

T.C.
BOLU ABANT İZZET BAYSAL ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANABİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME BİLİM DALI



ANGOFF VE DOĞRUDAN UZLAŞMA STANDART
BELİRLEME YÖNTEMLERİNİN KARŞILAŞTIRILMASI

YÜKSEK LİSANS TEZİ

GAMZE ÖZSOY DEMİREL

TEZ DANIŞMANI
Prof. Dr. Zekeriya NARTGÜN

BOLU, AĞUSTOS - 2023

KABUL VE ONAY SAYFASI

Gamze ÖZSOY DEMİREL tarafından hazırlanan “**ANGOFF VE DOĞRUDAN UZLAŞMA STANDART BELİRLEME YÖNTEMLERİNİN KARŞILAŞTIRILMASI**” adlı tez çalışması jürimiz tarafından Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı’nda Yüksek Lisans Tezi olarak oy birliği ile kabul edilmiştir. 27/07/2023

Jüri Üyeleri

İmza

Danışman
Prof. Dr. Zekeriya NARTGÜN
Bolu Abant İzzet Baysal Üniversitesi

.....

Üye
Doçent Dr. Sami PEKTAŞ
Niğde Ömer Halisdemir Üniversitesi

.....

Üye
Dr. Öğr. Üyesi Alperen YANDI
Bolu Abant İzzet Baysal Üniversitesi

.....

Lisansüstü Eğitim Enstitüsü Onayı

Prof. Dr. İbrahim KÜRTÜL
Lisansüstü Eğitim Enstitüsü Müdürü

ETİK BEYAN

Bolu Abant İzzet Baysal Üniversitesi, Lisansüstü Eğitim Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu bildirir,

aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Teze ilişkin 10/08/2023 tarihinde Turnitin adlı intihal tespit programından enstitü müdürlüğünce belirlenen filtrelemeler uygulanarak alınmış olan benzerlik raporuna göre, tezin benzerlik oranı % 16 olarak tespit edilmiştir.

Bu çalışma için İnsan Araştırmaları Etik Kurulundan 2023/133 sayısı ile etik izin alınmıştır.

Bu çalışma için Bolu İl Milli Eğitim Müdürlüğü'nden E-39307281-605.01-75832991 sayısı ile araştırma izni alınmıştır.

.....
GAMZE ÖZSOY DEMİREL

ÖZET

**ANGOFF VE DOĞRUDAN UZLAŞMA STANDART BELİRLEME
YÖNTEMLERİNİN KARŞILAŞTIRILMASI
YÜKSEK LİSANS TEZİ
GAMZE ÖZSOY DEMİREL
BOLU ABANT İZZET BAYSAL ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANABİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME BİLİM DALI
(TEZ DANIŞMANI: PROF. DR. ZEKERİYA NARTGÜN)**

BOLU, AĞUSTOS - 2023

xi + 59

Bu araştırmada Angoff ve Doğrudan Uzlaşma standart belirleme yöntemleri ürettikleri kesme puanları ve çalışılan uzmanların yöntemlere ilişkin görüşleri açısından karşılaştırılmıştır.

Araştırmanın öğrenci verileri ortaokul 5. Sınıf matematik dersi “Sayılar ve İşlemler” öğrenme alanına yönelik 24 maddelik bir başarı testinin 52 öğrenciye uygulanmasıyla elde edilmiştir. Yöntemlerin uygulanması ve değerlendirilmesi için 10 uzmanla çalışılmıştır. Uzmanların kararları sonucunda Angoff yöntemine dair kesme puanı 52,04, Doğrudan Uzlaşma yöntemine dair kesme puanı 49,17 olarak hesaplanmıştır. Angoff yöntemine göre öğrencilerin %52’si, Doğrudan Uzlaşma yöntemine göre öğrencilerin %58’i başarılı sayılmıştır. Yüzde değerleri arasındaki farklılık z-testi ile sorgulanmış ve farklılığın anlamlı olmadığı tespit edilmiştir. Yöntemlere göre öğrencilerin başarılı-başarısız olma durumları arasında, Cohen’in Kappa katsayısı 0,88, Spearman Brown Sıra farkları korelasyon katsayısı 0,89 hesaplanarak mükemmel uyum ve yüksek ilişki tespit edilmiştir. Yöntemlere göre minimum geçme puanları (MGP) arasındaki ilişki Pearson Momentler Çarpımı Korelasyon Katsayısı ile araştırılmış ve değerin 0,80 olduğu görülerek MGP’lerin pozitif yönde yüksek ilişkili olduğu sonucuna ulaşılmıştır. Uzmanların yöntemlere dair görüşleri betimsel olarak incelendiğinde Doğrudan Uzlaşma yöntemine yönelik daha yüksek olumlu görüş geliştirildiği sonucuna ulaşılmıştır.

ANAHTAR KELİMELER: Standart Belirleme, Kesme Puanı, Angoff Yöntemi, Doğrudan Uzlaşma Yöntemi

ABSTRACT

A COMPARISON OF ANGOFF AND DIRECT CONSENSUS STANDARD SETTING METHODS

MA THESIS

GAMZE ÖZSOY DEMİREL

BOLU ABANT İZZET BAYSAL UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DEPARTMENT OF EDUCATIONAL SCIENCES

(SUPERVISOR: PROF. DR. ZEKERİYA NARTGÜN)

BOLU, AUGUST 2023

xi + 59

In this study, Angoff and Direct Consensus standard-setting methods were compared in terms of the cut-off scores they produced and the opinions of the experts on the methods.

The student data of the study were obtained by applying a 24-item achievement test for the "Numbers and Operations" of 5th grade mathematics to 52 students. 10 experts were employed to implement and evaluate the methods. The cut-off point for the Angoff method was calculated as 52.04, Direct Consensus method was calculated as 49.17. According to the Angoff method 52%, according to the Direct Consensus method 58% of the students were considered successful. The difference between the percentage values was analyzed with the z-test and it was not significant. According to the methods, a perfect compatibility and a high correlation were determined by calculating Cohen's Kappa coefficient of 0.88 and Spearman's coefficient of 0.89 between students' successful-unsuccessful states. The relationship between the minimum passing scores (MPS) according to the methods was investigated with the Pearson's r, and the value was 0.80, and it was concluded that the MPSs were positively correlated. When the opinions of the experts on the methods were analyzed descriptively, it was concluded that a higher positive opinion was developed for the Direct Consensus method.

KEYWORDS: Standard setting, Cut-off Score, Angoff Method, Direct Consensus Method

İÇİNDEKİLER

	<u>Sayfa</u>
KABUL VE ONAY SAYFASI	iii
ETİK BEYAN.....	iv
ÖZET.....	v
ABSTRACT	vi
İÇİNDEKİLER	vii
TABLO LİSTESİ	ix
KISALTMA VE SEMBOLLER LİSTESİ	x
TEŞEKKÜR	xi
1. GİRİŞ.....	1
1.1 Problem Durumu	1
1.2 Problem Cümlesi	5
1.2.1 Alt Problemler	5
1.3 Araştırmanın Amacı.....	5
1.4 Araştırmanın Önemi	5
1.5 Sayıtlar.....	6
1.6 Sınırlılıklar.....	6
1.7 Tanımlar ve Kısaltmalar	7
2. ARAŞTIRMANIN KURAMSAL TEMELİ VE İLGİLİ	
ARAŞTIRMALAR.....	8
2.1 Kuramsal Temeller	8
2.1.1 Standart Belirleme	8
2.1.2 Angoff Yöntemi.....	9
2.1.3 Doğrudan Uzlaşma Yöntemi	11
2.2 İlgili Araştırmalar	14
2.2.1 Angoff Yöntemine Dayalı Araştırmalar	14
2.2.2 Doğrudan Uzlaşma Yöntemine Dayalı Araştırmalar	15
3. YÖNTEM.....	17
3.1 Araştırmanın Modeli.....	17
3.2 Çalışma Grubu	17
3.3 Veri Toplama Araçları.....	18
3.3.1 Matematik Başarı Testi.....	18
3.3.2 Uzman Görüşleri Formu	23
3.3.2.1 Angoff Yöntemi Uygulama Süreci	23
3.3.2.2 Doğrudan Uzlaşma Yöntemi Uygulama Süreci	24
3.3.3 Standart Belirleme Yöntem ve Süreçlerini Değerlendirme Formu ..	26
3.3.4 Verilerin Çözümlemesinde Kullanılan İstatistiksel Teknikler	26
4. BULGULAR	29
4.1 Birinci Alt Probleme Ait Bulgular.....	29
4.2 İkinci Alt Probleme Ait Bulgular	31

4.3 Üçüncü Alt Probleme Ait Bulgular	31
4.4 Dördüncü Alt Probleme Ait Bulgular	33
5. TARTIŞMA	37
6. SONUÇLAR VE ÖNERİLER.....	39
6.1 Sonuçlar	39
6.1.1 Birinci Alt Problemin Çözümlemesiyle Elde Edilen Sonuçlar	39
6.1.2 İkinci Alt Problemin Çözümlemesiyle Elde Edilen Sonuçlar	39
6.1.3 Üçüncü Alt Problemin Çözümlemesiyle Elde Edilen Sonuçlar	39
6.1.4 Dördüncü Alt Problemin Çözümlemesiyle Elde Edilen Sonuçlar..	39
6.2 Öneriler	40
6.2.1 Araştırma sonuçlarından çıkarılan öneriler	40
6.2.2 Araştırmacılara yönelik öneriler	40
7. KAYNAKLAR.....	42
8. EKLER.....	45
EK 1 Matematik Başarı Testi	45
EK 2 Uzman Görüşleri Formu	49
EK 3 Standart Belirleme Yöntem ve Süreçlerini Değerlendirme Formu.....	57
EK 4 Yöntemlere Göre Uzman Kararları	59

TABLO LİSTESİ

Sayfa

Tablo 1 MEB 4.Sınıf ve Ortaokul Seviyesi Puan Sistemi.....	2
Tablo 2 MEB Ortaöğretim Seviyesi Puan Sistemi.....	3
Tablo 3 Angoff Yöntemi Uygulama Örneği	10
Tablo 4 Doğrudan Uzlaşma Yöntemi Uygulama Örneği.....	12
Tablo 5 Angoff ve Doğrudan Uzlaşma Yöntemlerinin Özellikleri.....	13
Tablo 6 Deneme Testine Ait Betimsel İstatistikler	19
Tablo 7 Deneme Testi Sonuçlarına Göre Madde İstatistikleri	20
Tablo 8 Nihai Form Madde Analizleri	22
Tablo 9 Nihai Form Betimsel İstatistikler	22
Tablo 10 Karne notları ve Test Puanlarına İlişkin Korelasyon Katsayısı	23
Tablo 11 Angoff Yöntemi Toplantı Akışı.....	24
Tablo 12 Doğrudan Uzlaşma Yöntemi Toplantı Akışı.....	25
Tablo 13 Bağımlı İki Oran Testi İçin Frekans ve Oranlar.....	27
Tablo 14 Cohen'in Kappa Katsayısı Değer Aralıkları	28
Tablo 15 Yöntemlere Ait Uzman MGP'leri ve Testin Kesme Puanları.....	29
Tablo 16 Yöntemlere Göre Başarılı Sayılan Öğrenci Sayısı ve Yüzdesi (n=52)30	
Tablo 17 Başarılı Sayılan Öğrenci Sayısı, Yüzdesi ve z Değerleri.....	30
Tablo 18 Yöntemler Arasındaki Korelasyon Katsayıları	31
Tablo 19 MGP'lere Ait Betimsel İstatistik ve Normallik Testi Sonuçları.....	32
Tablo 20 MGP'ler Arasındaki Pearson Korelasyon Katsayıları.....	32
Tablo 21 MGP Ortalamaları Arasındaki Farkın Testi.....	33
Tablo 22 Angoff ve Doğrudan Uzlaşma Yöntemlerine İlişkin Uzman Görüşleri33	

KISALTMA VE SEMBOLLER LİSTESİ

MEB	: Milli Eğitim Bakanlığı
MGP	: Minimum Geçme Puanı
p	: Anlamlılık Düzeyi
$\bar{x}$	: Aritmetik Ortalama
pj	: Madde Güçlük İndeksi
rjx	: Madde Ayırt Edicilik İndeksi



TEŞEKKÜR

Yüksek lisans sürecimde her zaman yanımda olduğunu hissettiren ve yaşadığım her olumsuzlukta bana umut veren, çok saygı ve minnet duyduğum değerli danışman hocam Prof. Dr. Zekeriya NARTGÜN'e çok teşekkür ederim.

Sahip olduğum tüm güzellikleri borçlu olduğum annem Nurhayat ÖZSOY'a ve babam Salih ÖZSOY'a çok teşekkür ederim.

Tez sürecimde bana ilham veren ve desteğini benden hiç esirgemeyen ablam Aysel DEMİREL'e çok teşekkür ederim.

Yüreğimin en neşeli en yaramaz parçası olan ve mutsuz olmama asla izin vermeyen kızım Koza İda'ya çok teşekkür ederim.

Yol arkadaşım Suat DEMİREL'e sonsuz teşekkür ederim.

1. GİRİŞ

Bu bölümde araştırmanın problem durumu, amacı, önemi, sayıtları, sınırlılıkları ve tanımlarına yönelik açıklamalara yer verilmiştir.

1.1 Problem Durumu

Bilim ve teknoloji alanında yaşadığımız hızlı değişimler sosyal yaşamı daha karmaşık bir hale getirmektedir. Toplumların kalkınmak ve kaliteli yaşamak için bu değişime uyum sağlaması gerekmektedir. Bu durum kamu ve özel sektör kurum ve kuruluşlarının çağdaş bilgiyle donatılmış kalifiye insan gücüne ihtiyaç duymasına sebep olmuştur. Sosyal yaşamda hissedilen bu ihtiyaç eğitimden beklentileri arttırmış, eğitimin kalkınmada üstleneceği rolünü karmaşıktırmıştır. Bu durum eğitimi daha çağdaş daha yaygın ve daha aktif olmaya zorlamaktadır (Fidan ve Baykul, 1994).

Ertürk, (2013) eğitimin ne olduğu sorusuna “bireyin davranışında kendi yaşantısı yoluyla ve kasıtlı olarak istendik yönde bir değişim oluşturma süreci” tanımını yapmıştır. Eğitim süreci sonunda birey elde edilmesi amaçlanan davranışları yeterli düzeyde gösterebilir, yetersiz düzeyde gösterebilir, hiç göstermeyebilir hatta istenmeyen yeni bir davranış bile gösterebilir. Bu olasılıklar eğitim sürecinin çıktılarını kontrol etmeyi gerektirir (Baykul ve Tertemiz, 2004). Eğitimin verimliliğini ve etkililiğini kontrol etmek ölçme ve değerlendirme ile mümkündür. Toplumun her kesimi eğitimcilerden yapılan eğitimin etkili ve verimli olup olmadığı konusunda bilgi talep etmektedir. Bu talebin artması ölçme ve değerlendirme sürecini gün geçtikçe daha önemli hale getirmektedir (Mehrens ve Lehmann, 1973). Ölçme en genel tanımıyla “Bir niteliğin gözlenip gözlem sonuçlarının sayı veya sembollerle ifade edilmesini kapsayan işlemidir” (Baykul, 2021). Ölçme işlemi sonrasında yapılan değerlendirme ise ölçme sonuçlarının bir ölçütle kıyaslanması ve karar verilmesi işlemidir (Turgut, 1993). Ölçme sonuçlarının işevuruk bir anlam ifade etmesi için ölçüt ile karşılaştırılması gerekmektedir. Standart belirleme bu ölçütün belirlenme işlemidir. (Özçelik, 1992). Ölçme aracının güvenilir ve geçerli olması, güvenilir bir ölçüt seçilmesi ve ölçütün doğru kullanılması değerlendirme sonucunda verilen kararın doğru olması için gerekli koşullardır. Bu nedenle değerlendirmelerde ölçüt yani standart belirleme son derece önem arz etmektedir (Çetin, 2011; Kubiszyn ve Borich, 2007).

Standart belirleme sürecinde belli performans seviyeleri oluşturulur. Performans seviyeleri iki veya daha fazla sayıda olabilir. Herhangi bir performans seviyesi için yeterli olan öğrenci ile o performans seviyesine yetersiz kalan öğrenci arasında oluşan sınır kesme puanı ya da geçme puanı olarak adlandırılır. Kesme puanının altında kalan öğrenciler belirlenen üst performans seviyesi için yetersiz kabul edilir ve bir alttaki seviyeye yerleştirilir (Hambleton, 1998). Kesme puanı testin uygulandığı grubun başarısından etkileniyorsa bağıl standart, grubun başarısından etkilenmiyorsa mutlak standart olarak kabul edilir. Mutlak bir standart belirlemek bağıl standardı belirlemekten daha karmaşık ve uzun bir süreçtir (Livinston ve Zieky, 1982). Performans seviyelerinin sayısı eğitim kurumlarının ya da test merkezlerinin testi uygulama amacına bağlı olarak değişebilir. Lisans, sertifika, geçti-kaldı, başarılı-başarısız kararları verilecekse iki seviye belirlenir. Bu iki performans seviyesini puan cetvelinde birbirinden ayıracak olan bir kesme puanı belirlenir (Cizek, 2012).

Yurtdışında birçok ülkede büyük eğitim kurumları kendi standartlarını belirlemektedir. Örneğin Amerika Birleşik Devletlerinde bazı eyaletlerde İleri Düzey – Yeterli – Temel olmak üzere 3 adet performans seviyesi belirlenirken bazılarında ise İleri Düzey – Yeterli – Temel – Temelin Altında olmak üzere 4 adet performans seviyesi belirlenmektedir (Cizek ve Bunch, 2007). Ülkemizde ise Millî Eğitim Bakanlığına bağlı resmî ve özel ilköğretim kurumlarının 4. Sınıf ve ortaokul seviyesindeki 100'lük puan sistemi aşağıda gösterilmiştir. Tablo 1’de, üst performan seviyesinin “Başarılı”, alt performans seviyesinin “Başarısız” olarak belirlendiği ve bu iki seviyeyi birbirinden ayıran kesme puanının 45 olduğu görülmektedir (Meb, 2014).

Tablo 1 MEB 4.Sınıf ve Ortaokul Seviyesi Puan Sistemi

Puan Aralığı	Performans Seviyesi (Karar)
0-44,99	Başarısız
45-100	Başarılı

Millî Eğitim Bakanlığına bağlı resmî ve özel ortaöğretim kurumlarında 100'lük puan sistemi aşağıda gösterilmiştir. Tablo 2’de, 5 performans seviyesi

oluşturulduğu ve bu seviyelerin 50, 60, 70 ve 85 olmak üzere 4 kesme puanı ile sınırlandırıldığı görülmektedir (Meb, 2013).

Tablo 2 MEB Ortaöğretim Seviyesi Puan Sistemi

Puan Aralığı	Performans Seviyesi (Karar)
0-49,99	Geçmez
50,00-59,99	Geçer
60,00-69,99	Orta
70,00-84,99	İyi
85,00-100	Pekiyi

Literatürde standart belirlemek için kullanılabilecek yöntemi açıklayan birçok alternatif yaklaşım bulunmaktadır. Seçilen bir yöntemin uygulanmasıyla elde edilen standardın kıyaslanacağı başka bir standart ortada olmadığı için bu yaklaşımlar doğrulukları açısından değerlendirilemez. Yöntemler tutarlılık, uygunluk, makullük açısından değerlendirilebilir (Kane, 1998). Aynı test üzerinde farklı yöntemler kullanıldığında farklı sonuçlar verebilir (Irwin, 2007). Kesme puanındaki en ufak bir değişim öğrenciler hakkında geçme-kalma, hak etme-hak etmeme kararları verilen bir testte sonuçları büyük ölçüde değiştirebilmektedir (Kane, 1998). Literatürde kesme puanını belirlemek için aynı test için mümkünse birden fazla yöntemin kullanılması önerilmektedir. Birden fazla yöntem kullanıldığında benzer kesme puanları elde ediliyorsa bu kesme puanının savunulabilirliğini yükseltir. Yöntemlerden elde edilen kesme puanları birbirinden farklı ise dikkat edilmesi gereken bir dizi sonuç çıkmış demektir (Irwin, 2007).

Kesme puanlarını belirlemek için kullanılan yöntemler iki kategoride ele alınabilir. Birinci kategori test merkezli yöntemlerdir (Jeager, 1989; Beuk, 1984). Test merkezli yöntemlerde uzmanlar test maddelerini inceleyerek bu maddelerin gerektirdiği performans kriterlerine göre kesme puanına karar verir (Irwin, 2007). İkinci kategori ise öğrenci merkezli yöntemlerdir (Jeager, 1989; Beuk, 1984). Bu yöntemlerde ise uzmanlar teste katılan öğrencileri inceleyerek ve öğrenciler hakkında yargıda bulunarak kesme puanına karar verir. Öğrenci merkezli yöntemlerde uzmanlar öğrencilerin test puanı hakkında önceden bilgi sahibi olmazlar. Öğrencilerin bilgi, beceri ve yeteneklerine göre yargıda bulundukları için

öğrenci merkezli yöntemlerde uzmanların teste katılan öğrenci grubuna aşina olması gerekmektedir (Irwin, 2007).

Her standart belirleme yönteminin kendi içerisinde avantaj ve dezavantajları bulunmaktadır (Çetin, 2011). Güvenilir kesme puanları üreten bir yöntem uzmanlar tarafından kullanışlı bulunmayabilir ya da elde edilen kesme puanları testin amacına uygun olmayabilir. Eğitim kurumları ve test merkezleri, kesme puanını hesaplamak üzere seçtiği yöntemin testin amacına hizmet etme derecesini bilme ihtiyacı duymaktadır. Seçilen yöntemle hesaplanan kesme puanının öğrencinin performans yeterliliği hakkında verilecek kararı değiştirebilecek olması eğitimcilerin yöntemler hakkında bilgi sahibi olmasını ve seçim kararını titizlikle vermesini gerektirmektedir. Seçilen yöntemle elde edilen kesme puanının uygunluğu, kararlılığı ve kullanışlılığı testin amacına, testin yapısına, ölçülmek istenen özelliğin yapısına, uzmanların alan yeterliğine ve uzmanların bilişsel özelliklerine göre değişkenlik gösterebilir. Bu nedenle yöntem tercihi yapılırken literatürdeki yöntemlerin araştırılması, incelenmesi ve mevcut duruma en uygun olan yöntemin seçilmesi gerekmektedir.

Ulusal literature bakıldığında standart belirleme yöntemlerinden biri olan Doğrudan Uzlaşma yönteminin araştırılmamış bir yöntem olarak kalması ülkemizdeki test merkezlerinin ve özel eğitim kurumlarının bu yöntemle ilgili bilgisini sınırlandırmaktadır. Doğrudan Uzlaşma yöntemi uluslararası literatüre kullanışlılığı bakımından avantajlı görülen yeni bir yöntem olarak tanıtılmıştır. Eğitimin genel amaçları ve stratejileri ülkeden ülkeye değişkenlik gösterebildiği için Doğrudan Uzlaşma yöntemi ülkemizdeki eğitim stratejilerine uygun testlerle ve eğitim sistemimizde çalışılan uzmanlarla denenmeye muhtaç bir yöntem olmuştur. Bu araştırmada özel eğitim kurumlarına ve test merkezlerine bir çok araştırmaya tabi tutulmuş Angoff yönteminin bilinirliğinden yola çıkarak Doğrudan Uzlaşma yöntemini tanıtmak amacıyla Angoff ve Doğrudan Uzlaşma yöntemleriyle elde edilen kesme puanları ve uzmanların yöntemlere dair değerlendirmeleri karşılaştırılmıştır.

1.2 Problem Cümlesi

Angoff ve Doğrudan Uzlaşma standart belirleme yöntemleriyle belirlenen kesme puanları arasında nasıl bir ilişki vardır?

1.2.1 Alt Problemler

Araştırmanın problem cümlesine yanıt bulmak amacıyla aşağıda verilen 4 alt problem oluşturulmuştur.

1. Yöntemlerin ürettiği kesme puanlarına göre üst performans seviyesine yerleşerek başarılı sayılan öğrenci yüzdeleri arasında anlamlı bir farklılık var mıdır?
2. Yöntemler arasında öğrencilerin başarılı-başarısız olarak sınıflandırılmaları açısından uyum var mıdır?
3. Yöntemler arasında minimum geçme puanları (MGP) açısından uyum var mıdır?
4. Angoff ve Doğrudan Uzlaşma standart belirleme sürecine katılan uzmanların, yöntemlere ilişkin değerlendirmeleri nasıldır?

1.3 Araştırmanın Amacı

Bu araştırmanın amacı Angoff ve Doğrudan Uzlaşma yöntemleri ile elde edilen kesme puanlarını karşılaştırmak, bu kesme puanlarıyla yapılan değerlendirme sonuçları arasındaki farkı veya uyumu ortaya koymak ve yöntemlerin uygulama sürecini uzmanların görüşleriyle açıklamaktır.

1.4 Araştırmanın Önemi

Bu araştırmaya konu olan yöntemlerden Angoff yöntemi eğitim testlerinde standart belirlemek için sıklıkla başvurulmuş en popüler yöntemdir (Kane, 1994; Mehrens, 1995; Shen, 2001). Doğrudan Uzlaşma yöntemi ise Angoff yöntemine getirilen uzmanlara bilişsel zorluklar yüklediği eleştirileri dikkate alınarak geliştirilen ve Angoff yönteminin uygulamasında gereken süreyi kısaltması planlanan bir yöntem olmasına rağmen konu olduğu araştırma sayısının yetersiz olduğu görülmektedir (Sireci, Hambleton ve Pitoniak, 2004; Akt. Pitoniak, 2003). Araştırmacı tarafından yapılan ulusal literatür taramasında Doğrudan Uzlaşma yöntemi ile ilgili bir çalışmaya rastlanmamıştır. Uluslararası literatür taramasında ise Doğrudan Uzlaşma yöntemi ile ilgili yapılan 4 çalışmaya rastlanmıştır.

Bu çerçevede bu araştırmanın çok sayıda araştırmaya tabi tutulan ve bilinirliği yüksek Angoff yöntemi ile Doğrudan Uzlaşma yöntemini karşılaştırarak ülkemizdeki özel eğitim kurumlarına ve test merkezlerine, Doğrudan Uzlaşma yöntemini tanıttacağı, bu yöntemin mevcut eğitim sistemimizdeki kullanılabilirliği ve kesme puanlarına etkisi hakkında bilgi sağlayacağı düşünülmektedir. Diğer taraftan, araştırma, Doğrudan Uzlaşma yönteminin prosedürlerinin Angoff yönteminin prosedürlerinde görülen uzmanların bilişsel açıdan zorluk yaşaması dezavantajını telafi edip etmediği yönünde bilgi sağlaması bakımından önemli görülmektedir.

1.5 Sayıtlar

Kullanılan matematik testinin öğrencilerin gerçek performansını yansıttığı varsayılmaktadır.

Uzmanların kesme puanlarını hesaplarken yöntemleri doğru ve hassas bir şekilde kullandığı varsayılmaktadır.

1.6 Sınırlılıklar

Araştırma prosedürlerindeki benzerlikler dikkate alınarak karşılaştırılması yapılan Angoff ve Doğrudan Uzlaşma standart belirleme yöntemleri ile sınırlandırılmıştır.

Araştırma ortaokul 5. Sınıf seviyesinde öğrenim gören 101 öğrenci ile sınırlandırılmıştır.

Araştırma araştırmacının uzmanlık alanı dikkate alınarak matematik dersi ile sınırlandırılmıştır.

Araştırma süresi 2022-2023 eğitim öğretim yılı ile sınırlandırılmıştır.

Araştırma Milli Eğitim Bakanlığı 5. Sınıf Matematik Öğretim Programına göre 2022-2023 eğitim öğretim yılı güz dönemini kapsayan “Sayılar ve İşlemler” öğrenme alanı ile sınırlandırılmıştır.

Araştırma, Angoff yönteminin çoktan seçmeli maddelerden oluşan testlerde kullanılmasının daha uygun olduğu görüşünden yola çıkılarak, araştırmacı tarafından hazırlanan çoktan seçmeli 24 maddeden oluşan başarı testi ile sınırlandırılmıştır.

Arařtırma 10 uzmanın kararları ile sınırlandırılmıřtır.

1.7 Tanımlar ve Kısaltmalar

Kesme Noktası: Uygulanan bir testte belli bir performans seviyesi için yeterli olan öğrenci ile yeterli olmayan öğrenciyi ayırt eden sınır performans noktası.

Uzman: Angoff ve Doğrudan Uzlaşma yöntemleri ile ilgili eğitim alan ve bu yöntemlerle standart belirleme süreçlerine katılan öğretmenler.

Minimum geçme puanı (MGP): Bir uzmanın bireysel kararlarına göre hesaplanan öğrencilerin testte başarılı sayılmaları için almaları gereken en düşük puan.

Minimum yeterlik seviyesi: Testte başarılı olabilmek için öğrencinin sahip olması gereken en düşük yeterlik sınırı.

2. ARAŞTIRMANIN KURAMSAL TEMELİ VE İLGİLİ ARAŞTIRMALAR

2.1 Kuramsal Temeller

Bu bölümde araştırmanın dayandığı kuramsal temeller verilmiştir.

2.1.1 Standart Belirleme

Standart belirleme, bir ölçekten alınabilecek puanları belli performans seviyelerine ayırma ve bu seviyelerin gerektirdiği performansa göre tanımlamasını yapma işlemidir. Performans seviyelerinin sınırları kesme puanları ile belirlenir (Zwick, Şentürk ve Wang, 2001). Standart belirlemede kesme puanları belirlenirken sistemli kurallar ya da tanımlanmış bir sürecin izlenmesi gerekmektedir (Cizek, 2001). Standart belirlenirken her ne kadar kurallar ve süreç tanımlanmış olsa da ölçme işleminin politikasını belirleyen kurum ve kuruluşların, testi geliştiren kişi ya da merkezlerin ve ölçme uzmanlarının kararları da işe koşulmalıdır. Bu koşullarda oluşturulan kesme puanı daha kullanışlıdır ve daha güçlüdür (Bejar, 2008). Standartı belirlemek nicelikten çok niteliksel bir karardır ve nitelikli bir kesme puanı belirlemek uzman görüşü gerektirir (Wright & Grosse, 1993). Nitelikli bir kesme puanı elde edebilmek için görüşüne başvuru uzmanların kararları tutarlı ve güvenilir olmalıdır (Pitoniak, 2003). Eğer kesme puanı belirleme sürecinde hatalar veya eksiklikler olmuşsa, teste girenler hakkında yapılan değerlendirmelerin ve verilen kararların doğruluğu tartışılabilir. Bu yüzden standart belirleme test geliştirme sürecinin kilit noktalarından biridir (Bejar, 2008)

Hambleton, (2013) standart belirleme sürecinde yapılması gereken iş ve işlemleri şu şekilde sıralamıştır:

1. Testin uygulandığı grubu temsil edebilecek büyüklükte bir uzman grubu oluşturulması
2. Kesme puanını hesaplamak için kullanılacak yöntemin seçilmesi, ihtiyaç duyulan materyallerin hazırlanması ve toplantı programının netleştirilmesi
3. Performans seviyelerinin açıklamalarının belirlenmesi
4. Uzman grubunun seçilen yöntemle ilgili eğitilmesi
5. Testin maddelerinin performans beklentilerinin derlenmesi (Örneğin uzmanlar bir performans seviyesine yerleşen öğrenciden beklenen performansı tahmin edebilmelidir.)

6. Uzmanların kendi fikirlerini açıklama ve tartışma olanağı bulabileceği bir açık oturum düzenlenmesi ve istatistik verilerinin (örneğin, madde zorluk indeksi, madde karakteristik eğrileri, madde ayırt edicilik indeksleri, çeldiricilerin analizi) elde edilerek bu verilerin göz önünde bulundurulması

7. Uzmanların değerlendirme sonuçlarının tekrar derlenmesi ve gerekirse geri bildirimler sağlanması veya tartışılması

8. Uzmanların nihai kesme puanının elde edilmesi

9. Elde edilen kesme puanı ile yapılan değerlendirmenin sonuç verilerinin uzman grubuna açıklanması

10. Standart belirleme sürecinin ve elde edilen standardın son kez değerlendirilmesi ve nihai sonuca ulaşılması

11. Standart belirleme sürecinin ve elde edilen standardın geçerlik kanıtlarının belgelendirilmesi

Hambleton, (2013) tarafından standart belirleme sürecinin ikinci adımına yerleştirilen yöntem seçimi son derece önemlidir. Literatür incelendiğinde standart belirleme yöntemine ilişkin bir çok yaklaşım olduğu görülmektedir.

Akyıldız, (2015) belli başlı standart belirleme yöntemlerinin sınıflandırmasını şu şekilde yapmıştır:

Test merkezli yöntemler:

1. Angoff Yöntemi
2. Doğrudan Uzlaşma Yöntemi
3. İşaretleme (Bookmark) Yöntemi
4. Evet/Hayır (Yes/No) Yöntemi
5. Ebel Yöntemi
6. Nedelsky Yöntemi

Öğrenci Merkezli Yöntemler:

1. Sınır Grup Yöntemi
2. Zıt Gruplar Yöntemi
3. İş Gövdesi (Body of Work) Yöntemi

2.1.2 Angoff Yöntemi

Literatüre William Angoff tarafından 1971 yılında kazandırılmış bir yöntemdir. Bu yöntemde alan uzmanlarından minimum yeterlik seviyesindeki bir

öğrencinin testteki bir maddeyi doğru cevaplama olasılığını yüzde olarak tahmin etmesi beklenir. Maddeleri tek tek inceleyen uzmanlar her madde için bir olasılık tahmininde bulunur.

Uzmanlar minimum yeterlik seviyesinde bulunan bir öğrenciyle ilgili olasılık kararı verme konusunda zorluk yaşayabilir. Bu durumda minimum yeterlik seviyesinde olan 100 öğrenciden kaçının soruya doğru yanıt verebileceği sorularak düşünce sistemi değiştirilebilir (Livingston ve Zieky, 1982). Burada dikkat edilmesi gereken diğer bir nokta da test maddelerinin doğru yanıtlarının da uzmanlara verilir verilmeyeceğidir. Livingston ve Zieky, (1982) doğru yanıtların verilmesinin uzmanların maddeleri daha kolay algılamasına yol açabileceğini, doğru yanıtların verilmemesinin ise uzmanlarda maddeleri doğru yanıtlama baskısı oluşturabileceğini ve zaman kaybına yol açacağını ileri sürmüşlerdir. Bunun çözümünün ise testin önceden verilerek doğru yanıtları bulmaları için uzmanlara zaman tanımak olduğunu belirtmişlerdir.

Testi inceleyerek doğru yanıtları belirleyen uzmanlar daha sonrasında öğrenci sayısını göz önünde bulundurarak maddeleri tek tek inceler ve her madde için 100 öğrenciden kaçının o maddeye doğru yanıt verebileceği yönündeki kararını belirtir. Bu karar yüzde değeri olarak kabul edilir. Her uzmanın testteki maddeler için bireysel olarak verdiği yüzde değerleri toplanır ve minimum geçme puanı belirlenir. Minimum geçme puanları uzmanların bireysel olarak belirledikleri kesme puanını oluşturur. Uzmanların belirledikleri kesme puanlarının ortalaması alınarak testin nihai kesme puanı belirlenmiş olur. Tablo 3'te 5 maddeden oluşan 100 puanlık bir testin kesme puanı Angoff yöntemiyle hesaplanmıştır.

Tablo 3 Angoff Yöntemi Uygulama Örneği

Uzmanlar	Madde-1	Madde-2	Madde-3	Madde-4	Madde-5	MGP
Uzman-1	60	80	50	55	90	67
Uzman-2	45	70	35	60	75	57
Uzman-3	65	80	40	70	70	65
Uzman-4	50	70	50	55	80	61
Uzman-5	45	75	55	60	65	60
Kesme Puanı						62

Tablo 3’te uzmanların “Minimum yeterlik seviyesinde olan bir öğrenci bu maddeyi yüzde kaç olasılıkla doğru yanıtlar?” ya da “Minimum yeterlik seviyesinde olan 100 öğrenciden kaç tanesi bu maddeye doğru yanıt verebilir?” sorularından birine verdiği yanıtlar görülmektedir. Uzman-1 maddeler için olasılık değerlerini sırasıyla %60, %80, %50, %55 ve %90 olarak belirlemiştir. Bu olasılık değerlerinin aritmetik ortalaması alındığında Uzman-1’in kişisel kesme puanı olan minimum geçme puanı (MGP) 67 bulunur. Tüm uzmanların MGP’leri hesaplandıktan sonra MGP’lerin ortalaması alındığında testin nihai kesme puanına ulaşılır.

Angoff yöntemi psikometrik özellikler ve uygulanabilirlik açısından diğer standart belirleme yöntemlerine göre daha dengeli ve doğru sonuçlar vermektedir (Berk, 1986). Angoff yönteminin veri toplaması, veri analizi ve açıklaması diğer yöntemlere göre daha basittir (Mills and Melican, 1988). Angoff yönteminin çok güçlü araştırmalara tabi tutulduğu ve kullanılan en yaygın yöntem olduğu gözlenmiştir (Pitoniak, 2002). Buna karşılık Angoff yöntemi, minimum yeterlik seviyesindeki öğrenci grubunu kavramsallaştırarak maddeleri doğru cevaplama olasılıklarını tahmin etmenin uzmanlara büyük bir bilişsel yük getirdiği gerekçesiyle eleştirilmiştir (Shepard, Glaser, Linn ve Bohmstedt, 1993; Buckendahl, Smith, Impara ve Plake, 2002). Uzmanlarla yapılan bir odak grup çalışmasında da uzmanların minimum yeterlik seviyesindeki öğrenci grubunu kavramsallaştırmaktan kaçındığı ve bu görevi bilişsel olarak zor ve karmaşık buldukları sonucuna ulaşılmıştır (Hein ve Skaggs, 2009). Angoff yöntemine getirilen diğer bir eleştiri de uzmanların toplantı sürecinde hesaplamalara hakim olmaması durumudur. Uzmanlardan alınan kararlar nihai karar olarak kabul edilir ve kesme puanını hesaplama işlemini uzmanlara açıklama kararı toplantı yöneticisinin inisiyatifindedir. Bu keyfilik uzmanların kesme puanı üzerindeki etkisini kısıtlar (Pitoniak, 2003).

2.1.3 Doğrudan Uzlaşma Yöntemi

Literatüre Sireci, Hambleton, Huff ve Jodoin tarafından 2000 yılında kazandırılan bir yöntemdir (Keuning, Straat ve Feskens, 2017). Yöntemin isminde bulunan “doğrudan” kelimesi ile panelistlerin gerçek sınav ölçeğiyle çalışması ve nihai kesme puanı üzerinde doğrudan etkiye sahip olması kastedilirken “uzlaşma” kelimesi ile testin kesme puanı üzerinde uzmanların fikir birliğine varması

kastedilmektedir (Pitoniak, 2002). Bu yöntemde kesme puanı belirlenecek test incelenerek bölümlere ayrılır ve içerik olarak testin tamamından daha homojen olan alt testler oluşturulur. Uzmanlar testi incelerken madde madde ilerlemek yerine bu alt testleri inceleyerek minimum yeterli seviyesindeki bir öğrencinin bu alt testte başarılı kabul edilmesi için kaç maddeyi doğru cevaplaması gerektiğini belirtirler. Uzmanların her bir alt testte belirlediği puanlar toplanarak bireysel olarak tahmin ettikleri kesme puanlarına ulaşılır. Belirlenen bu kesme puanlarının farklılaşması durumunda toplantı yapılarak fikir birliğine ulaşılması amaçlanır (Cizek ve Bunch, 2007). Doğrudan Uzlaşma yönteminin üstünlüklerinden biri de uzmanlara uygulama sürecinin içerisinde geri dönüt verilmesi, yapılan hesaplama sonuçlarının açıklanması ve diğer uzmanların verdiği puanlar hakkında bilgi verilmesidir (Pitoniak, 2003). Toplantı yöneticileri uzmanları aldıkları kararlar konusunda tekrar düşünmesi ve revize etmesi için yönlendirebilir. Süreç sonunda fikir birliği sağlanırsa nihai kesme puanı belirlenmiş olur. Fikir birliği sağlanamazsa bireysel kesme puanlarının ortalaması alınır (Sireci, Hambleton, Huff ve Jodoin, 2000; Akt. Pitoniak, 2003). Tablo 4’te her biri 5 maddeden oluşan 4 alt-teste bölünmüş 100 puanlık bir testin kesme puanı Doğrudan Uzlaşma yöntemiyle hesaplanmıştır.

Tablo 4 Doğrudan Uzlaşma Yöntemi Uygulama Örneği

Alt Testler	Uzman-1	Uzman-2	Uzman-3	Uzman-4	Uzman-5	Ortalama
Alt Test-1	2	3	1	4	3	2,6
Alt Test-2	4	3	2	3	4	3,2
Alt Test-3	4	5	5	4	4	4,4
Alt Test-4	2	2	2	2	2	2
Toplam	12	13	10	13	13	12,2
MGP	60	65	50	65	65	61
Kesme Puanı						61

Tablo 4’te uzmanların her alt test için “Minimum yeterlik seviyesindeki bir öğrencinin bu alt testte başarılı kabul edilmesi için maddelerden kaçını doğru cevaplama gerekir?” sorusuna verdiği yanıtlar görülmektedir. Uzman-1 minimum yeterlik seviyesindeki bir öğrencinin birinci alt testten 2 maddeyi, ikinci alt testten 4 maddeyi, üçüncü alt testten 4 maddeyi, dördüncü alt testten 2 maddeyi doğru cevaplama gerektiğini belirtmiştir. Uzman-1’in verdiği kararlara göre minimum yeterlik seviyesindeki bir öğrencinin 20 maddelik ana testten başarılı sayılması için toplamda 12 maddeye doğru cevap vermesi gerektiği sonucuna ulaşılmıştır. Her madde 5 puan değerinde olduğu için Uzman-1’e ait MGP değeri 60 olarak hesaplanır. Doğrudan Uzlaşma yönteminde hesaplanan MGP’ler birbirinden farklılaşırsa fikir birliğine varmak için toplantı düzenlenir. Fikir birliği sağlanamazsa Tablo 4’teki gibi MGP’lerin ortalaması alınarak testin nihai kesme puanı belirlenmiş olur.

Doğrudan Uzlaşma yöntemi minimum yeterlik seviyesindeki öğrenci grubunu kavramsallaştırmayı gerektirmez. Bu durum uzmanların üzerindeki bilişsel yükü kaldırır. Uzmanlar nihai kesme puanını hesaplamada doğrudan etkiye sahiptir (Sireci, Hambleton ve Pitoniak, 2004; Akt. Cizek ve Bunch, 2007). Uzmanlar testi değerlendirirken madde madde ilerlemek yerine bölüm bölüm ilerlediği için bu yöntemin bir avantajı da diğer yöntemlere göre daha hızlı olmasıdır (Pitoniak, 2002).

Pitoniak, (2003) Angoff ve Doğrudan Uzlaşma yöntemlerinin özellikleri arasındaki benzerlik ve farklılıkları Tablo 5’teki gibi açıklamıştır.

Tablo 5 Angoff ve Doğrudan Uzlaşma Yöntemlerinin Özellikleri

Özellik	Angoff Yöntemi	Doğrudan Uzlaşma Yöntemi	Özellik	Angoff Yöntemi	Doğrudan Uzlaşma Yöntemi
Uzmanlar testin amacını bilir.	Evet	Evet	Uzmanlara maddelerin doğru yanıtları verilir.	Hayır	Hayır
Uzmanlar testin içeriğine aşinadır.	Evet	Evet	Uzmanlar maddeler üzerinde tartışır.	Evet	Hayır

Uzmanlar testi önceden inceler.	Evet	Evet	Uzmanlara maddelerin istatistiksel bilgileri verilebilir.	Evet	Evet
Odak noktası minimum geçme seviyesindeki öğrencidir.	Evet	Evet	Uzmanların verdikleri karar sonucunda elde edilen kesme puanı tartışılabilir.	Bazen	Evet
Test alt testlere bölünür.	Hayır	Evet	Uzmanlar farklı kararlar verirse fikir birliğine varmaları için teşvik edilebilir.	Hayır	Evet
Uzmanlara nihai kesme puanı açıklanır.	Bazen	Evet	Uzmanlar toplantı bitmeden kararlarını gözden geçirip değiştirebilir.	Bazen	Evet

2.2 İlgili Araştırmalar

Bu bölümde Angoff ve Doğrudan Uzlaşma yöntemine dayalı yapılan araştırmalara ve bu araştırmaların sonuçlarına yer verilmiştir.

2.2.1 Angoff Yöntemine Dayalı Araştırmalar

Norcini ve arkadaşları (1987), Angoff yöntemiyle verilen uzman kararlarının tutarlılık derecesini ölçmek istemişlerdir. Aynı test üzerinde farklı zamanlarda 3 ayrı toplantı yapılarak 3 kesme puanı elde edilmiş ve bu puanların karşılaştırılması yapılmıştır. Belirlenen kesme puanlarının farklılaşmadığı görülmüş ve araştırma sınırları içerisinde Angoff yöntemi tutarlı bulunmuştur.

Maurer ve arkadaşları, (1991) çalışılan uzman grubunun akademik yeterliğinin Angoff yöntemiyle elde edilen kesme puanlarının güvenilirliğine etkisini incelemişlerdir. Araştırmada psikoloji alanında hazırlanan 80 maddelik çoktan seçmeli bir test kullanılmıştır. Psikoloji alanında lisans seviyesinden henüz mezun olmayan öğrencilerden oluşan 20 uzmanın belirlediği kesme puanları ile psikoloji alanında en az lisans mezunu olan deneyimli 20 uzmanın belirlediği kesme puanları güvenilirlik açısından incelenmiş ve akademik yeterliği ve deneyimi yüksek olan uzman grubunun daha güvenilir kesme puanları ürettiği sonucuna ulaşılmıştır. Araştırmada Angoff yöntemini kullanan yeterliği yüksek 9 ila 11 uzmanın, yeterliği

düşük 20 uzmandan daha doğru ve güvenilir kesme puanları ürettiği sonucuna ulaşılmıştır.

Impara ve Plake (1998), yaptıkları araştırmada Angoff yöntemini kullanan 26 fen bilgisi öğretmeninin yüzde tahminlerinin başarısını incelemişlerdir. Çalışmada 50 maddelik bir test ilköğretim 6. Sınıf düzeyinde eğitim gören 724 öğrenciye uygulanmıştır. Öğretmenlerden “Minimum yeterlik seviyesindeki bir grup öğrencinin performansı” ve “Testin uygulandığı öğrencilerin tamamının mevcut performansı” kavramlarını düşünerek iki ayrı tahminde bulunmaları istenmiştir. Öğretmenlerin testin uygulandığı öğrencilerin mevcut performanslarına odaklandıklarında daha doğru tahminler yaptığı belirlense de yüzde tahminlerinin her iki durumda da yeterince yüksek doğrulukta kararlar olmadığı sonucuna ulaşılmıştır.

Hurtz ve Hertz (1999), araştırmalarında Angoff yöntemi ile elde edilen kesme puanının geçerli olması için karar verecek uzman grubunun kaç kişiden oluşması gerektiğini incelemişlerdir. Birbirinden farklı 8 adet lisans sınavı üzerinden farklı uzman sayıları ile çalışma yürütülmüş ve kesme puanları genellenebilirlik teorisi ile hesaplanmıştır. Araştırma sonunda optimal uzman sayısının 10 ile 15 aralığında olması gerektiği bazı durumlarda bu sayının daha az olabileceği sonucuna varılmıştır.

Clauser ve arkadaşları, (2008) araştırmalarında uzmanların Angoff yöntemi ile verdikleri kararların kararlılığını ölçmek, uzmanların kararlarıyla elde edilen madde güçlük indeksi ile gerçek madde güçlük indekslerini karşılaştırmak ve uzmanların kararları üzerine yapılan tartışmaların bireysel puanları nasıl etkilediğini incelemek istemişlerdir. Elde ettikleri bulgulara göre uzmanlar zor olan maddelerin zorluk derecesini, kolay olan maddelerin kolaylık derecesini var olandan daha düşük seviyede tuttukları belirtilmiştir. Uzman kararlarının tartışılmasının ise uzmanların bireysel puanları arasındaki farkı azalttığı sonucuna ulaşılmıştır.

2.2.2 Doğrudan Uzlaşma Yöntemine Dayalı Araştırmalar

Pitoniak (2002) tarafından yapılan bir çalışmada Doğrudan Uzlaşma yöntemi ile Madde Kümesi yöntemi karşılaştırılmıştır. Her yöntem için iki toplantı yapılarak yöntemler uygulanmış ve her toplantı için farklı 4 uzman olmak üzere 16

uzmanla çalışılmıştır. Yöntemlerin uygulanması için 35 maddelik çoktan seçmeli bir lisans testi kullanılmıştır. Çalışma bulgularına göre Doğrudan Uzlaşma yöntemi ile üretilen kesme puanlarının Madde Kümesi yöntemi ile üretilen kesme puanlarına göre anlamlı ölçüde düşük çıktığı ve Madde Kümesi yönteminin daha tutarlı kesme puanları ürettiği sonucuna ulaşılmıştır. Uzmanlarla yöntemler üzerine yapılan görüşmelerden elde edilen verilere göre uzmanların daha kolay ve daha kısa sürede uygulanabildiği gerekçesiyle Doğrudan Uzlaşma yöntemini tercih ettikleri açıklanmıştır.

Pitoniak, (2003) yaptığı çalışmada muhasebe alanında farklı madde türleriyle hazırlanan bir test üzerinde 4 standart belirleme yöntemini araştırmıştır. Testin yapılandırılmamış maddelerden oluşan ve bilgisayar ortamında uygulanan bölümü İş Sınıflandırma ve Analitik yöntemlerini karşılaştırmak için kullanılırken, çoktan seçmeli 45 maddeden oluşan bölümü Madde Kümesi ve Doğrudan Uzlaşma yöntemlerini karşılaştırmak için kullanılmıştır. 28 uzmanın kararıyla Madde Kümesi ve Doğrudan Uzlaşma yöntemleri kullanılarak elde edilen kesme puanları üç farklı yaklaşım kullanılarak karşılaştırılmış, Doğrudan Uzlaşma yönteminin daha yüksek puanlar ürettiği fakat farklılığın anlamlı olmadığı görülmüştür. Araştırmada Madde Kümesi yönteminin daha tutarlı kesme puanları ürettiği, Doğrudan Uzlaşma yönteminin ise daha kısa sürede tamamlandığı belirtilmiştir.

Sireci, Hambleton ve Pitoniak, (2004) yaptıkları araştırmada Doğrudan Uzlaşma yöntemi, Angoff yöntemi ve Madde Kümesi yöntemi olmak üzere üç standart belirleme yöntemlerden elde edilen kesme puanları, yöntemlerin uygulama kolaylığı ve yöntemlerin uygulama süresi açısından karşılaştırmışlardır. Doğrudan Uzlaşma yöntemi Angoff yönteminden %20, Madde Kümesi yönteminden %67 daha kısa sürede uygulanmıştır. Araştırmacılar uzmanların Doğrudan Uzlaşma yöntemini uygulamada Angoff yöntemine göre daha az zorlandıkları sonucuna varmıştır. Kesme puanları ise Angoff yönteminde 50 puanlık iki test için 34 ve 31 hesaplanırken, Doğrudan Uzlaşma yönteminde 34 ve 35 hesaplanmıştır. Angoff ve Doğrudan Uzlaşma yöntemleri ile hesaplanan kesme puanları arasında anlamlı bir farklılık görülmezken Madde Kümesi yöntemine ait kesme puanlarının diğer iki yöntemden farklılaştığı belirtilmiştir (Akt. Cizek ve Bunch, 2007)

3. YÖNTEM

3.1 Araştırmanın Modeli

Bu araştırma Angoff ve Doğrudan Uzlaşma yöntemleriyle elde edilen kesme puanlarının karşılaştırılmasına dayanmaktadır. Bir, iki ya da daha çok sayıdaki değişkenin birlikte değişim derecesini belirlemeyi amaçlayan araştırma modelleri ilişkisel türden araştırma modelleridir (Karasar, 2007). Bu yönüyle bu araştırma nicel yönelimli ilişkisel türden bir araştırma modelidir. Betimsel araştırmalar ise uygulamaların, fikirlerin veya yöntemlerin arasındaki ilişkinin tanınması, karşılaştırılması ve yorumlanması yoluyla "Mevcut durum nedir?" gibi çok önemli bir soruyu cevap arar (Devries, 1973). Bu araştırma her iki yönetime ilişkin betimsel istatistikleri elde etme, değişkenler arasındaki ilişkinin varlığını ve derecesini tespit etme ve sonuçları yorumlama yönüyle betimsel bir çalışmadır. Ayrıca araştırma, uzmanların görüşlerinin alınması yönüyle “gözlem, görüşme ve doküman analizi gibi nitel veri toplama tekniklerinin kullanıldığı, algıların ve olayların doğal ortamda gerçekçi ve bütüncül bir biçimde ortaya konması için nitel bir sürecin izlendiği araştırma” olarak tanımlanan nitel araştırma özelliği de taşımaktadır (Yıldırım ve Şimşek, 2021, s. 39).

3.2 Çalışma Grubu

Bu çalışmada iki farklı gruptan veri toplanmıştır. Birinci grup testin uygulandığı öğrencilerden oluşurken ikinci grup yöntemleri test üzerinde uygulayarak karar ve görüşlerini belirten uzmanlardan oluşmuştur. Her iki çalışma grubuna da amaca uygun örnekleme yöntemiyle ulaşılmış olup süreç içerisinde gönüllülük esasına uygun davranılmıştır. Öğrenci grubu Bolu ili Yunus Emre Ortaokulu 5. sınıf düzeyinde öğrenim gören 54 kız 47 erkek öğrenci olmak üzere toplam 101 öğrenciden oluşmaktadır.

Standart belirleme sürecinin kabul edilebilir güvenirlikte sonuçlar vermesi için uzman grubu birçok bileşeni karşılayabilecek nitelik ve nicelikte olmalıdır (Kane, 1994). Kabul edilebilir güvenirlikte sonuçlar elde etmek için uzman sayısının 5 ila 10 arasında olması gerekmektedir (Norcini ve Shea, 1997). Ayrıca Maurer ve ark. (1991), farklı sayıda uzmanlarla Angoff yöntemini kullanarak yaptığı güvenirlik çalışmasında 9 ila 11 uzmanın yeterince güvenilir sonuçlar ürettiğini bulmuştur. Fowell ve ark. (2008) ‘e göre seçilen standart belirleme

yöntemi tek oturumdan oluşuyorsa (uzmanlara ilk değerlendirmelerini değiştirme şansı verilmiyorsa) uzman grubunun 10 kişiden oluşması güvenilirlik için yeterliyken, iki veya daha çok oturumdan oluşuyorsa (uzmanlara ilk değerlendirmelerini değiştirme şansı veriliyorsa) uzman grubunun en az 6 kişiden oluşması güvenilirlik için yeterlidir.

Mevcut çalışmada kullanılan Angoff ve Doğrudan Uzlaşma yöntemlerinin prosedürleri ve literatürdeki araştırmalar göz önünde bulundurularak gerekli bilgi, beceri ve deneyime sahip 6 kadın 4 erkek olmak üzere 10 uzmanla çalışma yürütülmüştür. Uzman grubu en az 9 yıllık mesleki kıdeme ve en az 2 yıllık 5. Sınıf seviyesinde öğretmenlik deneyimine sahip ortaokul matematik öğretmenlerinden oluşmuştur. 8 uzman lisans derecesine, 2 uzman yüksek lisans derecesine sahiptir. Uzmanların tamamı MEB'e bağlı ortaokul kurumlarında matematik öğretmeni olarak görev yapmaktadır. Her iki standart belirleme yönteminde de aynı uzmanlarla çalışılmıştır.

3.3 Veri Toplama Araçları

Araştırmada öğrencilere ait verileri elde etmek için matematik başarı testi (EK 1) geliştirilmiştir. Bu test aynı zamanda uzmanların yöntemlere ilişkin kararlarını vermesi için de kullanılmıştır. Uzmanların yöntemlere ilişkin vermiş olduğu kararları elde etmek için araştırmacı tarafından hazırlanan “Uzman Görüşleri Formu” (EK 2) kullanılmıştır.

Çalışılan uzmanların Angoff ve Doğrudan Uzlaşma yöntemlerini değerlendirmeleri için uzmanlara her iki yönetime ilişkin değerlendirme formu (EK 3) uygulanmıştır.

3.3.1 Matematik Başarı Testi

Araştırma kapsamında kullanılan matematik başarı testi (EK 1) Milli Eğitim Bakanlığı 5. Sınıf Matematik Öğretim Programına uygun olarak “Sayılar ve İşlemler” öğrenme alanine ait “Doğal Sayılar”, “Doğal Sayılarla İşlemler”, “Kesirler”, “Kesirlerle İşlemler”, “Ondalık Gösterim” ve “Yüzdelere” olmak üzere 6 alt öğrenme alanını kapsayacak şekilde hazırlanmıştır. Testin hazırlanma aşamasında hedef davranışlarla içerik ilişkisinin belirlenmesi için belirtke tablosu hazırlanmıştır. Belirtke tablosunda bulunan 33 hedef davranışın her biri için 3 deneme maddesi hazırlanmış ve 99 maddelik deneme testi oluşturulmuştur. 5. Sınıf

düzeyinde eğitim vermekte olan alanında uzman 3 matematik öğretmenin maddelere ve testin kapsamına ilişkin görüşü alınmıştır. Uzmanların görüşü doğrultusunda 5. Sınıf düzeyinde bir öğrencinin dikkat süresi göz önünde bulundurularak madde sayısının kapsam geçerliğini düşürmeyecek şekilde azaltılmasına karar verilmiştir. Bu amaçla uzmanlarla birlikte her bir hedef davranış için hazırlanan 3 deneme maddesi kendi içlerinde karşılaştırılarak incelenmiş ve 2 maddeye düşürülmüştür. Maddeler incelenirken maddenin hedef davranışı ölçme derecesi, başka bir hedef davranışla binişik olup olmama durumu, öğrencilerin gelişim düzeyine uygunluğu gibi kriterler dikkate alınmıştır. Ayrıca testin yazım yanlışları ve anlatım bozukluklarından arınık olması için alanında uzman bir Türkçe öğretmene inceletilip görüşü sorulmuştur. Maddelerdeki maddi hatalar ve yazım yanlışları da giderilerek 66 maddelik deneme testi hazırlanmıştır. Hazırlanan 66 maddelik deneme testi öğrencilerin dikkat süreleri göz önünde bulundurularak madde sıraları değiştirilmeden 3 bölüme ayrılmıştır. Testin her bölümü öğrencilere birer gün arayla bir ders saati içerisinde uygulanmıştır.

Deneme testi uygulanırken çalışma grubunu oluşturan 101 öğrenci rastgele küme örnekleme yöntemi ile iki gruba bölünmüştür. İlk grup 49 öğrenciden oluşmuş olup deneme testinin uygulandığı gruptur. İkinci grup ise 52 öğrenciden oluşmuştur ve bu gruba deneme testinin madde istatistiklerine göre oluşturulan nihai test uygulanmıştır.

Deneme testinin uygulanması sonucunda ulaşılan öğrenci cevapları ve testin cevap anahtarı Excel programına girilmiş ve Test Analysis Program (version 19.1.4)'a aktararak analiz edilmiştir. Deneme testine ait betimsel istatistikler Tablo 6'da verilmiştir.

Tablo 6 Deneme Testine Ait Betimsel İstatistikler

Madde Sayısı	66	Ortalama Güçlük İndeksi	0,47
Öğrenci Sayısı	49	Ortalama Ayırt Edicik İndeksi	0,42
Aritmetik Ortalama	31,24	KR-20	0,91
Medyan	29	Çarpıklık	0,87
Mod	23	Basıklık	0,04
Standart Sapma	11,92	En Düşük Puan	15

Tablo 6 incelendiğinde 66 maddelik deneme testinin aritmetik ortalamasının 31,24 değerini aldığı görülmektedir. Testin ortalama madde güçlük indeksine bakıldığında orta güçlükte olduğu söylenebilir. KR-20 değeri 0,75'ten büyük olduğu için testin güvenilir olduğu söylenebilir (Murphy ve Davidshofer, 1988).

Aşağıda verilen Tablo 7'de deneme testinin madde istatistikleri belirtilmiştir. Nihai test oluşturulurken maddelerin madde ayırt edicilik indeksleri (r_{jx}) ile madde güçlük indeksleri (P_j) göz önünde bulundurulmuştur.

Tablo 7 Deneme Testi Sonuçlarına Göre Madde İstatistikleri

Madde	Madde Güçlük İndeksi (P _j)	Madde Ayırt Edicilik İndeksi (r _{jx})	Madde	Madde Güçlük İndeksi (P _j)	Madde Ayırt Edicilik İndeksi (r _{jx})
1	0,90	0,31	34	0,35	0,29
2	0,80	0,38	35	0,49	0,29
3	0,80	0,47	36	0,24	0,46
4	0,31	0,52	37	0,47	0,50
5	0,82	0,54	38	0,20	0,39
6	0,88	0,38	39	0,39	0,30
7	0,33	0,30	40	0,27	0,31
8	0,57	0,70	41	0,27	0,10
9	0,69	0,33	42	0,49	0,49
10	0,80	0,10	43	0,69	0,56
11	0,63	0,48	44	0,45	0,36
12	0,41	0,23	45	0,86	0,09
13	0,63	0,56	46	0,61	0,56
14	0,35	0,37	47	0,65	0,47
15	0,49	0,48	48	0,37	0,52
16	0,41	0,66	49	0,82	0,17
17	0,27	0,59	50	0,47	0,23
18	0,53	0,78	51	0,69	0,41
19	0,39	0,44	52	0,49	0,85
20	0,57	0,57	53	0,65	0,48
21	0,37	0,38	54	0,73	0,55
22	0,24	0,18	55	0,43	0,65
23	0,57	0,42	56	0,39	0,59
24	0,51	0,43	57	0,33	0,51
25	0,49	0,43	58	0,31	0,38

26	0,31	0,37	59	0,45	0,49
27	0,35	-0,14	60	0,43	0,45
28	0,43	0,29	61	0,29	0,39
29	0,31	0,67	62	0,49	0,71
30	0,33	0,03	63	0,27	0,45
31	0,37	0,36	64	0,33	0,39
32	0,24	0,24	65	0,31	0,37
33	0,24	0,53	66	0,29	0,30

Tablo 7 incelendiğinde madde ayırt edicilik indekslerinin -0,14 ile 0,85 arasında değiştiği görülmektedir. Madde güçlük indeksi ise 0,20 ile 0,90 arasında değerler almıştır. Nihai forma alınabilecek maddelerin madde ayırt edicilik indekslerinin en az 0,30 olması gerekmektedir (Tekin,2019). Bu koşul göz önünde bulundurularak 10, 12, 22, 26, 27, 28, 30, 34, 35, 39, 41, 45, 49, 50, 58, 64 numaralı maddeler madde ayırt edicilik indeksleri sınır değer olarak kabul edilen 0,30'dan küçük olduğu gerekçesiyle nihai teste alınmamıştır. Kapsam geçerliğini korumak için geriye kalan maddeler arasından içeriği aynı hedef davranışla eşleşen maddeler gruplandırılmış ve içlerinden madde ayırt ediciliği yüksek olanlar seçilmiştir. Başarı testlerindeki maddelerin çoğunluğu orta güçlükte olmalı, bunun yanında testte hem kolay hem de zor maddelere yer verilmelidir (Turgut ve Baykul, 2019). Analiz sonuçlarına göre çoğunluğu madde güçlük indeksi 0,50 civarında olan maddeler olmak üzere kolay ve zor maddeler de teste alınmıştır. Testin öğrenciler tarafından bir ders saati içerisinde tamamlanabilmesi için kritik olmayan hedef davranış ölçen maddeler teste alınmamıştır. Kritik olmayan hedef davranışların seçimi sırasında testi inceleyen 3 matematik öğretmeni ile çalışılmıştır. Seçim sırasında bir önceki hedef davranışın tekrarı niteliğinde olan ya da bir sonraki hedef davranışla da ölçülebilen hedef davranışlar kritik olmayan hedef davranışlar olarak belirlenmiştir. Sonuç olarak nihai teste 3, 4, 5, 6, 8, 13, 15, 20, 23, 24, 25, 31, 36, 37, 38, 40, 42, 43, 46, 52, 59, 60, 62, 64 numaralı maddeler seçilmiştir.

Testin kendisinden daha homojen alt testlere bölünmesi için maddeler ölçtüğü hedef davranışın ait olduğu alt öğrenme alanına göre gruplandırılmıştır. Testin kapsadığı 6 alt öğrenme alanına maddelerin eşit dağıldığı görülmüştür. Sonuç olarak 24 maddelik nihai test her biri 4 maddeden oluşan 6 alt teste bölünmüştür.

Oluşturulan 24 maddelik nihai test 5. Sınıf seviyesindeki 2. grup öğrencilere uygulanarak testin madde istatistikleri ile betimsel istatistikleri elde edilmiştir.

Tablo 8 Nihai Form Madde İstatistikleri

Madde	Madde Güçlük İndeksi (Pj)	Madde Ayırt Edicilik İndeksi (rjx)	Madde	Madde Güçlük İndeksi (Pj)	Madde Ayırt Edicilik İndeksi (rjx)
1	0,85	0,29	13	0,35	0,66
2	0,29	0,46	14	0,48	0,93
3	0,90	0,22	15	0,54	0,72
4	0,85	0,43	16	0,29	0,32
5	0,69	0,71	17	0,62	0,79
6	0,67	0,71	18	0,87	0,30
7	0,60	0,79	19	0,56	0,72
8	0,58	0,44	20	0,56	0,79
9	0,50	0,73	21	0,58	0,86
10	0,69	0,37	22	0,65	0,64
11	0,56	0,65	23	0,58	0,93
12	0,21	0,20	24	0,38	0,59

Tablo 9 Nihai Form Betimsel İstatistikler

Madde Sayısı	24	Testin Ortalama Güçlüğü	0,58
Öğrenci Sayısı	52	Testin Ortalama Ayırtedicilik İndeksi	0,59
Aritmetik Ortalama	13,83	KR-20	0,88
Medyan	13	Çarpıklık	-0,01
Mod	22	Basıklık	-1,10
Standart Sapma	5,74	En Düşük Puan	1
Varyans	32,99	En Yüksek Puan	24

Tablo 9 incelendiğinde 24 maddelik matematik başarı testinin aritmetik ortalamasının 13,83 hesaplandığı görülmektedir. Testin ortalama güçlük civarında olduğu ve ayırt edicilik düzeyinin yeterli olduğu söylenebilir. Testin KR-20 değerinin 0,88 olduğu görülmektedir. KR-20 güvenirlik katsayısı 0,75'ten büyük olan testlerin güvenilir olduğu düşünülürse matematik başarı testi güvenilir bir test olarak kabul edilebilir. (Murphy ve Davidshofer, 1988).

Matematik başarı testinin geçerlik çalışması için testin uygulandığı 52 öğrencinin 2021-2022 eğitim öğretim yılı matematik karne notları ile testten aldığı puanlar karşılaştırılmış ve testin ölçüt geçerliği değerlendirilmiştir. Öğrencilerin testten aldığı puanlar ile karne notlarının puan aralıklarının farklı olması sebebiyle Spearman Brown Sıra Farkları Korelasyon Katsayısı kullanılmıştır.

Tablo 10 Karne notları ve Test Puanlarına İlişkin Korelasyon Katsayısı

Karşılaştırılan Yöntemler	Karne Notları	Test Notları
Karne Notları	-	0,71**
Test Puanları	-	-

** p < 0,01

Tablo 10'da görüldüğü üzere korelasyon katsayısı 0,01 anlamlılık düzeyinde 0,71 hesaplanmıştır. Bu sonuç test puanları ile karne notları arasında %99 olasılıkla anlamlı ve yüksek bir ilişki olduğunu göstermektedir. Bu durumda testin ölçüt geçerliğinin yeterli olduğunu söylenebilir.

3.3.2 Uzman Görüşleri Formu

Uzman grubunun hazırlanan 24 maddelik matematik başarı testini inceleyerek kararlarını belirtmesi için araştırmacı tarafından Uzman Görüşleri Formu (EK 2) hazırlanmıştır. Uzmanlar Angoff ve Doğrudan Uzlaşma yöntemlerinin prosedürlerini aynı test üzerinde uygulayarak kararlarını vermişlerdir. Hazırlanan form 6 bölümden oluşmaktadır. Her bölüm matematik başarı testinin 4 maddelik bir alt testini içermektedir. Bu bölümlerde maddelerin iki yanına kutucuklar eklenmiş ve uzmanlardan Angoff yöntemine ait yanıtlarını sol kutucuğa Doğrudan Uzlaşma yöntemine ait yanıtlarını sağ kutucuğa yazmaları istenmiştir. Uzmanlar Angoff yöntemini uygulama aşamasında maddelerin her biri için ayrı yanıtlar verirken, Doğrudan Uzlaşma yöntemini uygulama aşamasında 4 maddelik alt testin tamamı için tek bir yanıt belirtmişlerdir. Her iki yönteme ilişkin sürecin nasıl işlediği aşağıda verilmiştir.

3.3.2.1 Angoff Yöntemi Uygulama Süreci

Araştırma hakkında önceden kısaca bilgi verilen ve matematik başarı testini inceleyen gönüllü 10 uzmanla toplantı başlatılmıştır. Gruba öncelikle standart belirleme yöntemi ve kesme puanı hakkında bilgilendirme yapılmıştır. Sonrasında Angoff yönteminin nasıl uygulanması gerektiği örneklerle açıklanmış ve

uzmanlardan gelen sorular cevaplanmıştır. Uzman Görüşleri Formu dağıtılarak uzmanlardan testteki her maddeyi tek tek inceleyip “Minimum yeterlik seviyesinde olan 100 öğrenciden kaç tanesi bu maddeye doğru yanıt verebilir?” sorusuna yanıt vermesi ve yanıtını maddenin solundaki kutucuğa yazması istenmiştir. Testteki 24 maddenin tamamını inceleyerek kararlarını veren uzmanlara düzeltme yapabilmeleri için zaman verilmiştir. Uzmanlar son kararlarını verdikten sonra toplantı sona ermiştir. Angoff yönteminin uygulanması için yapılan toplantının akışı aşağıda açıklanmıştır.

Tablo 11 Angoff Yöntemi Toplantı Akışı

Zaman Aralığı	Yapılan Etkinlik
12:00 - 12:10	Toplantı başlatıldı. Standart belirleme ve kesme puanı kavramları uzmanlara açıklandı.
12:10 - 12:25	Angoff yönteminin uygulanması hakkında bilgi verildi. Örnekler üzerinden uzmanların soruları cevaplandı.
12:25 - 12:45	Uzman Görüşleri formu dağıtıldı ve uzmanların yanıtlarını forma belirtmeleri için süre verildi.
12:45 - 12:50	Uzmanların “Minimum yeterlik seviyesindeki 100 öğrenci” kavramı ile ilgili soruları üzerine tekrar açıklama yapıldı.
12:50 - 13.00	Uzmanlar yapılan açıklamalar üzerine kararlarını gözden geçirip formu tamamladı. Toplantı bitirildi.

Toplantı sona erdikten sonra uzmanların verdiği kararlar (EK 4) yüzde olarak toplanmış ve soru sayısına (24) bölünerek her uzmana ait MGP hesaplanmıştır. Uzman MGP’lerinin aritmetik ortalaması alınarak testin nihai kesme puanına ulaşılmıştır.

3.3.2.2 Doğrudan Uzlaşma Yöntemi Uygulama Süreci

Uzmanlarla yapılan Angoff yönteminin uygulama toplantısı sona erdikten sonra uzmanlara dinlenmeleri için süre verilmiştir. Bu sürenin sonunda Doğrudan Uzlaşma yöntemi için tekrar toplanılmış ve süreç başlatılmıştır. Standart belirleme ve kesme puanı kavramlarına hakim olan uzmanlara Doğrudan Uzlaşma yöntemi ve uygulaması hakkında açıklama yapılmıştır. Örnek uygulamalar sunulmuş ve

uzmanların soruları cevaplanmıştır. Her uzmana daha önce kullanmış olduğu kendisine ait Uzman Görüşleri Formu tekrar dağıtılmıştır. Formda her biri 4 maddeden oluşan 6 alt öğrenme alanı bulunmaktadır. Uzmanlardan her alt öğrenme alanı için “Minimum yeterlik seviyesindeki bir öğrencinin bu alt öğrenme alanında başarılı kabul edilmesi için kaç maddeye doğru yanıt vermesi gerekir?” sorusuna yanıt vermeleri ve yanıtlarını alt öğrenme alanlarının sağında bulunan kutucuğa belirtmeleri istenmiştir. Testteki 6 alt öğrenme alanını inceleyerek kararlarını veren uzmanların her alt öğrenme alanı için verdiği kararlar toplanarak her uzmana ait MGP’ler ve testin kesme puanı hesaplanmıştır. Bu hesaplama sonucunda uzman MGP’lerinin farklılaştığı görülmüştür. Yöntem prosedürü gereği uzmanlara kararları hakkında çeşitli sorular sorularak açık oturum başlatılmış ve uzmanların fikir birliğine varmaları amaçlanmıştır. Testin hesaplanan kesme puanına ilişkin uzmanların görüşleri alınmıştır. Uzmanlar her alt öğrenme alanı ile ilgili fikir ve görüşlerini açıkladıktan sonra bazı uzmanlar kararlarını değiştirmiştir. Bu durumda 2, 5 ve 6 numaralı alt öğrenme alanlarında fikir birliği sağlanırken 1, 3 ve 4 numaralı alt öğrenme alanlarında fikir birliği sağlanamamıştır. Doğrudan Uzlaşma yönteminin uygulanması için yapılan toplantının akışı aşağıda açıklanmıştır.

Tablo 12 Doğrudan Uzlaşma Yöntemi Toplantı Akışı

Zaman Aralığı	Yapılan Etkinlik
13:30 - 13:40	Toplantı başlatıldı. Doğrudan Uzlaşma yöntemi hakkında bilgi verildi. Örnekler üzerinden uzmanların soruları yanıtlandı.
13:40 – 13:50	Uzman Görüşleri formu dağıtıldı ve uzmanların yanıtlarını forma belirtmeleri için süre verildi.
13:50 - 13:55	Uzmanların formu tamamlamaları üzerine MGP’ler ve testin önerilen kesme puanı hesaplandı.
13:55 - 14:05	MGP’lerin farklı değerler alması üzerine açık oturum başlatılarak fikir ve görüşler alındı.
14:05 - 14:10	Açık oturum sona erdi ve uzmanlara kararlarını gözden geçirmeleri için süre verildi.
14:10 - 14:15	Uzmanların nihai kararları (EK 4) ile testin kesme puanı hesaplandı ve toplantı bitirildi.

3.3.3 Standart Belirleme Yöntem ve Süreçlerini Değerlendirme Formu

Standart belirleme sürecinde çalışılan uzmanların Angoff ve Doğrudan Uzlaşma yöntemlerine ilişkin değerlendirme verilerini elde etmek için Standart Belirleme Yöntem ve Süreçlerini Değerlendirme Formu (EK 3) hazırlanmıştır. Form Hambleton'ın (2013) ve Cizek, Bunch ve Koons'un (2004) geliştirdiği standart belirleme süreçlerini değerlendirme anketlerinden yararlanılarak geliştirilmiştir. Formda standart belirleme sürecine ve kullanılan yöntemin uygulama aşamalarına dair olumlu görüş ifadeleri içeren 9 madde bulunmaktadır. Maddeler “Hiç Katılmıyorum”, “Katılmıyorum”, “Kararsızım”, “Katılıyorum” ve “Tamamen Katılıyorum” olmak üzere 5 yanıt seçeneğine sahiptir.

Yöntemlere dair uzman görüşlerinin ayrı ayrı elde edilebilmesi için geliştirilen form her iki yöntem için maddeler ve yanıt seçenekleri değiştirilmeden yalnızca yöntem isimleri değiştirilerek düzenlenmiştir. Böylece uzmanlara Form-1 ve Form-2 olmak üzere iki ayrı form verilmiştir. Form-1 Angoff yöntemine dair, Form-2 Doğrudan Uzlaşma yöntemine dair aynı olumlu görüş ifadelerini içermektedir.

3.3.4 Verilerin Çözümlemesinde Kullanılan İstatistiksel Teknikler

Araştırma sürecinde elde edilen verilerin çözümlemesinde çeşitli istatistiksel tekniklerden yararlanılmıştır. Bu teknikler Test Analysis Program (version 19.1.4) aracılığıyla uygulanmıştır.

Birinci alt probleme ilişkin olarak yöntemlerin ürettiği kesme puanlarına göre üst performans seviyesine yerleşerek başarılı sayılan öğrenci yüzdelerinin arasındaki farklılık bağımlı iki oran arasındaki farkın testi ile çözümlenmiş ve farkın anlamlılığı z testi kullanılarak sınanmıştır. İki oran arasındaki fark test edilirken z testi örneklem büyüklüğünün 30 ve üzeri olduğu durumlarda, t testi örneklem büyüklüğünün 30'dan küçük olduğu durumlarda tercih edilmelidir (Takiar, 2023). Mevcut araştırmada örneklem büyüklüğü N=52 olduğu için z istatistiğine başvurulmuştur. Çözümleme yapılırken Tablo 13'te yer alan gözlem sonuçlarından yararlanılmıştır.

Tablo 13 Bağımlı İki Oran Testi İçin Frekans ve Oranlar

		Frekans					Oran		
		Yöntem 2					Yöntem 2		
		Başarısız	Başarılı	Toplam			Başarısız	Başarılı	Toplam
Yöntem 1	Başarılı	A	B	A+B	Yöntem 1	Başarılı	a	b	P1
	Başarısız	C	D	C+D		Başarısız	c	d	Q1
	Toplam	A+C	B+D	N		Toplam	Q2	P2	1,00

Tablon 13'ün frekanslara ait bölümünde yöntem 1'e göre başarılı sayılırken yöntem 2'ye göre başarısız sayılan öğrenci sayısı A ile, her iki yönteme göre başarılı sayılan öğrenci sayısı B ile, her iki yönteme göre başarısız sayılan öğrenci sayısı C ile, yöntem 1'e göre başarısız sayılırken yöntem 2'ye göre başarılı sayılan öğrenci sayısı D ile ifade edilmiştir.

Tablo 13'ün oranlara ait bölümünde ise frekans değerlerinin öğrenci sayısı olan N ile bölümünden elde edilen yüzde değerleri gösterilmektedir. Yöntem 1'e göre başarılı sayılırken yöntem 2'ye göre başarısız sayılan öğrenci yüzdesi a ile, her iki yönteme göre başarılı sayılan öğrenci yüzdesi b ile, her iki yönteme göre başarısız sayılan öğrenci yüzdesi c ile, yöntem 1'e göre başarısız sayılırken yöntem 2'ye göre başarılı sayılan öğrenci yüzdesi d ile gösterilmektedir. Yöntem 1'e göre başarılı sayılan öğrenci yüzdesi P1 ile, yöntem 2'ye göre başarılı sayılan öğrenci yüzdesi P2 ile ifade edilmiştir.

Yüzde değerleri arasındaki farklılığın standart hata formülü aşağıda verilmiştir:

$$S_{P1-P2} = \sqrt{\frac{a+d}{N}}$$

Z istatistiği ise iki yüzde değeri arasındaki farkın hesaplanan standart hata değerine bölünmesiyle elde edilir:

$$Z = \frac{P1-P2}{\sqrt{\frac{a+d}{N}}}$$

İkinci alt probleme ilişkin elde edilen verilerin çözümlenmesinde Cohen'in Kappa istatistiği ve Spearman Brown Sıra Farkları korelasyon katsayısı tekniklerinden yararlanılmıştır. Cohen'in Kappa istatistiği iki puanlayıcının birbirinden bağımsız olarak yaptıkları değerlendirmenin uyumluluk derecesini belirlemek için kullanılır (Cohen, 1960). Landis ve Koch (1977) Cohen'in Kappa katsayısının alabileceği değerler için Tablo 14'teki sınıflandırmayı yapmıştır:

Tablo 14 Cohen'in Kappa Katsayısı Değer Aralıkları

Katsayı Aralığı	Uyum Derecesi
0,00 – 0,20	Düşük
0,21 – 0,40	Makul
0,41 – 0,60	Orta
0,61 – 0,80	Yüksek
0,81 – 1,00	Mükemmel

Üçüncü alt probleme ilişkin elde edilen verilerin çözümlenmesinde MGP'lerin normalliğinin sınanması amacıyla Shapiro Wilk testine başvurulmuştur. Bir veri grubu 50'den az veri içerdiğinde bu grubun normalliğinin araştırılması Shapiro Wilk testi ile yapılabilir. (Razali ve Wah, 2011). Normallik testi sonuçları dikkate alınarak MGP'ler arasındaki ilişki parametrik bir yöntem olan Pearson Momentler Çarpımı korelasyon katsayısı ile test edilmiştir. Yöntemlerin MGP ortalamaları arasındaki farklılık ise MGP'ler normal dağılım gösterdiği için bağımlı gruplar t testi ile sorgulanmıştır.

Dördüncü alt problem uzmanların yöntemlere ilişkin değerlendirmelerinin betimsel olarak incelenmesi yoluyla çözümlenmiştir.

4. BULGULAR

4.1 Birinci Alt Probleme Ait Bulgular

Yöntemlerin ürettiği kesme puanlarına göre üst performans seviyesine yerleşerek başarılı sayılan öğrenci yüzdeleri arasında anlamlı bir farklılık var mıdır?

Bu alt problemin çözümlenmesi amacıyla 10 uzmandan Angoff ve Doğrudan Uzlaşma yöntemlerini kullanarak 24 soruluk başarı testi hakkında karar vermeleri istenmiştir. Her uzmanın verdiği kararların ortalaması alınarak uzmanlara ait MGP'ler hesaplanmıştır. Teste ait kesme puanına bu MGP'lerin ortalaması alınarak ulaşılmıştır. Her bir yönteme ait hesaplanan MGP'ler ve kesme puanları Tablo 15'te belirtilmiştir.

Tablo 15 Yöntemlere Ait Uzman MGP'leri ve Testin Kesme Puanları

Uzmanlar	Angoff Yöntemi	Doğrudan Uzlaşma Yöntemi
Uzman-1	46,25	45,84
Uzman-2	54,58	45,84
Uzman-3	75,41	58,34
Uzman-4	81,04	58,34
Uzman-5	29,16	50
Uzman-6	28,33	41,67
Uzman-7	51,25	45,84
Uzman-8	53,95	50
Uzman-9	50,41	50
Uzman-10	50	45,84
Kesme Puanı	52,04	49,17

Tablo 15 incelendiğinde Angoff yöntemi ile belirlenen kesme puanının (52,04) Doğrudan Uzlaşma yöntemi ile belirlenen kesme puanından (49,17) yüksek olduğu görülmektedir. Uzmanların test maddeleri için verdiği kararlardaki değişimler MGP'leri etkilemiş, dolayısıyla hesaplanan kesme puanlarında farklılık oluşmuştur. Bu farklılık belirlenen kesme puanlarına göre başarılı sayılan öğrenci

sayısını ve yüzdesini etkilemiştir. Yöntemlere göre başarılı sayılan öğrenci sayısı ve yüzdesi Tablo 16’da verilmiştir.

Tablo 16 Yöntemlere Göre Başarılı Sayılan Öğrenci Sayısı ve Yüzdesi (n=52)

Yöntemler	Kesme Puanı	Başarılı Sayılan Öğrenci Sayısı	Başarılı Sayılan Öğrenci Yüzdesi
Angoff	52,04	27	%52
Doğrudan Uzlaşma	49,17	30	%58

Tablo 16 incelendiğinde Angoff yöntemiyle hesaplanan kesme puanının üzerinde puan alarak üst performans seviyesine yerleşen ve başarılı sayılan öğrencilerin yüzdesinin (0,52) Doğrudan Uzlaşma yöntemiyle hesaplanan kesme puanının üzerinde puan alarak üst performans seviyesine yerleşen ve başarılı sayılan öğrencilerin yüzdesinden (0,58) düşük olduğu tespit edilmiştir. Yöntemlere göre başarılı sayılan öğrencilerin yüzdeleri arasında oluşan bu fark bağımlı iki oran arasındaki farkın testi ile araştırılmış ve farkın anlamlı olup olmadığına z testi ile bakılmıştır.

Tablo 17 Başarılı Sayılan Öğrenci Sayısı, Yüzdesi ve z Değerleri

Yöntemler	Kesme Puanı	Başarılı Sayılan Öğrenci Sayısı	Başarılı Sayılan Öğrenci Yüzdesi	z
Angoff Yöntemi	52,04	27	%52	
Doğrudan Uzlaşma Yöntemi	49,17	30	%58	-1,172

Tablo 17’de gösterilen z değeri 0,05 anlamlılık düzeyinde araştırılmış ve 0,083 değerini aldığı görülerek anlamlı bulunmamıştır. Angoff ve Doğrudan Uzlaşma yöntemlerine göre hesaplanan kesme puanları üzerinde puan alarak üst performans seviyesine yerleşen ve başarılı sayılan öğrenci yüzdeleri arasında anlamlı bir farklılık görülmemektedir.

4.2 İkinci Alt Probleme Ait Bulgular

Yöntemler arasında öğrencilerin başarılı-başarısız olarak sınıflandırılması açısından uyum var mıdır?

Bu alt problemin çözümlenmesi amacıyla sürekli değişken olan öğrenci puanları 1-0 şeklinde yapay ikili değişkenlere dönüştürülmüştür. Kesme puanının üzerinde puan alan öğrencilere 1, kesme puanının altında puan alan öğrencilere 0 değeri verilerek Cohen'in Kappa katsayısı hesaplanmıştır. Öğrenci puanları kategorik değişkene dönüştürüldüğü için Spearman Brown Sıra Farkları korelasyon değerleri de hesaplanmıştır.

Tablo 18 Yöntemler Arasındaki Korelasyon Katsayıları

Yöntemler	Kappa	Spearman
Angoff Yöntemi	0,88**	0,89**
Doğrudan Uzlaşma Yöntemi		

** $p < 0,01$

Kappa katsayısı 0,81-1,00 aralığında bir değere sahip olduğunda neredeyse mükemmel bir uyumdan bahsedilebilir (Landis ve Koch, 1977). Tablo 18 incelendiğinde Kappa katsayısının 0,88 değerini aldığı görülmektedir. Bu durumda iki yöntem arasında öğrencilerin başarılı-başarısız olarak sınıflandırılması açısından 0,01 hata payıyla mükemmel bir uyum olduğu söylenebilir. Spearman Brown sıra farkları korelasyon değeri ise -1 ile +1 arasında değişim gösterir (Hauke ve Kossowski, 2011). Tablo 18'deki korelasyon değeri (0,89) göz önünde bulundurulduğunda iki yöntem arasında öğrencilerin başarılı-başarısız olarak sınıflandırılması bakımından %99 olasılıkla ($p < 0,01$) pozitif yönelimli yüksek bir korelasyon olduğu görülmektedir.

4.3 Üçüncü Alt Probleme Ait Bulgular

Yöntemler arasında minimum geçme puanları (MGP) açısından uyum var mıdır?

Bu alt problemin çözümü için uzmanlara ait MGP'lerin betimsel istatistikleri hesaplanmış ve MGP'lerin normal dağılım gösterip göstermediği sınanmıştır.

Tablo 19 MGP'lere Ait Betimsel İstatistik ve Normallik Testi Sonuçları

Betimsel İstatistikler	Angoff Yöntemi	Doğrudan Uzlaşma Yöntemi
N	10	10
Ortalama	52,04	49,17
Medyan	50,84	47,92
Tepedeğer	28,33	45,84
Standart Sapma	16,75	5,49
Varyans	280,43	30,09
Çarpıklık Katsayısı	0,35	0,82
Basıklık Katsayısı	0,09	-0,05
Shapiro Wilk İstatistiği	0,9	0,86
Shapiro Wilk Sig.	0,24	0,7

Tablo 19 incelendiğinde Angoff yönteminin merkezi eğilim ölçülerinden ortalama ve medyan birbirine yakın değerler alırken tepedeğerin farklılaştığı gözlenmektedir. Doğrudan Uzlaşma yönteminde ise üç merkezi eğilim ölçüsü de yakın değerler almıştır. MGP'lerin basıklık ve çarpıklık değerleri ele alındığında normallik koşulu olarak kabul edilen -1 ile +1 aralığında oldukları görülmektedir (Morgan, Leech, Gloeckner ve Barrett, 2004). Shapiro Wilk significant değeri 0,05'ten büyükse veri grubu normal dağılmaktadır. Her iki yöntemin Shapiro Wilk significant değerlerinin 0,05'ten büyük olduğu görülmektedir. MGP'ler betimsel istatistikler ve normallik test sonuçlarına göre normal dağılım gösterdiği için MGP'ler arasındaki ilişkiye ve anlamlılık düzeyine bakmak için parametrik bir yöntem tercih edilerek Pearson Momentler Çarpımı Korelasyon Katsayısı hesaplanmıştır.

Tablo 20 MGP'ler Arasındaki Pearson Korelasyon Katsayıları

Karşılaştırılan Yöntemler	Angoff Yöntemi	Doğrudan Uzlaşma Yöntemi
Angoff Yöntemi	-	0,80**
Doğrudan Uzlaşma Yöntemi	-	-

**p < 0,01

Tablo 20 incelendiğinde Angoff ve Doğrudan Uzlaşma yöntemlerine ait MGP'lerin arasında 0,01 anlamlılık düzeyinde pozitif yönde yüksek bir ilişki olduğu görülmektedir.

Bu alt problemin çözülmesi için ortalamalar da incelenmek istenmiş ve Angoff yönteminin MGP ortalaması (52,04) ile Doğrudan Uzlaşma yönteminin MGP ortalaması (49,17) arasındaki farkın anlamlılığına bağımlı gruplar t testi ile bakılmıştır.

Tablo 21 MGP Ortalamaları Arasındaki Farkın Testi

MGP'ler	$\bar{x}$	t
Angoff Yöntemi MGP'leri	52,04	0,71
Doğrudan Uzlaşma Yöntemi MGP'leri	49,17	

Yapılan analizde p değerinin 0,496 olduğu görülmüş ve Angoff ve Doğrudan Uzlaşma yöntemlerine ait MGP ortalamaları arasında anlamlı bir farklılık bulunmamıştır.

4.4 Dördüncü Alt Probleme Ait Bulgular

Angoff ve Doğrudan Uzlaşma standart belirleme sürecine katılan uzmanların, yöntemlere ilişkin değerlendirmeleri nasıldır?

Bu alt problemin çözümlenmesi için standart belirleme sürecine katılarak Angoff ve Doğrudan Uzlaşma yöntemlerini deneyimleyen uzmanlara süreç sonunda yöntemleri ve uygulama süreçlerini değerlendirmeleri için her biri 9 maddeden oluşan Form-1 ve Form-2 olmak üzere 2 form uygulanmıştır. Uygulanan formlardaki maddeler ve uzmanların maddelere verdiği yanıtlar Tablo 22'de frekans olarak verilmiştir.

Tablo 22 Angoff ve Doğrudan Uzlaşma Yöntemlerine İlişkin Uzman Görüşleri

Görüşler	Yöntem		Hiç Katılmıyorum	Katılmıyorum	Kararsızım	Katılıyorum	Tamamen Katılıyorum
1.Yöntem ile ilgili bilgilendirme yeterliydi.	Angoff	f	-	-	-	3	7
	Doğrudan Uzlaşma					4	6
2.Yöntemi uygulayabilmem	Angoff	f	1	1	-	2	6

için verilen talimatlar açıktı.	Doğrudan Uzlaşma					3	7
3.Örnek uygulamalar yöntemi daha net anlamamı sağladı.	Angoff	f	-	1	1	3	5
	Doğrudan Uzlaşma					1	9
4.Yöntemi uygulamaya geçme konusunda kendimi hazır hissettim.	Angoff	f	-	1	1	3	5
	Doğrudan Uzlaşma					2	8
5.Yöntemi uygularken kendimi rahat hissettim.	Angoff	f	1	1	3	-	5
	Doğrudan Uzlaşma				1	1	8
6.Yöntemi zaman açısından kullanışlı buldum.	Angoff	f	2	2	-	2	4
	Doğrudan Uzlaşma					2	8
7.Yöntemi bilişsel açıdan kolay anlaşılır buldum.	Angoff	f	1	1	2	2	4
	Doğrudan Uzlaşma					3	7
8.Yöntem sürecinde soru sorma ve düşünme fırsatım oldu.	Angoff	f	-	1	1	3	5
	Doğrudan Uzlaşma					1	9
9.Yöntemi kullanarak bulduğum puanlara güveniyorum.	Angoff	f	-	2	2	3	3
	Doğrudan Uzlaşma				1	2	7

Formun ilk maddesinde uzmanlara yöntemler hakkındaki bilgilendirmenin yeterli olduğu görüşüne katılma derecesi sorulmuştur. Bu maddeye Angoff yöntemi için 7 uzman “Tamamen Katılıyorum”, 3 uzman “Katılıyorum” yanıtını vermiştir. Doğrudan Uzlaşma yöntemi için 6 uzman “Tamamen Katılıyorum”, 4 uzman “Katılıyorum” yanıtını vermiştir. Uzmanlardan bu maddeye “Kararsızım”, “Katılmıyorum” ve “Hiç Katılmıyorum” yanıtını veren olmamıştır. Uzmanların yöntemler hakkındaki bilgilendirmenin yeterli olduğu görüşüne sahip oldukları görülmektedir.

İkinci maddede yöntemlerin uygulanma sürecine ilişkin verilen talimatların açık olduğu görüşüne uzmanların ne derecede katıldığı araştırılmıştır. Angoff yöntemi için 6 uzmandan “Tamamen Katılıyorum”, 2 uzmandan “Katılıyorum”, 1 uzmandan “Katılmıyorum”, 1 uzmandan “Hiç Katılmıyorum” yanıtı alınmıştır. Aynı madde Doğrudan Uzlaşma yöntemine ilişkin sorulduğunda 7 uzman “Tamamen Katılıyorum”, 3 uzman “Katılıyorum” yanıtını vermiştir. İki yönteme

ait verilen yanıtlar incelendiğinde uzmanların Doğrudan Uzlaşma yönteminin uygulama talimatlarını daha açık ve anlaşılır bulduğu tespit edilmiştir.

Üçüncü maddede örnek uygulamaların yöntemin daha net anlaşılmasını sağladığı görüşüne Angoff yöntemi için 5 uzman “Tamamen Katılıyorum” yanıtını verirken Doğrudan Uzlaşma yöntemi için 9 uzman “Tamamen Katılıyorum” yanıtını vermiştir. Uzmanların örnek uygulamaları Doğrudan Uzlaşma yöntemi üzerinde daha etkili bulduğu görülmektedir.

Dördüncü maddede uzmanlara yöntemi uygulamaya geçme konusunda kendisini ne kadar hazır hissettiği sorulmuştur. Angoff yöntemini uygulamaya geçme konusunda 1 uzman “Kararsız” olduğunu belirtirken, 1 uzman “Katılmıyorum” yanıtı vererek bu konuda kendisini hazır hissetmediğini belirtmiştir. Aynı madde Doğrudan Uzlaşma yöntemi için sorulduğunda 8 uzman “Tamamen Katılıyorum”, 2 uzman “Katılıyorum” yanıtını vermiştir. Yanıtlar incelendiğinde Angoff yöntemini uygulamaya geçme konusunda kendisini hazır hissetmeyen uzmanların olduğu fakat bu uzmanların Doğrudan Uzlaşma yöntemini uygulama konusunda kendisini hazır hissettiği görülmektedir.

Beşinci maddede uzmanların yöntemi uygulama sırasında kendisini ne derecede rahat hissettiği sorulmuştur. Yöntemi uygularken kendisini rahat hissettiği görüşüne, Doğrudan Uzlaşma yöntemi için 8 uzman “Tamamen Katılıyorum” yanıtını verirken, Angoff yöntemi için 5 uzman “Tamamen Katılıyorum” yanıtını vermiştir. Yanıtlar incelendiğinde Doğrudan Uzlaşma yönteminin Angoff yöntemine göre daha fazla uzman tarafından rahat uygulandığı tespit edilmiştir.

Altıncı maddede uzmanların yöntemleri zaman açısından değerlendirmeleri istenmiştir. Angoff yönteminin zaman açısından kullanışlı olduğu görüşüne 2 uzman “Hiç Katılmıyorum”, 2 uzman “Katılmıyorum” yanıtını vermiştir. Aynı maddeden Doğrudan Uzlaşma yöntemine ilişkin alınan yanıtlar da incelendiğinde uzman çoğunluğunun Angoff yöntemini zaman açısından kullanışlı bulmadıkları sonucuna varılmıştır.

Yedinci madde ile uzmanların yöntemleri bilişsel açıdan ne derecede kolay anlaşılır bulduğu ölçülmek istenmiştir. Bu madde Angoff yöntemine ilişkin sorulduğunda 1 uzman “Hiç Katılmıyorum”, 1 uzman “Katılmıyorum” 2 uzman

“Kararsızım”, 2 uzman “Katılıyorum ve 4 uzman “Tamamen Katılıyorum” yanıtını vermiştir. Aynı madde Doğrudan Uzlaşma yöntemine ilişkin sorulduğunda ise uzmanların tamamı “Tamamen Katılıyorum” veya “Katılıyorum” yanıtını vermiştir. İki yöntem karşılaştırıldığında Doğrudan Uzlaşma yönteminin uzmanlar tarafından bilişsel açıdan daha kolay anlaşılır bulunduğu görülmektedir.

Sekizinci maddede yöntemin uygulanma sürecinde soru sorma ve düşünme fırsatı olduğu görüşüne uzmanların ne derecede katıldığı sorulmuştur. Uzmanların tamamı Doğrudan Uzlaşma yöntemini uygularken soru sorabildiğini ve düşünme fırsatı olduğunu belirtmiştir. Bu maddeyi Angoff yöntemi için 1 uzman “Katılmıyorum” şeklinde yanıtlarken 1 uzman “Kararsızım” şeklinde yanıtlamıştır.

Dokuzuncu madde ile uzmanlara yöntemleri kullanarak verdikleri kararlara ne derecede güven duydukları sorulmuştur. Bu maddeye Angoff yöntemine ilişkin “Tamamen Katılıyorum” yanıtını veren 3 uzman olduğu görülmektedir. Doğrudan Uzlaşma yöntemine ilişkin “Tamamen Katılıyorum” yanıtı 7 uzman tarafından verilmiştir. Yanıtlar incelendiğinde uzmanların çoğunluğunun Doğrudan Uzlaşma yöntemi ile bulduğu puanlara güvendiği, Angoff yöntemi ile bulduğu puanlara güvenen uzman sayısının ise daha az olduğu tespit edilmiştir.

5. TARTIŞMA

Angoff ve Doğrudan Uzlaşma standart belirleme yöntemlerinin karşılaştırılmasının yapıldığı bu çalışmada birinci alt probleme ilişkin bulgular incelendiğinde her iki yöntemin de benzer kesme puanları ürettiği görülmektedir. Bu sonuçlar Sireci, Hambleton ve Pitoniak'ın (2004) çalışmasıyla örtüşmektedir. Ayrıca iki yöntemin kesme puanlarının üzerinde puan alarak üst performans seviyesine yerleşen ve başarılı sayılan öğrenci yüzdeleri de birbirine yakın sonuçlar vermiştir. Bunun sebebinin kesme puanlarının birbirine benzer olmasından kaynaklandığı yorumu yapılabilir. Irwin, (2007)'e göre birden fazla yöntem kullanıldığında benzer kesme puanları elde ediliyorsa bu kesme puanının savunulabilirliğini yükseltir.

İkinci ve üçüncü alt problemlere ait bulgular incelendiğinde iki yöntem arasında öğrencilerin başarılı-başarısız olarak sınıflandırılması bakımından mükemmel uyum olduğu ve uzmanlara ait MGP'lerin yüksek düzeyde ilişkili olduğu görülmektedir. Yöntemlerin işleyiş sürecindeki benzerliklerin bu sonuçları desteklemiş olabileceği düşünülebilir. Uzmanların her iki yöntemde “minimum yeterlik seviyesindeki öğrenci” kavramına odaklanarak kararlarını vermesi öğrencilerin sınıflandırılması bakımından oluşan mükemmel uyumu ve MGP'lerin arasındaki yüksek ilişkiyi açıklayabilir. Öte yandan yöntemlerin ikisinin de test merkezli olması sebebiyle uzmanların test maddelerine odaklanarak aynı madde için benzer kararlar vermiş olabileceği düşünülmektedir.

Dördüncü alt problemin bulguları incelendiğinde standart belirleme sürecinde çalışılan uzmanların Doğrudan Uzlaşma yöntemine Angoff yöntemine göre daha yüksek olumlu görüş geliştirdiği görülmüştür. Angoff yöntemine ilişkin bahsedilen “minimum yeterlik seviyesinde 100 öğrenci” tanımının uzmanlara yüklediği bilişsel zorluk sebebiyle uzmanların Doğrudan Uzlaşma yöntemini daha anlaşılır ve daha kolay uygulanabilir bulduğu düşünülebilir.

Angoff yöntemi maddeleri tek tek inceleyip her madde için bir karar verilmesini gerektirmektedir. Doğrudan Uzlaşma yönteminde ise alt testler bütün halinde incelenip her alt test için bir karar verilmesi gerekmektedir. Uzmanların iki

yöntem arasındaki bu uygulama farklılığından etkilenecek Doğrudan Uzlaşma yöntemini daha kullanışlı bulduğu düşünülmektedir.

Mevcut araştırmada uzmanlar ilk olarak Angoff yöntemini ardından Doğrudan Uzlaşma yöntemini uygulamışlardır. Angoff yönteminin uygulama aşamasında uzmanlar; standart belirleme, kesme puanı, minimum yeterlik düzeyi gibi kavramları tanımış ve Doğrudan Uzlaşma yönteminin uygulamasına geçildiğinde bu kavramlar konusunda kendilerini rahat hissetmiş olabilir. Bu durumun uzmanların Doğrudan Uzlaşma yöntemine karşı daha yüksek olumlu görüş geliştirmesinde etkili olabileceği düşünülmektedir.

İki yöntemin kesme puanlarının, başarılı öğrenci yüzdelerinin ve minimum geçme puanlarının benzer olduğu düşünülürse kesme puanı belirlemek için Doğrudan Uzlaşma yönteminin kullanılması kullanışlılık ve ekonomiklik açısından önerilebilir. Doğrudan uzlaşma yönteminin Angoff yöntemine göre sınırlılığı ise kullanılacak testin kendisinden daha homojen alt testlere bölünebilecek özellikte olması koşuludur.

6. SONUÇLAR VE ÖNERİLER

Araştırmanın bu bölümünde bulgulardan yola çıkılarak ulaşılan sonuçlar ve bu sonuçların dikkate alınmasıyla yapılan öneriler yer almaktadır.

6.1 Sonuçlar

Araştırmanın bulgularından elde edilen sonuçlar alt problemlerin sırası dikkate alınarak aşağıda açıklanmıştır.

6.1.1 Birinci Alt Problemin Çözümlemesiyle Elde Edilen Sonuçlar

24 Soruluk matematik başarı testinin kesme puanı 10 uzman tarafından Angoff ve Doğrudan Uzlaşma yöntemine göre belirlenmiştir. Angoff yöntemine göre hesaplanan kesme puanının (52,04) Doğrudan Uzlaşma yöntemine göre hesaplanan kesme puanından (49,17) daha yüksek değer aldığı görülmüştür. Kesme puanları arasında oluşan bu farklılık başarılı sayılan öğrenci sayılarının ve başarılı sayılan öğrenci yüzdelerinin farklı değerler almasına sebep olmuştur. İki yöntem arasındaki bu farklılığın anlamlı olup olmadığı test edilmiş ve farklılık anlamlı bulunmamıştır. Bu bulgudan hareketle iki yöntemin ürettiği kesme puanları ile yapılan değerlendirmeler sonucunda başarılı sayılan öğrenci yüzdelerinin benzer olduğu sonucuna ulaşılmıştır.

6.1.2 İkinci Alt Problemin Çözümlemesiyle Elde Edilen Sonuçlar

Angoff ve Doğrudan Uzlaşma yöntemleri arasında öğrencilerin başarılı-başarısız sınıflandırılması bakımından mükemmel derecede uyum ve pozitif yönlü yüksek derecede ilişki olduğu sonucuna ulaşılmıştır.

6.1.3 Üçüncü Alt Problemin Çözümlemesiyle Elde Edilen Sonuçlar

Yöntemlere ait MGP'ler arasındaki korelasyon Pearson Momentler Çarpımı Korelasyon Katsayısı kullanılarak test edilmiş ve MGP'ler arasında yüksek düzeyde uyum olduğu sonucu ortaya çıkmıştır. MGP'lerin ortalamaları arasındaki farklılık sınanmış ve farklılığın anlamlı olmadığı görülmüştür.

6.1.4 Dördüncü Alt Problemin Çözümlemesiyle Elde Edilen Sonuçlar

Uzmanların yöntemlere ilişkin görüşleri iki yöntemin karşılaştırmasını yapmak amacıyla betimsel olarak incelenmiştir. Uzman grubunun tamamı uygulama toplantısında yapılan bilgilendirmeleri her iki yöntem için de yeterli bulmuştur. Buna rağmen Angoff yöntemiyle ilgili verilen talimatların açık

olmadığını düşünen uzmanların olduğu görülmüştür. Formdaki maddelerin içeriğinde bulunan “örnek uygulamalarla yöntemi daha net anlama”, “uygulamaya geçme konusunda hazır hissetme”, “uygulama sırasında rahat hissetme”, “yöntemi zaman açısından kullanışlı bulma”, “uygulama sırasında soru sorma fırsatı bulabilme” ve “uygulama sonucunda bulunan puanlara güven duyma” durumları açısından yöntemler karşılaştırıldığında Doğrudan Uzlaşma yöntemine ilişkin olumlu görüş bildiren uzman sayısının Angoff yöntemine ilişkin olumlu görüş bildiren uzman sayısından daha yüksek olduğu tespit edilmiştir.

6.2 Öneriler

6.2.1 Araştırma sonuçlarından çıkarılan öneriler

- Angoff ve Doğrudan Uzlaşma yöntemleri ile elde edilen kesme puanlarının üzerinde puan alarak üst performans seviyesine yerleşen ve başarılı sayılan öğrenci yüzdeleri, kesme puanına göre başarılı-başarısız olarak sınıflandırılma durumları ve MGP’leri uyumlu olduğu için herhangi bir testte iki yöntem birlikte kullanılacaksa Angoff ve Doğrudan Uzlaşma yöntemlerinden sadece birinin seçilmesi ekonomiklik açısından önerilebilir.
- Doğrudan Uzlaşma yönteminin Angoff yöntemine göre uzmanlar tarafından daha kolay anlaşılması ve daha kısa sürede tamamlanması sebebiyle homojen alt testlere bölünebilecek yapıda olan testlerde kullanışlılık açısından Doğrudan Uzlaşma yönteminin seçilmesi önerilebilir.

6.2.2 Araştırmacılara yönelik öneriler

- Bu araştırmada standart belirleme yöntemlerinden Angoff ve Doğrudan Uzlaşma yöntemleri karşılaştırılmıştır. Benzer bir araştırmada yöntemler çoğaltılabilir ya da değiştirilebilir.
- Uzmanlar yöntemleri kullanırken bir önceki yöntemde verdiği kararı hatırlayarak bundan etkilenebilir. Bu etkinin derecesini tespit etmek için yöntem uygulamaları arasındaki zaman aralığı genişletilerek sonuçlar karşılaştırılabilir.
- Uzmanların alanlarındaki deneyim durumlarına ve görev yaptıkları kuruma göre verdikleri kararlar karşılaştırılabilir.

- Ülkemizde MEB ve OSYM tarafından uygulanan testlerin kesme puanı ile bu testlerin Doğrudan Uzlaşma yöntemiyle elde edilen kesme puanları karşılaştırılabilir.
- Kullanılan testteki maddelerin madde güçlüğü ile uzmanların o maddeye dair verdiği kararlar karşılaştırılabilir.
- Bu araştırma 5. sınıf düzeyinde çoktan seçmeli soru tipinde geliştirilen matematik başarı testi üzerinden yürütülmüştür. Sınıf seviyesi, soru tipi ya da disiplin değiştirilerek benzer bir araştırma yapılabilir.



7. KAYNAKLAR

(Bu tez çalışmasında APA 7th Edition atıf sistemi kullanılmıştır.)

- Akyıldız, M. (2015). Açık ve uzaktan öğretimde ölçme ve değerlendirme politikaları. *Açık öğretim Uygulamaları ve Araştırmaları Dergisi*, 1(3), 8-25.
- Baykul, Y., & Tertemiz, N. (2004). İlköğretim birinci, ikinci ve üçüncü sınıf matematik programı üzerine bir değerlendirme. *Eğitim ve Bilim*, 29(131)
- Baykul, Y. (2021). Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması. Pegem Akademi.
- Berk, R. A. (1986). A consumer's guide to setting performance standards on criterion-referenced tests. *Review of Educational Research*, 56, 137-172
- Bejar, I.I., (2008). Standard setting: What is it? Why is it important? *R&D Connections, Educational Testing Service*, No.7.
- Buckendahl, C. W., Smith, R. W., Impara, J. C., & Plake, B. S. (2002). A comparison of Angoff and Bookmark standard setting methods. *Journal of Educational measurement*, 39(3), 253-263.
- Clauser, B. E., Harik, P., Margolis, M. J., McManus, I. C., Mollon, J., Chis, L., & Williams, S. (2008). An empirical examination of the impact of group discussion and examinee performance information on judgments made in the Angoff standard-setting procedure. *Applied Measurement in Education*, 22(1), 1-21.
- Beuk, C. H. (1984). A method for reaching a compromise between absolute and relative standards in examinations. *Journal of Educational Measurement*, 21(2), 147-152
- Cizek, G. J. (2001). *Setting performance standards. Concepts, methods, and perspectives*, 2001.
- Cizek, G.J., & Bunch, M. B. (2007). *Standard Setting: A Guide to Establishing and Evaluating Performance Standards on Tests*. Sage Publication
- Cizek, G.J., (2012). An introduction to contemporary standard setting: Concepts, characteristics, and contexts. *Setting performance standards: Foundations, methods, and innovations*, 3-14.
- Cizek, G. J., Bunch, M. B., & Koons, H. (2004). Setting performance standards: Contemporary methods. *Educational measurement: issues and practice*, 23(4), 31-31.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1), 37-46.
- Çetin, S. (2011). İşaretleme ve angoff standart belirleme yöntemlerinin karşılaştırılması [Comparison of Bookmark and Angoff Standard Setting Methods]. *Tez Numarası*, 308506.
- Devries, T. (1973) *Descriptive Research and Problem Solving. Methods and Techniques of Educational Research*, 99-107.
- Ertürk, S. (2013). Eğitimde “program” geliştirme (6. baskı). Ankara: Edge Akademi Yayınları, 42.
- Fowell, S. L., Fewtrell, R., & McLaughlin, P. J. (2008). Estimating the minimum number of judges required for test-centred standard setting on written assessments. Do discussion and iteration have an influence?. *Advances in health sciences education*, 13(1), 11-24.
- Fidan, N., & Baykul, Y. (1994). İlköğretimde temel öğrenme ihtiyaçlarının karşılanması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 10(10).
- Hambleton, R.K. (1998). Setting performance standards on achievement tests: Meeting the requirements of Title I. In *Handbook for the development of performance standards: Meeting*

- the requirements of Title I (pp. 97-114). Washington, DC: Council of Chief State School Officers.
- Hambleton, R. K. (2013). Setting performance standards on educational assessments and criteria for evaluating the process. In *Setting performance standards* (pp. 103-130). Routledge
- Hauke, J., & Kossowski, T. (2011). Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data. *Quaestiones geographicae*, 30(2), 87-93.
- Hein, S. F., & Skaggs, G. E. (2009). A qualitative investigation of panelists' experiences of standard setting using two variations of the bookmark method. *Applied Measurement in Education*, 22(3), 207-228.
- Hurtz, G. M., & Hertz, N. R. (1999). How many raters should be used for establishing cutoff scores with the Angoff method? A generalizability theory study. *Educational and Psychological Measurement*, 59(6), 885-897.
- Impara, J. C., & Plake, B. S. (1998). Teachers' ability to estimate item difficulty: A test of the assumptions in the Angoff standard setting method. *Journal of Educational Measurement*, 35(1), 69-81
- Irwin, P. M. (2007). An alternative examinee-centered standard setting strategy. The University of Nebraska-Lincoln.
- Jeager, R. M. (1989). Certification of student competence. R. L. Linn (ed), *Educational measurement* (ss. 485-514) içinde. New York: American Council on Education/Macmillan
- Kane, M. (1994). Validating the performance standards associated with passing scores. *Review of Educational Research*, 63(4), 425-461.
- Kane, M. T. (1998). Choosing between examinee-centered and test-centered standard-setting methods. *Educational Assessment*, 5(3), 129-145
- Karasar, N. (2007). Bilimsel araştırma yöntemi: kavramlar, ilkeler, teknikler. Nobel yayın dağıtım
- Keuning, J., Straat, J. H., & Feskens, R. C. (2017). The data-driven direct consensus (3dc) procedure: A new approach to standard setting. *Standard Setting in Education: The Nordic Countries in an International Perspective*, 263-278.
- Kubiszyn, T. ve Borich G. (2007). *Educational testing and measurement classroom application and practice*. USA: Wiley Jossey-Bass Education.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.
- Livingston, S.A., Zieky J. M., (1982). *Passing Scores: A Manual for Setting Standards of Performance on Educational and Occupational Tests*, ETS
- Maurer, T. J., Alexander, R. A., Callahan, C. M., Bailey, J. J., & Dambrot, F. H. (1991). Methodological and psychometric issues in setting cutoff scores using the Angoff method. *Personnel Psychology*, 44(2), 235-262.
- Meb(2013). Milli Eğitim Bakanlığı Ortaöğretim Kurumları Yönetmeliği. Ankara. Resmî Gazete Tarihi: 07.09.2013 Resmî Gazete Sayısı: 28758.
- Meb(2014). Milli Eğitim Bakanlığı Ortaöğretim Kurumları Yönetmeliği. Ankara. Resmî Gazete Tarihi: 26.07.2014 Resmî Gazete Sayısı: 29072.
- Mehrens, W. A. (1995). Methodological issues in standard setting for educational exams. In *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments*, 221-263.
- Mehrens, W. A. L., & Irvin, J. (1973). *Measurement And Evaluation In Education And Psychology*/William A. Mehrens And Irvin J. Lehmann (No. Lb1051. M43 1991.).

- Morgan, G. A., Leech, N. L., Gloeckner, G. W., & Barrett, K. C. (2004). SPSS for introductory statistics: Use and interpretation. Psychology Press.
- Murphy, K. R., & Davidshofer, C. O. (1988). Psychological testing. Principles, and Applications, Englewood Cliffs, 18.
- Mills, C. N., & Melican, G. J. (1988). Estimating and adjusting cutoff scores: Features of selected methods. *Applied Measurement in Education*, 1, 261–275.
- Norcini, J. J., Lipner, R. S., Langdon, L. O., & Strecker, C. A. (1987). A comparison of three variations on a standard - setting method. *Journal of Educational Measurement*, 24(1), 56-64.
- Norcini, J. J., & Shea, J. A. (1997). The credibility and comparability of standards. *Applied Measurement in education*, 10(1), 39-59.
- Özbek, H., & Keskin, S. (2007). Standart sapma mı yoksa standart hata mı. *Van tıp dergisi*, 14(2), 64-67.
- Özçelik, D. A. (1992). Eğitim Programları ve Öğretimi (Genel Öğretim Yöntemi). Ankara: ÖSYM Yayınları.
- Pitoniak, M. J. (2002). Investigation of two standard setting methods for a licensure examination. University of Massachusetts Amherst, 9-10.
- Pitoniak, M. J. (2003). Standard setting methods for complex licensure examinations. University of Massachusetts Amherst.
- Razali, N. M., & Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of statistical modeling and analytics*, 2(1), 21-33.
- Shen, L. (2001). A comparison of angoff and rasch model based item map methods in standard setting. Paper Presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.
- Shepard, L., Glaser, R., Linn, R., & Bohrnstedt, G. (1993). Setting performance standards for student achievement. National Academy of Education.
- Takiar, R. (2023). The Validity Of T-Test, Mann-Whitney Test And Z Test For Testing Significant Differences Between Two Sample Means When The Sample Size Is Between 10 And 30. *Bulletin Of Mathematics And Statistics Research*, 11(3).
- Tekin, H. (2019). Eğitimde ölçme ve değerlendirme (27. bs.). Yargı Yayınevi.
- Turgut, F. M. (1993). Eğitimde ölçme ve değerlendirme metodları,(Measurement and evaluation methods in education), dokuzuncu baskı. Ankara: Saydam Matbaacılık.
- Turgut, M. F., & Baykul, Y. (2019). Eğitimde ölçme ve değerlendirme (8.Baskı). Ankara: Pegem Akademi.
- Yıldırım, A., & Şimşek, H. (2021). Sosyal Bilimlerde Nitel Araştırma Yöntemleri (12. Baskı). Ankara: Seçkin Yayıncılık.
- Wright, B. D., & Grosse, M. (1993). How to set standards. *Rasch Measurement Transactions*, 7(3), 315-316.
- Zwick, R., Senturk, D., Wang, J., & Loomis, S. C. (2001). An investigation of alternative methods for item mapping in the National Assessment of Educational Progress. *Educational Measurement: Issues and Practice*, 20(2), 15-25.

8. EKLER

EK 1 Matematik Başarı Testi

AÇIKLAMA

Sevgili Öğrenciler,

Bu test matematik dersindeki öğrenme düzeyinizi ölçmek amacıyla hazırlanmıştır. Testte matematik dersinin ‘Sayılar ve İşlemler’ ünitesinden 24 adet soru bulunmaktadır. Her sorudaki 4 seçenekten sadece bir tanesi doğru cevaptır. Soruları dikkatlice okuyunuz ve doğru bulduğunuz seçeneği işaretleyiniz. Testin cevaplama süresi 40 dakikadır.

Lütfen soruları boş bırakmamaya özen gösteriniz.

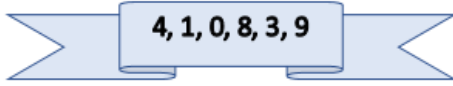
Başarılar dilerim.

Ad-Soyad:

Sınıf:

Numara:

1.



Yukarıdaki rakamları birer kez kullanarak yazılabilecek altı basamaklı en küçük doğal sayı aşağıdakilerden hangisidir?

- A) 104 389 B) 130 489
C) 103 489 D) 401 389

2.

"751 326 487" Doğal sayısında basamak değeri en büyük rakam **▲**, sayı değeri en küçük rakam **■** 'dir.

Buna göre **▲** + **■** kaçtır?

- A) 7 B) 8 C) 9 D) 14

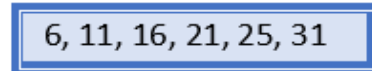
3.

7, 11, 15, 19, A, 27, B

Yukarıdaki sayı örüntüsüne göre A + B toplamı kaçtır?

- A) 52 B) 54 C) 55 D) 56

4.



Sayı örüntüsünde kuralı bozan sayı kaçtır?

- A) 11 B) 16 C) 21 D) 25

5.



Yukarıdaki toplama işlemine göre A + B + C + D kaçtır?

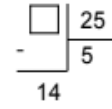
- A) 24 B) 27 C) 29 D) 30

6.

96 öğrenciden oluşan bir öğrenci grubunu geziye götürmek için araç kiralanacaktır. Kiralanacak araçlardan birine en fazla 16 öğrenci binebilmektedir. Buna göre bu grup için en az kaç araç kiralanmalıdır?

- A) 5 B) 6 C) 7 D) 8

7.



Yandaki bölme işlemine göre, bölünen sayı kaçtır?

- A) 125 B) 139 C) 145 D) 150

8.

$$A = (45 \times 3) + 20$$

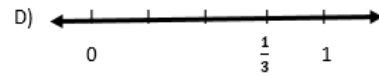
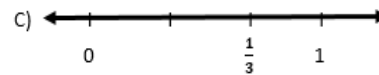
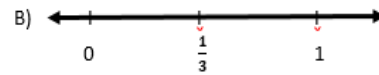
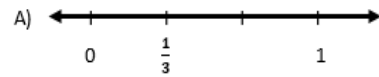
$$B = 45 + (3 \times 20)$$

Yukarıda verilenlere göre A + B kaçtır?

- A) 250 B) 255 C) 260 D) 265

9.

$\frac{1}{3}$ Birim kesri aşağıdaki sayı doğrularının hangisinde doğru şekilde gösterilmiştir?



10.

$$\frac{1}{20} < K < \frac{1}{9}$$

Yukarıda verilenlere göre K yerine aşağıdakilerden hangisi yazılabilir?

- A) $\frac{1}{6}$ B) $\frac{1}{8}$ C) $\frac{1}{19}$ D) $\frac{1}{21}$

11.

$4\frac{1}{6}$ Kesri aşağıdakilerden hangisine eşittir?

- A) $\frac{23}{6}$ B) $\frac{24}{6}$ C) $\frac{25}{6}$ D) $\frac{26}{6}$

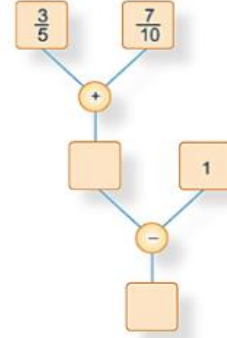
12.

$\frac{3}{8}$	$\frac{5}{9}$	$\frac{2}{3}$
---------------	---------------	---------------

Yukarıda verilen kesirlerin büyükten küçüğe doğru sıralanışı hangi şıkta doğru olarak verilmiştir?

- A) $\frac{3}{8} > \frac{5}{9} > \frac{2}{3}$ B) $\frac{3}{8} > \frac{2}{3} > \frac{5}{9}$
C) $\frac{5}{9} > \frac{3}{8} > \frac{2}{3}$ D) $\frac{2}{3} > \frac{5}{9} > \frac{3}{8}$

13.



İşlem ağacında sıra ile verilen işlemler yapıldığında en alttaki kutuya hangi kesir yazılabilir?

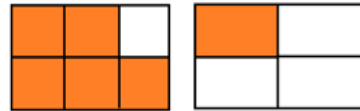
- A) $\frac{13}{10}$ B) $\frac{6}{5}$ C) $\frac{9}{5}$ D) $\frac{3}{10}$

14.

$\frac{7}{4} - \frac{3}{8}$ işleminin sonucu kaçtır?

- A) $\frac{11}{4}$ B) $\frac{11}{8}$ C) $\frac{4}{8}$ D) $\frac{10}{4}$

15.



1.

2.

Yukarıda 1. Şekilde boyalı bölgenin temsil ettiği kesir A, 2. Şekilde boyalı bölgenin temsil ettiği kesir B ile gösteriliyor. Buna göre (A - B) farkı kaçtır?

- A) $\frac{7}{12}$ B) $\frac{8}{12}$ C) $\frac{9}{12}$ D) $\frac{10}{12}$

16.



Sezgin 120 sayfalık bir kitabın önce $\frac{7}{12}$ 'sini, sonra $\frac{1}{6}$ 'sini okumuştur.

Buna göre Sezgin'in okuması gereken kaç sayfa kalmıştır?

- A) 20 B) 30 C) 40 D) 50

17.

96,075

Yukarıda verilen ondalık gösterimin okunuşu aşağıdakilerden hangisidir?

- A. Dokuz tam binde altı yüz yetmiş beş
B. Doksan altı tam binde yetmiş beş
C. Doksan altı tam yüzde yetmiş beş
D. Doksan altı tam onda yetmiş beş

18.

Sekiz tam binde dört yüz dört

Yukarıda okunuşu verilen ondalık gösterim aşağıdakilerden hangisidir?

- A) 80,404 B) 84,04 C) 8,404 D) 8,44

19.



Karıncaların taşıdığı ondalık gösterim ile ilgili aşağıdakilerden hangisi doğru değildir?

- A) "Beş tam binde altı yüz on sekiz" şeklinde okunur.
B) $5\frac{618}{1000}$ şeklinde gösterebiliriz.
C) 1 rakamının basamak değeri 0,01'dir.
D) 6 rakamının basamak değeri 6'dır.

20.

64,057 Ondalık gösteriminin kesir çizgisi ile gösterimi aşağıdakilerden hangisidir?

- A) $64\frac{57}{100}$ B) $64\frac{57}{10}$ C) $64\frac{570}{1000}$ D) $64\frac{57}{1000}$

21.

$\frac{13}{25}$ Kesrinin yüzde sembolü ile gösterilişi aşağıdakilerden hangisidir?

- A) %13 B) %26 C) %25 D) %52

22.

Aşağıdaki gösterimlerden üçünün değeri birbirine eşit birinin değeri farklıdır. Farklı olan hangisidir?

- A) %34 B) $\frac{17}{50}$ C) 3,4 D) $\frac{34}{100}$

23.

Kesir	Yüzde Sembolüyle Gösterimi
$\frac{42}{100}$	%36
$\frac{9}{20}$	%42
$\frac{14}{25}$	%45
	%56

Yukarıdaki kesirler ile yüzde sembolüyle gösterimleri eşleştirilecektir. Bu eşleştirme sonucunda hangi yüzde gösterimi açığa kalır?

- A) %36 B) %42 C) %45 D) %56

24.

"0,4" Ondalık gösteriminin yüzde sembolü ile gösterimi aşağıdakilerden hangisidir?

- A) %40 B) %4 C) %20 D) %50

EK 2 Uzman Görüşleri Formu

UZMAN GÖRÜŞLERİ FORMU

Değerli Öğretmenler,

Yüksek Lisans Tez çalışmasında kullanılmak üzere geliştirilen 24 sorudan oluşan başarı testi için iki farklı değerlendirme yapmanıza ihtiyaç duyulmaktadır. Yaptığınız değerlendirmeler sonucunda teste ait iki farklı kesme puanı belirlenecektir. Kesme puanı başarılı öğrenci ile başarısız öğrenciyi birbirinden ayıran sınır puanıdır. Başarı testi, “Sayılar ve İşlemler” öğrenme alanına ait “Doğal Sayılar”, “Doğal Sayılarla İşlemler”, “Kesirler”, “Kesirlerle İşlemler”, “Ondalık Gösterim” ve “Yüzdeler” olmak üzere 6 adet alt öğrenme alanını kapsamaktadır.

Değerlendirmenizi içtenlikle yapmanızı rica eder, katkılarınız için şimdiden teşekkür ederim.

Gamze ÖZSOY DEMİREL

AİBÜ Eğitimde Ölçme ve Değerlendirme Bilim Dalı Yüksek Lisans Öğrencisi

Değerlendirme 1:

Angoff yöntemiyle kesme puanını belirlemek için sizden, başarı testindeki her bir soru için “Sizce matematik dersinde geçme-kalma sınırı ya da sınır grup olarak ifade edilen minimum yeterlik düzeyindeki 100 öğrenciden kaç öğrenci bu soruyu doğru cevaplayabilir?” sorusuna cevap vermeniz beklenmektedir. Kararınızı sorunun solunda bulunan kutucuğa yazınız.

Örnek:

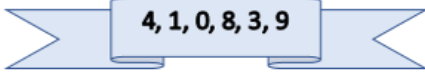
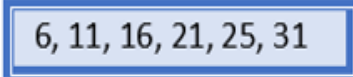
Değerlendirme 1	Sorular	Alt Öğrenme alanı	Değerlendirme 2
65	Arda'nın 14 kalemi vardır. Arda kaç kalem daha alırsa kalemlerinin sayısı 3' ün küpüne eşit olur? A) 13 B) 18 C) 9 D) 27	Doğal Sayılarla İşlemler	
80	Aşağıdaki işlemlerden hangisi doğrudur? A) $4^2 + 3^1 = 17$ B) $5^2 - 3^2 = 4$ C) $2^3 \times 8^1 = 64$ D) $5^3 : 1^3 = 5$		
40	$(3^2 + 4^2) - 5^2$ İşleminin sonucu aşağıdakilerden hangisidir? A) 3 B) 2 C) 1 D) 0		

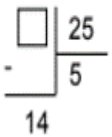
Değerlendirme 2:

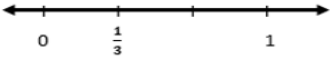
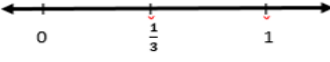
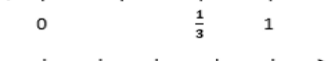
Doğrudan Uzlaşma yöntemiyle kesme puanını belirlemek için sizden, başarı testindeki her biri 4 sorudan oluşan 6 alt öğrenme alanını inceleyip “Sizce matematik dersinde geçme-kalma sınırı ya da sınır grup olarak ifade edilen minimum yeterli seviyesindeki bir öğrencinin bu alt öğrenme alanında başarılı kabul edilmesi için kaç maddeye doğru yanıt vermesi gerekir?” sorusuna cevap vermeniz beklenmektedir. Kararınızı alt öğrenme alanına ait soruların sağında bulunan kutucuğa yazınız.

Örnek:

Değerlendirme 1	Sorular	Alt Öğrenme alanı	Değerlendirme 2
	Arda'nın 14 kalemi vardır. Arda kaç kalem daha alırsa kalemlerinin sayısı 3' ün küpüne eşit olur? A) 13 B) 18 C) 9 D) 27	Doğal Sayılarla İşlemler	2
	Aşağıdaki işlemlerden hangisi doğrudur? A) $4^2 + 3^1 = 17$ B) $5^2 - 3^2 = 4$ C) $2^3 \times 8^1 = 64$ D) $5^3 : 1^3 = 5$		
	$(3^2 + 4^2) - 5^2$ İşleminin sonucu aşağıdakilerden hangisidir? A) 3 B) 2 C) 1 D) 0		

Değerlendirme 1	SORULAR	Alt Öğrenme Alanı	Değerlendirme 2
	1.  4, 1, 0, 8, 3, 9 Yukarıdaki rakamları birer kez kullanarak yazılabilecek altı basamaklı en küçük doğal sayı aşağıdakilerden hangisidir? A) 104 389 B) 130 489 C) 103 489 D) 401 389	DOĞAL SAYILAR	
	2. "751 326 487" Doğal sayısında <u>basamak değeri</u> en büyük rakam ▲, <u>sayı değeri</u> en küçük rakam ■ 'dir. Buna göre ▲ + ■ kaçtır? A) 7 B) 8 C) 9 D) 14		
	3. 7, 11, 15, 19, A, 27, B Yukarıdaki sayı örüntüsüne göre A + B toplamı kaçtır? A) 52 B) 54 C) 55 D) 56		
	4.  6, 11, 16, 21, 25, 31 Sayı örüntüsünde kuralı bozan sayı kaçtır? A) 11 B) 16 C) 21 D) 25		

Değerlendirme 1	SORULAR	Alt Öğrenme Alanı	Değerlendirme 2
	5.  Yukarıdaki toplama işlemine göre A + B + C + D kaçtır? A) 24 B) 27 C) 29 D) 30	DOĞAL SAYILARLA İŞLEMLER	
	6. 96 öğrenciden oluşan bir öğrenci grubunu geziye götürmek için araç kiralanacaktır. Kiralanacak araçlardan birine en fazla 16 öğrenci binebilmektedir. Buna göre bu grup için en az kaç araç kiralanmalıdır? A) 5 B) 6 C) 7 D) 8		
	7.  Yandaki bölme işlemine göre, bölünen sayı kaçtır? A) 125 B) 139 C) 145 D) 150		
	8. $A = (45 \times 3) + 20$ $B = 45 + (3 \times 20)$ Yukarıda verilenlere göre A + B kaçtır) A) 250 B) 255 C) 260 D) 265		

Değerlendirme 1	SORULAR	Alt Öğrenme Alanı	Değerlendirme 2
	9. $\frac{1}{3}$ Birim kesri aşağıdaki sayı doğrularının hangisinde doğru şekilde gösterilmiştir? A)  B)  C)  D) 	KESİRLER	
	10. $\frac{1}{20} < K < \frac{1}{9}$ Yukarıda verilenlere göre K yerine aşağıdakilerden hangisi yazılabilir? A) $\frac{1}{6}$ B) $\frac{1}{8}$ C) $\frac{1}{19}$ D) $\frac{1}{21}$		
	11. $4\frac{1}{6}$ Kesri aşağıdakilerden hangisine eşittir? A) $\frac{23}{6}$ B) $\frac{24}{6}$ C) $\frac{25}{6}$ D) $\frac{26}{6}$		
	12. $\frac{3}{8}$ $\frac{5}{9}$ $\frac{2}{3}$ Yukarıda verilen kesirlerin büyükten küçüğe doğru sıralanışı hangi şıkta doğru olarak verilmiştir? A) $\frac{3}{8} > \frac{5}{9} > \frac{2}{3}$ B) $\frac{3}{8} > \frac{2}{3} > \frac{5}{9}$ C) $\frac{5}{9} > \frac{3}{8} > \frac{2}{3}$ D) $\frac{2}{3} > \frac{5}{9} > \frac{3}{8}$		

Değerlendirme 1	SORULAR	Alt Öğrenme Alanı	Değerlendirme 2
	13. İşlem ağacında sıra ile verilen işlemler yapıldığında en alttaki kutuya hangi kesir yazılabilir? A) $\frac{13}{10}$ B) $\frac{6}{5}$ C) $\frac{9}{5}$ D) $\frac{3}{10}$	KESİRLERLE İŞLEMLER	
	14. $\frac{7}{4} - \frac{3}{8}$ işleminin sonucu kaçtır? A) $\frac{11}{4}$ B) $\frac{11}{8}$ C) $\frac{4}{8}$ D) $\frac{10}{4}$		
	15. 1. 2. Yukarıda 1. Şekilde boyalı bölgenin temsil ettiği kesir A, 2. Şekilde boyalı bölgenin temsil ettiği kesir B ile gösteriliyor. Buna göre (A - B) farkı kaçtır? A) $\frac{7}{12}$ B) $\frac{8}{12}$ C) $\frac{9}{12}$ D) $\frac{10}{12}$		
	16. Sezgin 120 sayfalık bir kitabın önce $\frac{7}{12}$'sini, sonra $\frac{1}{6}$'sini okumuştur. Buna göre Sezgin'in okuması gereken kaç sayfa kalmıştır? A) 20 B) 30 C) 40 D) 50		

Değerlendirme 1	SORULAR	Alt Öğrenme Alanı	Değerlendirme 2
	17. 96,075 Yukarıda verilen ondalık gösterimin okunuşu aşağıdakilerden hangisidir? A. Dokuz tam binde altı yüz yetmiş beş B. Doksan altı tam binde yetmiş beş C. Doksan altı tam yüzde yetmiş beş D. Doksan altı tam onda yetmiş beş	ONDALIK GÖSTERİM	
	18. Sekiz tam binde dört yüz dört Yukarıda okunuşu verilen ondalık gösterim aşağıdakilerden hangisidir? A) 80,404 B) 84,04 C) 8,404 D) 8,44		
	19.  Karıncaların taşıdığı ondalık gösterim ile ilgili aşağıdakilerden hangisi doğru değildir? A) "Beş tam binde altı yüz on sekiz" şeklinde okunur. B) $5\frac{618}{1000}$ şeklinde gösterebiliriz. C) 1 rakamının basamak değeri 0,01'dir. D) 6 rakamının basamak değeri 6'dır.		
	20. 64,057 Ondalık gösteriminin kesir çizgisi ile gösterimi aşağıdakilerden hangisidir? A) $64\frac{57}{100}$ B) $64\frac{57}{10}$ C) $64\frac{570}{1000}$ D) $64\frac{57}{1000}$		

Değerlendirme 1	SORULAR	Alt Öğrenme Alanı	Değerlendirme 2										
	21. $\frac{13}{25}$ Kesrinin yüzde sembolü ile gösterilişi aşağıdakilerden hangisidir? A) %13 B) %26 C) %25 D) %52	YÜZDELER											
	22. Aşağıdaki gösterimlerden üçünün değeri birbirine eşit birinin değeri farklıdır. Farklı olan hangisidir? A) %34 B) $\frac{17}{50}$ C) 3,4 D) $\frac{34}{100}$												
	23. <table><thead><tr><th>Kesir</th><th>Yüzde Sembolüyle Gösterimi</th></tr></thead><tbody><tr><td>$\frac{42}{100}$</td><td>%36</td></tr><tr><td>$\frac{9}{20}$</td><td>%42</td></tr><tr><td>$\frac{14}{25}$</td><td>%45</td></tr><tr><td></td><td>%56</td></tr></tbody></table> Yukarıdaki kesirler ile yüzde sembolüyle gösterimleri eşleştirilecektir. Bu eşleştirme sonucunda hangi yüzde gösterimi açıkta kalır? A) %36 B) %42 C) %45 D) %56		Kesir	Yüzde Sembolüyle Gösterimi	$\frac{42}{100}$	%36	$\frac{9}{20}$	%42	$\frac{14}{25}$	%45		%56	
Kesir	Yüzde Sembolüyle Gösterimi												
$\frac{42}{100}$	%36												
$\frac{9}{20}$	%42												
$\frac{14}{25}$	%45												
	%56												
	24. “0,4” Ondalık gösteriminin yüzde sembolü ile gösterimi aşağıdakilerden hangisidir? A) %40 B) %4 C) %20 D) %50												

EK 3 Standart Belirleme Yöntem ve Süreçlerini Değerlendirme Formu

STANDART BELİRLEME YÖNTEM VE SÜREÇLERİNİ DEĞERLENDİRME FORMU-1

Değerli Öğretmenler,

Bu form Angoff standart belirleme yöntemine yönelik görüşlerinizi almak amacıyla hazırlanmıştır.

Her bir ifadede belirtilen görüşe katılma derecenize göre “Tamamen Katılıyorum”, “Katılıyorum”, “Kararsızım”, “Katılmıyorum”, “Hiç Katılmıyorum” seçeneklerinden size uygun olan seçeneği işaretleyiniz.

Formdan elde edilen bilgiler standart belirleme yöntem ve süreçlerini değerlendirmek amacıyla kullanılacaktır. Form üzerine isim belirtmeniz gerekmemektedir. Katkılarınız için teşekkür ederim.

İFADELER	Hiç Katılmıyorum	Katılmıyorum	Kararsızım	Katılıyorum	Tamamen Katılıyorum
Angoff Yöntemi ile ilgili bilgilendirme yeterliydi.					
Angoff Yöntemi'ni uygulayabilmem için verilen talimatlar açıktı.					
Örnek uygulamalar, Angoff Yöntemi'ni daha net anlamamı sağladı.					
Angoff Yöntemi'ni uygulamaya geçme konusunda kendimi hazır hissettim.					
Angoff Yöntemi'ni uygularken kendimi rahat hissettim.					
Angoff Yöntemi'ni zaman açısından kullanışlı buldum.					
Angoff Yöntemi'ni bilişsel açıdan kolay anlaşılır buldum.					
Angoff Yöntemi sürecinde soru sorma ve düşünme fırsatım oldu.					
Angoff Yöntemi'ni kullanarak bulduğum puanlara güveniyorum.					

STANDART BELİRLEME YÖNTEM VE SÜREÇLERİNİ DEĞERLENDİRME FORMU-2

Değerli Öğretmenler,

Bu form Doğrudan Uzlaşma standart belirleme yöntemine yönelik görüşlerinizi almak amacıyla hazırlanmıştır.

Her bir ifadede belirtilen görüşe katılma derecenize göre “Tamamen Katılıyorum”, “Katılıyorum”, “Kararsızım”, “Katılmıyorum”, “Hiç Katılmıyorum” seçeneklerinden size uygun olan seçeneği işaretleyiniz.

Formdan elde edilen bilgiler standart belirleme yöntem ve süreçlerini değerlendirmek amacıyla kullanılacaktır. Form üzerine isim belirtmeniz gerekmemektedir. Katkılarınız için teşekkür ederim.

İFADELER	Hiç Katılmıyorum	Katılmıyorum	Kararsızım	Katılıyorum	Tamamen Katılıyorum
Doğrudan Uzlaşma Yöntemi ile ilgili bilgilendirme yeterliydi.					
Doğrudan Uzlaşma Yöntemi'ni uygulayabilmem için verilen talimatlar açtı.					
Örnek uygulamalar, Doğrudan Uzlaşma Yöntemi'ni daha net anlamamı sağladı.					
Doğrudan Uzlaşma Yöntemi'ni uygulamaya geçme konusunda kendimi hazır hissettim.					
Doğrudan Uzlaşma Yöntemi'ni uygularken kendimi rahat hissettim.					
Doğrudan Uzlaşma Yöntemi'ni zaman açısından kullanışlı buldum.					
Doğrudan Uzlaşma Yöntemi'ni bilişsel açıdan kolay anlaşılır buldum.					
Doğrudan Uzlaşma Yöntemi sürecinde soru sorma ve düşünme fırsatım oldu.					
Doğrudan Uzlaşma Yöntemi'ni kullanarak bulduğum puanlara güveniyorum.					

EK 4 Yöntemlere Göre Uzman Kararları

Angoff Yöntemi Uzman Kararları

ANGOFF	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s15	s16	s17	s18	s19	s20	s21	s22	s23	s24
UZMAN 1	60	40	70	60	45	40	40	45	70	30	50	20	30	45	40	35	60	65	50	50	40	40	40	45
UZMAN 2	80	80	90	90	60	60	70	70	60	50	70	30	30	50	50	30	40	40	35	40	60	50	40	35
UZMAN 3	70	75	90	90	85	65	85	95	85	60	80	70	85	85	65	70	90	90	75	65	60	60	65	50
UZMAN 4	85	75	90	95	65	75	80	90	70	60	90	75	80	90	70	70	90	95	90	95	80	80	80	75
UZMAN 5	40	15	30	10	10	10	15	45	15	10	35	40	30	45	35	5	40	50	30	40	50	35	20	45
UZMAN 6	20	15	75	80	55	40	45	15	20	10	30	10	15	15	15	5	30	30	20	35	35	20	20	25
UZMAN 7	90	60	90	90	50	40	60	60	30	40	50	20	30	40	40	20	90	90	30	30	70	50	30	30
UZMAN 8	85	20	70	75	50	65	70	40	80	20	85	30	15	70	20	40	70	90	40	35	70	65	50	40
UZMAN 9	60	40	80	80	45	30	70	50	60	30	50	60	30	30	50	20	60	80	80	40	60	30	40	35
UZMAN 10	75	60	70	60	55	45	45	70	65	60	30	40	30	40	40	15	70	65	75	30	60	50	20	30

Doğrudan Uzlaşma Yöntemi Uzman Kararları

Doğrudan Uzlaşma	1. Alt Öğrenme Alanı	2. Alt Öğrenme Alanı	3. Alt Öğrenme Alanı	4. Alt Öğrenme Alanı	5. Alt Öğrenme Alanı	6. Alt Öğrenme Alanı
UZMAN 1	2	2	2	1	2	2
UZMAN 2	3	2	1	1	2	2
UZMAN 3	3	2	2	3	2	2
UZMAN 4	3	2	2	3	2	2
UZMAN 5	2	2	2	2	2	2
UZMAN 6	2	2	1	1	2	2
UZMAN 7	3	2	1	1	2	2
UZMAN 8	3	2	2	1	2	2
UZMAN 9	3	2	2	1	2	2
UZMAN 10	2	2	2	1	2	2