konuların yetiştirilememesi veya bu dönemlerde öğrencilerin okula devam etmeme durumlarının öğrencilerin öğretim programının son kazanımlarını yeterince edinmemelerinde etkili olduğu düşünülmektedir. Bunun yanında, uygulama yapılan sınıflardaki öğretmenler tarafından özellikle kesirler alt öğrenme alanındaki kazanımların öğrencilerce somutlaştırmada güçlük çekilen kavramlar içerdiğinden ve aynı zamanda dönem sonuna denk geldiğinden yeterince edinilemediği belirtilmektedir.

Test maddelerinin ayırıcılık gücü genel bir ifadeyle madde ile ölçülen davranışa sahip olan ve olmayan cevaplayıcıları ayırma gücü olarak tanımlanmaktadır (Özçelik, 2010; Turgut ve Baykul, 2012). Araştırmada testlerin uygulanması sonucu elde edilen veriler madde istatistik programında analiz edilerek maddelere ait ayırıcılık indeksleri ortaya çıkarılmıştır. Bu indeksler maddelerin testte yoklanan kazanımları edinen ve edinmeyen öğrencileri birbirinden ayırabilme gücü olarak değerlendirilmiştir.

Turgut ve Baykul'a göre (2012) madde ayırıcılık gücü belirli bir davranışa sahip olanla olmayanı ayırdığından maddenin ölçmesi istenen davranışı ölçüp ölçmediği ile de ilgilidir, dolayısıyla madde ayırıcılık indeksi madde geçerliği olarak da adlandırılabilir. Çift serili korelasyon analizi ile hesaplanan madde-toplam korelasyonu göz önünde bulundurularak seçilmiş maddelerden oluşan bir test ve testin her bir maddesi cevaplayıcılar arasındaki ayrımı ortaya koyabildiğinden iç tutarlılığa sahip olduğu söylenebilir (Anastasi, 1988). İç tutarlılık, testin yapı geçerliğine ilişkin bir gösterge olarak kabul gördüğünden madde ayırıcılık gücü indeksleri testin geçerliği hakkında bilgi vermektedir. Dolayısıyla, testi oluşturan maddelerin

ayırıcılık gücü indekslerinin yüksek olması maddelerin ve toplamda testin tamamının geçerliğinin de yüksek olduğu şeklinde yorumlanabilir.

Geliştirilen testin deneme uygulaması yapıldıktan sonra asıl forma konulacak maddelerin seçiminde genel olarak aşağıdaki ayırıcılık indeksi ölçütleri kullanılmaktadır:

Ayırıcılık indeksi

0,19'un altında olan maddeler testten çıkarılmalı,

0,20 –0,29 arasında olan maddeler düzeltilmeli,

0,30 ve üzerinde olan maddeler herhangi bir düzeltmeye gerek olmadan teste konur (Turgut ve Baykul, 2012).

Ayırıcılık gücünün 0,40 ve üzerinde olması maddelerin yoklanan kazanımı edinen ve edinmeyenleri ayırma amacına oldukça iyi bir şekilde hizmet ettiği şeklinde yorumlanmaktadır (Üçgül Öcal, 2011).

Madde seçiminde ideal olarak yüksek ayırıcılık indeksi ile birlikte 0,50 civarında güçlük indeksine sahip olan maddeleri seçmekten yana davranılmalıdır (Anastasi, 1988; Üçgül Öcal, 2011). Böylece katılımcılar arasında maddeyi doğru cevaplayanların testin bütününde başarılı olan katılımcılar olduğu, dolayısıyla bilen ve bilmeyeni tam olarak ayıran istendik niteliklere sahip maddelerden oluşan ortalama güçlük düzeyinde bir test oluşturulmaktadır. Madde seçiminde ayırıcılık gücü katsayısına öncelik verilmektedir (Turgut ve Baykul, 2012). Bu önceliğin gerekçesi, madde ayırıcılık gücü testin ölçme amacına hizmet etme derecesi olan geçerlik özelliği ile doğrudan ilişkili olduğundan dolayı taşıdığı önemdir. Bu çalışmada geliştirilecek olan testlerin amacı başarı ölçme olduğundan, test maddeleri ile ölçülmesi amaçlanan kazanımları edinmiş olan ve olmayan katılımcıları ayıran maddelerin seçilmesi için bu durum göz önünde bulundurularak maddelerin öncelikle ayırıcılık indekslerine dikkat edilmiştir. Ancak, testin asıl formunun ortalama güçlüğünü ayarlayabilmek amacıyla ayırıcılık gücü 0,30'un altına düşmeyecek şekilde en ayırıcı maddeler yerine daha az ayırıcı olanlardan yana seçim yapılabilmektedir (Turgut ve Baykul, 2012). Bulgulara göre bazı kazanımlar için yazılan birkaç madde arasında ayırıcılık

gücüindeksi yüksek olmasına rağmen güçlük indeksi çok düşük ya da yüksek bulunduğu ayıricılık indeksleri için daha önce belirtilen ölçütler dahilinde kalınarak güçlük düzeyine göre seçim yapılmıştır. Örneğin, 1. sınıf testi deneme formunun 7 ve 8. maddelerinden ayıricılık gücü indeksi 0,46 ve güçlük indeksi 0,84 olan 7. maddenin güçlük indeksi oldukça yüksek olduğundan testin güçlük ortalamasını dengelemek için bunun yerine ayıricılık gücü indeksi 0,35 ve güçlük indeksi 0,59 olan 8. madde seçilmiştir.

İlgili araştırmalarda aynı kazanımı ölçmeye yönelik maddeler arasından testlerin asıl formlarına madde seçiminde ayıricılık gücü indeksi için 0,30 değerini alt sınır olarak belirleyen çalışmaların (Başer, 1996; Üçgül Öcal, 2011; Üçüncü, 2010) yanına bu sınır değerinin 0,20 ya da 0,25'e düşürüldüğü çalışmalar da (Başer, 1996; Gönen, vd., 2011; Gucken, 1986; Tosun ve Taşkesenligil, 2011) görülmüştür. Üçgül Öcal'ın (2011) çalışmasında geliştirilen testteki maddelerin ayıricılık değerleri 0,39-0,80 arasındadır. Bu araştırmada 1-4. sınıf testlerinin asıl formlarına konmak üzere seçilen maddelerin ayıricılık gücü indekslerinin 0,35-0,94 arasında bulunduğu göz önüne alındığında literatürdeki test geliştirme çalışmalarıyla benzer madde seçim ölçütü kullanıldığı ve ilgili pek çok çalışmadakinden yüksek düzeyde ayıricılık gücü indekslerine sahip maddelerin seçildiği görülmektedir.

Araştırma bulgularına göre, 1. sınıf testinde yalnızca bir maddenin ayıricılık gücü indeksi oldukça düşük (0,03) bulunmuş, bunun dışındaki maddelerin ayıricılık indekslerinin 0,35 ile 0,91 arasında değerler aldığı görülmüş ve testin asıl formu için seçilen maddelerin ayıricılık gücü indeksleri 0,35 ile 0,91 aralığında olmuştur. Asıl formdaki maddelerin 12'sinin (%92,3) ayıricılık indeksi 0,40'ın üzerinde 1'inin (%7,7) ise ayıricılık indeksi 0,35'tir. Buna göre geliştirilen 1. sınıf testinin ayıricılık gücü oldukça yüksek maddelerden oluştuğu söylenebilir. Bu durum Turgut ve Baykul'un (2012) ayıricılık gücü ile madde geçerliği arasında kurdukları bağ açısından değerlendirildiğinde testin yapıgeçerliliğinde kanıtı şeklinde yorumlanmıştır.

İkinci sınıf testi bulguları, testin deneme formunda bulunan maddelerin ayıricılık indekslerinin 0,53 ile 0,94 arasında olduğunu göstermiştir. Testin asıl formu için seçilen maddelerin ayıricılık gücü indeksleri ise 0,65 ile 0,94

aralığındadır. Ayırıcılık gücü indekslerinin oldukça yüksek değerlere sahip olması testin yapı geçerliği için de bir gösterge olarak yorumlanmıştır.

Üçüncü sınıf testinin deneme formundaki maddelerden 35 ve 36. maddelerin ayırıcılık indeksi oldukça düşük (0,19 ve 0,18) bulunmuş, bunun dışındaki maddelerin ayırıcılık indekslerinin 0,35 ile 0,94 arasında değerler aldığı görülmüştür. Testin asıl formu için seçilen maddelerin ayırıcılık gücü indeksleri de 0,35 ile 0,94 aralığında bulunmaktadır. Asıl formdaki maddelerden yalnızca 1'inin (%6,2) ayırıcılık indeksi 0,35, diğerlerinininki (%93,8) 0,54 ve üzerindedir. Üçüncü sınıf testinin madde ayırıcılık indeksleri oldukça tatmin edici değerlerde maddelerden oluşmasından dolayı yapı geçerliğinin yüksek olduğu sonucuna varılmıştır.

Dördüncü sınıf testinin deneme formundaki maddelerden 37 ve 38. maddelerin ayırıcılık indeksi düşük (0,15 ve 0,22) bulunmuş, bunun dışındaki maddelerin ayırıcılık indekslerinin 0,33 ile 0,87 arasında değerler aldığı görülmüştür. Testin asıl formu için seçilen maddelerin ayırıcılık gücü indeksleri de 0,33 ile 0,87 aralığındadır. Asıl forma konulan maddelerden yalnızca 1'inin (%4,1) ayırıcılık indeksi 0,33, diğerlerinininki (%95,9) 0,49 ve üzerindedir. Dördüncü sınıf testinin ayırıcılık indeksleri yüksek maddelerden oluşmasından dolayı yapı geçerliğinin yüksek olduğu çıkarımı yapılmıştır.

Testlerdeki maddelere ait madde istatistikleri değerlendirilerek güçlük ve ayırıcılık indeksleri açısından testlere konulmaya uygun olan maddelerin seçilmesiyle testler son haline getirilmiştir. Testlerin asıl formlarında 1-4. sınıf düzeylerinde sırasıyla 13, 15, 16 ve 24 madde bulunmaktadır. Turgut ve Baykul (2012) bir testte başka bir konu için ön koşul niteliğinde olan bir konuya ait maddeye öncelik verilmesini; ayrıca maddelerin kolaydan zora doğru sıralanmasını önermektedir. Testlerin asıl formları oluşturulurken bu durumlar dikkate alınarak ön koşul niteliğinde olduğu düşünülen kazanımlara ait maddelere öncelik verilmek üzere diğer maddeler güçlük indekslerine göre yüksekten düşüğe doğru, başka bir deyişle maddeler kolaydan zora doğru sıralanmıştır.

Sonuç olarak 1-4. sınıf düzeylerinde matematik dersi sayılar öğrenme alanındaki kazanımlar çerçevesinde yeterli düzeyde geçerli ve güvenilir başarı testleri elde edilmiştir. Geliştirilen testlere ait birer örnek soru EK-5'te sunulmaktadır.

BÖLÜM 6

ÖNERİLER

Çalışma sonunda geliştirilen testler, 1-4. sınıf öğrencilerinin matematik dersi sayılar öğrenme alanındaki kazanımlar çerçevesindeki başarılarının belirlenmesine ihtiyaç duyulan farklı çalışmalarda kullanılabilir. Araştırmamanın bulgu ve sonuçları geliştirilen testlerin sınıf düzeyi ve sosyoekonomik düzeyler arasındaki başarı farkını ortaya çıkarabilecek nitelikte olduğunu gösterdiğinden bu tür değişkenler açısından başarı farklılıklarını incelemeyi amaçlayan başka çalışmalarda bu testlerden yararlanılabilir.

Ayrıca, matematik eğitimi alanında özel öğrenme güçlükleri yaşayan veya özel eğitime ihtiyaç duyan bireylerin belirlenmesi amacıyla kullanılabilir.

Geliştirilen testler 1-4. sınıflarda uygulanmakta olan öğretim programındaki kazanımlar temelinde olduğundan bu sınıf düzeylerinde eğitim veren öğretmenler tarafından bir eğitim-öğretim yılı sonunda öğrencilerin erişti düzeylerinin ya da eğitim- öğretim yılının başında hazırbulunuşluklarının belirlenmesinde yararlanılabilir.

Bu çalışmada geliştirilen testler ile ölçülen kritik kazanımlar belirlenirken, testler kısa cevaplı maddelerden oluşturulduğundan ve çok sayıda katılımcıya toplu olarak uygulanacağından ölçülmesi bireysel uygulama gerektiren zihinden işlem, tahmin gibi becerileri kapsayan kazanımlar alınmamıştır. Yeni çalışmalarda, bu türden becerilerin de ölçülebileceği testler geliştirilebilir.

Geliştirilen testlere ait geçerlik ve güvenirlik göstergeleri testlerin deneme uygulamalarından elde edilen veriler üzerinden ortaya çıkarılmıştır. Bu testler farklı örneklemeler üzerinde yeniden uygulanarak geçerlik ve güvenirlik çalışmaları yapılabilir.

KAYNAKÇA

- Akdeniz, M. (2003). *Kısa Cevaplı ve Çoktan Seçmeli Testlerin Madde ve Test İstatistiklerinin Karşılaştırılması*. Yüksek lisans tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Anastasi, A. (1988). *Psychological Testing*. New York: Macmillan Publishing.
- Ansley, T. (1997). The Role of Standardized Achievement Tests in Grades K-12. In G. D. Phye (Ed.), *Handbook of Classroom Assessment. Learning, Achievement and Adjustment*. California: Academic Press.
- Aydın, B. (1993). *Seçme Gerektiren Test Maddeleri ile Kısa Cevap Gerektiren Test Maddelerinin Psikometrik Özellikleri ve Öğrenci Başarısı Bakımından Karşılaştırılması*. Yüksek lisans tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Aydın Gürler, S. (2011). 6. Sınıf Fen Ve Teknoloji Programı "Maddenin Tanecikli Yapısı Ve Özellikleri" Ünitesindeki Öğrenci Kazanımlarının Gerçekleşme Düzeyinin Belirlenmesi. Yüksek lisans tezi, Fırat Üniversitesi Fen Bilimleri Enstitüsü, Elazığ.
- 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.
- Büyüköztürk, Ş. (2007). *Sosyal Bilimler İçin Veri Analizi El Kitabı*. Ankara: Pegem A Yayıncılık.
- Büyüköztürk, Ş., Çokluk, Ö., & Köklü, N. (2011). *Sosyal Bilimler İçin İstatistik*. Ankara: Pegem Akademi.
- 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.
- Deering, T. E. (1988). The Missouri Mastery and Achievement Tests: An Innovation in Test Development. *American Secondary Education*, 17(2), 20-24.

- Demirel, Ö. (2010). *Kuramdan Uygulamaya Eğitimde Program Geliştirme*. Ankara: Pegem Akademi.
- Dursun, Ş. (2004). Öğrencilerin Matematikte Başarısını Etkileyen Faktörler: Matematik Öğretmenlerinin Görüşleri Bakımından. *Gazi Eğitim Fakültesi Dergisi*, 24(2), 217-230.
- Ebel, R. L. (1965). *Measuring Educational Achievement*. New Jersey: Prentice-Hall Education Series.
- Ertürk, S. (1972). *Eğitimde Program Geliştirme*. Ankara: Yelkentepe Yayınları.
- 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.
- Gucken, J. E. (1986). *The Construction, Validation, and Use Of a Diagnostic Test in Percent for High School Students*. Temple University, Ann Arbor.
- Jitendra, A. K., Parker, R., & Kameenui, E. J. (1997). Aligning the Basal Curriculum and Assessment in Elementary Mathematics: The Exploratory Development of Curriculum-Valid Tests. *Assessment for Effective Intervention*, 22(2), 101-127.
- Kane, M. (2006). Content Related Validity Evidence in Test Development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of Test Development*. New Jersey: Lawrence Erlbaum Associates.
- Kılıç, D., & Sağlam, N. (2009). Öğrencilerin Genetik Kavramları Anlamalarını Belirlemeye Yönelik İki Aşamalı Bir Testin Geliştirilmesi. *Eurasian Journal of Educational Research*(36), 227-244.
- Koç, N. (1985). Standart Başarı Testlerinin, Bir Eğitim Sisteminde Verilen Çeşitli Kararlardaki Yeri ve Önemi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 17(1-2), 19-30.
- Koç, N. (1987). Öğrenci Gelişiminin Değerlendirilmesi-Okul Test Programları. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 19(1-2), 281-294.

- Koç, N. (1993). Eğitim Sistemimizde Ölçme ve Değerlendirme Alanındaki Gelişmeler. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 25(2), 387-407.
- Lichtenberger, E. O., & Smith, D. R. (2005). *Essentials of WIAT II and KTEA II Assessment Essentials of Psychological Assessment*. New Jersey: John Willey & Sons.
- MEB. (2009). *İlköğretim Matematik Dersi 1-5. Sınıflar Öğretim Programı*. Ankara: MEB.
- Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). *Trends in International Mathematics and Science Study (TIMSS) 2011 Assessment Frameworks*. Boston College: TIMMS & PIRLS International Study Center.
- Özçelik, D. A. (2010). *Test Hazırlama Kılavuzu*. Ankara: Pegem Akademi.
- Paek, P., Holland, P. W., & Suppes, P. (1999). Development and Analysis of a Mathematics Aptitude Test for Gifted Elementary School Students. *School Science and Mathematics*, 99(6), 338-347.
- Perry, L. B., & McConney, A. (2010). Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003. *Teachers College Record*, 112(4), 1137-1162.
- Sanchez, W. B., & Ice, N. F. (2005). Strike a Balance in Assessment. Web: <http://www.nctm.org/news/content.aspx?id=624> adresinden 25 Mayıs 2012'de alınmıştır.
- Sanlı Richards, N. (2004). *İlköğretim Sekizinci Sınıf Düzeyi İçin Akademik Yetenek Testi Geliştirmeye İlişkin Bir Çalışma*. Doktora tezi, Ankara Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Seber, G. (2001). *Çoklu Zekâ Alanlarında Kendini Değerlendirme Ölçeğinin Geliştirilmesi*. Yüksek lisans tezi, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Şencan, H. (2005). *Sosyal ve Davranışsal Ölçümlerde Güvenilirlik ve Geçerlik*. Ankara: Seçkin Yayıncılık.
- Şirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, 75(3), 417-453.

- Şirin, A. (2008). Oluşturmacılığın Kuramsal Temelleri. *Marmara Coğrafya Dergisi*, 17, 196-207.
- Tosun, C., & Taşkesenligil, Y. (2011). Revize Edilmiş Bloom'un Taksonomisine Göre Çözümler Ve Fiziksel Özellikleri Konusunda Başarı Testinin Geliştirilmesi: Geçerlik Ve Güvenirlik Çalışması. *Kastamonu Eğitim Dergisi*, 19(2), 499-522.
- Turgut, M. F., & Baykul, Y. (2012). *Eğitimde Ölçme ve Değerlendirme*. Ankara: Pegem Akademi.
- Üçgöl Öcal, İ. (2011). *İlköğretim 6. Sınıf Matematik Dersi Sayılar Öğrenme Alanı İçin Bilişsel Hazırbulunuşluk Testinin Geliştirilmesi*. Yüksek lisans tezi, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Üçüncü, K. (2010). *İlköğretim (2-5. sınıf) Öğretmenlerinin Çarpma İşlemi Öğretimine İlişkin Görüşleri Ve Öğrencilerin Kazanımlara Ulaşma Düzeyleri*. Yüksek lisans tezi, Gazi Üniversitesi, Ankara.

Ek-1**1. Sınıf Matematik Dersi Sayılar Öğrenme Alanı
Başarı Testi Belirtke Tablosu**

Madde No	Alt Öğrenme Alanı	Kazanım	Bilişsel Düzey	Soru Sayısı	p_{jx}	r_{jx}
1	Doğal Sayılar	100 içinde ileriye doğru birer ve onar ritmik sayar.	Kavrama	1	0,95	0,42
2	Doğal Sayılar	20 içinde geriye birer sayar.	Kavrama	1	0,95	0,54
3	Doğal Sayılar	Miktarı 20'ye kadar olan bir grup nesneye karşılık gelen sayıyı rakamlarla yazar.	Kavrama	1	0,96	0,40
4	Doğal Sayılar	20'ye kadar olan bir sayıya karşılık gelen çokluğu belirler.	Kavrama	1	0,92	0,45
5	Doğal Sayılarla Toplama İşlemi	Toplamları 20'ye kadar olan iki doğal sayının toplamını bulur.	Uygulama	1	0,86	0,63
6	Doğal Sayılarla Toplama İşlemi	Toplamları 20'yi geçmeyen iki doğal sayının toplandığı bir işlemde, toplam ile toplananlardan biri verildiğinde verilmeyen toplananı bulur.	Analiz	1	0,73	0,69
7	Doğal Sayılarla Toplama İşlemi	Toplamları 10 veya 20 olan sayı ikililerini belirler.	Analiz	1	0,78	0,91
8	Doğal Sayılarla Toplama İşlemi	Doğal sayılarla toplama işlemini gerektiren problemleri çözer.	Uygulama	1	0,84	0,89
9	Doğal Sayılarla Çıkarma İşlemi	20'ye kadar olan iki doğal sayının farkını bulur.	Uygulama	1	0,80	0,72
10	Doğal Sayılarla Çıkarma İşlemi	Doğal sayılarla çıkarma işlemini gerektiren problemleri çözer.	Uygulama	1	0,76	0,78
11	Doğal Sayılar	Sıra bildiren ilk yirmi sayıyı kullanır.	Kavrama	1	0,59	0,35
12	Kesirler	Yarım ve bütün arasındaki ilişkiyi açıklar.	Kavrama	1	0,55	0,61

13	Doğal Sayılarla Çıkarma İşlemi	Bir çıkarma işleminde verilmeyen eksilen veya çıkanı bulur.	Analiz	1	0,40	0,52
----	--------------------------------	---	--------	---	------	------

Ek-2**2. Sınıf Matematik Dersi Sayılar Öğrenme Alanı
Başarı Testi Belirtke Tablosu**

Madde No	Alt Öğrenme Alanı	Kazanım	Bilişsel Düzey	p_{jx}	r_{jx}
1	Doğal Sayılarla Toplama İşlemi	Toplamları 100'e kadar olan doğal sayıların eldesiz toplama işlemini yapar.	Uygulama	0,95	0,80
2	Doğal Sayılar	100'den küçük doğal sayıların basamaklarını adlandırır, basamaklardaki rakamların basamak değerlerini belirtir.	Bilgi	0,77	0,76
3	Doğal Sayılar	100'den küçük en çok dört doğal sayıyı büyükten küçüğe veya küçükten büyüğe doğru sıralar.	Kavrama	0,82	0,74
4	Doğal Sayılarla Toplama İşlemi	Eldeli toplama işlemini yapar.	Uygulama	0,73	0,65
5	Doğal Sayılar	Sayı örüntüleri oluşturur.	Kavrama	0,72	0,78
6	Doğal Sayılarla Toplama İşlemi	Doğal sayılarla toplama işlemini gerektiren problemleri çözer.	Uygulama	0,87	0,83
7	Doğal Sayılarla Çarpma İşlemi	10'a kadar olan doğal sayıları 2, 3, 4 ve 5 sayılarıyla çarpar.	Bilgi	0,84	0,94
8	Doğal Sayılarla Çıkarma İşlemi	100'den küçük ve onluk bozmayı gerektirmeyen iki doğal sayının farkını bulur.	Uygulama	0,78	0,68
9	Doğal Sayılarla Çıkarma İşlemi	Onluk bozmayı gerektiren iki doğal sayının farkını bulur.	Uygulama	0,67	0,85
10	Doğal Sayılarla Çıkarma İşlemi	Doğal sayılarla yapılan bir çıkarma işleminde verilmeyen eksileni veya çıkanı belirler.	Analiz	0,71	0,78
11	Doğal Sayılarla	Doğal sayılarla toplama ve çıkarma	Uygulama	0,63	0,82

	Çıkarma İşlemi	işlemlerini gerektiren problemleri çözer.			
12	Doğal Sayılarla Çarpma İşlemi	Doğal sayılarla çarpma işlemini gerektiren problemleri çözer.	Uygulama	0,55	0,77
13	Doğal Sayılarla Toplama İşlemi	İki doğal sayının toplandığı işlemde verilmeyen toplananı belirler.	Analiz	0,53	0,77
14	Doğal Sayılarla Bölme İşlemi	En çok 20 nesneyi kalansız olarak 2, 3, 4 ve 5 gruba eşit olarak paylaştırarak her gruptaki nesne sayısını belirtir.	Uygulama	0,51	0,77
15	Kesirler	Bütün, yarım ve çeyrek arasındaki ilişkiyi açıklar	Kavrama	0,45	0,67

Ek-3**3. Sınıf Matematik Dersi Sayılar Öğrenme Alanı
Başarı Testi Belirtke Tablosu**

Madde No	Alt Öğrenme Alanı	Kazanım	Bilişsel Düzey	p_{jx}	r_{jx}
1	Doğal Sayılar	Üç basamaklı doğal sayıların basamak adlarını, basamaklarındaki rakamların basamak değerlerini belirtir.	Bilgi	0,75	0,80
2	Doğal Sayılar	1000'den küçük en çok beş doğal sayıyı, büyükten küçüğe veya küçükten büyüğe doğru sembol kullanarak sıralar.	Kavrama	0,84	0,80
3	Doğal Sayılar	Bir örüntüdeki ilişkiyi belirler ve örüntüyü genişletir.	Kavrama	0,80	0,78
4	Doğal Sayılarla Çarpma İşlemi	Çarpım tablosunu oluşturur.	Bilgi	0,79	0,73
5	Doğal Sayılarla Toplama İşlemi	Toplamları en çok üç basamaklı olan doğal sayılarla eldesiz ve eldeli toplama işlemini yapar.	Uygulama	0,74	0,67
6	Doğal Sayılar	Üç basamaklı doğal sayıları okur ve yazar.	Bilgi	0,71	0,66
7	Doğal Sayılarla Çıkarma İşlemi	En çok üç basamaklı doğal sayılarla çıkarma işlemini yapar.	Uygulama	0,68	0,83
8	Doğal Sayılarla Bölme İşlemi	İki basamaklı doğal sayıları bir basamaklı doğal sayılara böler.	Uygulama	0,64	0,89
9	Doğal Sayılarla Çarpma İşlemi	Eldeli çarpma işlemini yapar.	Uygulama	0,58	0,73
10	Kesirler	Paydası en çok iki basamaklı doğal sayı olan en çok üç kesri karşılaştırır ve sıralar.	Kavrama	0,57	0,35
11	Doğal Sayılarla Bölme İşlemi	Biri bölme olacak şekilde iki işlem gerektiren problemleri çözer.	Uygulama	0,54	0,90
12	Doğal	Doğal sayılarla çarpma	Uygulama	0,53	0,78

	Sayılarla Çarpma İşlemi	işlemini gerektiren problemleri çözer.			
13	Kesirler	Bir çokluğun belirtilen kesrin birimi kadarını belirler.	Uygulama	0,49	0,86
14	Doğal Sayılarla Toplama İşlemi	Doğal sayılarla toplama işlemini gerektiren problemleri çözer.	Uygulama	0,48	0,85
15	Doğal Sayılar	Tek ve çift doğal sayıları belirtir.	Bilgi	0,46	0,54
16	Doğal Sayılarla Çıkarma İşlemi	Doğal sayılarla toplama ve çıkarma işlemlerini gerektiren problemleri çözer.	Uygulama	0,42	0,94

Ek-4**4. Sınıf Matematik Dersi Sayılar Öğrenme Alanı
Başarı Testi Belirtke Tablosu**

Madde No	Alt Öğrenme Alanı	Kazanım	Bilişsel Düzey	p_{jx}	r_{jx}
1	Doğal Sayılar	Bir örüntüyü sayılarla ilişkilendirir ve eksik olan bölümü tamamlar.	Kavrama	0,88	0,54
2	Kesirler	Eşit paydalı en çok dört kesri, büyükten küçüğe veya küçükten büyüğe doğru sıralar.	Kavrama	0,81	0,33
3	Kesirler	Payı ve paydası en çok iki basamaklı doğal sayı olan kesirleri, kesrin birimlerinden elde ederek isimlendirir.	Analiz	0,78	0,76
4	Doğal Sayılarla Toplama İşlemi	En çok dört basamaklı doğal sayılarla toplama işlemini yapar.	Uygulama	0,78	0,57
5	Ondalık Kesirler	İki ondalık kesri karşılaştırarak aralarındaki ilişkiyi büyük, küçük veya eşit sembolüyle gösterir.	Kavrama	0,78	0,49
6	Kesirlerle Çıkarma İşlemi	Paydaları eşit kesirlerle çıkarma işlemini yapar.	Uygulama	0,69	0,77
7	Doğal Sayılar	En çok altı basamaklı doğal sayıları sıralar.	Kavrama	0,69	0,59
8	Kesirlerle Toplama İşlemi	Paydaları eşit kesirlerle toplama işlemini yapar.	Uygulama	0,67	0,77
9	Kesirlerle Çıkarma İşlemi	Kesirlerle toplama ve çıkarma işlemlerini gerektiren problemleri çözer.	Uygulama	0,67	0,75
10	Doğal Sayılarla Bölme İşlemi	Üç basamaklı doğal sayıları en çok iki basamaklı doğal sayılara böler.	Uygulama	0,61	0,86
11	Doğal Sayılarla Çıkarma İşlemi	Doğal sayılarla çıkarma işlemini gerektiren problemleri çözer.	Uygulama	0,60	0,79
12	Ondalık Kesirler	Ondalık kesirleri virgöl kullanarak yazar.	Kavrama	0,60	0,87
13	Doğal Sayılar	4, 5 ve 6 basamaklı doğal sayıların bölüklerini ve basamaklarını, basamaklarındaki	Bilgi	0,57	0,69

		rakamların basamak değerlerini belirtir.			
14	Ondalık Kesirler	Ondalık kesirlerin tam kısmını, kesir kısmını ve basamak adlarını belirtir.	Bilgi	0,56	0,79
15	Doğal Sayılar	4, 5 ve 6 basamaklı doğal sayıları okur ve yazar.	Bilgi	0,53	0,73
16	Doğal Sayılarla Çarpma İşlemi	Çarpımı en çok beş basamaklı doğal sayı olacak şekilde iki doğal sayıyla çarpma işlemini yapar.	Uygulama	0,49	0,68
17	Doğal Sayılarla Çıkarma İşlemi	En çok dört basamaklı doğal sayılarla çıkarma işlemini yapar.	Uygulama	0,48	0,68
18	Doğal Sayılarla Çarpma İşlemi	Doğal sayılarla çarpma işlemini gerektiren problemleri çözer.	Uygulama	0,47	0,64
19	Doğal Sayılarla Bölme İşlemi	İki adımlı işlemleri yapar.	Uygulama	0,54	0,74
20	Doğal Sayılarla Bölme İşlemi	Doğal sayılarla bölme işlemini gerektiren problemleri çözer.	Uygulama	0,48	0,81
21	Kesirler	Bir çokluğun belirtilen bir basit kesir kadarını belirler.	Uygulama	0,45	0,87
22	Doğal Sayılar	Doğal sayıları en yakın onluğa veya yüzlüğe yuvarlar.	Kavrama	0,44	0,73
23	Doğal Sayılarla Toplama İşlemi	Doğal sayılarla toplama işlemini gerektiren problemleri çözer.	Uygulama	0,40	0,68
24	Kesirler	Payları eşit, paydaları birbirinden farklı en çok dört kesri, büyükten küçüğe veya küçükten büyüğe doğru sıralar.	Kavrama	0,38	0,77

Ek-5

1-4. Sınıf Matematik Dersi Sayılar Öğrenme Alanı Başarı Testlerine Ait Örnek Sorular*

*Geliştirilen testlerin asıl formları, soruların yaygın ve amacının dışında kullanımının önlenmesi için bu çalışmada verilmemiştir. İstendiği takdirde araştırmacıya esra.fidan9@gmail.com adresinden ulaşarak testlerin alınıp kullanılması mümkündür.

1.SINIF TESTİNE AİT ÖRNEK SORU

Aşağıdaki kutunun içine 14 tane yuvarlak çiziniz. ($p_j = 0,92$; kolay)

Cevap: Kutunun içine çizilmiş 14 yuvarlak (veya 14 nesne)

2. SINIF TESTİNE AİT ÖRNEK SORU

Günler	Sayfa sayıları
1. gün	8
2. gün	12
3. gün	16
4. gün	20
5. gün	?
6. gün	28

Yandaki tablo Tuğçe'nin farklı günlerde okuduğu kitap sayfa sayılarını göstermektedir. Sayfa sayıları belli bir düzende arttığına göre Tuğçe 5. gün kaç sayfa kitap okumuştur? Yazınız. ($p_j = 0,72$; kolay)

Cevap: 24

3. SINIF TESTİNE AİT ÖRNEK SORU

Bir okul kütüphanesi için ilk gün 45, ikinci gün 33 kitap toplanmıştır. Bu kitaplar kütüphanedeki 6 rafa eşit olarak dizildiğinde her rafta kaç kitap olur? ($p_j = 0,54$; orta güçlükte)

Cevap: 13

4. SINIF TESTİNE AİT ÖRNEK SORU

Sezer bir kitabın pazartesi günü $\frac{1}{8}$ 'ini, Salı günü de $\frac{2}{8}$ 'sini okumuştur. Sezer kitabın kaçta kaçını okumuştur? ($p_j = 0,67$; kolay)

Cevap: $\frac{3}{8}$