

**EVET/HAYIR, EBEL VE İŞARETLEME STANDART
BELİRLEME YÖNTEMLERİNİN KARŞILAŞTIRILMASI**

**COMPARING YES/NO, EBEL AND BOOKMARK
STANDARD SETTING METHODS**

Sevda KORKMAZ

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı İçin
Öngördüğü

Yüksek Lisans Tezi

olarak hazırlanmıştır.

2015

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼'ne,

Sevda KORKMAZ'ın hazırladıđı “Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemlerinin karşılaştırılması” başlıklı bu çalışma j¼rimiz tarafından **Eđitim Bilimleri Anabilim Dalı, Eđitimde Ölçme ve Deđerlendirme Bilim Dalı'nda Yüksek Lisans Tezi** olarak kabul edilmiştir.

Başkan Prof. Dr. Selahattin GELBAL

Üye
(Danışman) Dr. Zafer ÇEPNİ

Üye Yrd. Doç. Dr. Devrim ERDEM KEKLİK

Üye Yrd. Doç. Dr. Gülden KAYA UYANIK

Üye Yrd. Doç. Dr. Sevda ÇETİN

ONAY

Bu tez Hacettepe Üniversitesi Lisansüstü Eđitim-Öđretim ve Sınav Yönetmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri üyeleri tarafından 03 / 07/ 2015 tarihinde uygun gör¼lmüş ve Enstitü Yönetim Kurulunca / / tarihinde kabul edilmiştir.

Prof. Dr. Berrin AKMAN
Eđitim Bilimleri Enstitüsü M¼d¼r¼

EVET/HAYIR, EBEL VE İŞARETLEME STANDART BELİRLEME YÖNTEMLERİNİN KARŞILAŞTIRILMASI

Sevda KORKMAZ

ÖZ

Bu araştırmada Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleriyle elde edilen geçme puanları ve uzmanlardan alınan kararlar çeşitli açılardan karşılaştırılmıştır.

Araştırmanın öğrenciler ve uzmanlardan oluşan iki farklı veri grubu vardır. Araştırmada veri toplama aracı olarak 9. sınıf “Kümeler” ünitesinin sorularından oluşan 20 soruluk çoktan seçmeli test geliştirilmiştir. Ayrıca her bir uzmanın üç farklı standart belirleme yöntemine göre verdiği kararları almak için bilgilendirme formları ve sürecin değerlendirilebilmesini sağlayan anket kullanılmıştır. Araştırma kapsamında geliştirilen başarı testini 505 9. sınıf öğrencisi cevaplamış ve 30 uzmandan bu başarı testindeki maddeler için üç yöntemle göre ayrı ayrı görüşler alınmıştır.

Araştırmada, üç farklı standart belirleme yöntemine göre başarılı sayılan öğrenci yüzdeleri arasında fark olup olmadığı bağımlı iki oran arasındaki farkın testiyle, bu farkın anlamlılığı ise z testiyle sınanmıştır. Üç yöntemle elde edilen geçme puanlarına göre öğrencilerin puanları başarılı-başarısız olarak sınıflandırılmış ve yöntemler arasındaki uyuma Cohen'in Kappa istatistiğiyle bakılmıştır. Uzmanların Ebel yönteminde verdikleri madde güçlük değerleri ile gerçek güçlük değerleri arasındaki ilişkiyi bulabilmek için Pearson momentler çarpımı korelasyon katsayısına bakılmıştır. Her bir yöntem için uzman kararları arasındaki uyuma Fleiss'in Kappa istatistiği ve Kendall'in uyuşma katsayısı ile bakılmıştır. Uzmanlar; deneyim durumu, cinsiyet ve uzmanlık alanlarına göre ayrılmış ve uzman özelliklerine göre belirlenen geçme puanlarının değişip değişmediği bağımsız gruplar t testi ile incelenmiştir. Araştırmada ayrıca uzmanların standart belirleme süreciyle ilgili görüşleri, yapılan anketle belirlenmiştir.

Araştırmanın sonucunda üç farklı standart belirleme yöntemine göre belirlenen geçme puanları ve belirlenen geçme puanlarına göre başarılı sayılan öğrenci yüzdeleri arasında istatistiksel olarak anlamlı farklar olduğu bulunmuştur. Bu üç yöntemle elde edilen geçme puanlarına göre öğrencilerin başarılı-başarısız olarak

sınıflanma durumları arasında Evet/Hayır ve Ebel yöntemleri arasında makul, Evet/Hayır ve İşaretleme yöntemleri arasında mükemmel, İşaretleme ve Ebel yöntemleri arasında makul bir uyum olduğu bulunmuştur. Ebel yönteminde uzmanların maddelere verdikleri güçlük değerleri ile gerçek veriden hesaplanan güçlük değerleri arasında istatistiksel olarak anlamlı, pozitif ve yüksek düzeyde bir ilişki olduğu görülmüştür. Evet/Hayır ve Ebel yöntemiyle belirlenen geçme puanları açısından uzmanlar arasında orta derecede uyum varken İşaretleme yönteminde uzmanlar arasındaki uyumun makul düzeyde olduğu görülmüştür. Uzmanların deneyim durumuna, cinsiyetine ve çalıştıkları pozisyona göre belirlenen geçme puanları arasındaki farklar incelenmiş ancak istatistiksel olarak anlamlı farklılıklar bulunamamıştır.

Uzmanların standart belirleme süreciyle ilgili görüşlerini almak için yapılan anketin sonucunda sürecin tamamıyla ilgili bilgilendirmenin, yapılan tanımların ve verilen örneklerin yeterli olduğu sonucuna ulaşılmıştır. Uzmanların çoğu Ebel yönteminde zorlanırken en rahat karar verilen yöntem Evet/Hayır yöntemi olarak belirlenmiştir.

Anahtar sözcükler: Standart belirleme yöntemleri, Evet/Hayır yöntemi, Ebel yöntemi, İşaretleme yöntemi, uzman özellikleri, geçme puanı

Danışman: Dr. Zafer ÇEPNİ, Hacettepe Üniversitesi, Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı

COMPARING YES/NO, EBEL AND BOOKMARK STANDARD SETTING METHODS

Sevda KORKMAZ

ABSTRACT

In this study, passing scores obtained through Yes/No, Ebel and Bookmark standard setting methods as well as decisions made by the experts have been compared in various dimensions.

The study has two different data groups composed of students and experts. In the study, a multiple choice test of 20 questions composed of the questions of 9th year class' "Sets" unit has been developed as a data collection tool. Moreover; information forms have been used in order to get the decisions made by each expert in accordance with three different standard setting methods; also and a survey has been used in order to be able to make an assessment about the process. 505 9th year students have answered the achievement test developed within the scope of the study and separate opinions have been obtained from 30 experts in accordance with these three methods for the items in this achievement test.

In the study, any difference among the percentages of the so-called successful students in accordance with three different standard setting methods has been examined through testing the difference between two dependent ratios while the significance of this difference has been examined through z test. In accordance with the passing scores obtained through three methods, the scores of the students have been classified as successful and unsuccessful and the consistency between the methods has been examined through Cohen's Kappa statistics. Pearson product-moment correlation coefficient has been referred in order to find the relationship between item difficulty values and real difficulty values provided by the experts in Ebel method. For each method, the consistency among the expert decisions has been examined through Fleiss' Kappa statistics and Kendall's coefficient of concordance. Experts have been divided in accordance with their experience, gender and fields of specialization and any change in the passing scores determined by the expert characteristics has been examined through t test

of independent groups. Furthermore; opinions of the experts about standard setting process have been determined through a survey carried out in the study.

As a result of the study, it has been found that there are statistically significant differences between the passing scores and the percentages of the so-called successful students in accordance with these scores determined in accordance with three different standard determination methods. When the students classified as successful-unsuccessful in accordance with the passing scores obtained through these three methods, there is reasonable consistency between Yes/No and Ebel methods; there is an excellent consistency between Yes/No and Bookmark methods and finally there is a reasonable consistency between Bookmark and Ebel methods. It has been found that there is a statistically significant, positive and a high level of relationship between difficulty values assigned by the experts to the items in Ebel method and difficulty values calculated out of the real data. It has been realized that while there is an adequate level of consistency among the experts in terms of passing score determined Yes/No and Ebel methods, there is a reasonable level of consistency among the experts in Bookmark method. Differences among the passing score determined in accordance with the experience, gender and working positions of the experts have been examined; but no statistically significant differences have been found.

As a result of the survey carried out in order to get the opinions of the experts about the standard setting process, it has been concluded that information, descriptions and examples regarding the whole process are adequate. While the majority of the experts have had difficulty in Ebel method, it has been determined that Yes/No method is the easiest method to apply.

Keywords: Standard setting methods, Yes/No method, Ebel method, Bookmark method, properties of experts, passing score

Advisor: Dr. Zafer ÇEPNİ, Hacettepe University, Department of Educational Sciences, Division of Educational Measurement and Evaluation

ETİK BEYANNAMESİ

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

İmza
Sevda KORKMAZ

Canım aileme

TEŞEKKÜR

Tezimin gelişmesinde önemli katkılar sağlayan danışmanım Dr. Zafer ÇEPNİ'ye ve sorduğum her soruya cevap veren, önerileriyle tezime katkıda bulunan Yrd. Doç. Dr. Sevda ÇETİN'e,

Yüksek lisans eğitimim boyunca çok şey öğrendiğim, akademik yaşantımda kendimi geliştirmemde büyük katkıları olan hocalarım Prof. Dr. Selahattin GELBAL'a, Prof. Dr. Hülya KELECİOĞLU'na ve Doç. Dr. Nuri DOĞAN'a,

Tezimi okuyarak görüşleri ve eleştirileriyle çalışmama katkı sağlayan tüm jüri üyelerime,

Eğitimde ölçme ve değerlendirme bölümü ile tanışmamı sağlayan, beni bu alana yönlendirip akademik hayata atılmamda çok büyük desteği olan Doç. Dr. İbrahim KEKLİK ve desteğinin yanı sıra tezime önerileriyle katkı sağlayan Yrd. Doç. Dr. Devrim ERDEM KEKLİK'e,

Tezimin giriş kısmında benden hiçbir yardımını esirgemeyen ve manevi olarak tüm süreçte yanımda olan arkadaşım Ömer KARATAY'a,

Tezimin veri toplama kısmında emeği olan ve tüm süreçte manevi olarak bana hep destek olan tüm arkadaşlarıma,

Bugünlere gelmemde büyük payı olan, hayatımın her alanında benden desteğini esirgemeyip hep arkamda olan, her yaşadığım moral bozukluğunda daha da güçlenmemi sağlayıp beni teşvik eden canım annem Gönül DATLI'ya, canım babam Hasan DATLI'ya ve canım kardeşim Seda DATLI'ya,

Hayatıma girdiği her andan itibaren tüm yaşadığım zorluklarda benimle yürüyen, elimi tutan ve sabrı, anlayışı ile beni hep destekleyen, iyi ki varsın dediğim biricik eşim Cem KORKMAZ'a çok teşekkür ederim.

İÇİNDEKİLER

ÖZ	iii
ABSTRACT	v
ETİK BEYANNAMESİ	vii
TEŞEKKÜR	ix
İÇİNDEKİLER.....	x
TABLolar DİZİNİ	xiii
ŞEKİLLER DİZİNİ.....	xiv
SİMGELER VE KISALTMALAR DİZİNİ	xiv
1. GİRİŞ	1
1.1. Problem Durumu	1
1.2. Araştırmanın Amacı ve Önemi:	5
1.3. Problem Cümlesi:	6
1.3.1. Alt Problemler:	6
1.4. Sayıtlılar:	6
1.5. Sınırlılıklar:	6
1.6. Tanımlar:	7
1.7. Araştırmanın Kuramsal Temeli	7
1.7.1. Standart Belirleme Süreci	7
1.7.1.1. Standart Belirleme Sürecinde Dikkat Edilmesi Gereken Önemli Noktalar	7
1.7.2. Standart Belirleme Yöntemlerinin Sınıflandırılması.....	9
1.7.3. Standart Belirleme Yöntemleri	10
1.7.3.1. Angoff Yöntemi.....	10
1.7.3.2. Evet/ Hayır Yöntemi	11
1.7.3.3. Ebel Yöntemi	12
1.7.3.4. İşaretleme Yöntemi	13
1.7.3.1. Madde Tepki Kuramı	18
2. İLGİLİ ARAŞTIRMALAR.....	21
2.1. Yurt İçinde Yapılan Standart Belirleme Yöntemleri İle İlgili Çalışmalar.....	21
2.2. Yurt Dışında Yapılan Standart Belirleme Yöntemleri İle İlgili Çalışmalar	24
2.3. İlgili Araştırmalar Özet	26
3. YÖNTEM.....	28
3.1. Araştırmanın Yöntemi.....	28
3.2. Katılımcılar	28
3.2.1. Katılımcıların Özellikleri	28
3.2.2. Katılımcılarla İlgili Demografik Bilgiler	28

3.3. Veri Toplama Araçları.....	28
3.3.1. Kümeler Başarı Testi	29
3.3.2. Bilgilendirme Formu	30
3.4. Veri Toplama Araçlarının Uygulanışı	30
3.5. Verilerin İşlenmesi ve Çözülmesi.....	33
3.5.1. Tek Boyutluluk	33
3.5.2. Yerel Bağımsızlık	33
3.5.3. Şans Başarısı	34
4. BULGULAR VE TARTIŞMA	39
4.1. Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemlerine göre belirlenen geçme puanlarının üzerinde puan alan öğrenci yüzdesi arasında fark var mıdır?	39
4.2. Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleri arasında öğrencilerin başarılı-başarısız olarak sınıflandırılması durumları bakımından uyum var mıdır?	42
4.3. Uzmanların Ebel yönteminde verdikleri güçlük derecesi kararı ile maddelerin gerçek güçlük değerleri arasındaki ilişki nasıldır?	43
4.4. Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemlerinin her biri için verilen uzman kararları arasındaki uyum nasıldır?	44
4.5. Uzman özelliklerine göre Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleriyle belirlenen geçme puanları farklılık göstermekte midir?	44
4.5.1. Uzmanların deneyim duruma göre Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleriyle belirlenen geçme puanları farklılık göstermekte midir?	48
4.5.2. Uzmanların cinsiyetlerine göre Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleriyle belirlenen geçme puanları farklılık göstermekte midir?	49
4.5.3. Uzmanların çalıştıkları pozisyonlarına göre Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleriyle belirlenen geçme puanları farklılık göstermekte midir?	49
4.6. Evet/Hayır, Ebel ve İşaretleme standart belirleme süreçlerinde çalışılan uzmanların standart belirleme süreçlerine ilişkin görüşleri nasıldır? ..	50
5. SONUÇ ve ÖNERİLER	53
5.1. Sonuçlar	53
5.2. Öneriler.....	54
5.2.1. Araştırma Sonuçlarından Çıkan Öneriler	54
5.2.2. Araştırmacılara Dönük Öneriler	55
KAYNAKÇA.....	57
EKLER DİZİNİ	61
EK 1. ETİK KURUL İZİN MUAFİYET FORMU.....	62
EK 2. KÜMELER BAŞARI TESTİ	63

EK 3. EVET/HAYIR FORMU	67
EK 4. EBEL FORMU	73
EK 5. BİLGİLENDİRME FORMU	74
EK 6. ANKET MADDELERİ.....	76
EK 7. EVET/HAYIR YÖNTEMİ İÇİN ALINAN KARARLAR.....	78
EK 8. İŞARETLEME YÖNTEMİ İÇİN KESTİRİLEN MADDE PARAMETRELERİ..	79
EK 9. DOĞRULAYICI FAKTÖR ANALİZİ ÖLÇME MODELİ.....	81
EK 10. ORJİNALLİK RAPORU	82
ÖZGEÇMİŞ	83

TABLolar DİZİNİ

Tablo 1: Angoff Yönteminde Geçme Puanı Belirlenmesi.....	10
Tablo 2: Ebel Yönteminin Güçlük ve Uygunluk Düzeyleri.....	12
Tablo 3: Ebel Yöntemine Göre Geçme Puanı Hesaplanması	13
Tablo 4: Deneme Uygulaması Sonuçlarına Göre Test İstatistikleri	29
Tablo 5: Nihai Test Sonuçlarına Göre Test İstatistikleri.....	29
Tablo 6: Test Uygulama Sonuçlarına Göre Test İstatistikleri.....	30
Tablo 7: Düşük Yetenek Grubundaki Öğrencilerin Testteki Zor Maddeleri Doğru Cevaplama Yüzdeleri	34
Tablo 8: Evet/Hayır ve Ebel Yöntemleri İçin Elde Edilen MGP'ler	39
Tablo 9: İşaretleme Yöntemi ile Belirlenen Sayfa ve Madde Numarası ve Θ Değerleri.....	40
Tablo 10: Yöntemlerle Belirlenen Geçme Puanları, Başarılı Kabul Edilen Öğrenci Sayıları ve Yüzdeleri	41
Tablo 11: Geçme Puanları, Başarılı Öğrenci Sayısı ve Yüzdesi, z değerleri	41
Tablo 12: Yöntemler arasındaki uyum için hesaplanan katsayılar	42
Tablo 13: Uzmanların Maddelere Verdikleri Güçlük Değerlerinin Ortalamaları ve Gerçek Madde Güçlükleri İle İlgili Betimleyici İstatistikler	43
Tablo 14: Yöntemlere Göre Uzmanlar Arasındaki Uyum.....	44
Tablo 15: Uzmanların Deneyim Durumuna Göre Belirlenen Geçme Puanlarının Ortalamaları Arasındaki Farkın Testi.....	48
Tablo 16: Uzmanların Cinsiyetlerine Göre Belirlenen Geçme Puanlarının Ortalamaları Arasındaki Farkın Testi.....	49
Tablo 17: Uzmanların Cinsiyetlerine Göre Belirlenen Geçme Puanlarının Ortalamaları Arasındaki Farkın Testi.....	49
Tablo 18: Uzmanların Standart Belirleme Sürecindeki Unsurlarla İlgili Görüşleri ..	50
Tablo 19: Uzmanların öğrenci performansı sınıflandırırken kullanılan unsurların önem düzeyi ile ilgili görüşleri.....	51

ŞEKİLLER DİZİNİ

Şekil 1: Testin Q-Q Grafiği.....	31
Şekil 2: Frekans ve yüzdelerle ait Yöntem 1 ve Yöntem 2'ye göre geçen-kalan öğrenci sayıları	35
Şekil 3: Evet/Hayır Geçme Puanlarının Q-Q Grafiği.....	45
Şekil 4: Ebel Geçme Puanlarının Q-Q Grafiği	46
Şekil 5: Deneyim Durumuna Göre Belirlenen İşaretleme Geçme Puanlarının Q-Q Grafiği.....	47
Şekil 6: Cinsiyete Göre Belirlenen İşaretleme Geçme Puanlarının Q-Q Grafiği	47
Şekil 7: Çalışılan Pozisyona Göre Belirlenen İşaretleme Geçme Puanlarının Q-Q Grafiği.....	48

SİMGELER VE KISALTMALAR DİZİNİ

MGP: Bir uzmanın test için belirlediği minimum geçme puanı

KTK: Klasik Test Kuramı

MTK: Madde Tepki Kuramı

BDL: İşaretleme yönteminin güçlük noktasıdır.

RP: Cevaplama Olasılığı, belirli bir yetenek düzeyindeki kişinin maddeye doğru cevap verme olasılığıdır.

U: Uzman

m: Madde

1. GİRİŞ

Bu bölümde araştırmanın problem durumuna, amacına, önemine, problem cümlesine, alt problemlerine, sayıltı ve sınırlılıklarına ve kuramsal temeline yer verilmiştir.

1.1. Problem Durumu

Eğitimin ne olduğu konusunda geçmişten günümüze pek çok tanımlar yapılsa da bunlar içinde en fazla kabul gören tanımlardan biri eğitimin “bireyin davranışında kendi yaşantısı yoluyla ve kasıtlı olarak istendik bir değişme meydana getirme süreci” olarak kabul edildiği tanımlamadır (Ertürk, 1994). Bu tanımlamada dikkat edilmesi gereken hususlar bireyin yaşantısı yoluyla ve istendik davranış değişiklikleri meydana getirme durumlarıdır. Diğer bir ifadeyle eğitime dair verilerin ne denli edinildiğini öğrenebilmek için gözlenebilir davranış değişikliklerinin olması gerekliliği eğitimi eğitim yapan, amacına ulaştığını gösteren bir gerekliliktir. Bunu kalıcı izli davranış değişikliklerinin yaşanmaması durumunda eğitimin var olmadığını söyleyerek izah etmemiz de mümkündür. Eğitimin gerekliliklerinden olan kalıcı izli davranış değişikliklerinin olup olmadığını anlamak ise bir ölçme ve değerlendirme sürecinin ortaya çıkmasına neden olur. Ölçme işlemi ve bu işlemin devamında yapılan değerlendirme kullanılan eğitim sistemine dair dönütlerle iyileştirmeler yapılmasını sağlayacak bir sistemdir (Baykul, 2000).

Ölçme en genel tanımıyla “Bir niteliğin gözlenip gözlem sonuçlarının sayı veya sembollerle ifade edilmesini kapsayan işlemdir” (Baykul, 2000 s. 91). Ayrıca unutmamak gerekir ki ölçme işlemi değerlendirme yapabilmek için bir ön gerekliliktir. Diğer bir ifade ile ölçme olmadan değerlendirmenin olması mümkün değildir. Ölçme sonrasında gerçekleştirilen değerlendirme işlemi ise ölçme sonuçlarının bir ölçütle kıyaslanarak karara varılmasını kapsamaktadır (Turgut, 1993). Burada üçüncül bir kavram olarak ölçüt kavramı karşımıza çıkmaktadır. Ölçüt, en basit tanımıyla ölçme sonuçlarının karşılaştırıldığı değerlerdir. Ölçüt, mutlak ve bağıl ölçüt olmak üzere ikiye ayrılır. Bağıl ölçüt, bir grup dikkate alınarak belirlenirken mutlak ölçüt, önceden belirlenmiş bir değerdir (Livingston ve Zieky, 1982, s. 10 – 11). Bu ölçütlere göre yapılan değerlendirmeler de ölçütün adını alır. Mutlak ölçütle yapılan değerlendirmeye mutlak değerlendirme, bağıl ölçütle

yapılan deęerlendirmeye ise baęıl deęerlendirme adı verilir. Örneęin; “Testi çözenlerden puanı en yüksek 20 kiři sertifikayı almaya hak kazanacaktır.” denildięinde baęıl deęerlendirme, “Testi çözen bireylerden 100 soru üzerinden en az 80’ini doęru yapanlar sertifikayı almaya hak kazanacaktır.” denildięinde ise mutlak deęerlendirme yapılmıř olur. Baęıl deęerlendirmelerdeki ölçüt standart puan, kesme puanı, geçme puanı veya standart olarak adlandırılır. Bu puanların belirlenmesi süreci standart belirleme süreci olarak adlandırılır (Özçelik, 1992). Standart belirlemedeki standart kelimesi genellikle performans standardının kısaltılmıřı olarak kullanılır. Standart kelimesi ile ölçülmek istenen kapsam ya da yetenek için gerekli olan minimum bilgi ve donanım kastedilir (Hambleton ve Pitoniak, 2006).

Standart belirleme süreciyle belirlenen kesme puanları ile öğrenciler çeřitli yeterlilik seviyelerine göre sınıflandırmaya tabi tutulabilirler. Bu sınıflandırmalar en genel hâliyle yeterli ve yetersiz olarak iki ana grup şeklinde olmakla beraber bu grupların arasında belirlenecek olan yeterliliklere göre ara basamaklar da yerleřtirmek mümkündür. Ancak bu basamaklar oluşturulurken bir keyfiyet durumundan ziyade bireylerin gerçek anlamda saęlamaları gereken bir takım yeterliliklerin dikkate alınması gerekmektedir. Nitekim 1-100 arasındaki puanlama şekillerinde her bireye bir puan üzerinden tanımlama yapılabilmeyle beraber 100 puan üzerinden 1 puanlık bir farkın belirleyici olmayacaęı da açıktır. Burada anlatılmak istenen esas nokta standart belirleme işleminde yeterli-yetersiz sınıflamasının arasına oturtulacak olan ara basamakların ölçülen özellięe ve sınıflama amacına göre hazırlanması gereklilięidir. Bu bağlamda Türkiye’de hemen her eğitim kademesinde karřımıza çıkacak olan bir standart puana göre başarı sınıflaması mevcuttur. Bu sınıflamaya göre 100 puan üzerinden yapılan örnek bir puanlama ve bu puanlamaya göre yapılan sınıflama ařaęıda verilmiřtir.

0-24	0	Yetersiz
25-44	1	Zayıf
45-54	2	Geçer
55-69	3	Orta
70-84	4	İyi
85-100	5	Pekiyi

Yükseköğrenim verilen eğitim kurumlarında yapılabilecek örnek bir puanlama aşağıda verilmiştir.

F – Zayıf

D – Geçer

C – Orta

B – İyi

A – Pekiyi

Verilen örneklerde yer alan sınıflandırmalar, kesme puanların belirlenmesinden sonra ortaya çıkar. İş yerine işçi alırken, mezuniyet derecesi verilirken, kredi puanları verilirken, bireyleri belli özelliklere göre sınıflarken (geçti-kaldı, başarılı-başarısız, yeterli-yetersiz) ve daha birçok durumda kesme puanları kullanılır. Sadece bireyler değil, okullar, ilçeler, iller de standart kesme puanları kullanılarak değerlendirilebilir, karşılaştırılabilir (Hambleton ve Pitoniak, 2006).

Yukarıda bahsedilen bu süreç standart belirleme sürecidir. Ancak durum, verilen örneklerde olduğu kadar basit bir şekilde ortaya çıkmamaktadır. Standart belirleme süreci ülkesel, ekonomik vb. birçok faktör de göz önüne alınarak sürdürülen ve ulusal anlamda en uygun bölümlendirmelerin yapılması gerekliliğini kapsayan bir süreçtir.

Standart belirleme süreci ile ilgili olarak araştırmacılar psikometrinin istatistiksel, politik ve kültürel yönlerini bir araya getiren bir süreç olduğundan bahsederek, sürecin ciddiyetine ve karışık yapısına dikkat çekmişlerdir (Cizek, 2001). Bunun yanı sıra standart belirleme süreci için performansın iki ya da daha fazla düzeyini ayırt ederek, puan atayabilmek için uygun kural ve işlemleri takip etme süreci olduğundan bahsedenler de vardır (Cizek, 2001). Hambleton (2001) ise standart

belirleme sürecini; kişileri seviyelerine göre ayıran performans standartlarını, test puanları ölçeğinde sayısal yani işevuruk hâle getirme süreci olarak tanımlamaktadır.

Standart belirleme sürecinde temel nokta belirlenecek standardın hizmet edeceği amaçtır. Diğer bir ifadeyle uygulanan testten elde edilen puanların ne amaçla kullanılacağı konusu standart belirleme sürecinin temelini oluşturur. Nitekim amaç doğrultusunda hazırlanmayan standart belirleme süreçlerinde uygulanan testlerin herhangi bir işe yarayabileceğini söylemek oldukça güçtür. Eğitim kurumları içinde uygulanan testlere dair standart belirleme genel olarak öğrenci başarılarını karşılaştırmaya yöneliktir. Bu durumda da bireylerin başarılı–başarısız ve bu aralıkta belirlenen seviyelere göre sınıflandırılmaları söz konusudur.

Yukarıda da değinildiği gibi standart belirlemede standardın hizmet edeceği amaç oldukça önemlidir. Ancak bunun yanı sıra en az amaç kadar önemli olan bir diğer husus standart belirlemede kullanılacak olan yöntemin seçimidir. Yöntem seçimi esnasında tek boyutlu olarak düşünmek birçok eksikliğe neden olabileceği için çok boyutlu düşünmek gerekmektedir. Yöntem belirlenmesi konusunda araştırmacılar genel olarak aşağıdaki ana başlıklar üzerinde durmaktadırlar (Berk, 1986).

- Öncelikle belirlenecek yöntem, kullanılacak ölçme aracına uygun olmalıdır.
- Belirlenen yöntemle elde edilen test kesme puanları gerçekçi, gerçeğe yakın olmalıdır. Bu kesme puanı sadece sınav puanları ile değil, performansa yönelik değerlendirmelerin de yapılmasıyla elde edilebilir ki belirlenen yöntem bu ihtiyacı da karşılar nitelikte olmalıdır.
- Belirlenen yöntemden elde edilecek olan veriler farklı alanlarda da kullanılabileceği için yöntemin buna uygun hale getirilmesi gerekliliği doğabilir. Bu nedenle de belirlenen yöntemin istatistiksel olarak üzerinde çalışılabilecek bir yöntem olması gerekmektedir.
- Yöntem seçiminde en önemli konulardan biri uygulamada kolaylık sağlayan bir yöntemin belirlenmesidir. Yöntemin uygulanmasının kolay olmasının yanı sıra tamamlanması da uzun zaman almamalı ve her basamak kendi içinde ve basamaklar arasında sistematik bir özellik taşımalıdır.

- Belirlenen yöntemle dair hesaplamalar, ilgili programlar tarafından hesaplanabilecek nitelikte olmakla beraber, yöntemin uygulayıcıları da hesaplamaları yapabilmelidir.

Standart belirleme öznel elementlere fazlasıyla bağlı olduğundan çok eleştirilmektedir (Hambleton ve Pitoniak, 2006). Standart belirlemede sonuçlar uzman kararına bağlı olduğundan genelleme yapabilmek çok zordur. Popham (1978) standart belirleme sürecini genellemenin hatalı bir davranış olduğunu ifade etmiştir. Ancak; aynı büyüklükte iki farklı gruba aynı uzman grubu ile değerlendirmeler yaptırılıp sonuçlar birbirini desteklerse genelleme yapılabilir. Ya da tek panel yapıp, panelin altında alt paneller oluşturulabilir. Belirlenecek olan standart belirleme puanı bulunan geçme puanlarının ortalaması alınarak hesaplanır (Hambleton ve Pitoniak, 2006).

1.2. Araştırmanın Amacı ve Önemi:

Bu araştırma kapsamında alanyazında birçok kez kullanılmış olan Evet/Hayır yöntemi, pek fazla tercih edilmemiş olan Ebel yöntemi ve Türkiye’de yeni bir yöntem olarak sayılabilen İşaretleme yöntemi karşılaştırılmıştır. Çalışma kapsamında bu yöntemlere ait geçme puanlarının nasıl olduğu, geçme puanlarının başarılı sayılan ve sayılmayan öğrenci yüzdelerini nasıl etkilediği, uzmanların Ebel yönteminde tahmin ettikleri madde güçlükleri ile gerçek madde güçlük değerleri arasındaki uyumun ne düzeyde olduğu, uzmanların belirledikleri geçme puanları arasındaki uyumun hangi düzeyde olduğu belirlenmeye çalışılmıştır. Çalışmada ayrıca uzman özelliklerinin uzman kararlarını nasıl etkilediği de araştırılmıştır.

Alanyazın tarandığında Türkiye’de standart belirleme konusu ile ilgili yapılan çalışma sayısının oldukça az olduğu görülmektedir. Ebel ve İşaretleme yöntemlerinin kullanıldığı çalışma sayısının da çok az sayıda olduğu görülmekle birlikte Evet/Hayır, Ebel ve İşaretleme yöntemlerinin üçünün aynı anda karşılaştırıldığı çalışmaya rastlanmamıştır. Ayrıca yapılan araştırmalarda uzman özelliklerine göre belirlenen geçme puanlarının kıyaslanması işi de pek rastlanan bir durum değildir. Yapılan çalışmanın, Evet/Hayır, Ebel ve İşaretleme yöntemlerinin karşılaştırılması ve belirlenen geçme puanlarının uzman özelliklerine göre kıyaslanması açısından alanyazına faydalı bilgiler sağlayacağı

düşünülmektedir. Çalışma, bahsedilen bu yönleriyle gelecek çalışmalara ışık tutması açısından önemlidir.

1.3. Problem Cümlesi:

Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemlerinin geçme puanlarına etkisi nedir?

1.3.1. Alt Problemler:

1. Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemlerine göre belirlenen geçme puanlarının üzerinde puan alan öğrenci yüzdesi arasında manidar bir fark var mıdır?
2. Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleri arasında öğrencilerin başarılı-başarısız olarak sınıflandırılması durumları bakımından uyum var mıdır?
3. Uzmanların Ebel yönteminde verdikleri güçlük derecesi kararı ile maddelerin gerçek güçlük değerleri arasındaki ilişki nasıldır?
4. Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemlerinin her biri için verilen uzman kararları arasındaki uyum nasıldır?
5. Uzman özelliklerine göre Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleriyle belirlenen geçme puanları farklılık göstermekte midir?
6. Evet/Hayır, Ebel ve İşaretleme standart belirleme süreçlerinde çalışılan uzmanların standart belirleme süreçlerine ilişkin görüşleri nasıldır?

1.4. Sayıtlar:

1. Uzmanların test maddeleri hakkındaki kararlarını, öğrencilerin öğrenme düzeylerini dikkate alarak verdikleri varsayılmıştır.
2. Testi çözen öğrencilerin cevaplarının öğrenme düzeylerini yansıttığı varsayılmıştır.

1.5. Sınırlılıklar:

Araştırma, 9. sınıf kümeler ünitesinin sorularından oluşan bir test ile sınırlandırılmıştır.

1.6. Tanımlar:

Kesme Puanı/Geçme Puanı: İstenilen performans düzeyinde olan öğrenci ile yeterli düzeyde olmayan öğrenciyi ayırt edebilen uygun performans noktasını gösterir.

Kümeler ünitesinin geçme kalma sınırı: Kümeler tanımı ve kümelerle işlemler kazanımlarını kavrayarak bu öğrenmelerini kümeler ünitesiyle ilgili konularda uygulayabilecek öğrencilerle uygulayamayacak öğrenciler arasındaki sınırdır.

Minimum Geçme Puanı: İstenilen performans düzeyinde olan öğrenci ile yeterli düzeyde olmayan öğrenciyi ayırt edebilmek için gerekli olan en küçük puana denir.

1.7. Araştırmanın Kuramsal Temeli

Bu bölümde standart belirleme süreci, standart belirleme yöntemleri ve Madde Tepki Kuramı anlatılacaktır.

1.7.1. Standart Belirleme Süreci

Amacı belirlenen bir standart belirleme sürecinde 9 adım bulunmaktadır (Hambleton ve Pitoniak, 2006):

1. Standart belirleme yönteminin seçilmesi
2. Uygulama grubunun belirlenmesi
3. Performans kategorilerinin ya da geçme puanının tanımlanması
4. Seçilen metodun kullanılabilmesi için uzmanların eğitilmesi
5. Madde oranlarının toplanması ve betimsel istatistiklerin hesaplanması
6. Panel tartışması oluşturulması ve uzmanlara geri bildirim verilmesi
7. Uzmanların verdiği kararların toplanması ve bu karara göre performans standartlarının veya geçme puanının belirlenmesi
8. Süreç hakkında uzmanlardan bilgi alınması
9. Geçerlik kanıtlarının sağlanması ve bunun teknik olarak raporlanması

1.7.1.1. Standart Belirleme Sürecinde Dikkat Edilmesi Gereken Önemli Noktalar

Standart belirleme süreci, uygulanırken bazı önemli noktalara dikkat edilmelidir:

- Standart belirleme yöntemi seçilirken; kullanılan testteki madde tipine, yöntemin gerekliliklerine, standart belirleme süreci için ayrılan zamana, uzmanların tecrübelerine dikkat edilmelidir. Çalışma sonucunun karşılaştırılabilir olması için iki ya da üç yöntem kullanılması önerilir (Hambleton ve Pitoniak, 2006).
- Uzmanlar seçilirken kullanılan standart belirleme yöntemine göre seçilmelidir. Bazı yöntemlerde (sınır grup yöntemi) uzmanların grubu tanınması gerekirken bazı yöntemlerde (Angoff) konu kapsamının ve öğrenci profilinin bilinmesi yeterlidir (Hambleton ve Pitoniak, 2006). Uzmanlar öğretmenler, program geliştirme uzmanları, eğitim politikacılarından seçilebilir (Hambleton ve Pitoniak, 2006). Uzman sayısı etnik alt yapı, yaş, cinsiyet, iş deneyimi, kullanılan yöntemle ilgili tecrübe açısından çeşitlilik olacak şekilde belirlenmelidir (Hambleton ve Pitoniak, 2006). Daha genel sonuçlara ulaşmak için uzman sayısının 15-30 arası olması önerilmektedir (Hambleton ve Pitoniak, 2006).
- Süreçte kullanılan performans kategorileri ya da geçme puanı uzmanların anlayabileceği şekilde tanımlanmalıdır. Uzmanların ilk verdikleri kararları tartışabileceği ve gerekirse kendi kararlarını değiştirebilecekleri tartışma ortamı yaratılmalı ve uzmanlara süreç boyunca geri bildirim verilmelidir. Geri bildirimler ve tartışma sayesinde alınan kararların birbirine yakınlığı genellikle artmaktadır (Hambleton ve Pitoniak, 2006).
- Standart belirleme süreci, yapılan çalışmada tüm adımlarıyla detaylı bir şekilde anlatılmalıdır. Kullanılan yöntemle bakılmaksızın standart belirleme sürecinin savunulabilir ve geçerli olması gerekir. Bu yüzden süreç açık ve net, sistematik ve makul olmalıdır (Hambleton ve Pitoniak, 2006).
- Standart belirleme sürecinde; geçme puanı uzman kararına bağlı olduğu için uzman özellikleri çalışmada raporlanmalıdır ve uzmanlar süreç başlamadan önce yöntem ve süreç ile ilgili yeterince eğitilmelidir. Etkili bir uzman eğitimi için aşağıdaki önemli noktalara dikkat edilmelidir (Hambleton ve Pitoniak, 2006).
 - Süreç, görev ve amaçlar açıklanmalıdır.
 - Takip edilecek standart belirleme yöntemi adımları açıklanmalıdır.

- Puanlama anahtarı açıklanmalı ve anlaşılması sağlanmalıdır.
- Doldurulacak formlar açıklanmalı ve benzer formlarla alıştıırma yapılmalıdır.
- Süreçte kullanılan veriler açıklanmalıdır.
- Konu kapsamı anlatılmalıdır.
- Performans kategorileri isteniyorsa kategorilerin tanımları ve aralıkları açıklanmalıdır.
- Sınır değeri öğrenci kavramı açıklanmalıdır.
- Standart ya da standart altı koşullar belirlenmelidir.

1.7.2. Standart Belirleme Yöntemlerinin Sınıflandırılması

Hambleton, Jeager, Plake ve Mills (2000) standart belirleme yöntemleri için yeni bir sınıflama önermişlerdir. Önerilen bu sınıflamada yöntemler 4 gruba ayrılmıştır.

1. Test maddeleri ve puanlama anahtarı için karar verilen yöntemler:

Bu sınıftaki yöntemlere Angoff yöntemi ve çeşitlemeleri, Nedelsky yöntemi, Ebel yöntemi, Jeager yöntemi, İşaretleme yöntemi birer örnektir.

2. Öğrenci görevi için karar verilen yöntemler

Madde madde yaklaşımları, bütüncül yaklaşımlar ve karışık yaklaşımlar bu sınıflama için birer örnektir.

3. Puan profilleri için karar verilen yöntemler

Baskın profil yöntemi, madde kümeleme yöntemi bu sınıflamaya uygun birer örnektir.

4. Öğrenci hakkında karar verilen yöntemler

Sınır grup yöntemi, karşıt grup yöntemi bu sınıflamaya birer örnektir.

Yukarıda verilen sınıflamanın yanı sıra standart belirleme yöntemleri, uzun süre öğrenci ve test merkezli olmak üzere ikiye ayrılarak sınıflanmıştır. Test merkezli yöntemlerde uzmanların görevi test içindeki maddeleri inceleyip karar vermektir. Uzmanlar, öğrencileri sınır değerde düşünerek karar verir. Öğrenci merkezli yöntemlerde, uzmanlardan belli bir öğrenci grubunu düşünerek karar vermeleri beklenir.

1.7.3. Standart Belirleme Yöntemleri

Bu çalışmada kullanılan test merkezli yöntemler ve bu yöntemlerin her birine göre geçme puanı hesaplaması açıklanacaktır.

1.7.3.1. Angoff Yöntemi

1971 yılında William H. Angoff tarafından önerilen yöntem en sık ve yaygın kullanılan standart belirleme yöntemidir (Hambleton ve Pitoniak, 2006). Bu yöntemde uzmanlar maddenin bütününe odaklanıp geçme-kalma sınırındaki bir öğrencinin maddeyi doğru cevaplama olasılığını belirlerler. Bu olasılığı belirlerken uzmanlar tek bir öğrenciyi düşünmekte zorlanıyorsa geçme-kalma sınırındaki 100 kişilik bir öğrenci grubunu düşünerek bunlardan kaçının maddeye doğru cevap vereceğini belirleyebilirler (Livingston ve Zieky, 1982). Uzmanların verdikleri tahminin değeri 0 ve 1 değerleri arasında yer alır. 0, sınırdaki öğrencinin soruyu doğru cevaplayamayacağını belirtirken 1, öğrencinin maddeyi kesin doğru cevaplayacağını belirtir. Sorular kolaylaştıkça öğrencilerin soruyu doğru cevaplama olasılığı artacaktır. Eğer soru çoktan seçmeliyse bu olasılık en az şans başarısıyla soruyu doğru cevaplama olasılığı olan $1/(\text{seçenek sayısı})$ oranı kadar olacaktır (Çetin, 2011).

Sınavda her bir uzmanın maddelere verdiği doğru cevaplama yüzdelerinin ortalaması uzmanın belirlemiş olduğu geçme puanını belirtir. Uzmanların belirledikleri geçme puanlarının ortalaması ise testin geçme puanını oluşturur. Angoff yöntemiyle geçme puanı belirlenmesine yönelik bir örnek aşağıda verilmiştir.

Tablo 1: Angoff Yönteminde Geçme Puanı Belirlenmesi

<i>Maddeler</i>						
<i>Uzmanlar</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>Geçme Puanı</i>
1	0.45	0.60	0.75	0.50	0.40	0.54
2	0.60	0.50	0.80	0.40	0.60	0.58
3	0.50	0.40	0.70	0.30	0.50	0.48
4	0.40	0.50	0.75	0.70	0.45	0.56
5	0.60	0.65	0.65	0.60	0.45	0.59
6	0.50	0.55	0.70	0.50	0.55	0.56

Geçme Puanı: 0.55

Tablo 1’de 5 maddelik bir sınavın her bir maddesine 6 uzmanın verdiği cevaplama yüzdeleri belirtilmiştir. Tablodan görüldüğü gibi geçme puanı sütunu her bir

uzmanın maddelere verdiği yüzdelerin ortalamalarını belirtmektedir. Yani her bir uzmanın sınav için belirlediği geçme puanını belirtmektedir. Her bir uzmanın belirlediği geçme puanının ortalaması ise sınavın geçme puanını oluşturur. Tablo 1'e göre bu sınavın geçme puanı 55.16 olarak hesaplanmıştır.

Angoff yönteminde uzmanlar maddenin bütününe odaklandığından yöntemin tüm sınav tiplerine uygulanabiliyor olması oldukça önemli bir avantajdır. Geçme-kalma sınırındaki öğrenciler için test maddelerinin zorluğunu değerlendirmek uzmanlar için zor olabileceğinden bu yöntem eleştirilmektedir (Shepard, Glaser, Linn ve Bohrnstedt,1993).

Angoff, (1988) uzmanlar arasında uzlaşma olmaması durumunda bunun 2 nedeni olabileceğini söylemiştir:

1. Uzmanların geçme-kalma sınırındaki öğrenci kavramından aynı anlamı çıkarmamaları
2. Bu kavramdan aynı anlamı çıkarsalar bile test maddelerinde doğru olasılık tahmini yapamamaları

1.7.3.2. Evet/ Hayır Yöntemi

1997 yılında J.C. Impara ve B.S. Plake tarafından geliştirilen Angoff çeşitlemelerinden biridir. Bu yöntemde uzmanlardan Angoff yöntemindeki gibi maddenin bütünü düşünerek, geçme-kalma sınırındaki bir öğrencinin maddeyi doğru cevaplayıp cevaplayamayacağına karar vermeleri beklenir. Uzman, geçme-kalma sınırındaki bir öğrencinin maddeyi doğru cevaplayacağını düşünüyorsa maddeye 1 puan, maddeyi doğru cevaplayamayacağını düşünüyorsa maddeye 0 puan vererek geçme puanını belirler. Her bir uzmandan alınan geçme puanlarının ortalaması sınavın geçme puanını oluşturur.

Evet/Hayır yöntemini uygulamak Angoff yöntemine göre daha basittir. Angoff yönteminde uzmanlar sınırdaki öğrencilerin yüzde kaçının maddeye doğru cevap vereceğini tahmin ederken zorlanmaktadır. Evet/Hayır yöntemiyle karşılaşılan bu zorluğun giderildiği savunulmaktadır (Impara ve Plake, 1998). Chinn ve Hertz (2002) çalışmalarında uzmanların Evet/Hayır yönteminde daha kolay karar verdiklerini fakat buldukları geçme puanlarının süreç birkaç kez tekrarlandığında Angoff yöntemine göre daha az kararlı olduğunu görmüşlerdir.

1.7.3.3. Ebel Yöntemi

1972 yılında R. L. Ebel tarafından geliştirilmiş, uzmanlara birbirini takip eden bazı görevler yükleyen bir yöntemdir. Öncelikle uzmanlar sınavdaki tüm maddeleri teker teker inceler ve bu maddeleri iki özelliğine göre sınıflarlar. Ebel, uzmanlardan maddeleri zor, orta ve kolay olmak üzere üç güçlük derecesine; gerekli, önemli, kabul edilebilir ve tartışılabilir olmak üzere dört uygunluk düzeyine göre sınıflamalarını beklemiştir. Bir başka deyişle Ebel yönteminde uzmanların her bir maddeyi güçlük ve uygunluk düzeylerine göre sınıflamaları beklenir. Yani uzmanlardan Tablo 2'deki 12 (3 x 4) hücrelik tabloyu doldurmaları beklenir. Daha sonra uzmanlardan doldurdukları bu tablonun her bir hücresindeki soruları geçme-kalma sınırındaki öğrencilerin yüzde kaçının cevaplayabileceği sorusuna cevap vermeleri beklenir.

Her bir hücredeki soru sayısı ile belirlenen yüzdenin çarpımının toplanması ile hücrelerin puanı bulunur. Hücrelerin puanlarının toplamının sınavdaki madde sayısına bölümü ile uzmanların geçme puanları elde edilir. Sınavın geçme puanı ise uzmanların her birinden elde edilen geçme puanlarının ortalaması ile bulunur. Tablo 2'de Ebel yöntemine göre 20 maddelik bir testteki her bir maddenin güçlük ve uygunluk düzeylerine göre sınıflanması görülmektedir.

Tablo 2: Ebel Yönteminin Güçlük ve Uygunluk Düzeyleri

	UYGUNLUK			GÜÇLÜK		
		<i>Kolay</i>		<i>Orta</i>		<i>Zor</i>
<i>Gerekli</i>		Sorular: 1, 3, 6 Yüzde: 90		Sorular: 4, 7, 15 Yüzde: 80		Sorular: 17, 20 Yüzde: 70
<i>Önemli</i>		Sorular: 2, 12,19. Yüzde: 90		Sorular: 5, 11, 18 Yüzde: 80		Sorular: 9 Yüzde: 75
<i>Kabul Edilebilir</i>		Sorular: 13 Yüzde: 60		Sorular: 16 Yüzde: 65		Sorular:10 Yüzde: 45
<i>Tartışılabilir</i>		Sorular: 8 Yüzde: 30		Sorular: 14 Yüzde: 10		Sorular: Yüzde:

Geçme puanının hesaplanabilmesi için öncelikle tablonun her bir hücresinin puanları belirlenmelidir.

Tablo 3: Ebel Yöntemine Göre Geçme Puanı Hesaplanması

<i>Kategori</i>	<i>Doğru Cevaplanma Yüzdesi</i>	<i>Soru sayısı</i>	<i>Hücre Puanı</i>
Gerekli			
<i>Kolay</i>	90	3	$0.90 \times 3 = 2.70$
<i>Orta</i>	80	3	$0.80 \times 3 = 2.40$
<i>Zor</i>	70	2	$0.70 \times 2 = 1.40$
Önemli			
<i>Kolay</i>	90	3	$0.90 \times 3 = 2.70$
<i>Orta</i>	80	3	$0.80 \times 3 = 2.40$
<i>Zor</i>	75	1	$0.75 \times 1 = 0.75$
Kabul Edilebilir			
<i>Kolay</i>	60	1	$0.60 \times 1 = 0.60$
<i>Orta</i>	65	1	$0.65 \times 1 = 0.65$
<i>Zor</i>	45	1	$0.45 \times 1 = 0.45$
Tartışılabilir			
<i>Kolay</i>	30	1	$0.30 \times 1 = 0.30$
<i>Orta</i>	10	1	$0.10 \times 1 = 0.10$
<i>Zor</i>	0	0	0
			Toplam: 14.45

Tabloda görüldüğü gibi örneğin gerekli-orta kategorisinde bulunan 3 maddenin doğru cevaplanma olasılığı 0.80 olarak belirlendiği için bu kategorinin olduğu hücrenin puanı $0.80 \times 3 = 2.40$ olarak hesaplanır. Aynı hesaplama tüm kategoriler için yapıldıktan sonra testin geçme puanı 14.45 olarak hesaplanmıştır.

Ebel yöntemi hem iki hem de çok kategorili maddelerde kullanılabilir. Çok kategorili maddelere Likert tipi maddeler örnek verilebilirken iki kategorili maddelere veri matrisi 1/0 olan maddeler örnek verilebilir. Ebel yönteminde belirlenen güçlük ve uygunluk boyutlarının ilişkili bulunabilmesi oldukça zordur. Bu yüzden bu yöntemi kullanmak diğer bahsedilen yöntemlere göre uzmanların anlaması ve uygulaması açısından daha zordur (Hambleton ve Pitoniak, 2006).

1.7.3.4. İşaretleme Yöntemi

Angoff yönteminin bazı eksikliklerini gidermek amacıyla 1996 yılında Lewis, Mitzel ve Green tarafından önerilmiştir. Yöntem, çoktan seçmeli sorularla yapılandırılmış cevaplar içeren soruların bir arada bulunduğu sınavlarda kullanabilmekte, uzmanların iş yükünü hafifleterek karar vermelerini kolaylaştırmakta, geçme puanlarını belirlemede uzman kararlarını ölçme modelleriyle birleştirebilmekte ve

test içeriğini performans seviyesi tanımlarıyla birleştirebilmektedir (Mitzel, Lewis, Patz ve Green, 2001).

İşaretleme yönteminin temeli Madde Tepki Kuramına dayanarak maddelerin haritalanmasıdır (Lin, 2006). Öğrencilerin seviyesi ya da yeterliğini belirlemek için testteki soruların güçlükleri MTK kullanılarak kestirilir. Maddeler bir ölçek üzerinde madde güçlüklerine göre sıralanır. En basit olan madde ilk sırada en zor olan madde son sırada olacak şekilde sıralanma işlemi gerçekleştirilir. Bu yöntemde uzmanlardan her bir madde hakkında yargıda bulunması istenmez. Bu yöntemde uzmanlardan beklenen şey kolaydan zora doğru sıralanmış soru kitapçığında işaretleme yaparak geçme puanını belirlemektir. İşaretleme yapılan kitapçıkta işaretlenen soru ve öncesindeki sorular öğrencilerin yapabileceği soruları belirtirken işaretlenen sorudan sonraki sorular öğrencilerin yapamayacağı soruları belirtmektedir.

İşaretleme yapılan kitapçığa “sıralı soru kitapçığı” denir. Bu kitapçığın her bir sayfasında sadece bir madde bulunur. Aşağıda sıralı madde kitapçığından örnek bir sayfa bulunmaktadır. Çalışmada kullanılan testte 2. sırada olan madde, kılavuzun 1. sayfasında yer almaktadır. Sayfada madde kökü ve seçenekler birlikte bulunmaktadır. Sayfanın sol alt köşesinde maddenin doğru cevabı, sayfanın başında belirlenen cevaplama oranına göre maddeyi doğru cevaplamak için gerekli olan yetenek düzeyi bulunmalıdır.

1
Madde 2
Maddeyi % 67 olasılıkla doğru cevaplandırabilmek için gerekli olan yetenek düzeyi: -1.06092
Rakamları farklı 4 basamaklı en büyük doğal sayının rakamlarından oluşan kümenin liste biçiminde gösterimi aşağıdakilerden hangisinde doğru olarak verilmiştir?
A) {9, 8, 7, 6} B) {9, 8, 7, 5} C) {9, 7, 5, 3}

İşaretleme yönteminin tercih edilmesinin birçok nedeni vardır (Hambleton ve Pitoniak, 2006). Bunlardan birincisi cevabı öğrencinin yazdığı sorularla çoktan seçmeli soruların bir arada olduğu sınavlarda kullanılabilmesidir. İkincisi uzmanlar için karmaşıklığın, iş yükünün azalmasıdır. Örneğin 20 maddelik bir sınav için 3 performans düzeyi belirlenecekse Angoff ya da Evet/Hayır yönteminde uzmanın $20 \times 3 = 60$ olasılık değeri belirlemesi gerekirken, işaretleme yönteminde aynı sıralı kitapçık üzerinden ilk kesme puanı belirlendikten sonra diğer kesme puanları diğer maddeler incelenerek belirlenir. Böylece karar veren uzmanlar her bir kesme puanı için aynı maddeleri tekrar tekrar inceleyemezler.

Üçüncüsü, sınav içeriği ile performans seviyelerinin birleştirilebilmesi, dördüncüsü istenen öğrenci profilinin daha iyi anlaşılmasının sağlanması, beşincisi etkili bir şekilde birden fazla kesme puanı belirlenebilmesi, altıncısı uzmanların zamanı daha etkili bir şekilde kullanmasının sağlanması ve düşük standart hatalı kesme puanları elde edilebilmesidir (Lin, 2006). Diğer yöntemlerde uzmanlar öncelikle istenilen öğrenci profilini anlamak daha sonra sınavdaki her bir maddeyi tek tek incelemek zorundadır; fakat işaretleme yönteminde sınavın tamamı sıralı soru kitapçığı ile incelendiğinden uzmanların karar vermesi daha kolaydır.

İşaretleme yönteminin avantajlarının yanı sıra yöntemin temeli MTK'ya dayandığından tek boyutluluk, yerel bağımsızlık ve kestirim metodu olarak en çok kullanılabilirlik metodu kullanılıyorsa normallik sayılıtları vardır. Ayrıca bazı maddelerin kılavuzda farklı yerlerde olması gerektiğini düşünen uzmanlar olabildiği gibi çok farklı maddelerde kesme puanı belirleyen uzmanlar da olabilir (Skaggs ve Tessema, 2001).

Uzmanlar, maddeleri öğrencilerin cevaplayıp cevaplayamayacağına göre kategorilerine ayırıp bir kesme puanı belirlerken bir olasılık kararı vermek zorundadırlar. Bu değer madde karakteristik eğrisinde θ_k yetenek düzeyindeki kişinin i maddesine doğru cevap verme olasılığına (RP) denk gelir (Çetin, 2011). Cevaplama olasılığı genellikle 0.67 ya da 0.50 olarak alınır fakat 0.50 ile 0.80 arasındaki değerler de kullanılmaktadır (Huyhn, 2006). Yüksek RP değerleri yüksek performans değerleri için kullanılmalıdır. Cevaplama olasılığı olarak kullanılacak değer madde bilgi fonksiyonu ile ilgilidir. Huyhn'a göre doğru cevabın madde bilgi fonksiyonunun maksimum olduğu noktayı dikkate almak önemlidir; kestirim 3 parametrelili lojistik model kullanılarak yapılacaksa RP değeri 0.67

olduğunda madde bilgi fonksiyonu maksimum değerini alır, kestirim, 2 parametrelili lojistik model ya da 1 parametrelili lojistik model kullanılarak yapılacaksa RP değeri 0.50 olduğunda bilgi fonksiyonu maksimum değerini alır. Wang (2003) yaptığı çalışmada RP değeri 0.50 olduğunda Rasch modelinin bilgi fonksiyonunun maksimum değer alacağını söylemektedir. Bazı araştırmacılar uzmanların 0.50 cevaplama olasılığının yeterli düzeyi belirlemede düşük olabileceğini savunurlar. Williams ve Schulz (2005) RP değerinin en az 0.67 olması gerektiğini savunurlar.

Madde güçlüğü ve RP değeri bindiğinde, RP değerine eşit bir başarı için ihtiyaç duyulan yetenek düzeyi belirlenebilir. Bu yetenek düzeyi işaretleme yönteminin güçlük noktasıdır (Çetin, 2011). Belirlenen kestirim modeline göre yetenek düzeyi hesaplaması aşağıdaki gibidir:

Öncelikle her madde için madde güçlüğü MTK kullanılarak hesaplanır ve sıralı madde kitapçığı bu güçlük değerlerine göre en düşük yetenek düzeyi değerinden en yüksek yetenek düzeyi değerine yani maddeler kolaydan zora olacak şekilde oluşturulur. Bu yetenek düzeyi, işaretleme yönteminin güçlük noktasıdır (bookmark difficulty location BDL) (Çetin, 2011).

Beretvas (2004), 1 Parametrelili Lojistik Model ve 2 Parametrelili Lojistik Model için güçlük değeri, BDL, hesaplamalarını aşağıdaki gibi yapmıştır.

1 Parametrelili Lojistik Model için BDL aşağıdaki gibi hesaplanır:

1 Parametrelili Lojistik Model Cevaplama Olasılığı $RP = \frac{2}{3}$ ve θ değeri için

$P(X = 1|\theta) = \frac{2}{3}$ hesaplanmalıdır.

$$P(X = 1|\theta) = \frac{e^{(\theta - \hat{b})}}{1 + e^{(\theta - \hat{b})}} = \frac{2}{3}, \text{ denkleminde gerekli işlemler ve sadeleştirmeler}$$

yapıldığında denklem, $\ln 2 = \theta - \hat{b}$ denklemine dönüşür.

1 Parametrelili Lojistik Model Cevaplama Olasılığı $RP = \frac{2}{3}$ için BDL,

$$\theta_{1PL} = \ln 2 + \hat{b} \text{ şeklinde bulunur.}$$

En genel haliyle 1 Parametrelili Lojistik Model Cevaplama Olasılığı herhangi bir RP değeri (X) için BDL,

$$\theta_{1PL} = \ln\left(\frac{X}{1-X}\right) + \hat{b} \text{ şeklinde bulunur.}$$

2 Parametrelili Lojistik Model için BDL aşağıdaki gibi hesaplanır:

2 Parametrelili Lojistik Model Cevaplanma Olasılığı $RP = \frac{2}{3}$ ve θ değeri için

$$P(X = 1|\theta) = \frac{2}{3} \text{ hesaplanmalıdır.}$$

$$P(X = 1|\theta) = \frac{1}{1 + e^{-D\hat{a}(\theta - \hat{b})}} = \frac{2}{3}, \text{ denkleminde gerekli işlemler ve sadeleştirmeler}$$

yapıldığında denklem, $\ln 2 = D\hat{a}(\theta - \hat{b})$ denklemine dönüşür.

2 Parametrelili Lojistik Model Cevaplanma Olasılığı $RP = \frac{2}{3}$ için BDL,

$$\theta_{2PL} = \left(\frac{1}{D\hat{a}}\right) \ln 2 + \hat{b} \text{ şeklinde bulunur.}$$

En genel haliyle 2 Parametrelili Lojistik Model Cevaplanma Olasılığı herhangi bir RP değeri (X) için BDL,

$$\theta_{2PL} = \left(\frac{1}{D\hat{a}}\right) \ln\left(\frac{X}{1-X}\right) + \hat{b} \text{ şeklinde bulunur.}$$

1 Parametrelili Lojistik Model kullanıldığında maddeleri yetenek düzeyine göre sıralamakla madde güçlüklerine göre sıralamak arasında bir fark yoktur, her iki durumda da aynı sıralı soru kitapçığı elde edilir. Fakat 3 Parametrelili Lojistik Model ya da 2 Parametrelili Lojistik Model kullanıldığında madde güçlüğü ve yetenek seviyelerine göre yapılan sıralama farklı olacaktır. Çünkü bu modellerde yetenek (θ), madde güçlüğü ve madde ayırt ediciliğinin bir fonksiyonu olarak hesaplanır (Çetin, 2011).

İşaretleme yönteminde, kesme puanı belirleme yönteminin temeli MTK'ya dayandığından bir sonraki bölümde MTK'dan kısaca bahsedilecektir.

1.7.3.1. Madde Tepki Kuramı

Madde Tepki Kuramı aşağıda belirtilen 2 temel esasa dayanır (Hambleton, Swaminathan ve Rogers; 1991).

1. Sınava giren bireylerin bir test maddesindeki performansları gizil özellik ya da yetenek denilen faktörle açıklanabilir.

2. Sınava giren bireylerin madde performansı ile madde performansı altındaki özellikler seti arasında madde karakteristik fonksiyonu ya da madde karakteristik eğrisi denilen monoton artan bir fonksiyon tanımlanmıştır. Bu fonksiyonda gizil özelliğin seviyesi artarken maddenin doğru cevaplanma oranı da artar.

MTK modelleri, KTK'nın aksine yanlılanabilen modellerdir. Daha önceden denenmiş maddelerde madde parametreleri gruba bağımlı değilken yetenek kestirimi de teste bağımlı değildir. Farklı madde setlerinden kestirilen yetenek ve farklı gruplardan elde edilen madde parametreleri aynı olacaktır, MTK'da madde ve test parametreleri (ölçme hatası hariç) değişmezdir. MTK standart hatayı her bir bireyin yetenek kestirimi için hesaplarırken KTK bütün sınava girenler için tek bir standart hata değeri hesaplar (Hambleton, Swaminathan ve Rogers; 1991).

MTK uygulanabilmesi için (1) tek boyutluluk, (2) yerel bağımsızlık varsayımlarının sağlanması gerekir (Hambleton, Swaminathan ve Rogers; 1991).

Tek Boyutluluk: Testte yer alan maddelerin tek bir yeteneği ölçmeye yönelik olması anlamına gelir. Kesin bir şekilde sağlandığını söylemek oldukça zordur. Çünkü test çözülürken kaygı, motivasyon, süre gibi bir çok başka özellik karışır. Bu yüzden testin baskın bir özelliği ölçüp ölçmediğine bakılmalıdır (Hambleton, Swaminathan ve Rogers; 1991).

Yerel Bağımsızlık: Belirli bir yetenek düzeyinde olan bireylerin bir maddeyi cevaplarırken başka maddelere verdiği cevaplardan bağımsız davranmasıdır. Tek boyutluluk varsayımı sağlanırsa, yerel bağımsızlık da sağlanmış olur. (Lord, 1980).

Kestirilen parametre sayısının ve madde karakteristik fonksiyonun matematiksel formu için farklılaşan birçok MTK modeli vardır. Tüm MTK modelleri sınava giren bireyler için bir ya da daha çok parametre içerir ve tüm MTK uygulamalarının ilk adımı bu parametreleri kestirmektir. Madde karakteristik eğrisini tanımlayan bu parametreler b, a ve c parametreleridir. Genelde (-2,2) aralığında değer alan b parametresi güçlük yani maddenin doğru cevaplanma oranını belirtirken ($-\infty$, ∞)

aralığında değer alan a parametresi ayırt edicilik, c parametresi de şans başarısını belirtmektedir (Hambleton, Swaminathan ve Rogers; 1991).

En yaygın kullanılan MTK modeli bir parametrelili lojistik regresyondur. (Hambleton, Swaminathan ve Rogers; 1991). Madde karakteristik eğrisi aşağıdaki eşitlik ile gösterilir:

$$P_i(\theta) = \frac{e^{(\theta-b_i)}}{1+e^{(\theta-b_i)}} \quad i = 1,2,3,\dots,n$$

$P_i(\theta)$: Tesadüfi seçilen θ yetenek düzeyindeki kişinin i maddesini cevaplama olasılığı (Bu değer 0 ile 1 arasındadır.)

b_i : i maddesinin güçlüğü

1 parametrelili lojistik regresyon tüm maddelerin eşit ayırıcı olduğu varsayımına dayanmaktadır.

MTK modellerinden yaygın olarak kullanılan bir başka model 2 parametrelili lojistik regresyondur. 1 parametrelili modelin eşit ayırt edicilik varsayımının karşılanmadığı durumlarda kullanılabilmesi amacıyla geliştirilmiştir. Madde karakteristik eğrisi aşağıdaki eşitlik ile gösterilir:

$$P_i(\theta) = \frac{e^{D a_i(\theta-b_i)}}{1+e^{D a_i(\theta-b_i)}} \quad i = 1,2,3,\dots,n$$

Bu eşitlikte 1 parametrelili lojistik regresyonda verilen eşitlikten farklı olarak,

D: 1.7

a_i : i maddesinin ayırıcılığı

eklenmiştir.

İki parametrelili modele ek olarak şans parametresi eklenerek üç parametrelili model oluşturulmuştur. Üç parametrelili modelin madde karakteristik eğrisi aşağıdaki eşitlik ile gösterilir:

$$P_i(\theta) = c_i + (1-c_i) \frac{e^{D a_i(\theta-b_i)}}{1+e^{D a_i(\theta-b_i)}} \quad i = 1,2,3,\dots,n$$

c_i : i maddesinin şans parametresi

Kestirilen parametre sayısı arttıkça modellerin karmaşıklığı da artmaktadır.

Bu çalışmada MTK, İşaretleme yönteminde madde parametrelerini belirlemede ve BDL (güçlük noktaları) hesaplamada kullanılmaktadır.

2. İLGİLİ ARAŞTIRMALAR

Bu bölümde, ilgili alanyazında yurt içinde ve yurt dışında yapılan standart belirleme yöntemlerinin karşılaştırılmasına dayanan araştırmalara ve bu araştırmaların sonucunda ulaşılan bulgulara yer verilmiştir.

2.1. Yurt İçinde Yapılan Standart Belirleme Yöntemleri İle İlgili Çalışmalar

Tanrıverdi (2006) çalışmasında standart belirleme yöntemlerinin geçme puanları üzerine etkisini araştırarak öğrencilerin geçti/kaldı durumlarını farklı standart belirleme yöntemlerine göre karşılaştırmıştır. Çalışmada 25'er maddelik Türkçe ve matematik testleri 129 ilköğretim 7. sınıf öğrencisine uygulanmıştır. Çalışmanın sonucunda Türkçe ve matematik dersi için Angoff, Nedelsky ve Evet/Hayır standart belirleme yöntemleri ile belirlenen geçme puanları üzerinde puan alan öğrenci yüzdeleri arasında anlamlı bir farklılık olduğu görülmüştür. Çalışmada Türkçe ve matematik testleri için yöntemler arasında uyum bulunmuştur. Ayrıca çalışmada geçme puanı belirlerken standart belirleme yöntemlerinden hangisinin seçileceğinin öğrenci grubuna, derse, kullanılan testin yapısına ve öğretmenlere bağlı olarak değişebileceği sonucuna varılmıştır.

Taşdelen (2009) Nedelsky ve Angoff yöntemleri ile belirlenen kesme puanlarının genellenebilirlik kuramı ile karşılaştırdığı çalışmasında Ankara ilinde görev yapan 40 uzman ile çalışmıştır. Araştırmanın sonucunda yöntemlerle belirlenen kesme puanları arasında anlamlı bir farklılık olduğu görülmüştür. Genellenebilirlik kuramı kullanıldığında yöntemlerin birlikte ele alındığı tümüyle çaprazlandırılmış modelde her iki yöntemde de maddeler arası farklılıkların olduğu ve uzmanlar arasında tutarlılık olduğu bulunmuştur. Çalışmada ayrıca her iki yöntemle de standart belirlerken 10 uzmanın yeterli olduğu sonucuna varılmıştır.

Ömür ve Selvi (2010) çalışmalarında farklı uzman sayıları kullanarak Angoff, Ebel ve Nedelsky yöntemleriyle elde edilen kesme puanlarının sınıflama tutarlığı indekslerini incelemiştir. Araştırmada 30 uzman ve 176 ilköğretim 8. sınıf öğrencisi ile çalışılarak uzman sayısına bağlı olarak belirlenen kesme puanlarında önemli bir değişim olmadığı sonucuna ve üç yöntemde üç ayrı uzman sayısında en yüksek sınıflama tutarlığı indeksinin Nedelsky yöntemiyle elde edildiği sonucuna varılmıştır.

Çetin (2011) çalışmasında bir üniversitenin yeterlik sınavının kesme puanı ile standart belirleme yöntemlerinden Angoff ve İşaretleme yöntemleri ile belirlenen kesme puanlarını karşılaştırmıştır. Çalışmada Angoff yönteminde uzmanların belirlediği madde cevaplama olasılıkları ile gerçek madde güçlükleri arasındaki korelasyon katsayısı incelenmiş ve bu katsayı 0.60 bulunmuştur. Araştırmada üniversitenin belirlediği kesme puanı ile Angoff ve İşaretleme yöntemleri ile belirlenen kesme puanları arasında anlamlı bir fark olduğu fakat Angoff ile İşaretleme yöntemleri ile belirlenen kesme puanları arasında anlamlı bir fark olmadığı bulunmuştur. Çalışmada farklı Madde Tepki Kuramı modelleri ve cevaplama olasılığı değerlerinin sıralı madde kitapçığındaki maddelerin yerlerini değiştirip değiştirmediği incelenmiş, bir parametrelili lojistik modelde herhangi bir değişiklik olmadığı iki parametrelili lojistik modelde kılavuzun 33 maddesinin değiştiği bulunmuştur.

Gündeğer (2012) araştırmasında Angoff, Evet/Hayır ve Ebel standart belirleme yöntemleri ile elde edilen kararları ve kesme puanlarını karşılaştırmıştır. Araştırmada, hazırlanan 20 soruluk bir matematik testi 489 öğrenci tarafından çözülmüş ve bu çalışmada 17 uzman kullanılmıştır. Araştırmanın sonucunda başarılı sayılan öğrenci yüzdeleri arasında farklılık olduğu ve belirlenen kesme puanları açısından yöntemler arasında genel bir uyum olmasına rağmen yöntemlerle belirlenen geçme puanlarının farklı olduğu bulunmuştur. Araştırmada uzmanlar arasındaki uyum da incelenmiş ve yöntemlere göre alınan uzman kararları arasında makul derecede bir uyum olduğu sonucuna ulaşılmıştır. Yöntemlere ait uzman kararlarının genellenebilirlik kuramı ile incelenmesiyle üç yöntemin maddelere ait farklılıkları fazla ortaya çıkaramadığı; uzmanlar arası tutarlılık bakımından en düşük sonucu veren yöntemin Ebel yöntemi olduğu, onu takiben Angoff yönteminin geldiği ve en yüksek uzman tutarlılığını Evet/Hayır yönteminin verdiği sonucuna ulaşılmıştır. Ayrıca üç yöntemin birlikte ele alındığı genellenebilirlik çalışmasında maddeler arası farklılığın ortaya çıkarılmadığı; uzmanlar arasında uyumun olduğu ve yöntemler arasında da büyük farklılığın olduğu sonucuna ulaşılmıştır. Ayrıca araştırmada Angoff ve Evet/Hayır yöntemleri ile kesme puanı belirlenirken 17 civarında uzmanın; Ebel yönteminde ise 22 civarında uzmanın yeterli olacağı sonucuna ulaşılmıştır.

Çukadar (2013) mutlak ve bağıl değerlendirmelerin karşılaştırıldığı çalışmasında 20 soruluk bir test 107 öğrenciye uygulanarak sonuçlar T puanına çevrilmiş ve 50 ile 60 ölçütleri dikkate alınarak öğrenciler başarılı-başarısız olarak sınıflanmıştır. Çalışmada 14 uzmandan alınan görüşler doğrultusunda Angoff ve Nedelsky yöntemlerine göre kesme puanları belirlenerek yine öğrenciler başarılı-başarısız olarak sınıflandırılmışlardır. Çalışmanın sonucunda başarılı öğrenci yüzdeleri, kullanılan değerlendirme yöntemlerine göre ikili olarak karşılaştırıldığında yöntemler arasında anlamlı farklılıklar bulunmuştur. Çalışmanın sonucunda mutlak ve bağıl değerlendirmeler sonucu alınan kararlarda anlamlı farklılıklar olduğu sonucuna ulaşılmıştır. Yöntemlerin uygunluk geçerliğinin incelenmesi amacıyla öğrencilerin dönem sonu notları kullanılmış ve uygunluk geçerliği açısından sadece Nedelsky yönteminin ders notlarıyla uyumlu olduğu, diğer yöntemlerin ise düşük uyuma sahip olduğu sonucuna ulaşılmıştır.

Taşdemir (2013) çalışmasında Evet/Hayır ve Nedelsky standart belirleme yöntemleri ve sınır değer saptama yöntemleri olan İşlem Karakteristik Eğrisi (ROC) ve Aralık Tahmini yöntemi ile belirlenen sınıflama tutarlılıklarının doğruluklarını incelemiştir. Araştırmada veri toplama aracı olarak Türkçe ve matematik sorularından oluşan 30 maddelik bir test, çalışma grubu olarak 211 öğrenci ve 40 uzman kullanılmıştır. Araştırmanın sonucunda en yüksek kesme puanı Evet/Hayır yöntemi ile elde edilmiş olup, yöntemler arasında ve başarılı sayılan öğrenci yüzdeleri arasında manidar bir farklılık olduğu sonucuna ulaşılmıştır. Ayrıca Evet/Hayır yönteminde uzman kararları arasında uyum olmadığı, Nedelsky yönteminde uzman kararları arasında uyum olduğu bulunmuştur. Araştırmada ROC analizinde belirlenen kesme puanlarıyla Aralık Tahmini yöntemine göre belirlenen kesme puanları uyumlu bulunmuştur. ROC analizinin, kesme puanı belirleme ve öğrencileri sınıflamada doğru karar verilmesinde yaygın olarak kullanılan Evet/Hayır ve Nedelsky yöntemleri gibi kullanılabileceği sonucuna ulaşılmıştır.

Demir (2014) Angoff, Ebel ve Nedelsky yöntemleri ile belirlenen kesme puanlarını karşılaştırdığı çalışmasında 245 öğrenciye 26 maddelik bir matematik testi uygulamıştır. Araştırmada belirlenen kesme puanlarına göre başarılı olan öğrenci yüzdeleri, öğrencilerin belirlenen kesme puanlarına göre başarılı-başarısız olma durumları ve uzmanların her bir yöntem için belirledikleri MGP'ler arasındaki uyum

kıyaslanmıştır. Araştırmanın sonucunda en yüksek kesme puanının Angoff yöntemiyle, en düşük kesme puanının Nedelsky yöntemiyle belirlendiği ve farklı standart belirleme yöntemlerine göre kesme puanının üzerinde puan alarak başarılı kabul edilen öğrenci yüzdelerinin anlamlı şekilde farklılaştığı görülmüştür. Angoff ve Ebel yöntemleriyle belirlenen kesme puanları arasında yüksek düzeyde anlamlı bir ilişki olduğu görülürken, Nedelsky ve Ebel yöntemleri arasında düşük düzeyde bir ilişki olduğu bulunmuştur. Ayrıca çalışmada üç yöntemle ait minimum geçme puanlarının ortalamaları arasında anlamlı fark bulunamamıştır.

2.2. Yurt Dışında Yapılan Standart Belirleme Yöntemleri İle İlgili Çalışmalar

Zahran (1981) uzman deneyiminin, madde güçlüğü'nün ve madde tipinin standart belirleme üzerindeki etkisini araştırdığı çalışmasında Nedelsky yöntemini kullanmıştır. Araştırmanın sonucunda deneyimsiz uzmanlar, deneyimli uzmanlara göre yüksek geçme puanı belirlerken, grafik ve ilişki içeren sorularda çoktan seçmeli sorulara göre uzmanların daha yüksek standart belirlediği sonucuna ulaşılmıştır. Ayrıca madde güçlük düzeyi, belirlenen kesme puanları ve uzmanlar arasındaki uyum arasında ilişki olduğu sonucuna ulaşılmıştır.

Livingston ve Zieky (1983) yaptığı çalışmalarında dört okulda sınır grup ve karşıt grup yöntemlerini Nedelsky yöntemi ile başka dört okulda sınır grup ve karşıt grup yöntemlerini Angoff yöntemi ile ve bu okullarda aynı yöntemleri kendi aralarında okuma ve matematiksel temel beceri testlerini kullanarak karşılaştırmışlardır. Yeterli ve yetersiz olarak sınıflandırılan öğrenci sayısı eşit olduğunda sınır grup ve karşıt grup yöntemlerinin benzer sonuçlar verdiği görülmüştür. Yetersiz olarak sınıflandırılan öğrenciler fazla olduğunda sınır grup yöntemi ile belirlenen geçme puanlarının karşıt grup yöntemi ile belirlenen geçme puanına göre daha düşük çıktığı, yetersiz olarak sınıflandırılan öğrenci sayısı az olduğunda daha yüksek çıktığı sonucuna ulaşılmıştır. Ayrıca okullar arası Nedelsky ve Angoff yöntemi karşılaştırıldığında tutarlı sonuçlar elde edilmediği görülmüştür.

Chang (1996) Angoff ve Nedelsky yöntemlerini karşılaştırdığı çalışmasında Nedelsky yönteminin, Angoff yöntemine göre daha düşük kesme puanları belirlediği ve uzmanlar arası uyumun Nedelsky yönteminde daha yüksek düzeyde olduğu sonuçlarına ulaşmıştır. Nedelsky yöntemi ile belirlenen kesme puanlarının öğrencilerin gerçek puanlarına daha yakın olduğu, Angoff yöntemi ile belirlenen

kesme puanlarının öğrencilerin gerçek puanlarından büyük oranda farklı olduğu sonucuna varılmıştır.

Shen (2001) çalışmasında Angoff yöntemi, Rasch modeli kullanılan madde haritalama yöntemi ve Hoffsee yöntemini karşılaştırmıştır. 23 uzman görüşü doğrultusunda Angoff yönteminde uzmanlar arası uyumun daha düşük olduğu, Angoff yönteminin belirlenen puanlar açısından daha fazla çeşitlilik gösterdiği sonucuna varılmıştır. Haritalama yöntemi ile belirlenen kesme puanı, Angoff yöntemi ile belirlenen kesme puanından daha yüksek bulunurken Hoffsee yöntemi ile belirlenen kesme puanından farklı olmadığı bulunmuştur.

Buckendahl, Smith, Impara, Plake (2002) Angoff ve işaretleme yöntemlerini karşılaştırdıkları çalışmalarında öğrencilere 7. sınıf matematik konularından oluşan 69 maddelik bir test uygulayarak kesme puanları belirlemişlerdir. Araştırma sonucunda her iki yöntem için de belirlenen kesme puanları arasında büyük farklılıklar gözlenmemiştir.

Green, Trimble ve Lewis (2003), İşaretleme yöntemi, Karşıt Gruplar yöntemi ve Jeager yöntemleriyle elde edilen kesme puanlarını karşılaştırmışlardır. Araştırmanın sonucunda en düşük kesme puanı İşaretleme yöntemiyle en yüksek kesme puanı Jeager yöntemiyle belirlenmiştir.

Yin, Schulz ve Sconing (2005) Angoff ve haritalama yöntemlerini karşılaştırdığı çalışmalarında 2005 NAEP 12. sınıf matematik sorularını kullanmışlardır. Belirlenen kesme puanları dağılımları incelendiğinde haritalama yöntemiyle belirlenen kesme puanlarının daha değişken olduğu ve normalden daha fazla saptığı bulunmuştur. Araştırmanın sonucunda haritalama yönteminin Angoff yöntemine göre daha düşük geçme puanı belirlediği bulunmuştur. Geçme puanları arasında oluşan fark, yöntemlerin uygulanma süreçlerinin farklı olmasına bağlanmıştır.

Davis, Buckhendal, Chin ve Gerrow (2008) Evet/Hayır ve İşaretleme yönteminin karşılaştırıldığı çalışmalarında 34 uzman ile çalışılmıştır. Çalışmada uzmanların 17'si önce İşaretleme yöntemiyle geçme puanı belirlerken diğer 17'si önce Evet/Hayır yöntemiyle geçme puanı belirlemişlerdir. Bu sürecin sonunda belirlenen geçme puanları arasında farklılık olduğu sonucuna ulaşılmıştır.

Olsen ve Smith (2008) modifiye edilmiş Angoff ve İşaretleme yöntemlerini karşılaştırdıkları çalışmalarında 120 çoktan seçmeli maddeden oluşan bir test kullanmışlardır. Araştırmanın sonucunda iki yöntem ile de benzer geçme puanları belirlendiği sonucuna ulaşılmıştır.

Peterson, Schulz ve Engelhard (2011) Angoff ve İşaretleme yöntemlerinin geçerlik ve güvenilirliklerini karşılaştırmak amacıyla 27 çalışma incelemiştir. İncelemeler sonucunda; İşaretleme yönteminin standart hatasının daha düşük olduğu, kesme puanları ve uzman değerlendirmeleri açısından İşaretleme yönteminin Angoff yöntemine göre daha geçerli ve kullanışlı olduğu sonuçlarına ulaşılmıştır.

Hsieh (2013) çalışmasında İngilizce konu kapsamı değerlendirilmesi amacıyla Evet/Hayır ve İşaretleme standart belirleme yöntemlerini karşılaştırmıştır. 32 uzmandan alınan görüşler doğrultusunda iki yöntemle belirlenen kesme puanları arasında büyük farklılık olduğu sonucuna ulaşılmıştır. Çalışmada İşaretleme yöntemi ile elde edilen puanların standart sapmasının Evet/Hayır yöntemi ile elde edilen puanlardan daha küçük olduğu sonucuna ulaşılmıştır. Ayrıca araştırmada her iki yöneme göre de uzmanlar arası uyum olduğu belirlenmiştir.

2.3. İlgili Araştırmalar Özet

Standart belirleme yöntemleri ve bu yöntemlerin karşılaştırılmasına yönelik yapılan çalışmaların sayısının yurt dışında, yurt içine göre daha fazla olduğu görülmektedir. Yöntemlerin işleyişi bakımından kolay olması nedeniyle yapılan çalışmaların çoğunda test merkezli standart belirleme yöntemlerinden Angoff ve Nedelsky yöntemlerinin kullanıldığı görülmektedir. Çalışmalarda, geçme puanları arasındaki farklılıklar ile ilgili yöntem kıyaslamaları ve uzmanlar arası uyum konularına odaklanılmıştır. Bunun yanında cevaplama olasılığının (RP) karşılaştırıldığı, uzman sayısının ve uzman özelliklerinin geçme puanlarına etkisinin araştırıldığı çalışmalara da az da olsa rastlanmaktadır. Ayrıca çalışmaların hemen hemen hepsinde veri toplama aracı olarak çoktan seçmeli test kullanılmıştır.

Farklı standart belirleme yöntemleri ile belirlenen geçme puanlarının kıyaslandığı çalışmalarda genellikle yöntem fark etmeksizin geçme puanlarının üstünde puan alan öğrenci yüzdeleri arasında manidar bir fark olduğu bulunmuştur. Yapılan çalışmaların çoğunda standart belirleme yöntemleriyle geçme puanı belirlenirken

10 uzmanın yeterli olacağı sonucuna ulaşılmıştır. Uzman sayısının geçme puanına etkisinin araştırıldığı çalışmalarda uzman sayısının 10'dan fazla olmasının belirlenen geçme puanını önemli bir ölçüde etkilemediği sonucuna ulaşılmıştır. Uzmanlar arasındaki uyumun incelendiği çalışmalarda genellikle uzmanlar arasında makul derecede uyum olduğu sonucuna ulaşılrken en yüksek uzman tutarlılığının Nedelsky yönteminde, en düşük uzman tutarlılığının Ebel yönteminde olduğu sonucuna varılmıştır.

Benzer bir araştırma olarak, standart belirleme yöntemleriyle belirlenen geçme puanlarının üstünde puan alan öğrenci yüzdesinin kıyaslandığı, uzman özelliklerinin geçme puanına etkisinin araştırıldığı ve her bir yöntem için uzmanlar arasındaki uyumun belirlendiği bu tez çalışması yapılmıştır.

3. YÖNTEM

3.1. Araştırmanın Yöntemi

İki ya da daha çok sayıdaki değişkenin arasındaki birlikte değişim derecesini belirlemeyi amaçlayan araştırma modelleri ilişkisel türden araştırma modelleridir (Karasar, 1982). Bu çalışmada üç farklı standart belirleme yöntemi ile belirlenen geçme puanları çeşitli açılardan karşılaştırılmıştır. Çalışma bu açıdan ilişkisel türden bir araştırmadır. Olayların, objelerin, varlıkların, kurumların, grupların ve çeşitli alanların “ne” olduğunu açıklayan araştırma modellerine betimsel araştırmalar denir. (Kaptan, 1977). Araştırmada Evet/Hayır, İşaretleme ve Ebel yöntemleri ile belirlenen geçme puanları incelendiğinden çalışma bu yönüyle betimsel bir araştırma niteliği taşımaktadır.

3.2. Katılımcılar

3.2.1. Katılımcıların Özellikleri

Araştırmanın öğrenciler ve uzmanlar olmak üzere iki farklı katılımcı grubu vardır. Araştırmada her bir yöntem için geçme puanı belirleyen 30 uzman ile çalışılmıştır. Çalışılan uzmanlar matematik öğretmenleri ve eğitsel tasarımcılardan oluşmaktadır. Uzmanlara elverişli örnekleme yöntemiyle ulaşılmıştır, uzman seçimi gönüllülük durumuna göre yapılmıştır.

Çalışmanın öğrenci grubunu 505 9. sınıf öğrencisi oluşturmuştur. Öğrenciler Ankara ağırlıklı olmak üzere Türkiye'nin çeşitli illerinde, farklı lise tiplerinde okuyan öğrencilerdir. Bu öğrencilere araştırmacı tarafından geliştirilmiş olan ve kümeler ünitesinin sorularından oluşan 25 soruluk bir matematik başarı testi uygulanmıştır.

3.2.2. Katılımcılarla İlgili Demografik Bilgiler

Araştırmada 10 erkek, 20 kadın uzmanla çalışılmıştır. Uzmanların 20'si öğretmen 10'u eğitsel tasarımcıdır. Uzmanların deneyimlerine bakıldığında 4 yıldan az tecrübesi olan 13, 4-7 yıl arası tecrübesi olan 9, 7 yıldan fazla tecrübesi olan 8 uzman olduğu görülmektedir.

3.3. Veri Toplama Araçları

Bu araştırmada öğrencilerden veri toplamak için araştırmacı tarafından geliştirilmiş, 9. sınıf matematik dersi kümeler konusuna yönelik hazırlanmış 25 maddelik bir başarı testi kullanılmıştır. (Ek 1)

Araştırmada uzmanlardan, hazırlanan bu başarı testindeki maddeleri kullanarak üç farklı standart belirleme yöntemine göre geçme puanı belirlemeleri istenmektedir. Bu amaçla Evet/Hayır Formu (Ek 2), Ebel Formu (Ek 3) ve İşaretleme Formu, Bilgilendirme Formu (Ek 4) ve anket (Ek 5) kullanılmıştır.

3.3.1. Kümeler Başarı Testi

Araştırmada kullanılan başarı testi, araştırmacı tarafından “Test Geliştirme Teknikleri” dersinde geliştirilen, matematik dersi kümeler alt öğrenme alanının 8 kazanımına yönelik beş seçenekli 40 soruluk bir testten 25 soru seçilerek oluşturulmuştur. Geliştirilen başarı testinin deneme uygulaması 2013 yılında 110 öğrenciye 1 ders saati boyunca uygulanmış ve maddeler ITEMAN programı kullanılarak analiz edilmiştir. Madde seçiminde ayırt edicilik gücü indeksi 0.30’dan büyük olan, madde güçlüğü 0.50 civarında olan ve çeldiricileri en iyi işleyen maddeler dikkate alınmıştır. Deneme uygulaması sonucunda testin betimsel istatistikleri Tablo 4’de verilmiştir.

Tablo 4: Deneme Uygulaması Sonuçlarına Göre Test İstatistikleri

	<i>Madde sayısı</i>	<i>Öğrenci sayısı</i>	<i>Standart sapma</i>	<i>Ortalama</i>	<i>Testin güçlüğü</i>	<i>KR-20</i>
Test Sonuçları	40	110	6.123	22.4	0.560	0.782

Tablo 4’den görüldüğü gibi deneme uygulaması sonuçlarından elde edilen iç tutarlık katsayısı 0.78’dir. Testin güçlüğü ise (0.56) orta düzeydedir.

Deneme uygulaması sonucunda seçilen maddelerin oluşturduğu nihai testin betimsel istatistikleri Tablo 5’te verilmiştir.

Tablo 5: Nihai Test Sonuçlarına Göre Test İstatistikleri

	<i>Madde sayısı</i>	<i>Öğrenci sayısı</i>	<i>Standart sapma</i>	<i>Ortalama</i>	<i>Testin güçlüğü</i>	<i>KR-20</i>
Test Sonuçları	25	110	4.658	12.582	0.503	0.771

Tablo 5’e göre nihai test sonuçlarından elde edilen iç tutarlık katsayısı 0.771’dir. Murhpy ve Davidshofer’a göre (1991) sınıf içi başarı testlerinde iç tutarlık katsayısının en az 0.75 olması gerekmektedir. Bu bakımdan oluşturulan nihai testin yeterli olduğu ifade edilebilir. Testin kapsam geçerliğini sağlamak amacıyla 3

alan uzmanından testin kapsam geçerliğinin yüksek olduğu görüşü alınmıştır. Nihai testin orta güçlükte (0.50) olduğu ifade edilebilir.

3.3.2. Bilgilendirme Formu

Uzmanların kararı ile belirlenecek geçme puanı için uzmanlara sürecin işleyişini anlatmak amacıyla hazırlanmıştır. Bilgilendirme formunda her bir yöntem için uzmanların yöntemlerde nasıl bir yol izleyerek formları dolduracakları anlatılmaktadır.

3.4. Veri Toplama Araçlarının Uygulanışı

Araştırmacı tarafından daha önceden geliştirilen, 25 soruluk matematik dersi “kümeler” konusuna yönelik, başarı testi bu araştırma kapsamında 505 9. sınıf öğrencisine bir ders saati süresi boyunca uygulanmıştır. Uygulamada öğrencileri motive etmek için öğretmenleri tarafından testten alınacak puanın sözlü notu olarak kullanılacağı söylenmiştir. 2014 yılının Ekim ayında yapılan uygulama sonuçlarına göre testin betimsel istatistikleri Tablo 6’da verilmiştir.

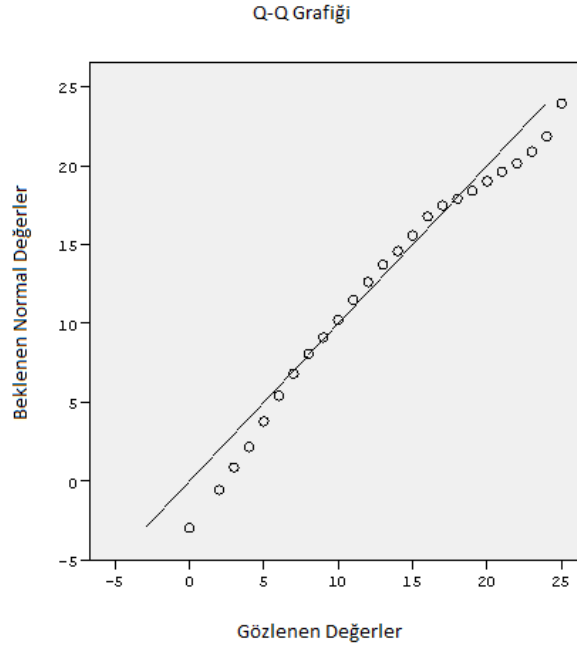
Tablo 6: Test Uygulama Sonuçlarına Göre Test İstatistikleri

<i>Test İstatistikleri</i>	
<i>Madde sayısı</i>	25
<i>Öğrenci sayısı</i>	505
<i>Aritmetik ortalama</i>	12.71
<i>Varyans</i>	26.3
<i>Standart sapma</i>	5.13
<i>Basıklık</i>	0.53
<i>Çarpıklık</i>	-0.07
<i>Minimum puan</i>	0
<i>Maksimum puan</i>	25
<i>Medyan</i>	12
<i>KR-20</i>	0.82
<i>Testin güçlüğü</i>	0.51
<i>Ortalama çift serili korelasyon</i>	0.55

Tablo 6’dan görüldüğü gibi 505 kişinin çözdüğü 25 maddelik testin ortalaması 12.71 bulunmuştur. Testin ortalama güçlüğü 0.51 olarak hesaplanmıştır. Buna göre testin orta güçlükte olduğu ifade edilebilir. Testin standart sapması 5.13, medyanı 12 bulunmuştur. Testin güvenilirliği 0.82 bulunduğundan testin iç tutarlılığın yeterli olduğu ifade edilebilir. (Murphy ve Davidshofer, 1991). Testin ortalama ayırıcılığı (0.55) testin çok iyi düzeyde ayırıcı olduğunu belirtmektedir (Crocker ve Algina, 1986).

Yapılan uygulamanın geçerlik çalışması için ulaşılan 210 kişinin 1. dönem karne notları ile uygulama sonucunda aldıkları puanları arasındaki korelasyona bakılmış ve korelasyon değeri 0.85 bulunmuştur ($p < 0.01$).

Testin verilerinin normal dağılıp dağılmadığını test etmek için sık kullanılan yöntemlerden biri olan Q-Q grafiği incelenmiştir. Bu grafikte veriler büyüklüklerine göre artan bir şekilde sıralanır ve beklenen normal dağılım değerlerine karşılık gelecek şekilde grafikleştirilir. Bu grafikte gözlenen değerler x ekseninde, beklenen değerler y ekseninde yer alır. Eğer dağılım normal dağılımdan sapma göstermiyorsa grafik köşegenler odağında bir doğru çizgisi üzerinde olur. Grafikteki noktalar bu çizgiden ne kadar uzaklaşırsa veriler de normal dağılımdan o kadar sapar (Mertler ve Vannatta, 2005).



Şekil 1: Testin Q-Q Grafiği

Test puanlarının Q-Q grafiği incelendiğinde test verilerinin köşegen çizgisi üzerinden pek fazla sapmadığı görülür. Bu durumda test puanlarının normale yakın dağıldığı ifade edilebilir.

25 soruluk test için 30 uzmanın her birinden üç farklı standart belirleme yöntemine göre kararlar alınmıştır. Uzmanlardan, bu üç yöntem için karar alma süreçleri aşağıdaki gibidir:

Evet/Hayır yöntemi için uzmanlardan “Geçme-kalma sınırındaki bir öğrenci (minimum yeterlik düzeyindeki bir öğrenci) test maddesini doğru cevaplayabilir mi?” sorusuna cevap vermesi beklenmektedir. Uzmanlardan öğrencinin doğru cevaplayabileceğini düşündüğü sorular için formdaki doğru cevaplayabilir kutucuğunu, doğru cevaplayamayacağını düşündükleri sorular için doğru cevaplayamaz kutucuğunu işaretlemeleri istenmiştir. Bu şekilde 25 soru için karar alındıktan sonra doğru cevaplayabilir kutucuğu 1, doğru cevaplayamaz kutucuğu 0 puan olarak kodlanmıştır. Her bir uzmanın işaretlediği kutucuklardan elde edilen puanlar toplanıp 4 ile çarpılarak her bir uzmanın 100 üzerinden verdikleri Minimum Geçme Puanları (MGP) hesaplanmıştır. Yönteme ait geçme puanı da MGP’lerin ortalamaları alınarak elde edilmiştir.

Ebel yöntemi için uzmanlardan sorularla ilgili “güçlük” ve “uygunluk” olmak üzere iki konuda karar vermeleri beklenmektedir. Her bir soru için uzmanlardan kolay, orta, zor ve gerekli, önemli, kabul edilebilir, tartışılabilir, seçeneklerinden uygun olanına karar verip hazırlanan formda ilgili hücreye yazmaları istenmiştir. Bu iki karar verildikten sonra geçme-kalma sınırındaki öğrencilerin (minimum yeterlik düzeyindeki öğrencilerin) yüzde kaçının hücredeki soruları doğru cevaplayabileceği yazılması istenmiştir. Elde edilen veriler hücrelerdeki soru sayıları ve yüzdeler ile çarpılarak, bu çarpımlar toplanmış ve her bir uzmana ait MGP’ler hesaplanmıştır. Yönteme ait geçme puanı da MGP’lerin ortalamaları alınarak elde edilmiştir.

İşaretleme yönteminde maddelerin yetenek düzeyleri, en çok olabilirlik kestirim metoduyla MTK kullanılarak güçlük ve ayırtedicilik parametreleri ile kestirilmiştir. İşaretleme yöntemi için hesaplanan yetenek düzeylerine göre testteki sorular kolaydan zora doğru sıralanarak sıralı madde kitapçığı oluşturulup uzmanlardan geçme-kalma seviyesindeki bir öğrencinin bu sorulardan en son hangi sayfadaki soruyu yapabileceğini işaretlemesi beklenir. Bu işaretleme yapıldıktan sonra her bir uzmanın belirlediği sorunun yetenek düzeyi toplanarak ortalama yetenek düzeyi elde edilir. Ortalama yetenek düzeyi her bir madde için İki Parametrelili Lojistik Regresyon formülünde yerine koyularak ham puanlar elde edilir. Testin nihai geçme puanı, elde edilen ham puanlar toplanarak hesaplanır.

3.5. Verilerin İşlenmesi ve Çözülmesi

MTK ve 2 Parametrelili Lojistik Model sayıtlarının incelenmesi ve alt problemlerin çözülmesi için arařtırmada Excel, BILOG MG 3, SPSS 15.0 ve LISREL 8.8 programları kullanılmıřtır.

3.5.1. Tek Boyutluluk

Tek boyutluluk; testin tek bir örtük özelliđi ölçmesi anlamına gelmektedir. Bir test çözülürken teste başka bir çok özellik de (kaygı, motivasyon, süre, vb.) karışacağından testin tek boyutlu olup olmadığı kesin bir şekilde kanıtlanamaz. Bu yüzden testin baskın bir özelliđi ölçüp ölçmediđine bakılmalıdır (Hambleton, Swaminathan ve Rogers, 1991).

Arařtırmada kullanılan testin tek boyutlu olup olmadığı asimptotik kovaryanslar matrisi üzerinden tek boyutlu bir ölçme modeli oluşturularak doğrulayıcı faktör analizi ile test edilmiştir. Doğrulayıcı faktör analizi model veri uyumu istatistikleri aşağıda verilmiştir. Ayrıca testin tek boyutluluđunun sınanmasında kullanılan ölçme modeli Ek 8'de verilmiştir.

Brown (2006) ve Hu ve Bentler (1999) tarafından belirtilen uyum indekslerine göre; RMSEA indeksinin 0.05 altında olması, SRMR indeksinin 0.08 altında olması, CFI ve NFI indeksinin 0.95'den yukarıda olması çok yakın model veri uyumuna işaret etmektedir.

$\chi^2(275) = 579.68$, $p < 0.00$, NFI = 1.00, CFI = 1.00, SRMR = 0.052 ve RMSEA = 0.045 (%90 güven aralıđı 0.045-0.056) olduğundan kullanılan ölçme modeli, istenen modelle çok yüksek derecede uymaktadır. Bu bakımdan arařtırma kapsamında uygulanan testin tek bir baskın özelliđi ölçtüđü ifade edilebilir.

3.5.2. Yerel Bađımsızlık

Belirli bir yetenek düzeyinde olan bireylerin bir maddeyi cevaplarken başka maddelere verdiđi cevaplardan bađımsız davranmasına yerel bađımsızlık denir. Bir testin maddelerinin yerel bađımsızlık özelliđini sağlayıp sağlamadığının tespit edilmesi için tek şart olan testin tek boyutluluđuna bakılmalıdır. Eğer bir test tek boyutlu ise bu testte yer alan maddelerin yerel bađımsızlık özelliđine sahip olduğu ifade edilebilir (Lord, 1980). Arařtırmada kullanılan test tek boyutlu olduğundan testin maddeleri de yerel bađımsızlık şartını sağlanmış olur.

3.5.3. Şans Başarısı

Şans başarısı, yanıtlayıcının çoktan seçmeli testte doğru cevabı ölçülmek istenen bilgiyle değil, tahmine giderek bulmasıdır (Turgut, 1993). Bir testte yer alan maddelere verilen cevapların, tahminle mi yoksa bilgiyle mi cevaplandırıldığının belirlenmesinde doğrusal olmayan madde-test puanı regresyon eğrilerinin kullanılmasını ya da düşük test puanı alan öğrencilerin zor test maddelerinin üzerindeki performansının incelenmesini dikkate alan iki farklı yöntem önerilmektedir (Hambleton ve Swaminathan, 2006).

Tablo 7: Düşük Yetenek Grubundaki Öğrencilerin Testteki Zor Maddeleri Doğru Cevaplama Yüzdeleri

<i>Madde numarası</i>	<i>Madde güçlük düzeyi</i>	<i>Doğru cevaplanma yüzdesi</i>
6	0.38	0.15
13	0.30	0.17
15	0.35	0.21
18	0.25	0.11
19	0.38	0.19

Bu çalışmada şans başarısının testin uygulandığı grupta minimum olduğunun gösterilmesi için düşük test puanı alan öğrencilerin zor test maddeleri üzerindeki performansının incelenmesi yöntemi kullanılmaktadır. Bunun için testin zor olan 5 maddesine düşük yetenek grubunda olan (testin uygulandığı grubun %27'lik alt kesimi) öğrencilerin yüzde kaçının doğru cevapladığı Tablo 7'de verilmiştir. Bu sayıların sağlanması için düşük yetenek grubundaki öğrencilerin zor maddeleri cevaplama oranlarının 0.20'nin altında olması gerekir. Bu incelemede, teorik olarak beş seçenekte tahminle doğru cevaplama olasılığı olan 0.20 sınır olarak alınmıştır. Cevap verme yüzdeleri göz önüne alındığında şans başarısının minimum düzeyde olacağı kabul edilebilir.

Bu çalışmada yetenek parametreleri, 2 Parametrelili Lojistik Model kullanılarak kestirilecektir. Araştırmanın her bir alt probleminin analizi için aşağıdaki adımlar izlenmiştir.

1. Araştırmanın birinci alt probleminin "Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemlerine göre belirlenen geçme puanlarının üzerinde puan alan öğrenci yüzdesi arasında fark var mıdır?" çözümünde; bağımlı iki oran/yüzde arasındaki farkın testi yapılmıştır. Bağımlı iki oran arasındaki farkın testi, yöntemlere göre geçme puanının üstünde puan alan ve alamayan öğrenci

sayılarına ve yüzdelerine ait 2x2'lik bir tablo yardımıyla hesaplanabilmektedir. Aşağıda 2x2'lik tablolara ve değerlerin anlamlarına yer verilmiştir (Ferguson,1971).

		Frekanslar		Yüzdeler		
		Yöntem 2		Yöntem 2		
		Kaldı	Geçti	Kaldı	Geçti	
Yöntem 1	Geçti	A	B	A + B		
	Kaldı	C	D	C + D		
		A + C	B + D	N		

Yöntem 1	Geçti	a	b	p_1
	Kaldı	c	d	q_1
		q_2	p_2	1.00

Şekil 2: Frekans ve yüzdelerine ait Yöntem 1 ve Yöntem 2'ye göre geçen-kalan öğrenci sayıları

Yukarıda frekanslara ait şekilde Yöntem 1 ve Yöntem 2'ye göre geçen-kalan öğrenci sayıları yer almaktadır.

A: Birinci yönteme göre geçip, ikinci yönteme göre kalan öğrenci sayısını,

B: Her iki yönteme göre geçen öğrenci sayısını,

C: Her iki yönteme göre kalan öğrenci sayısını,

D: İkinci yönteme göre geçip, birinci yönteme göre kalan öğrenci sayısını belirtmektedir.

Yüzdelerine ait verilen tabloda Yöntem 1 ve Yöntem 2'ye göre geçen-kalan öğrenci sayılarının tüm öğrenci sayısına bölümü olan geçen-kalan öğrenci oranları yer almaktadır.

a: Birinci yönteme göre geçip, ikinci yönteme göre kalan öğrenci oranını,

b: Her iki yönteme göre geçen öğrenci oranını,

c: Her iki yönteme göre kalan öğrenci oranını,

d: İkinci yönteme göre geçip, birinci yönteme göre kalan öğrenci oranını belirtmektedir.

Bağımlı örneklerde iki yüzde arasındaki farkın anlamlılık testinde z istatistiğinden faydalanılır. z istatistiği her iki yönteme göre geçen öğrenci oranları arasındaki

farkın oranlara ait standart sapmaya bölünmesiyle elde edilir. Bu durumda standart sapma,

$$S_{p_1-p_2} = \sqrt{\frac{a+d}{N}},$$

z istatistiği,

$$z = \frac{p_1 - p_2}{\sqrt{\frac{a+d}{N}}} \text{ olarak ifade edilir (Ferguson, 1971).}$$

505 öğrenciye ait verilen doğru cevaplar 1, yanlış ve boş cevaplar 0 kodlanarak Excel programına aktarılmıştır. Excel yardımıyla her bir öğrencinin testten aldıkları puanlar belirlenmiştir. Ardından üç yöntem için alınmış olan uzman kararları Excel'e girilerek her bir uzmana ait MGP'ler belirlenmiştir. Daha anlamlı olabilmesi için belirlenen MGP'ler 4 ile çarpılarak elde edilen geçme puanlarının 100 üzerinden olması sağlanmıştır. Her bir uzman için belirlenen MGP'lerin ortalaması alınarak testin her bir yönetime göre geçme puanı hesaplanmıştır. Yöntemler için hesaplanan geçme puanlarına göre başarılı sayılan öğrenci sayıları ve yüzdeleri belirlenmiştir. Yöntemler ikili olarak sırasıyla Evet/Hayır-Ebel, Evet/Hayır-İşaretleme ve Ebel-İşaretleme olmak üzere karşılaştırılarak üç farklı karşılaştırma için z istatistikleri hesaplanmıştır. Bu şekilde yöntemlere göre geçen öğrenci yüzdeleri arasında istatistiksel olarak anlamlı bir fark olup olmadığı sınıanmıştır.

2. Araştırmanın ikinci alt probleminin "Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleri arasında öğrencilerin başarılı-başarısız olarak sınıflandırılması durumları bakımından uyum var mıdır?" çözümünde öğrenci puanları üç yöntemden elde edilen geçme puanlarına göre başarılı öğrenciler 1, başarısız öğrenciler 0 olacak şekilde kodlanarak yapay ikili hâle getirilmiştir. Yöntemlerin ikili uyumlarına kategorik verilerin uyumunu ölçen Cohen'in Kappa istatistiği ile bakılmıştır. Cohen tarafından geliştirilen Kappa istatistiği, sınıflama düzeyinde puanlama yapan iki gözlemcinin yaptığı değerlendirmeler arasındaki uyum derecesini belirlemek için kullanılır. Bu istatistik -1 ile +1 arasında değer alır. Sıfır değeri tesadüfi uyumu, negatif değerler tesadüfi olmaktan daha kötü bir uyumu, pozitif değerler tesadüfi uyumdan daha iyi bir uyumu ve +1 değeri ise mükemmel uyumu temsil eder (Von Eye ve Mun, 2005). Uyuşma oranlarını belirten Kappa istatistiği ile ilgili farklı aralıklar önerilmiştir (Von Eye ve Mun, 2005).

Kappa sayısı 0.40 ile 0.75 arasında deęer aldığında makul bir uyuşma ve 0.75'ten büyük bir deęer aldığında da mükemmel uyuşma anlamına gelir (Şencan, 2005).

3. Araştırmanın üçüncü probleminin “Uzmanların Ebel yönteminde verdikleri güçlük derecesi kararı ile maddelerin gerçek güçlük deęerleri arasındaki ilişki nasıldır?” çözümünde uzmanların belirledięi madde güçlük dereceleri ile gerçek veriden hesaplanan güçlük dereceleri arasındaki korelasyon katsayıları incelenerek tutarlık düzeyine bakılmıştır.

4. Araştırmanın dördüncü alt probleminin “Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemlerinin her biri için verilen uzman kararları arasındaki uyum nasıldır?” çözümünde uzmanlar arasındaki uyumu belirlemek amacıyla uzmanların yöntemlere göre maddelere vermiş oldukları kararlar arasındaki ilişkiye bakılmıştır. Her bir yöntem için maddelere verilmiş olan kararlar arasındaki uyuma Evet/Hayır ve İşaretleme yöntemleri için ikiden fazla puanlayıcının kullanılması durumunda ve kategorik verilerin uyumunu belirlemeye yönelik olan Fleiss'in Kappa istatistięinden, Ebel yöntemi için Kendall'ın uyuşma katsayısından (W) yararlanılmıştır. Kendall W istatistięi, ikiden fazla puanlayıcının kullanılması durumunda, puanlayıcılar arası güvenirlilięin belirlenmesinde kullanılan istatistiksel bir tekniktir (Özdamar, 2010). Kendall'ın W katsayısı 0 (hiç uyuşma olmaması) ile 1 (tam olarak uyuşma olması) arasında deęer alır. Landis ve Koch (1977) elde edilen κ deęerlerini yorumlamak için aşıęıdaki aralık deęerlerini sunmuşlardır. κ katsayısı 0.0-0.20 aralığında olduęunda önemsiz uyuşmayı 0.21-0.40 aralığında olduęunda orta derecede uyuşmayı, 0.41-0.60 aralığında olduęunda büyük ölçüde uyuşmayı, 0.61-0.80 aralığında olduęunda önemli derecede uyuşmayı, 0.81-1.00 aralığında olduęunda neredeyse mükemmel uyuşmayı belirtmektedir.

5. Araştırmanın beşinci alt probleminin “Uzman özelliklerine göre Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleriyle belirlenen geçme puanları farklılık göstermekte midir?” çözümünde uzmanlar deneyim, cinsiyet ve çalıştıkları pozisyon özelliklerine göre sınıflanmış ve bu sınıflamalara göre uzmanların belirledikleri geçme puanları kıyaslanmıştır. Ayrıca geçme puanları arasındaki fark bağımsız gruplar t testi ile incelenmiştir.

6. Araştırmanın altıncı alt probleminin “Evet/Hayır, Ebel ve İşaretleme standart belirleme süreçlerinde çalışılan uzmanların standart belirleme süreçlerine ilişkin

görüşleri nasıldır?” çözümünde standart belirleme süreci bittikten sonra uzmanlara uygulanan anket kullanılarak, sonuçlar üç yöntem açısından yorumlanmıştır. Anket sonucunda her bir maddenin analizinde uzmanların görüşleri frekans olarak belirtilmiştir.

4. BULGULAR VE TARTIŞMA

Bu bölümde, alt problem sırasına göre verilmiş araştırma bulguları ve bu bulgularla ilgili değerlendirmeler yer almaktadır.

4.1. Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemlerine göre belirlenen geçme puanlarının üzerinde puan alan öğrenci yüzdesi arasında fark var mıdır?

Bu alt problemin çözümü için öncelikle 30 uzmandan her bir yöntem için alınan görüşler doğrultusunda geçme puanları belirlenmiştir. Evet/Hayır yöntemi, Ebel yöntemi ve İşaretleme yöntemine göre uzmanların verdikleri kararlar doğrultusunda her bir uzmana ait MGP'ler ve bu MGP'lerin ortalamaları alınarak oluşturulmuş testin geçme puanı, Tablo 8'de verilmiştir.

Tablo 8: Evet/Hayır ve Ebel Yöntemleri İçin Elde Edilen MGP'ler

<i>Uzman</i>	<i>Evet/Hayır Yöntemi</i>	<i>Ebel Yöntemi</i>	<i>Uzman</i>	<i>Evet/Hayır Yöntemi</i>	<i>Ebel Yöntemi</i>
1	44	62.76	16	44	43.8
2	44	47	17	52	49.88
3	60	42.2	18	64	52.8
4	68	42.4	19	48	37.2
5	76	58.4	20	32	32.6
6	64	65	21	44	43.6
7	64	45.2	22	52	51.96
8	56	52.6	23	72	47.6
9	52	31.84	24	68	54
10	64	45	25	48	40.6
11	44	31.2	26	44	37.4
12	72	29.76	27	64	50.6
13	56	55.6	28	72	75
14	92	57.4	29	44	41.2
15	44	42.6	30	84	49.6
Geçme Puanı				57.73	47.23

Uzmanların verdikleri kararlar doğrultusunda İşaretleme yöntemi için belirlenen ortalama Θ değeri Tablo 9'da verilmiştir.

Tablo 9: İşaretleme Yöntemi ile Belirlenen Sayfa ve Madde Numarası ve Θ Değerleri

Uzman	Sayfa numarası	Madde numarası	Θ	Uzman	Sayfa numarası	Madde numarası	Θ
1	9	12	0.2965	16	5	4	-0.1787
2	10	7	0.3130	17	2	20	-0.4541
3	1	2	-1.0608	18	18	6	1.4059
4	7	5	-0.0135	19	10	7	0.3130
5	13	21	0.6586	20	3	1	-0.3560
6	5	4	-0.1787	21	8	8	0.1166
7	9	12	0.2965	22	13	21	0.6586
8	9	12	0.2965	23	7	5	-0.0135
9	12	11	0.6348	24	8	8	0.1166
10	7	5	-0.0135	25	6	9	-0.1260
11	5	4	-0.1787	26	2	20	-0.4541
12	7	5	-0.0135	27	14	25	0.8333
13	15	17	0.8422	28	11	14	0.5047
14	7	5	-0.0135	29	8	8	0.1166
15	13	21	0.6586	30	6	9	-0.1260
Ortalama			Θ	0.1627			

Tablo 9'dan görüldüğü gibi uzmanların belirlemiş oldukları maddelere ve bu maddelerin Θ değerlerine göre ortalama Θ , 0.1627 olarak hesaplanmıştır. Bu adımdan sonra bulunan ortalama Θ değeri her bir madde için madde parametreleri ile beraber

$$P(X = 1|\theta) = \frac{1}{1 + e^{[-Da_i(\theta - b_i)]}}$$

formülünde yerine konularak ham puanlar elde edilmiştir. Bu ham puanların toplanmasıyla testin geçme puanı 55.59 olarak bulunmuştur.

Tablo 8 ve Tablo 9'da görüldüğü gibi yöntemlere göre belirlenen geçme puanları, uzmanların her bir yöntem için farklı süreç adımlarıyla karar vermesi nedeniyle farklılık göstermektedir. Evet/Hayır yöntemi ile belirlenen geçme puanı 57.73, Ebel yöntemi ile belirlenen geçme puanı 47.23 ve İşaretleme yöntemi ile belirlenen geçme puanı 55.59 bulunmuştur. Bu üç yöntemde en yüksek geçme puanı Evet/Hayır yöntemi ile elde edilirken en düşük geçme puanı Ebel yöntemi ile elde edilmiştir. Her üç yöntem ile belirlenen geçme puanlarına göre başarılı kabul edilen öğrenci sayısı ve yüzdesi Tablo 10'da verilmiştir.

Tablo 10: Yöntemlerle Belirlenen Geçme Puanları, Başarılı Kabul Edilen Öğrenci Sayıları ve Yüzdeleri

<i>Yöntem</i>	<i>Geçme Puanı</i>	<i>Başarılı Kabul Edilen Öğrenci Sayısı</i>	<i>Başarılı Kabul Edilen Öğrenci Yüzdesi</i>
<i>Evet/Hayır Yöntemi</i>	57.73	166	% 32
<i>Ebel Yöntemi</i>	47.23	376	% 74
<i>İşaretleme Yöntemi</i>	55.59	190	% 38

Tablo 10'dan görüldüğü gibi yöntemlerle belirlenen geçme puanlarına göre başarılı kabul edilen öğrenci sayıları ve yüzdeleri farklılık göstermektedir. Başarılı kabul edilen öğrenci sayısının en fazla olduğu yöntem en düşük geçme puanına sahip olan Ebel yöntemidir. Ebel yöntemini, İşaretleme yöntemi takip etmektedir. Başarılı kabul edilen öğrenci sayısının en az olduğu yöntem ise en yüksek geçme puanına sahip olan Evet/Hayır yöntemidir. Araştırmada bulunan bu sonuç Gündeğer (2012) tarafından yapılan çalışma ile örtüşmektedir. Ebel yöntemi ile elde edilen kesme puanın en düşük çıkma nedeni, yöntemin işleyişinin diğer yöntemlere göre uzmanlara daha karmaşık gelmesi olabilir. Impara ve Plake (1997) çalışmalarında Evet/Hayır yönteminde uzmanların daha kolay karar verdiği sonucuna varmışlardır. Bu çalışmada bulunan Evet/Hayır yöntemi ile en yüksek geçme puanı elde edilmesi sonucu Impara ve Plake (1997) tarafından yapılan çalışmayı destekler niteliktedir.

Yöntemlerle belirlenen geçme puanlarına göre başarılı kabul edilen öğrenci yüzdeleri arasındaki bu farklılıkların anlamlılığı bağımlı iki oran arasındaki farkın testi ile sınanmıştır.

Tablo 11: Geçme Puanları, Başarılı Öğrenci Sayısı ve Yüzdesi, z değerleri

<i>Karşılaştırılan Yöntemler</i>	<i>Geçme Puanı</i>	<i>Başarılı Öğrenci Sayısı</i>	<i>Başarılı Öğrenci Yüzdesi</i>	<i>z</i>
<i>Evet/Hayır Yöntemi</i>	57.73	166	% 32	
<i>Ebel Yöntemi</i>	47.23	376	% 74	-14.6*
<i>Evet/Hayır Yöntemi</i>	57.73	166	% 32	
<i>İşaretleme Yöntemi</i>	55.59	190	% 38	-5.5*
<i>Ebel Yöntemi</i>	47.23	376	% 74	13.5*
<i>İşaretleme Yöntemi</i>	55.59	190	% 38	

* $p < 0.01$ düzeyinde anlamlı

Evet/Hayır, Ebel ve İşaretleme yöntemleri ikili olarak karşılaştırıldığında belirlenen geçme puanlarına göre başarılı kabul edilen öğrenci yüzdeleri arasında istatistiksel olarak anlamlı farklılıklar olduğu Tablo 11'den görülmektedir. Bu farklılığın

nedeninin her bir yöntemde farklı süreçlerle geçme puanı belirlenmesi olabileceği düşünülmektedir.

4.2. Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleri arasında öğrencilerin başarılı-başarısız olarak sınıflandırılması durumları bakımından uyum var mıdır?

Bu alt problemin çözümü için öncelikle sürekli olan öğrenci puanları, her bir yöntemle belirlenen geçme puanına göre başarılı ve başarısız olarak sınıflandırılmıştır. Başarılı öğrenciler 1, başarısız öğrenciler 0 olmak üzere öğrenci puanları yapay ikili hâle getirilmiştir. Yöntemler arasındaki ikili uyumlara kategorik verilerdeki uyumu ölçen Cohen'in Kappa istatistiği ile bakılmıştır. Ayrıca veriler arasındaki korelasyonun hesaplanması için Pearson korelasyon katsayısı da hesaplanmıştır.

Tablo 12: Yöntemler arasındaki uyum için hesaplanan katsayılar

<i>Yöntemler</i>	<i>Kappa İstatistiği</i>	<i>Pearson Katsayısı</i>
<i>Evet/Hayır Yöntemi</i> <i>Ebel Yöntemi</i>	0.585*	0.643*
<i>Evet/Hayır Yöntemi</i> <i>İşaretleme Yöntemi</i>	0.905*	0.909*
<i>Ebel Yöntemi</i> <i>İşaretleme Yöntemi</i>	0.667*	0.707*

* $p < 0.01$

Tablo 12 incelendiğinde öğrencilerin başarılı-başarısız olarak sınıflanması durumu ikişer ikişer karşılaştırıldığında karşılaştırmaların tamamında 0.01 anlamlılık düzeyinde pozitif yönde bir ilişki olduğu görülmektedir. Öğrencilerin başarılı-başarısız olarak sınıflandırılmaları bakımından Evet/Hayır ve Ebel yöntemleri ve Ebel ve İşaretleme yöntemleri arasında makul bir uyum olduğu bulunurken, Evet/Hayır ve İşaretleme yöntemleri arasında mükemmel bir uyum olduğu sonucuna ulaşılmıştır. Evet/Hayır ve İşaretleme yöntemlerinde benzer olarak öğrenci seviyeleri düşünülerek geçme puanı belirlenmektedir. Bu yüzden Evet/Hayır ve İşaretleme yöntemleri arasındaki uyumun diğer karşılaştırmalardan daha yüksek olduğu düşünülmektedir. Ebel yöntemiyle geçme puanı belirlenirken diğer iki yöntemden farklı bir süreç izlendiğinden Ebel yönteminin diğer yöntemlerle uyumu makul düzeyde kalmaktadır. Demir (2014) çalışmasında yöntemler arasında yüksek düzeyde uyum bulunurken bu çalışmada Evet/Hayır ve İşaretleme yöntemlerinin arasındaki uyum dışında diğer uyum yüzdeleri makul düzeyde bulunmuştur. Çalışmada bulunan Evet/Hayır ve Ebel yöntemleri

arasındaki makul uyum Gündeğer (2012) tarafından yapılan çalışmayla desteklenmektedir.

4.3. Uzmanların Ebel yönteminde verdikleri güçlük derecesi kararı ile maddelerin gerçek güçlük değerleri arasındaki ilişki nasıldır?

Ebel yöntemine göre uzmanların belirledikleri madde güçlük değerlerinin ortalaması ile gerçek madde güçlük değerlerinin betimsel istatistikleri Tablo 13'te verilmiştir.

Tablo 13: Uzmanların Maddelere Verdikleri Güçlük Değerlerinin Ortalamaları ve Gerçek Madde Güçlükleri İle İlgili Betimleyici İstatistikler

	<i>N</i>	<i>Minimum</i>	<i>Maksimum</i>	<i>Ortalama</i>	<i>Standart sapma</i>	<i>r</i>
<i>Uzmanların Güçlük Dereceleri</i>	25	0.32	0.73	0.51	0.12	0.83
<i>Madde Güçlükleri</i>	25	0.25	0.84	0.51	0.15	

Tablo 13 incelendiğinde uzmanların testteki maddelere verdikleri güçlük değerlerinin ortalamasının en düşük 0.32 ve en yüksek 0.73 olduğu görülmüştür. Test puanlarından hesaplanan gerçek madde güçlükleri en düşük 0.25 ve en yüksek 0.84 olarak bulunmuştur. Uzmanların belirledikleri değerlerin gerçek değerlere yakın olduğu ifade edilebilirken uzmanların zor maddelerde maddelerin güçlük derecelerini gerçekten daha yüksek olarak belirledikleri bulunmuştur. Uzmanlara göre testin ortalama güçlüğü 0.51, standart sapması 0.12 bulunurken uygulanan testin sonucuna göre testin güçlüğü de 0.51, standart sapması 0.15 bulunmuştur.

Uzmanların belirledikleri madde güçlüklerinin ortalamaları ile gerçek madde güçlükleri arasındaki korelasyon katsayısı 0.83 olarak hesaplanmıştır. Bu katsayı, uzmanların belirledikleri madde güçlüklerinin ortalamaları ile gerçek madde güçlükleri arasındaki ilişkinin pozitif, yüksek düzeyde ve 0.01 düzeyinde anlamlı olduğu anlamına gelmektedir. Brandon (2004) uzman tahmin ortalamalarının gerçek madde güçlükleri ile yüksek düzeyde ilişkili olması gerektiğini savunmaktadır. Impara ve Plake (1998) çalışmalarında uzmanların belirledikleri ortalama madde güçlükleri ile gerçek madde güçlükleri arasındaki ilişkiyi 0.78 olarak belirlerken Çetin (2011) yaptığı çalışmasında uzmanların belirledikleri ortalama madde güçlükleri ile gerçek madde güçlükleri arasındaki ilişkiyi 0.60

olarak belirlemiştir. Bahsedilen bu üç çalışma bu çalışmanın sonucunu destekler niteliktedir.

4.4. Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemlerinin her biri için verilen uzman kararları arasındaki uyum nasıldır?

Bu alt problemin çözümünde uzman kararları arasındaki uyumu belirlemek için Evet/Hayır ve İşaretleme yöntemleri için Fleiss'in Kappa istatistiği, Ebel yöntemi için Kendal'ın W istatistiği hesaplanmıştır. Ayrıca uzman kararları arasındaki uyumu belirlemek amacıyla yöntemlerin her biri için yöntemlerle belirlenen MGP'lerin standart sapmaları hesaplanmıştır.

Tablo 14: Yöntemlere Göre Uzmanlar Arasındaki Uyum

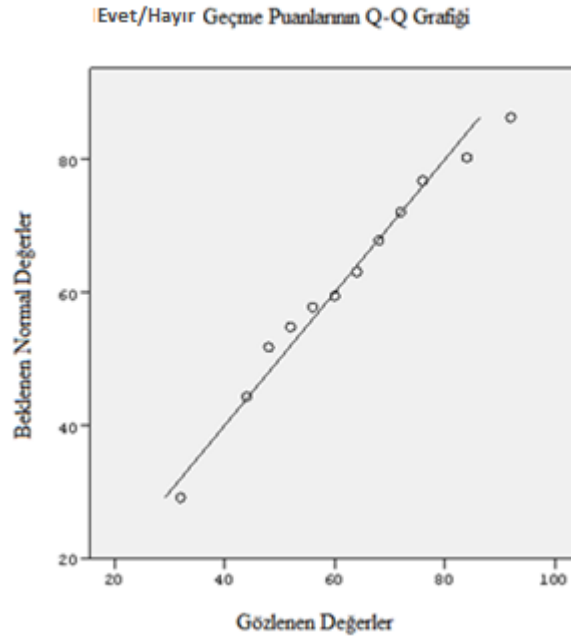
<i>Yöntem</i>	<i>Standart Sapma</i>	<i>Katsayı</i>	<i>Katsayının Anlamı</i>
<i>Evet/Hayır Yöntemi</i>	14.00	0.26	Orta derecede uyum
<i>Ebel Yöntemi</i>	10.46	0.26	Orta derecede uyum
<i>İşaretleme Yöntemi</i>	5.15	0.49	Makul düzeyde uyum

Evet/Hayır yönteminde uzmanların uyumu için belirlenen Fleiss'in Kappa istatistiği 0.26 bulurken İşaretleme yönteminde uzmanların uyumu için belirlenen Fleiss'in Kappa istatistiği 0.49 ve Ebel yönteminde uzmanların uyumu için belirlenen Kendal'ın W istatistiği 0.26 bulunmuştur. Bu durumda uzmanlar arasında Evet/Hayır ve Ebel yöntemlerinde orta derecede bir uyum olduğu görülürken İşaretleme yönteminde makul düzeyde bir uyum olduğu sonucuna varılabilir. Gündeğer (2012) çalışmasında Evet/Hayır yönteminde, Ebel yöntemine göre uzmanlar arasında daha yüksek uyum olduğu bulunurken bu çalışmada da benzer bir uyum yüzdesi bulunmuştur.

4.5. Uzman özelliklerine göre Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleriyle belirlenen geçme puanları farklılık göstermekte midir?

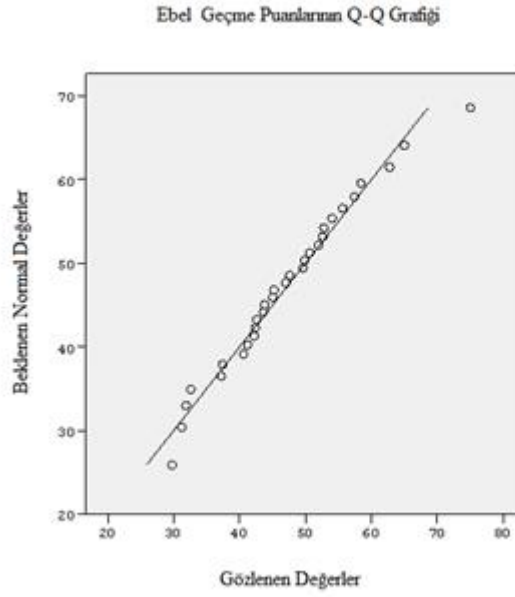
Bu alt problemin çözümünde öncelikle uzmanlar deneyim durumuna, cinsiyetine ve çalıştıkları pozisyona göre sınıflandırılmıştır. Uzmanların deneyim durumu 0-3 yıl arası çalışanlar için deneyimsiz, 4 yıl ve üzeri çalışanlar için deneyimli olarak kabul edilmiştir. Çalışmaya katılan uzman grubu çalıştıkları pozisyonları açısından uzman ve öğretmen olarak sınıflanmışlardır. Araştırmaya katılan uzman grubunun 13'ü deneyimsiz, 17'si deneyimli; 10'u erkek, 20'si kadın ve 20'si öğretmen, 10'u uzmandır (eğitsel tasarımcıdır).

30 veriden az olan gruplarda parametrik testlerin kullanılabilmesi için puanların normal dağılıp dağılmadığına bakılmalıdır. Çalışmada, puanların normalliği yüzdelik-yüzdelik (Q-Q) grafikleri ile bakılmıştır. Aşağıda verilen şekillerde görüldüğü gibi geçme puan dağılımlarının normal olduğu bulunduğundan analize parametrik yöntemlerle devam edilmiştir. Yöntemlerle belirlenen geçme puanlarının uzman özelliklerine göre karşılaştırılması bağımsız t testleri ile yapılmıştır.



Şekil 3: Evet/Hayır Geçme Puanlarının Q-Q Grafiği

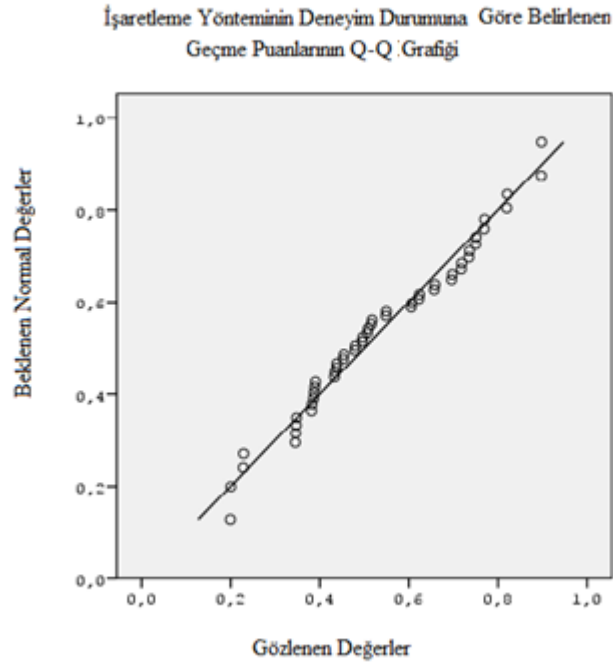
Uzmanların Evet/Hayır yöntemine göre belirledikleri geçme puanlarının Q-Q grafiği incelendiğinde puanların köşegen çizgisi üzerinden pek fazla sapmadığı görülür. O hâlde Evet/Hayır yöntemiyle belirlenen geçme puanlarının normal dağıldığı ifade edilebilir.



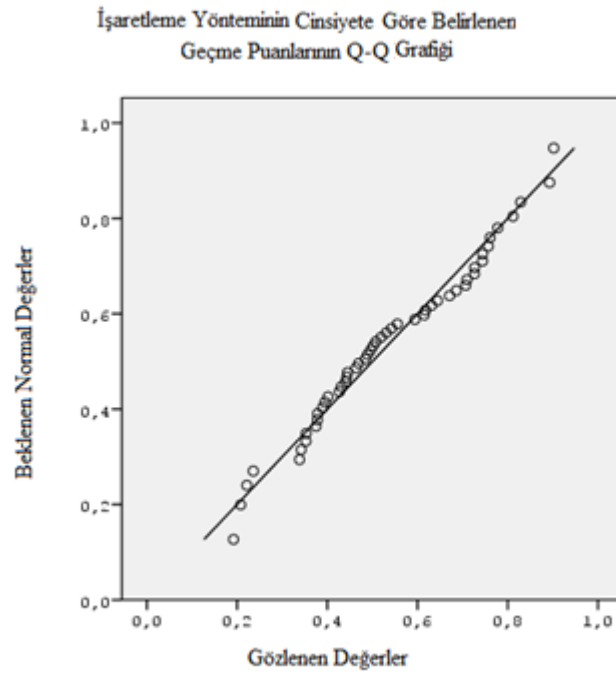
Şekil 4: Ebel Geme Puanlarının Q-Q Grafiđi

Uzmanların Ebel yöntemine göre belirledikleri geme puanlarının Q-Q grafiđi incelendiđinde puanların köşegen çizgisi üzerinden pek fazla sapmadığı görülür. Buna göre Ebel yöntemiyle belirlenen geme puanlarının normal dağıldığı ifade edilebilir.

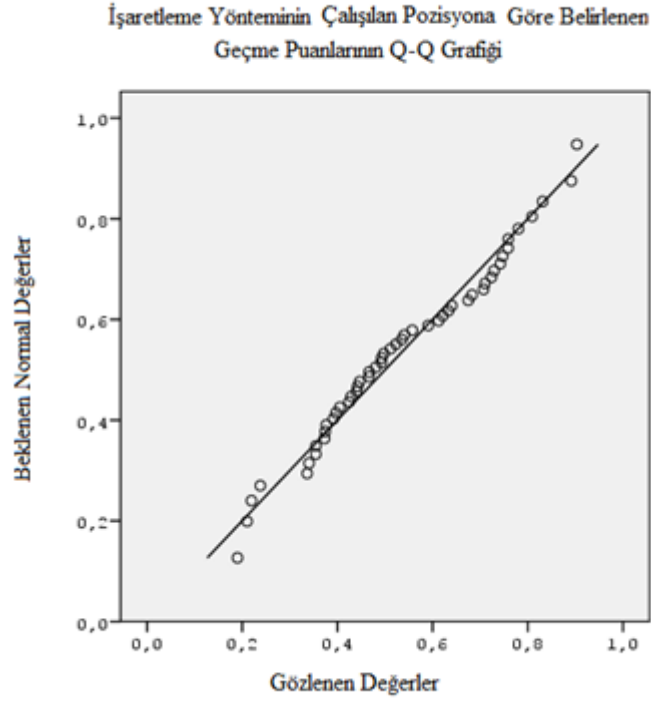
İşaretleme yöntemine göre hesaplanan geme puanları uzman grubu deđiştikçe madde parametreleri yeniden hesaplandıđından İşaretleme yönteminin geme puanları dağılımının normallik incelenmesi deneyim, cinsiyet ve çalışılan pozisyona göre ayrı ayrı yapılmıştır.



Şekil 5: Deneyim Durumuna Göre Belirlenen İşaretleme Geçme Puanlarının Q-Q Grafiği



Şekil 6: Cinsiyete Göre Belirlenen İşaretleme Geçme Puanlarının Q-Q Grafiği



Şekil 7: Çalışılan Pozisyona Göre Belirlenen İşaretleme Geçme Puanlarının Q-Q Grafiği

4.5.1. Uzmanların deneyim duruma göre Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleriyle belirlenen geçme puanları farklılık göstermekte midir?

Tablo 15: Uzmanların Deneyim Durumuna Göre Belirlenen Geçme Puanlarının Ortalamaları Arasındaki Farkın Testi

<i>Standart Belirleme Yöntemi</i>	<i>Deneyim Durumu</i>	<i>N</i>	<i>Ortalama</i>	<i>Standart Sapma</i>	<i>t</i>	<i>p</i>
<i>Evet/Hayır Yöntemi</i>	Deneyimli	17	56	11.31	0.77	0.45
	Deneyimsiz	13	60	17.12		
<i>Ebel Yöntemi</i>	Deneyimli	17	47.56	11.25	-0.2	0.84
	Deneyimsiz	13	46.78	9.75		
<i>İşaretleme Yöntemi</i>	Deneyimli	17	53.77	18.4	0.022	0.98
	Deneyimsiz	13	53.65	18.5		

Tablo 15'ten görüldüğü gibi uzmanların deneyimlerine göre belirledikleri geçme puanları sırasıyla deneyimli ve deneyimsiz uzmanlar için Evet/Hayır yönteminde 56 ve 60, Ebel yönteminde 47.56 ve 46.78, İşaretleme yönteminde 53.77 ve 53.65'tir. Yapılan t testi sonucunda uzmanların deneyim durumuna göre belirledikleri geçme puanları arasında istatistiksel olarak anlamlı bir farklılık bulunamamıştır. Zahran (1981) yaptığı çalışmada deneyimli uzmanların

deneyimsiz uzmanlara göre bariz şekilde düşük geçme puanı belirlediği sonucuna ulaşmıştır.

4.5.2. Uzmanların cinsiyetlerine göre Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleriyle belirlenen geçme puanları farklılık göstermekte midir?

Tablo 16: Uzmanların Cinsiyetlerine Göre Belirlenen Geçme Puanlarının Ortalamaları Arasındaki Farkın Testi

<i>Standart Belirleme Yöntemi</i>	<i>Cinsiyet</i>	<i>N</i>	<i>Ortalama</i>	<i>Standart Sapma</i>	<i>t</i>	<i>p</i>
<i>Evet/Hayır Yöntemi</i>	Kadın	20	56.8	14.31	-0.51	0.61
	Erkek	10	59.6	13.91		
<i>Ebel Yöntemi</i>	Kadın	20	46.31	10.15	-0.67	0.51
	Erkek	10	49.05	11.38		
<i>İşaretleme Yöntemi</i>	Kadın	20	54.82	18.4	0.44	0.66
	Erkek	10	52.56	18.5		

Tablo 16'dan görüldüğü gibi uzmanların cinsiyetlerine göre belirledikleri geçme puanları sırasıyla kadın ve erkek uzmanlar için Evet/Hayır yönteminde 56.8 ve 59.6, Ebel yönteminde 46.31 ve 49.05, İşaretleme yönteminde 54.82 ve 52.56'dır. Yapılan t testi sonucunda uzmanların cinsiyetlerine göre belirledikleri geçme puanları arasında istatistiksel olarak anlamlı bir farklılık bulunamamıştır.

4.5.3. Uzmanların çalıştıkları pozisyonlarına göre Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleriyle belirlenen geçme puanları farklılık göstermekte midir?

Tablo 17: Uzmanların Cinsiyetlerine Göre Belirlenen Geçme Puanlarının Ortalamaları Arasındaki Farkın Testi

<i>Standart Belirleme Yöntemi</i>	<i>Çalışılan Pozisyon</i>	<i>N</i>	<i>Ortalama</i>	<i>Standart Sapma</i>	<i>t</i>	<i>p</i>
<i>Evet/Hayır Yöntemi</i>	Öğretmen	20	55.8	14.06	-1.07	0.29
	Uzman	10	61.6	13.75		
<i>Ebel Yöntemi</i>	Öğretmen	20	46.1	10.54	-0.76	0.46
	Uzman	10	49.28	10.51		
<i>İşaretleme Yöntemi</i>	Öğretmen	20	54.86	18.4	-0.34	0.73
	Uzman	10	52.56	18.5		

Tablo 17'den görüldüğü gibi uzmanların çalıştıkları pozisyonlara göre belirledikleri geçme puanları sırasıyla öğretmen ve uzman için Evet/Hayır yönteminde 55.8 ve 61.6, Ebel yönteminde 46.1 ve 49.28, İşaretleme yönteminde 54.86 ve 52.56'dır. Yapılan t testi sonucunda uzmanların çalıştıkları pozisyona göre belirledikleri geçme puanları arasında istatistiksel olarak anlamlı bir farklılık bulunamamıştır. Kara ve Kelecioğlu (2015) yaptığı çalışmalarında öğretmen ve uzman olan

puanlayıcıların geçme puanlarına etkisini araştırılmış ve belirlenen geçme puanları arasında önemli bir farklılık bulunamamıştır. Benzer bir durum Norcini, Shea ve Kanya'nın (1988) yaptıkları çalışmalarında da söz konusudur. Bu çalışma, bahsedilen çalışmaları destekler niteliktedir.

4.6. Evet/Hayır, Ebel ve İşaretleme standart belirleme süreçlerinde çalışılan uzmanların standart belirleme süreçlerine ilişkin görüşleri nasıldır?

Uzmanlara standart belirleme sürecini ve kullanılan standart belirleme yöntemlerini değerlendirmeleri için Çetin'in (2011) doktora tez çalışmasında geliştirdiği anketin 10 sorusundan oluşan kısmı uygulanmıştır. Uzmanların standart belirleme sürecindeki unsurlarla ilgili görüşleri aşağıdaki tabloda özetlenmiştir.

Tablo 18: Uzmanların Standart Belirleme Sürecindeki Unsurlarla İlgili Görüşleri

<i>Anket maddeleri</i>	<i>Yeterli Değil</i>	<i>Kısmen Yeterli</i>	<i>Yeterli</i>	<i>Oldukça Yeterli</i>
İlgili test konusu alanı hakkında bilgilendirme	0 (% 0)	2 (% 7)	18 (% 60)	10 (% 33)
Standart belirleme süreci ile ilgili bilgilendirme	1 (% 3)	3 (% 10)	17 (% 56)	9 (% 30)
Geçme puanı tanımı ile ilgili bilgilendirme	0 (% 0)	1 (% 3)	16 (% 53)	13 (% 43)
Kitapçık ve puanlamalarla ilgili bilgilendirmeler	0 (% 0)	5 (% 17)	17 (% 56)	8 (% 27)
Yapılacak işlemlerle ilgili örnek bir uygulama	0 (% 0)	2 (% 6)	17 (% 56)	11 (% 37)

Uzmanlara, testin ilgili olduğu konu alanı hakkındaki bilgilendirmenin yeterliliği sorulmuş ve uzmanların ikisi bilgilendirmeyi kısmen yeterli bulurken, on sekizi yeterli, onu oldukça yeterli bulmuştur. Uzmanlara standart belirleme süreci ile ilgili bilgilendirmenin yeterliliği sorulmuş ve uzmanların biri bilgilendirmeyi yeterli bulmazken, üçü kısmen yeterli on yedisi yeterli, dokuzu oldukça yeterli bulmuştur.

Uzmanlardan biri geçme puanı tanımı ile ilgili bilgilendirmenin kısmen yeterli olduğunu düşünürken on altısı yeterli olduğunu ve on üçü bilgilendirmenin oldukça yeterli olduğunu düşünmektedir. Uzmanlardan beşi kitapçık ve puanlamalarla ilgili bilgilendirmenin kısmen yeterli olduğunu düşünürken on yedisi yeterli olduğunu ve sekizi bilgilendirmenin oldukça yeterli olduğunu düşünmektedir. Uzmanlardan ikisi yapılacak işlemlerle ilgili örnek bir uygulamanın kısmen yeterli olduğunu düşünürken on yedisi yeterli olduğunu ve on biri bilgilendirmenin oldukça yeterli olduğunu düşünmektedir. Verilen tablodan da görüldüğü gibi uzmanlar genel

olarak standart belirleme sürecindeki unsurlarla ilgili bilgilendirmeyi yeterli bulmuşlardır.

Uzmanlar anketin 5. sorusu olan “Çalışma başında kitapçıkların sınıflandırılması ve puanlamaya ilişkin yapılan bilgilendirmenin süresini yeterli buldunuz mu?” sorusuna istisnasız “bilgilendirme süresi tam gerektiği kadardı” cevabını vermişlerdir. Uzmanların öğrenci performansını sınıflarken kullandıkları unsurların önem düzeyi ile ilgili görüşleri aşağıda verilmiştir.

Tablo 19: Uzmanların Öğrenci Performansı Sınıflarken Kullanılan Unsurların Önem Düzeyi İle İlgili Görüşleri

<i>Anket maddeleri</i>	<i>Önemli Değil</i>	<i>Biraz Önemli</i>	<i>Önemli</i>	<i>Çok Önemli</i>
Geçme puanı tanımı	1 (% 3)	4 (% 13)	20 (% 67)	5 (% 17)
Soruların güçlük düzeyleriyle ilgili düşünceleriniz	0 (% 0)	1 (% 3)	22 (% 73)	7 (% 23)
Öğrencilerin düzeyiyle ilgili kendi düşünceleriniz	0 (% 0)	0 (% 0)	20 (% 67)	10 (% 33)
Deneyimleriniz	0 (% 0)	2 (% 6)	11 (% 37)	17 (% 56)

Uzmanlardan dördü öğrenci performansını sınıflarken geçme puanı tanımını biraz önemli, yirmisi önemli, 5'i çok önemli bulurken, uzmanlardan yirmi ikisi soruların güçlük düzeyleriyle ilgili düşüncelerini önemli, ikisi çok önemli bulmuştur. Öğrencilerin düzeyiyle ilgili kendi düşüncelerini önemli bulan uzman sayısı yirmi iken çok önemli bulan uzman sayısı ondur. Öğrenci performansını sınıflarken kendi deneyimlerini, önemli bulan on bir uzman, çok önemli bulan on yedi uzman vardır.

Uzmanlara karar verirken en çok zorlandıkları yöntem sorulduğunda uzmanların üçünün Evet/Hayır, yirmi birinin Ebel ve altısının İşaretleme yönteminde zorlandığı görülmüştür. Uzmanlara en rahat karar verdikleri yöntem sorulduğunda uzmanlardan yirmisi Evet/Hayır, beşi Ebel ve beşi İşaretleme yöntemini en rahat karar verdikleri yöntem olarak belirtmişlerdir. Uzmanlar sırasıyla Evet/Hayır, İşaretleme ve Ebel yöntemlerini en kolay uygulanabilen, en uygun ve anlaşılır buldukları yöntem olarak belirtmişlerdir.

Uzmanların kullanılan yöntemlerle ilgili fikirleri genel olarak incelendiğinde; uzmanların çoğunun Evet/Hayır yöntemini en kolay uygulanabilen, en uygun ve anlaşılır yöntem olarak bulduğu ve uzmanların karar vermede en çok zorlandıkları yöntemin Ebel yöntemi olduğu görülmektedir. Ebel yöntemiyle geçme puanı belirlerken uzmanlar soruların güçlük ve önem derecelerine de karar verdiğinden

bu yöntemle geçme puanı belirlemek diğer yöntemlere göre daha karmaşıktır. Uzmanların en çok Ebel yönteminde zorlanmalarının nedeni bu karmaşık süreç olabilir. Çalışmanın sonucunda uzmanların Evet/Hayır yöntemiyle daha kolay karar verdiği ve Evet/Hayır yöntemini daha uygulanabilir bulmaları sonucu Impara ve Plake'in (1997) yapmış oldukları çalışmaları ile örtüşmektedir.

5. SONUÇ ve ÖNERİLER

Bu bölümde araştırmanın bulgu ve yorumlarına dayalı olarak ulaşılan sonuçların özetine ve bu sonuçlardan yola çıkarak geliştirilen önerilere yer verilmiştir.

5.1. Sonuçlar

Araştırmada elde edilen sonuçlar alt problemlerin sırasına uygun olarak aşağıda özetlenmiştir.

1. Evet/Hayır, Ebel ve İşaretleme yöntemiyle belirlenen geçme puanlarının birbirinden farklı olduğu bulunmuştur. En yüksek geçme puanı Evet/Hayır yöntemi ile elde edilirken en düşük geçme puanı Ebel yöntemiyle elde edilmiştir. Geçme puanları arasındaki bu farklılık başarılı sayılan öğrenci sayılarını ve yüzdelerini etkilemiştir. Evet/Hayır yöntemiyle belirlenen geçme puanına göre öğrencilerin % 32'si, Ebel yöntemiyle belirlenen geçme puanına göre öğrencilerin % 74'ü, İşaretleme yöntemiyle belirlenen geçme puanına göre öğrencilerin % 38'i başarılı sayılmıştır. Evet/Hayır, Ebel ve İşaretleme yöntemleriyle belirlenen geçme puanlarının üzerinde puan alan öğrenci yüzdeleri arasındaki bu farklılığın istatistiksel olarak anlamlı olduğu sonucuna ulaşılmıştır.

2. Yöntemlerle belirlenen geçme puanlarına göre her bir yöntem için başarılı öğrenciler 1, başarısız öğrenciler 0 ile kodlanarak Evet/Hayır, Ebel ve İşaretleme yöntemleri arasındaki uyum ikişer ikişer karşılaştırılmış ve karşılaştırmaların tamamında pozitif bir ilişki olduğu sonucuna ulaşılmıştır. Öğrencilerin başarılı-başarısız olarak sınıflandırılmaları bakımından Evet/Hayır ve Ebel yöntemleri arasında makul, Evet/Hayır ve İşaretleme yöntemleri arasında mükemmel, İşaretleme ve Ebel yöntemleri arasında makul bir uyum olduğu bulunmuştur.

3. Ebel yönteminde uzmanların maddelere verdikleri güçlük değerleri ile gerçek veriden hesaplanan güçlük değerleri arasındaki ilişki düzeyini belirlemek için uzmanların madde güçlük değerleri ile gerçek madde güçlükleri arasındaki korelasyon katsayısına bakılmıştır. Uzmanların belirledikleri madde güçlüklerinin ortalamaları ile gerçek madde güçlükleri arasındaki korelasyon katsayısı 0.83 olarak hesaplanmıştır. Bu katsayıyla uzmanların belirledikleri madde güçlüklerinin ortalamaları ile gerçek madde güçlükleri arasında istatistiksel olarak anlamlı, pozitif ve yüksek düzeyde bir ilişki olduğu sonucuna ulaşılmıştır.

4. Evet/Hayır ve Ebel yönteminde belirlenen geçme puanları açısından uzmanlar arasında orta derecede uyum olduğu sonucu bulunurken İşaretleme yönteminde uzmanlar arasındaki uyumun makul düzeyde olduğu sonucuna ulaşılmıştır.

5. Uzman özelliklerine göre uzmanlar deneyim durumuna, cinsiyetine ve çalıştıkları pozisyona göre sınıflanmıştır. Evet/Hayır, Ebel ve İşaretleme yöntemlerinin her birine göre belirlenen geçme puanları arasındaki farklılık araştırılmış ve uzman özelliklerine göre belirlenen geçme puanları arasında istatistiksel olarak anlamlı farklılıklar bulunamamıştır.

6. Uzmanlara standart belirleme sürecini ve kullanılan standart belirleme yöntemlerini değerlendirmeleri için uygulanan anketin sonuçları aşağıda yorumlanmıştır.

Uzmanlara, testin ilgili olduğu konu alanı hakkındaki bilgilendirmenin yeterliliği sorulmuş ve bilgilendirmenin yeterli olduğu sonucuna ulaşılmıştır. Uzmanlara, standart belirleme süreci, geçme tanımı, kitapçık ve puanlamalarla ilgili bilgilendirmenin yeterliliği sorulmuş ve bilgilendirmelerin hepsinin yeterli olduğu sonucuna ulaşılmıştır. Anket sonucunda uzmanlara yapılacak işlemlerle ilgili anlatılan örnek uygulamanın da yeterli olduğu sonucuna ulaşılmıştır. Genel olarak uzmanların hemen hemen hepsinin, standart belirleme sürecinin tamamıyla ilgili yapılan bilgilendirmeleri yeterli bulduğu sonucuna ulaşılmıştır. Uzmanların hepsi kitapçıkların sınıflandırılması ve puanlamaya ilişkin yapılan bilgilendirmenin süresini yeterli bulmuşlardır.

Uzmanların büyük bir kısmının öğrenci düzeyini sınıflarken, geçme puanı tanımını, soruların güçlük dereceleriyle ilgili düşüncelerini ve kendi deneyimlerini büyük oranda çok önemli bulduğu sonucuna ulaşılmıştır. Uzmanların çoğunun karar verirken en çok Ebel yönteminde zorlandığı sonucuna ulaşılmıştır. Uzmanların çoğunun en rahat karar verdiği yöntem Evet/Hayır yöntemi olarak bulunmuştur. Uzmanlar, en kolay uygulanabilen, en uygun ve en anlaşılır buldukları yöntem olarak sırasıyla Evet/Hayır, İşaretleme ve Ebel yöntemini belirlemişlerdir.

5.2. Öneriler

5.2.1. Araştırma Sonuçlarından Çıkan Öneriler

1. Araştırmada en düşük geçme puanı Ebel yöntemiyle belirlenirken onu takiben İşaretleme yöntemi ve en yüksek geçme puanı Evet/Hayır yöntemi kullanılarak

belirlenmiştir. Bu durumda öğrencilerin ders başarıları Ebel yöntemiyle daha yüksek olurken, İşaretleme ve Evet/Hayır yöntemiyle daha düşük olmaktadır. Ders öğretmenlerinin bu durumu kullanarak uygulanan sınavın çeşidine ve amacına dönük olarak standart belirleme yöntemini seçmesi önerilebilir. Ders öğretmeni yüksek başarı istiyorsa Ebel yöntemini, kullanım amacına yönelik daha düşük başarı istiyorsa İşaretleme ve Evet/Hayır yöntemlerini kullanabilir.

2. Belirlenen geçme puanlarının yakın olması, öğrencilerin başarılı-başarısız olarak sınıflandırılmaları bakımından Evet/Hayır yöntemiyle İşaretleme yöntemi arasında mükemmel bir uyum olmasını sağlamıştır. Bu nedenle yapılan değerlendirmelerde iki yöntemin aynı anda kullanılacağı durumda zaman tasarrufu olması açısından yöntemlerden birinin kullanılması standart belirleme yöntemlerini uygulayacaklara önerilebilir.

3. Yapılan anket sonucunda en kolay uygulanabilen, en anlaşılır bulunan, en rahat karar verilen yöntem olarak uzmanların % 66'sı tarafından Evet/Hayır yöntemi belirlendiğinden standart belirleme sürecini uygulayacaklara Evet/Hayır yöntemini kullanmaları önerilebilir.

4. Belirlenen geçme puanlarına göre başarılı sayılan öğrenci yüzdeleri farklılaştığından yapılacak olan standart belirleme sürecinde birden fazla yöntem kullanılması standart belirleme sürecini uygulayacaklara önerilebilir.

5.2.2. Araştırmacılara Dönük Öneriler

1. Araştırmada test merkezli yöntemlerden Evet/Hayır, Ebel ve İşaretleme standart belirleme yöntemleri kullanılmıştır. Benzer bir araştırma farklı test merkezli standart belirleme yöntemleri kullanılarak ya da öğrenci merkezli standart belirleme yöntemleri kullanılarak da yapılabilir.

2. Araştırmada sadece çoktan seçmeli maddelerin olduğu bir sınav kullanılmıştır. Benzer araştırmalar farklı soru tiplerinin olduğu sınavlarla da yapılabilir.

3. Araştırmada sadece matematik dersinin kümeler konusunu içeren bir sınav kullanılmıştır. Benzer araştırmalar aynı öğrenme alanına ait farklı konuların birlikte bulunduğu sınavlarla ya da farklı derslere ait sınavlarla yapılabilir.

4. Uzmanların görüşlerinin ve değerlendirmelerinin alındığı daha detaylı bir çalışma yapılabilir.

5. Uzman sayısının, yöntemlerle belirlenen geçme puanını nasıl etkilediğinin araştırıldığı bir çalışma yapılması faydalı olabilir.
6. Standart belirleme sürecinde kullanılan sınavın cevap anahtarının ve sınavın sorularının parametrelerinin uzmanlara verilmesinin uzman kararını nasıl etkilediğinin araştırıldığı bir çalışma yapılabilir.
7. Benzer bir çalışma İşaretleme yöntemi için literatürdeki farklı cevaplama olasılık değerleri için yapılabilir.
8. Uzmanlardan yöntemler için farklı zamanlarda karar alınarak aynı zamanda alınan kararlarla karşılaştırmalar yapan bir araştırma yapılabilir.

KAYNAKÇA

- Angoff, W. H. (1988). Proposals for theoretical and applied development in measurement. *Applied Measurement in Education*, 1, 215-222.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Ankara: ÖSYM Yayınları.
- Beretvas, N.S. (2004). Comparison of Bookmark difficulty locations under different item response models. *Applied Psychological Measurement*, 28(1), 25-47.
- Berk, R. A. (1986). A consumer's guide to setting performance standards on criterion-referenced tests. *Review of Educational Research*, 56(1), 137-172.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: The Guilford Press.
- Buckendahl, C. W., Smith, W. R., Impara, J. C. ve Plake, B. S. (2002). A comparison of angoff and bookmark standard setting methods. *Journal of Educational Measurement*, 39(3), 253-263.
- Cizek, G. J. (2001). *Setting performance standards: Concepts, methods and perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Chang, L. (1996). *Comparison between the nedelsky and angoff standart-setting methods*. Paper presented at the Annual Meeting of the National Council on Measurement in Education. New York: USA.
- Chinn, R. N. ve Hertz, N. R. (2002), Alternative approaches to standard setting for licensing and certification examinations. *Applied Measurement in Education*, 15 (1), 1-14.
- Crocker, L. ve Algina, J. (1986). *Introduction to classical and modern test theory*. Florida: Holt, Rinehart and Winston, Inc.
- Çetin, S. (2011). *İşaretleme ve Angoff standart belirleme yöntemlerinin karşılaştırılması*. Doktora Tezi, Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Çukadar, İ. (2013). *Norm ve ölçüt dayanaklı değerlendirmelerin karşılaştırılmasına ilişkin bir çalışma*. Yüksek Lisans Tezi, Ankara: Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü.
- Davis-Becker, S. L., Buckendahl, C. W. ve Gerrow, J. (2011). Evaluating the Bookmark standard setting method: The impact of random item ordering. *International Journal of Testing*, 11(1), 24-37.
- Demir, O. (2014). *Angoff, Nedelsky ve Ebel standart belirleme yöntemleri ile belirlenen kesme puanlarının karşılaştırılması*. Yüksek Lisans Tezi, Bolu: Abant İzzet Baysal Üniversitesi Eğitim Bilimleri Enstitüsü.
- Ertürk, S. (1994). *Eğitimde program geliştirme* (8. Baskı). Ankara: Meteksan Yayınevi.
- Ferguson, G. A. (1959), *Statistical analysis in psychology and education*. New York: McGraw-Hill

- Green, D. R., Trimble, C. S. ve Lewis, D. M. (2003). Interpreting the results of three different standard setting procedures. *Educational Measurement: Issues and Practices*, 22(1), 22-32.
- Gündeğer, C. (2012). *Angoff, Evet/Hayır ve Ebel standart belirleme yöntemlerinin karşılaştırılması*. Yüksek Lisans Tezi, Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Hambleton R. K ve Pitoniak M. J. (2006). Setting performance standards. In: Brennan RL, editor. *Educational Measurement* 4(433-470). Westport, CT: Praeger Publishers.
- Hambleton, R. K. (2001). *Setting performance standards on educational assessments and criteria for evaluating the process*. In G. J.
- Cizek (Ed.), *Setting performance standards: Concepts, methods and perspectives* (89-116). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hambleton, R. K., Jaeger, R. M., Plake, B. S. ve Mills, C. (2000). Setting performance standards on complex educational assessments. *Psychological Measurement*, 24(4), 355-366.
- Hambleton, R. K., Swaminathan, H. ve Rogers, H. J. (1991). *Fundamentals of item response theory*. CA: Sage Publication, Newbury Park.
- Hsieh, M. (2013). Comparing Yes/No, Angoff and Bookmark standard setting methods in the context of english assessment. *Language Assessment Quarterly*, 10, 331-350.
- Hu, L. ve Bentler, P. M. (1999). *Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives*. Structural Equation Modeling, 1-55.
- Huyhn, H. (2006). A Clarification on the response probability criterion r_p 67 for standard settings based on Bookmark and item mapping, *Educational Measurement: Issues and Practice*.
- Impara, J. C. ve 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.
- Impara, J. C. ve Plake, B. S. (1997). Standard Setting: An Alternative Approach. *Journal of Educational Measurement*, 34(4), 353-366.
- Kaptan, S. (1977). *Bilimsel araştırma teknikleri*. Ankara: Tekişik Matbaası ve Rehber Yayınevi.
- Kara, Y. ve Kelecioğlu, H. (2015). Puanlayıcı Niteliklerinin Kesme Puanlarının Belirlenmesine Etkisinin Genellenabilirlik Kuramı'yla İncelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi* 6(1), 58-71
- Karasar, N. (1982). *Bilimsel araştırma yöntemi*. Ankara: Matbaş Matbaacılık ve Ambalaj Sanayi.
- Landis, J. R. ve Koch, G. G. (1997). Measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.

- Lin, J. (2006). The bookmark procedure for setting cut-scores and finalizing performance standards: Strengths and weaknesses. *Alberta journal of educational research*, 52, 36-52.
- Livingston, S. A. ve Zieky, M. J. (1983). *A comparative study of standard-setting methods*. Princeton: Educational Testing Service.
- Livingston, S. A. ve Zieky, M. J. (1982). *Passing scores: A manual for setting standards of performance on educational and occupational tests*. New Jersey: Educational Testing Service. (http://www.ets.org/Media/Research/pdf/passing_scores.pdf, 25.12.2014 tarihinde erişildi).
- Lord, F. (1980). *Applications of Item Response Theory to Practical Testing Problems*. New Jersey: Lawrence Erlbaum Associates Publishers.
- Mertler, Craig A. ve Rachel A. Vannatta. (2005). *Advanced and multivariate statistical methods: practical application and interpretation*. 3rd Edition. Glendale, CA: Pyrczak Publishing.
- Mitzel, H. C., Lewis, D. M., Patz, R. J. ve Green, D. R. (2001). The bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Standard setting: Concepts, methods and perspectives*. (pp. 249-281). Mahwah, NJ: Lawrence Erlbaum Associates.
- Murphy, K. R.; Davidshofer, C. O. (1991). *Psychological Testing: Principles and Applications*. New Jersey: Prantice Hall.
- Norcini, J. J., Shea, J. A. & Kanya, D. T. (1988). The effect of various factors on standard setting. *Journal of Educational Measurement*, 25(1), 57-65.
- Popham, W. J. (1978). As always, provocative. *Journal of Educational Measurement*, 15, 230-297.
- Olsen, J.B. ve Smith, R. (2008). Cross validating modified Angoff and Bookmark standard setting for a home inspection certification. *Paper presented at the Annual Meeting of the National Council on Measurement in Education*, NY.
- Ömür, S. ve Selvi, H. (2010). Angoff, Ebel ve Nedelsky yöntemleriyle belirlenen kesme puanlarının sınıflama tutarlılıklarının karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(2), 109-113.
- Özçelik, D. A. (1992). *Ölçme ve değerlendirme*. Ankara: ÖSYM Yayınları.
- Özdamar, K. (2004). *Paket programlar ile istatistiksel veri analizi (Çok değişkenli analizler)*. Eskişehir: Kaan Kitabevi.
- Peterson, C. H., Schulz, E. M. ve Engelhard Jr., G. (2011). Reliability and validity of bookmark-based methods for standard setting: Comparisons to Angoff-based methods in the National Assessment of Educational Progress. *Educational Measurement: Issues and Practice*, 30(2), 3-14.
- 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. A., Glaser, R., Linn, R. ve Bohrnstedt, G. (1993). *Setting performance standards for student achievement*. Stanford, CA: National Academy of Education.
- Skaggs, G. ve Tessema, A. (2001). *Item disorderliness with the Bookmark standard setting procedure*. Annual Meeting of the National Council on Measurement in Education, Seattle: WA.
- Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik*. Ankara: Seçkin Yayıncılık.
- Tanrıverdi, S. (2006). *Standart Belirleme Yöntemlerinin Geçme Puanları Üzerine Etkisi*. Yayımlanmamış yüksek lisans tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Taşdelen, G. (2009). *Nedelsky ve Angoff standart belirleme yöntemlerinin genellenebilirlik kuramı ile karşılaştırılmasına ilişkin bir araştırma*. Yüksek Lisans Tezi, Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Taşdemir, F. (2013). *Angoff (1-0), Nedelsky ve sınır değerleri saptama yöntemleri ile bir testin sınıflama doğruluklarının incelenmesi*. Doktora Tezi, Ankara: Ankara Üniversitesi Eğitim Bilimleri Enstitüsü.
- Turgut, M. F. (1993). *Eğitimde ölçme ve değerlendirme metotları*. Ankara: Saydam Matbaacılık.
- Von Eye, A. ve Mun, E. Y. (2005). *Analyzing Rater Agreement: Manifest Variable Methods*. Mahwah, NJ: Lawrence Erlbaum.
- Wang, N. (2003). Use of the Rasch IRT model in standard setting: An item mapping method. *Journal of Educational Measurement*, 40, 231-253.
- Williams, N. J. ve Schulz, E. M. (2005). *An investigation of response probability values used in standard setting*. Paper presented at the meeting of the National Council on Measurement in Education, Montreal, Quebec, Canada.
- Yin, P., Schulz, M. ve Sconing, J. (2005). A comparison of cut scores and cut score variability from Angoff-based and Bookmark procedure in standard setting. *Paper presented at the meeting of the National Council on Measurement in Education*, San Diego, CA.
- Zahran, A. E. A. H. (1981). *The Impact of Multiple-Choice Item Styles, Judge Experience and Item Taxonomy Level on Minimum Passing Standards and Inter Scorer Agreement*. Yayımlanmamış yüksek lisans tezi. The Ohio State University.

EKLER DİZİNİ

EK 1. ETİK KURUL İZİN MUAFİYET FORMU

Form 40

Tez Çalışması Etik Kurul İzin Muafiyeti Formu

03 / 07 / 2015

Hacettepe Üniversitesi
Eğitim Bilimleri Enstitüsü
Eğitim Bilimleri Anabilim Dalı Başkanlığı'na

Tez Başlığı / Konusu:	EVET/HAYIR, EBEL VE İŞARETLEME STANDART BELİRLEME YÖNTEMLERİNİN KARŞILAŞTIRILMASI
-----------------------	---

Yükarıda başlığı/konusu gösterilen tez çalışmam:

1. İnsan ve hayvan üzerinde deney niteliği taşımamaktadır.
2. Biyolojik materyal (kan, idrar vb. biyolojik sıvılar ve numuneler) kullanılmasını gerektirmemektedir.
3. Beden bütünlüğüne müdahale içermemektedir.
4. Gözlemsel ve betimsel araştırma (anket, ölçek/skala çalışmaları, dosya taramaları, vcr kaynakları taraması, sistem-model geliştirme çalışmaları) niteliğinde değildir.

Hacettepe Üniversitesi Etik Kurulları ve Komisyonlarının Yönergelerini inceledim ve bunlara göre tez çalışmamın yürütülebilmesi için herhangi bir Etik Kuruldan izin alınmasına gerek olmadığını; aksi durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.

Sevda Korkmaz
(Öğrencinin Fdi Soyadı, İmzası)

Öğrenci Bilgileri

Adı Soyadı	Sevda KORKMAZ
Öğrenci No	N11222528
Anabilim Dalı	Eğitim Bilimleri Anabilim Dalı
Programı	Eğitimde Ölçme ve Değerlendirme Bilim Dalı
Statüsü	<input checked="" type="checkbox"/> Yüksek Lisans <input type="checkbox"/> Doktora <input type="checkbox"/> Bütünleşik Dr.

Danışman Görüşü ve Onayı

Dr. Zafer ÇEPNİ
(İmza)
(Danışmanın Zımnı, Adı ve Soyadı)



EK 2. KÜMELER BAŞARI TESTİ

Adı Soyadı:

Sınıfı:

Okulu:

9. Sınıf 1. Dönem Matematik Karne Notu:

1- **Kümelerin gösterim yöntemlerine ilişkin aşağıdaki bilgilerden hangisi ya da hangileri doğrudur?**

1. Her küme Venn şeması ile gösterilebilir.
2. Her küme liste yöntemi ile gösterilebilir.
3. Her küme ortak özellik yöntemi ile gösterilebilir.

- A) Yalnız 1 B) Yalnız 2 C) Yalnız 3
D) 1 ve 3 E) 1 ve 2

2- **Rakamları farklı 4 basamaklı en büyük doğal sayının rakamlarından oluşan kümenin liste biçiminde gösterimi aşağıdakilerden hangisinde doğru olarak verilmiştir?**

- A) {9, 8, 7, 6} B) {9, 8, 7, 5}
C) {9, 7, 5, 3} D) {1, 2, 3, 4}
E) {1, 0, 2, 3}

3- **Sonlu, sonsuz ve boş kümeye ilişkin aşağıdaki ifadelerden hangileri doğrudur?**

- I. 7 ile bölünen sayıların oluşturduğu küme sonsuz kümedir.
- II. Rakamların kümesi sonsuz kümedir.
- III. Sıfırdan küçük tam sayıların kümesi sonlu kümedir.
- IV. İki basamaklı çift asal sayıların kümesi boş kümedir.

- A) I ve II B) I ve IV C) II ve IV
D) I, II, III E) III ve IV

4- I. $A = \{x : x \text{ çift tam sayı}\}$
II. $E = \{\text{Haftanın ş ile başlayan günleri}\}$
III. $S = \{\text{Türkiye'nin d ile başlayan illeri}\}$
Yukarıda verilen kümelere hangisi ya da hangileri boş küme ifade eder?

- A) I ve II B) I ve III C) II ve III
D) I, II, III E) Yalnız II

5- Bir kümenin kendisi hariç tüm alt kümelerine özalt kümeleri denir.

Aşağıdaki kümelerden hangisinin özalt küme sayısı 7'dir?

- A) {a} B) {a, b} C) {a, b, c}
D) {a, b, c, s} E) {a, b, c, s, k}

6- M ve N iki küme olmak üzere $s(M) = s(N) + 2$ 'dir.

M ve N kümelerinin alt küme sayıları toplamı 80 ise M kümesinin eleman sayısı kaçtır?

- A) 3 B) 4 C) 5 D) 6 E) 7

7- 3 elemanlı alt kümelerinin sayısı 5 elemanlı alt kümelerinin sayısına eşit olan küme kaç elemanlıdır?

- A) 11 B) 10 C) 9 D) 8 E) 7

8- $N = \{i, s, t, a, n, b, u, l\}$ kümesinin alt kümelerinin kaç tanesinde i ve n harfleri eleman olarak bulunur?

- A) 256 B) 128 C) 64 D) 32 E) 16

9- $D = \{f, k, e, r, \{t\}\}$ kümesi veriliyor.

Aşağıdaki kümelerden hangisi D kümesinin alt kümesi değildir?

- A) $\{f, k, e, r, \{t\}\}$ B) $\{f, k, r, \{t\}\}$
C) $\{f, k, e, \}$ D) $\{f, k, e, r, t\}$ E) $\{\{t\}\}$

10- $A \cap B \neq \emptyset$ olmak üzere $s(A \cup B) = 18$ ise A kümesinin en fazla kaç elemanı vardır?

- A) 18 B) 17 C) 14 D) 16 E) 15

11- $A: \{2, 3\}$, $B: \{-2, 2, 3\}$,
 $C: \{\}$,

$D: \{x: (x^2 - 4)(x - 3) = 0 \text{ ve } x \text{ tamsayı}\}$

$E: \{\text{sıfırdan küçük ilk 2 asal sayı}\}$

Yukarıda verilen kümelerden hangileri birbirine eşittir?

- A) A ile D B) A ile E C) B ile C
D) E ile B E) E ile C

12- $G = \{x: x \in \mathbb{R} \text{ ve } -1 \leq x \leq 5\}$

$H = \{x: x \in \mathbb{R} \text{ ve } -3 \leq x \leq 9\}$

kümelerinin kesişimi aşağıdakilerden hangisinde verilmiştir?

A) $\{-3, -4, -5, -6 \dots\}$

B) $\{-1, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

C) $\{-1, 0, 1, 2, 3, 4\}$

D) G

E) H

13- Bir mahalledeki ailelerin %85'inde bilgisayar, %35'inde tablet olduğu tespit edilmiştir.

Buna göre ailelerin en fazla yüzde kaçında hem bilgisayar hem tablet vardır?

- A) 20 B) 35 C) 50 D) 85 E) 120

14- E evrensel küme, A ve B herhangi iki küme olmak üzere aşağıdaki ifadelerden hangisi yanlıştır?

- A) $A \cap \emptyset = \emptyset$ B) $A \cap A = A$
C) $A \cap E = A$
D) $A \cap B = A$ ise $B \subset A$
E) $A \cap B = B \cap A$

15- E, evrensel küme ve E'nin 2 alt kümesi F ve G olmak üzere $s(G) + s(F') = 22$ ve $s(F) + s(G') = 12$ ise evrensel kümenin eleman sayısı kaçtır?

- A) 16 B) 17 C) 18 D) 19 E) 20

16- A,B herhangi iki küme

$s(A') = 12$, $s(A) = 10$, $s(A \cap B) = 3$,
 $s(A \cup B) = 16$ ise B' kümesinin eleman sayısı kaçtır?

- A) 9 B) 11 C) 13 D) 15 E) 17

17- A,B herhangi iki küme olmak üzere, aşağıdakilerden hangisi her zaman evrensel kümeyi belirtir?

- A) $(A \cap B) \cup B$ B) $A \cup B$
C) $(A \cup B) \cup (A \cap B)$ D) $A' \cup A$
E) $A \cup (B \setminus A)$

18- A ve B kümeleri E evrensel kümesinin alt kümeleridir.

$$[(A \cap B) \cup (A \cap B')] \cup (B \cap B')$$

işleminin sonucu aşağıdakilerden hangisidir?

- A) A B) B C) $A \cup B$ D) $A \cap B$ E) \emptyset

19- E, evrensel küme $E = \{x: x \text{ doğal sayı ve } x < 10\}$ ve $A \cap B = \{4, 6, 8, 9\}$ ise $A' \cup B'$ kümesi hangisidir?

- A) $\{1, 2, 3, 5, 11, 12, \dots\}$ B) $\{\}$
C) $\{1, 2, 3, 5, 7\}$ D) $\{0, 1, 2, 3, 5, 7\}$
E) $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$

20- $(A \setminus B) \cap (B \setminus A)$ aşağıdakilerden hangisine eşittir?

- A) \emptyset B) A C) B D) A' E) B'

21- 27 kişilik bir gruptaki öğrenciler İngilizce ve Almanca dillerinden en az birini bilmektedir. Bu sınıftaki öğrencilerden yalnız İngilizce bilenlerin sayısı, yalnız Almanca bilenlerin sayısına eşittir.

Almanca ve İngilizce dillerinin her ikisini bilenlerin sayısı yalnız İngilizce bilenlerin sayısına eşit olduğuna göre, Almanca bilenlerin sayısı kaçtır?

- A) 25 B) 22 C) 18 D) 15 E) 13

22- Bir toplulukta gözlüklü kızların sayısı gözlüksüz erkeklerin sayısına eşittir. Gözlüksüz kızların sayısı gözlüklü erkeklerin sayısından 12 fazladır.

Topluluktaki kızların sayısı erkeklerin sayısının 4 katı olduğuna göre bu toplulukta kaç kişi vardır?

A) 15 B) 18 C) 20 D) 25 E) 30

23- Futbol, basketbol ve voleyboldan en çok ikisini oynayanların bulunduğu 48 kişilik bir sınıfta, futbol ve basketbol oynayan 5, futbol ve voleybol oynayan 4, basketbol ve voleybol oynayan 3 kişi bulunmaktadır. Bu sınıfta basketbol oynayanların sayısı 20, voleybol oynayanların sayısı 14 'tür. **Buna göre sınıfta sadece futbol oynayan kaç kişi bulunmaktadır?**

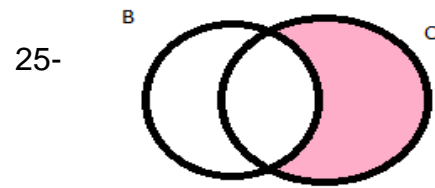
Test bitti. Teşekkür ederim ☺

A) 14 B) 17 C) 20 D) 22 E) 25

24- Piyano ve gitar kursuna giden öğrencilerden oluşan bir sınıfta piyano kursuna 20, gitar kursuna 16 kişi gitmektedir.

Her iki kursa 7 kişi gittiğine göre sınıfta kaç kişi vardır?

A) 24 B) 25 C) 27 D) 28 E) 29



Şekildeki taralı bölge A aşağıdakilerden hangisine eşittir?

A) $C \setminus B$ B) $B \cap C$ C) $C \cup B$

D) $C \cap B$ E) B

EK 3. EVET/HAYIR FORMU

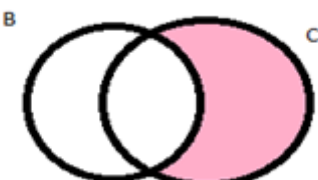
SORULAR	EVET/HAYIR YÖNTEMİ				
<p>1. Kümelerin gösterim yöntemlerine ilişkin aşağıdaki bilgilerden hangisi ya da hangileri <u>doğrudur</u>?</p> <p>1. Her küme Venn şeması ile gösterilebilir.</p> <p>2. Her küme liste yöntemi ile gösterilebilir.</p> <p>3. Her küme ortak özellik yöntemi ile gösterilebilir.</p> <p>A) Yalnız 1 B) Yalnız 2 C) Yalnız 3</p> <p style="text-align: center;">D) 1 ve 3 E) 1 ve 2</p>	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 80%;">Doğru cevaplayabilir.</td> <td style="width: 20%;"></td> </tr> <tr> <td>Doğru cevaplayamaz.</td> <td></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>2- Rakamları farklı 4 basamaklı en büyük doğal sayının rakamlarından oluşan kümenin liste biçiminde gösterimi aşağıdakilerden hangisinde <u>doğru</u> olarak gösterilmiştir?</p> <p>A){9, 8, 7, 6} B) {9, 8, 7, 5} C) {9, 7, 5, 3}</p> <p style="text-align: center;">D) {1, 2, 3, 4} E) {1, 0, 2, 3}</p>	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 80%;">Doğru cevaplayabilir.</td> <td style="width: 20%;"></td> </tr> <tr> <td>Doğru cevaplayamaz.</td> <td></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>3- Sonlu, sonsuz ve boş kümeye ilişkin aşağıdaki ifadelerden hangileri <u>doğrudur</u>?</p> <p>I. 7 ile bölünen sayıların oluşturduğu küme sonsuz kümedir.</p> <p>II. Rakamların kümesi sonsuz kümedir.</p> <p>III. Sıfırdan küçük tam sayıların kümesi sonlu kümedir.</p> <p>IV. İki basamaklı çift asal sayıların kümesi boş kümedir.</p> <p>A) I ve II B) I ve IV C) II ve IV</p> <p style="text-align: center;">D) I, II, III E) III ve IV</p>	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 80%;">Doğru cevaplayabilir.</td> <td style="width: 20%;"></td> </tr> <tr> <td>Doğru cevaplayamaz.</td> <td></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					

<p>4- I. $A = \{x : x \text{ çift tam sayı}\}$</p> <p>II. $E = \{\text{Haftanın ş ile başlayan günleri}\}$</p> <p>III. $S = \{\text{Türkiye'nin d ile başlayan illeri}\}$</p> <p>Yukarıda tanımlanan kümelerden hangisi ya da hangileri <u>boş küme</u> ifade eder?</p> <p>A) I ve II B) I ve III C) II ve III D) I, II, III E) Yalnız II</p>	<table border="1"> <tr> <td>Doğru cevaplayabilir.</td> <td></td> </tr> <tr> <td>Doğru cevaplayamaz.</td> <td></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>5- Bir kümenin kendisi hariç tüm alt kümelerine özalt kümeleri denir.</p> <p>Aşağıdaki kümelerden hangisinin özalt küme sayısı 7'dir?</p> <p>A) $\{a\}$ B) $\{a, b\}$ C) $\{a, b, c\}$</p> <p> D) $\{a, b, c, s\}$ E) $\{a, b, c, s, k\}$</p>	<table border="1"> <tr> <td>Doğru cevaplayabilir.</td> <td></td> </tr> <tr> <td>Doğru cevaplayamaz.</td> <td></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>6- M ve N iki küme olmak üzere $s(M) = S(N) + 2$ 'dir.</p> <p>M ve N kümelerinin alt küme sayıları toplamı 80 ise M kümesinin eleman sayısı kaçtır?</p> <p>A) 3 B) 4 C) 5 D) 6 E) 7</p>	<table border="1"> <tr> <td>Doğru cevaplayabilir.</td> <td></td> </tr> <tr> <td>Doğru cevaplayamaz.</td> <td></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>7- 3 elemanlı alt kümelerinin sayısı 5 elemanlı alt kümelerinin sayısına eşit olan küme kaç elemanlıdır?</p> <p>11 B) 10 C) 9 D) 8 E) 7</p>	<table border="1"> <tr> <td>Doğru cevaplayabilir.</td> <td></td> </tr> <tr> <td>Doğru cevaplayamaz.</td> <td></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					

<p>8- $N = \{i, s, t, a, n, b, u, l\}$ kümesinin alt kümelerinin kaç tanesinde i ve n harfleri eleman olarak bulunur?</p> <p>256 B) 128 C) 64 D) 32 E) 16</p>	<table border="1"> <tr> <td>Doğru cevaplayabilir.</td> <td></td> </tr> <tr> <td>Doğru cevaplayamaz.</td> <td></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>9- $D = \{f, k, e, r, \{t\}\}$ kümesi veriliyor.</p> <p>Aşağıdaki kümelerden hangisi D kümesinin alt kümesi <u>değildir</u>?</p> <p>A) $\{f, k, e, r, \{t\}\}$ B) $\{f, k, r, \{t\}\}$ C) $\{f, k, e, \}$ D) $\{f, k, e, r, t\}$ E) $\{\{t\}\}$</p>	<table border="1"> <tr> <td>Doğru cevaplayabilir.</td> <td></td> </tr> <tr> <td>Doğru cevaplayamaz.</td> <td></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>10- $A \cap B \neq \emptyset$ olmak üzere $s(A \cup B) = 18$ ise A kümesinin <u>en fazla kaç elemanı vardır</u>?</p> <p>18 B)17 C)14 D)16 E)15</p>	<table border="1"> <tr> <td>Doğru cevaplayabilir.</td> <td></td> </tr> <tr> <td>Doğru cevaplayamaz.</td> <td></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>11- $A = \{2, 3\}$, $B = \{-2, 2, 3\}$, $C = \{\}$, $D = \{x: (x^2 - 4) \cdot (x - 3) = 0 \text{ ve } x \text{ tam sayı}\}$ $E = \{\text{sıfırdan küçük ilk 2 asal sayı}\}$</p> <p>Yukarıda verilen kümelerden hangileri birbirine eşittir?</p> <p>A) A ile D B) A ile E C) B ile C D) E ile B E) E ile C</p>	<table border="1"> <tr> <td>Doğru cevaplayabilir.</td> <td></td> </tr> <tr> <td>Doğru cevaplayamaz.</td> <td></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					

<p>12- $G = \{x: x \in \mathbb{R} \text{ ve } -1 \leq x \leq 5\}$</p> <p>$H = \{x: x \in \mathbb{R} \text{ ve } -3 \leq x \leq 9\}$ kümelerinin kesişimi aşağıdakilerden hangisinde verilmiştir?</p> <p>A) $\{-3, -4, -5, -6 \dots\}$ B) $\{-1, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ C) $\{-1, 0, 1, 2, 3, 4\}$ D) G E) H</p>	<table border="1"> <tr> <td data-bbox="1050 277 1342 421">Doğru cevaplayabilir.</td> <td data-bbox="1342 277 1404 421"></td> </tr> <tr> <td data-bbox="1050 421 1342 564">Doğru cevaplayamaz.</td> <td data-bbox="1342 421 1404 564"></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>13- Bir mahalledeki ailelerin %85'inde bilgisayar, %35'inde tablet olduğu tespit edilmiştir.</p> <p>Buna göre ailelerin <u>en fazla</u> yüzde kaçında hem bilgisayar hem tablet vardır?</p> <p>20 B) 35 C) 50 D) 85 E) 120</p>	<table border="1"> <tr> <td data-bbox="1050 651 1342 795">Doğru cevaplayabilir.</td> <td data-bbox="1342 651 1404 795"></td> </tr> <tr> <td data-bbox="1050 795 1342 938">Doğru cevaplayamaz.</td> <td data-bbox="1342 795 1404 938"></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>14- E evrensel küme, A ve B herhangi iki küme olmak üzere aşağıdaki ifadelerden hangisi <u>yanlıştır</u>?</p> <p>A) $A \cap \emptyset = \emptyset$ B) $A \cap A = A$ C) $A \cap E = A$ D) $A \cap B = A$ ise $B \subset A$ E) $A \cap B = B \cap A$</p>	<table border="1"> <tr> <td data-bbox="1050 1014 1342 1158">Doğru cevaplayabilir.</td> <td data-bbox="1342 1014 1404 1158"></td> </tr> <tr> <td data-bbox="1050 1158 1342 1301">Doğru cevaplayamaz.</td> <td data-bbox="1342 1158 1404 1301"></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>15- E, evrensel küme ve E'nin 2 alt kümesi F ve G olmak üzere $s(G) + s(F') = 22$ ve $s(F) + s(G') = 12$ ise evrensel kümenin eleman sayısı kaçtır?</p> <p>A) 16 B) 17 C) 18 D) 19 E) 20</p>	<table border="1"> <tr> <td data-bbox="1050 1420 1342 1563">Doğru cevaplayabilir.</td> <td data-bbox="1342 1420 1404 1563"></td> </tr> <tr> <td data-bbox="1050 1563 1342 1706">Doğru cevaplayamaz.</td> <td data-bbox="1342 1563 1404 1706"></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>16- A, B herhangi iki küme $s(A') = 12$, $s(A) = 10$, $s(A \cap B) = 3$, $s(A \cup B) = 16$ ise B' kümesinin eleman sayısı kaçtır?</p> <p>A) 9 B) 11 C) 13 D) 15 E) 17</p>	<table border="1"> <tr> <td data-bbox="1050 1724 1342 1868">Doğru cevaplayabilir.</td> <td data-bbox="1342 1724 1404 1868"></td> </tr> <tr> <td data-bbox="1050 1868 1342 2011">Doğru cevaplayamaz.</td> <td data-bbox="1342 1868 1404 2011"></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					

<p>17- A,B herhangi iki küme olmak üzere, aşağıdakilerden hangisi <u>her zaman</u> evrensel kümeyi belirtir?</p> <p>A) $(A \cap B) \cup B$ B) $A \cup B$ C) $(A \cup B) \cup (A \cap B)$ D) $A' \cup A$ E) $A \cup (B \setminus A)$</p>	<table border="1"> <tr> <td data-bbox="1045 230 1342 371">Doğru cevaplayabilir.</td> <td data-bbox="1342 230 1409 371"></td> </tr> <tr> <td data-bbox="1045 371 1342 515">Doğru cevaplayamaz.</td> <td data-bbox="1342 371 1409 515"></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>18- A ve B kümeleri E evrensel kümesinin alt kümeleridir.</p> <p>$[(A \cap B) \cup (A \cap B')] \cup (B \cap B')$ işleminin sonucu aşağıdakilerden hangisidir?</p> <p>A) A B) B C) A \cup B D) $A \cap B$ E) \emptyset</p>	<table border="1"> <tr> <td data-bbox="1045 622 1342 763">Doğru cevaplayabilir.</td> <td data-bbox="1342 622 1409 763"></td> </tr> <tr> <td data-bbox="1045 763 1342 907">Doğru cevaplayamaz.</td> <td data-bbox="1342 763 1409 907"></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>19- E, evrensel küme $E = \{x : x \text{ doğal sayı ve } x < 10\}$ ve $A \cap B = \{4, 6, 8, 9\}$ ise $A' \cup B'$ kümesi hangisidir?</p> <p>A) $\{1, 2, 3, 5, 11, 12, \dots\}$ B) $\{\}$ C) $\{1, 2, 3, 5, 7\}$ D) $\{0, 1, 2, 3, 5, 7\}$ E) $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$</p>	<table border="1"> <tr> <td data-bbox="1045 922 1342 1064">Doğru cevaplayabilir.</td> <td data-bbox="1342 922 1409 1064"></td> </tr> <tr> <td data-bbox="1045 1064 1342 1207">Doğru cevaplayamaz.</td> <td data-bbox="1342 1064 1409 1207"></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>20- $(A \setminus B) \cap (B \setminus A)$ aşağıdakilerden hangisine eşittir?</p> <p>A) \emptyset B) A C) B D) A' E) B'</p>	<table border="1"> <tr> <td data-bbox="1045 1214 1342 1355">Doğru cevaplayabilir.</td> <td data-bbox="1342 1214 1409 1355"></td> </tr> <tr> <td data-bbox="1045 1355 1342 1498">Doğru cevaplayamaz.</td> <td data-bbox="1342 1355 1409 1498"></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>21- 27 kişilik bir gruptaki öğrenciler İngilizce ve Almanca dillerinden en az birini bilmektedir. Bu sınıftaki öğrencilerden yalnız İngilizce bilenlerin sayısı, yalnız Almanca bilenlerin sayısına eşittir.</p> <p>Almanca ve İngilizce dillerinin her ikisini bilenlerin sayısı yalnız İngilizce bilenlerin sayısına eşit olduğuna göre, Almanca bilenlerin sayısı kaçtır?</p> <p>A) 25 B) 22 C) 18 D) 15 E) 13</p>	<table border="1"> <tr> <td data-bbox="1045 1581 1342 1722">Doğru cevaplayabilir.</td> <td data-bbox="1342 1581 1409 1722"></td> </tr> <tr> <td data-bbox="1045 1722 1342 1865">Doğru cevaplayamaz.</td> <td data-bbox="1342 1722 1409 1865"></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					

<p>22- Bir toplulukta gözlüklü kızların sayısı gözlüksüz erkeklerin sayısına eşittir. Gözlüksüz kızların sayısı gözlüklü erkeklerin sayısından 12 fazladır.</p> <p>Topluluktaki kızların sayısı erkeklerin sayısının 4 katı olduğuna göre bu toplulukta kaç kişi vardır?</p> <p>A) 15 B) 18 C) 20 D) 25 E) 30</p>	<table border="1"> <tr> <td data-bbox="1038 280 1342 421">Doğru cevaplayabilir.</td> <td data-bbox="1342 280 1422 421"></td> </tr> <tr> <td data-bbox="1038 421 1342 562">Doğru cevaplayamaz.</td> <td data-bbox="1342 421 1422 562"></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>23- Futbol, basketbol ve voleyboldan en çok ikisini oynayanların bulunduğu 48 kişilik bir sınıfta, futbol ve basketbol oynayan 5, futbol ve voleybol oynayan 4, basketbol ve voleybol oynayan 3 kişi bulunmaktadır. Bu sınıfta basketbol oynayanların sayısı 20, voleybol oynayanların sayısı 14 'tür.</p> <p>Buna göre sınıfta sadece futbol oynayan kaç kişi bulunmaktadır?</p> <p>A) 14 B) 17 C) 20 D) 22 E) 25</p>	<table border="1"> <tr> <td data-bbox="1038 674 1342 815">Doğru cevaplayabilir.</td> <td data-bbox="1342 674 1422 815"></td> </tr> <tr> <td data-bbox="1038 815 1342 956">Doğru cevaplayamaz.</td> <td data-bbox="1342 815 1422 956"></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>24- Piyano ve gitar kursuna giden öğrencilerden oluşan bir sınıfta piyano kursuna 20, gitar kursuna 16 kişi gitmektedir.</p> <p>Her iki kursa 7 kişi gittiğine göre sınıfta kaç kişi vardır?</p> <p>A) 24 B) 25 C) 27 D) 28 E) 29</p>	<table border="1"> <tr> <td data-bbox="1038 1218 1342 1359">Doğru cevaplayabilir.</td> <td data-bbox="1342 1218 1422 1359"></td> </tr> <tr> <td data-bbox="1038 1359 1342 1500">Doğru cevaplayamaz.</td> <td data-bbox="1342 1359 1422 1500"></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					
<p>25-</p> <div style="text-align: center;">  </div> <p>Şekildeki taralı bölge A aşağıdakilerden hangisine eşittir?</p> <p>A) $C \setminus B$ B) $B \cap C$ C) $C \cup B$ D) $C \cap B$ E) $B \setminus C$</p>	<table border="1"> <tr> <td data-bbox="1038 1612 1342 1753">Doğru cevaplayabilir.</td> <td data-bbox="1342 1612 1422 1753"></td> </tr> <tr> <td data-bbox="1038 1753 1342 1895">Doğru cevaplayamaz.</td> <td data-bbox="1342 1753 1422 1895"></td> </tr> </table>	Doğru cevaplayabilir.		Doğru cevaplayamaz.	
Doğru cevaplayabilir.					
Doğru cevaplayamaz.					

EK 4. EBEL FORMU

UYGUNLUK	GÜÇLÜK		
	Kolay	Orta	Zor
Gerekli	Sorular: Yüzde:	Sorular: Yüzde:	Sorular: Yüzde:
Önemli	Sorular: Yüzde:	Sorular: Yüzde:	Sorular: Yüzde:
Kabul Edilebilir	Sorular: Yüzde:	Sorular: Yüzde:	Sorular: Yüzde:
Tartışılabilir	Sorular: Yüzde:	Sorular: Yüzde:	Sorular: Yüzde:

EK 5. BİLGİLENDİRME FORMU

UZMAN GÖRÜŞLERİ İÇİN ARAŞTIRMA FORMU

Değerli Uzmanlar ve Öğretmenler

Yüksek lisans tezim için yaptığım araştırmada, Matematik dersi 9. sınıf “Kümeler” ünitesinin kazanımlarından oluşan çoktan seçmeli test kullanılarak teste ait geçme puanı belirlenecektir. Bu geçme puanının belirlenmesi için üç farklı yöntemle göre değerlendirme yapmanıza ihtiyaç duyulmaktadır. Yöntemlere göre yapılacak adımlar aşağıda verilmiştir.

Evet/Hayır Yöntemi: Testteki her bir soru için “ Matematik dersinde geçme-kalma sınırındaki bir öğrenci (minimum yeterlik düzeyindeki bir öğrenci) soruyu doğru cevaplayabilir mi? ” sorusuna cevap vermeniz istenmektedir. Cevabınızı, size verilen formda sorunun yanındaki kutucuğu işaretleyerek veriniz.

Ebel Yöntemi: Bu değerlendirmede sizden sorularla ilgili iki karar vermeniz istenmektedir. Tüm soruları gözden geçirdikten sonra öncelikle soruların güçlük derecesine (kolay, orta, zor) karar veriniz. Daha sonra sorunun testteki diğer sorulara göre uygunluk (gerekli, önemli, kabul edilebilir, tartışılabilir) derecelerine karar veriniz. Bu iki karara göre sorunun numarasını size verilen formda uygun gördüğünüz hücrenin karşısındaki “ Sorular” kısmına yazınız. Hücrelerin hepsi dolu olmak zorunda değildir.

Bu iki kararı verip, formu doldurduktan sonra “Matematik dersinde geçme-kalma sınırındaki öğrencilerin (minimum yeterlik düzeyindeki bir öğrencilerin) yüzde kaç hücredeki soruları doğru cevaplayabilir?” sorusuna cevap verin. Cevabınızı formdaki “yüzde” kısmına yazın.

Örnek: Testin 1,6, ve 9. soruları için güçlük derecesine kolay uygunluk derecesine önemli kararını verdiniz. Daha sonra bu sorular için geçme-kalma sınırındaki öğrencilerin yüzde 55'inin doğru cevap verebileceği düşündünüz. Bu kararlarınızı formun aşağıda belirtilen hücrelerine yazmalısınız.

UYGUNLUK	GÜÇLÜK		
	<u>Kolay</u>	Orta	Zor
Gerekli	Sorular: Yüzde:	Sorular: Yüzde:	Sorular: Yüzde:
<u>Önemli</u>	Sorular:1, 6, 9 Yüzde: 55	Sorular: Yüzde:	Sorular: Yüzde:
Kabul Edilebilir	Sorular: Yüzde:	Sorular: Yüzde:	Sorular: Yüzde:
Tartışılabilir	Sorular: Yüzde:	Sorular: Yüzde:	Sorular: Yüzde:

İşaretleme Yöntemi: Bu değerlendirme için size verilen formda, testteki sorular kolaydan zora doğru sıralanmıştır. Sizden bu değerlendirme için “ Matematik dersinde geçme-kalma sınırındaki bir öğrencinin (minimum yeterlik düzeyindeki bir öğrencinin) kolaydan zora sıralanmış bu sorulardan en son hangisini cevaplayabileceğini işaretlemeniz beklenmektedir.

Örnek: “Geçme-kalma sınırındaki bir öğrenci en son 10. soruyu yani 10. sayfadaki soruyu cevaplayabilir, 11. sayfadaki soruyu cevaplayamaz” diye düşünen bir öğretmen 10. sayfayı işaretlemelidir.

Soruları inceledikten sonra değerlendirmeleri teker teker yapmanız önemle rica olunur.

Katılımınız ve özveriniz için çok teşekkür ederim.

Hacettepe Üniversitesi Eğitimde Ölçme ve Değerlendirme Bölümü

Yüksek Lisans Öğrencisi

Sevda KORKMAZ

EK 6. ANKET MADDELERİ

Değerli uzmanlar ve öğretmenler,

Bu anket, standart belirleme (geçme puanı belirleme) sürecine ilişkin görüşlerinizi almak amacıyla hazırlanmıştır. Görüşleriniz kesme puanı belirleme yöntemlerini ve süreci değerlendirmede dikkate alınacaktır. Sorulara içtenlikle cevap vermeniz çalışmanın sonuçlarının güvenilirliği açısından önemlidir. Ankete isminizi yazmanıza gerek yoktur. Araştırmada elde edilen veriler bilimsel amaçlar dışında hiçbir şekilde kullanılmayacaktır. Katkılarınız için teşekkür ederim.

Hacettepe Üniversitesi Eğitimde Ölçme ve Değerlendirme Bölümü

Yüksek Lisans Öğrencisi

Sevda KORKMAZ

1. **Cinsiyetiniz:** Kadın () Erkek ()

2. **Mesleğiniz:** Öğretmen () Uzman ()

3. **Mesleğinizde kaçınıcı yılınızdasınız?**

4. Standart belirleme (geçme puanı belirleme) sürecindeki unsurlarla ilgili olarak görüşlerinizi aşağıdaki tabloda uygun sütunu işaretleyerek belirtiniz.

	Yeterli Değil	Kismen Yeterli	Yeterli	Oldukça Yeterli
İlgili test konusu alanı hakkında bilgilendirme				
Standart belirleme süreci ile ilgili bilgilendirme				
Geçme puanı tanımı ile ilgili bilgilendirme				
Kitapçık ve puanlamalarla ilgili bilgilendirmeler				
Yapılacak işlemle ilgili örnek bir uygulama				

5. Çalışma başında kitapçıkların sınıflandırılması ve puanlamaya ilişkin yapılan bilgilendirmenin süresini yeterli buldunuz mu?

- () Bilgilendirme süresi tam gerektiği kadardı.
() Bilgilendirme süresi azdı, daha fazla süre ayrılmalıydı.
() Bilgilendirme süresi fazlaydı, daha kısa tutulabilirdi.

6. Öğrenci performansını sınıflarken kullandığınız unsurların önem düzeyini uygun sütunu işaretleyerek belirtiniz.

	Önemli Değil	Biraz Önemli	Önemli	Çok Önemli
Geçme puanı tanımı				
Soruların güçlük düzeyleriyle ilgili düşünceleriniz				
Öğrencilerin düzeyiyle ilgili kendi düşünceleriniz				
Deneyimleriniz				
Diğer (varsa ekleyiniz)				

7. Karar verirken en çok zorlandığınız yöntem hangisidir?

Yes/ No Yöntemi () Ebel Yöntemi () İşaretleme Yöntemi ()

8. En rahat karar verdiğiniz yöntem hangisidir?

Yes/ No Yöntemi () Ebel Yöntemi () İşaretleme Yöntemi ()

9. En uygun, anlaşılır, kullanılabilir bulduğunuz yöntem hangisidir?

Yes/ No Yöntemi () Ebel Yöntemi () İşaretleme Yöntemi ()

10. Yöntemler ve süreçle ilgili eklemek istediklerinizi lütfen belirtiniz.

EK 7. EVET/HAYIR YÖNTEMİ İÇİN ALINAN KARARLAR

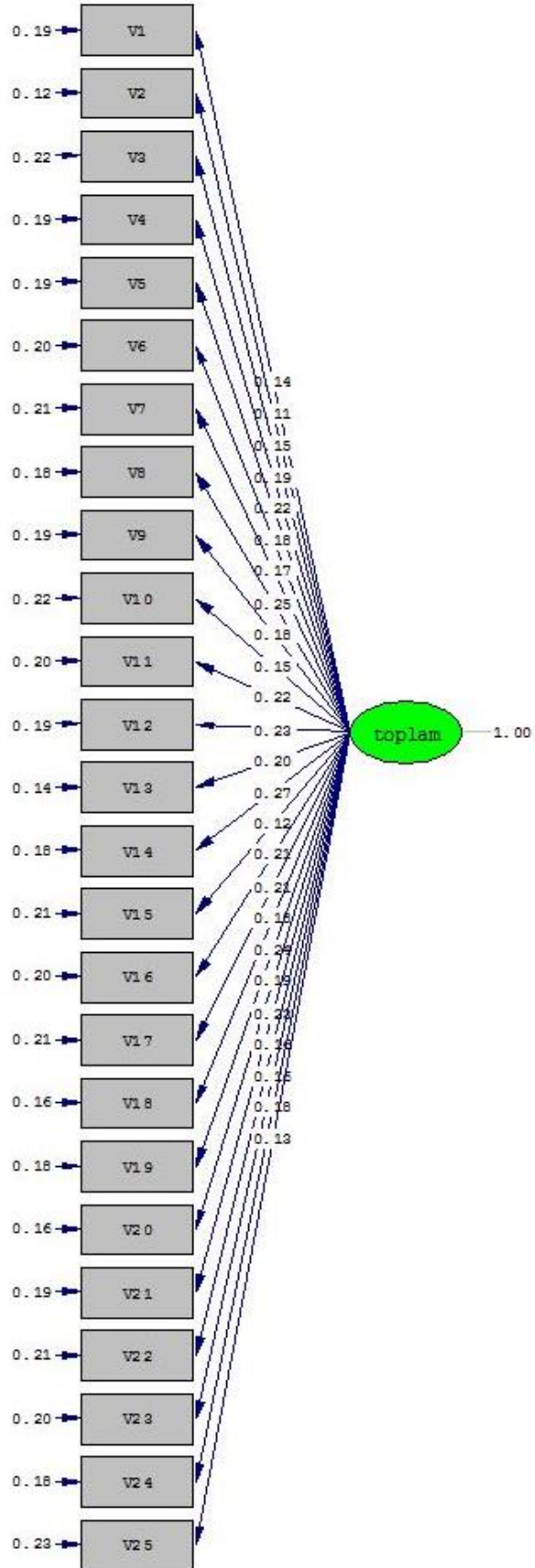
Uzmanlar	Maddeler																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
u1	1	1	0	1	1	0	0	0	1	0	0	1	0	1	0	0	0	0	1	1	0	0	0	1	1	
u2	1	1	1	1	1	0	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1
u3	1	1	0	1	1	0	1	0	0	0	0	1	1	1	0	1	1	0	1	1	0	0	1	1	1	
u4	1	1	1	1	1	0	0	0	1	0	0	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1
u5	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0	1	1	0	1	1	0	0	1	1	1	
u6	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	0	0	1	1	
u7	1	1	1	1	1	0	1	1	1	0	0	0	1	1	0	1	1	0	1	1	0	0	0	1	1	
u8	1	1	0	1	1	0	0	0	1	0	0	1	1	1	1	0	1	0	0	1	1	0	0	1	1	
u9	1	1	1	1	1	0	0	1	1	0	1	1	0	1	0	0	1	0	0	0	0	0	0	1	1	
u10	0	1	0	1	1	1	0	1	0	1	1	1	0	1	1	0	1	1	1	1	0	0	0	1	1	
u11	1	1	1	1	1	0	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1
u12	1	1	1	1	1	1	1	0	1	0	0	1	0	1	1	1	1	1	1	1	0	0	0	1	1	
u13	1	1	0	1	1	0	1	1	1	0	1	1	0	1	0	0	0	0	0	1	1	0	0	1	1	
u14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	
u15	1	1	0	1	0	0	1	0	1	0	1	1	0	0	1	0	0	0	0	1	0	0	0	1	1	
u16	1	1	1	1	1	0	0	1	0	0	0	1	0	1	0	0	0	0	0	1	0	0	1	1	0	
u17	0	1	0	1	1	0	1	1	1	0	0	0	0	1	1	0	1	0	0	1	1	0	0	1	1	
u18	0	1	0	1	1	0	1	1	1	0	1	1	0	1	1	0	1	0	0	1	1	0	1	1	1	
u19	1	1	0	1	1	0	0	1	1	0	0	0	1	1	0	0	1	0	0	1	0	0	0	1	1	
u20	0	0	0	1	0	0	1	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	1	1	
u21	1	1	0	1	1	0	0	0	0	0	0	0	0	1	1	1	1	0	0	1	0	0	0	1	1	
u22	0	1	0	1	1	0	0	0	0	0	0	0	1	0	1	0	1	1	0	1	1	1	1	1	1	
u23	0	1	1	1	1	0	0	0	1	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	
u24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1	1	0	0	0	1	
u25	1	1	0	1	1	0	0	0	0	0	1	1	0	1	0	0	1	0	1	1	0	0	0	1	1	
u26	1	1	0	1	1	0	1	1	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	1	
u27	0	1	0	1	1	1	0	1	0	1	0	1	0	1	0	1	1	0	0	1	1	1	1	1	1	
u28	1	1	0	1	1	0	1	1	1	1	0	1	1	0	0	0	1	0	1	1	1	1	1	1	1	
u29	1	1	1	1	1	1	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	
u30	1	1	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	

EK 8. İŞARETLEME YÖNTEMİ İÇİN KESTİRİLEN MADDE PARAMETRELERİ

<i>ITEM</i>	<i>INTERCEPT</i>	<i>SLOPE</i>	<i>THRESHOLD</i>	<i>LOADING</i>	<i>ASYMPTOTE</i>	<i>CHISQ</i>	<i>DF</i>
		<i>(a)</i>	<i>(b)</i>		<i>(c)</i>		
	<i>S.E.</i>	<i>S.E.</i>	<i>S.E.</i>	<i>S.E.</i>	<i>S.E.</i>	<i>(PROB)</i>	
ITEM0001	0.605 0.067*	0.530 0.088*	-1.143 0.181*	0.468 0.078*	0.000 0.000*	8.2 (0.4159)	8.0
ITEM0002	1.215 0.095*	0.753 0.112*	-1.614 0.197*	0.602 0.089*	0.000 0.000*	18.8 (0.0021)	5.0
ITEM0003	-0.181 0.056*	0.428 0.066*	0.424 0.147*	0.393 0.061*	0.000 0.000*	11.2 (0.1910)	8.0
ITEM0004	0.544 0.070*	0.711 0.100*	-0.764 0.107*	0.580 0.082*	0.000 0.000*	7.9 (0.3456)	7.0
ITEM0005	0.427 0.069*	0.765 0.100*	-0.558 0.087*	0.608 0.079*	0.000 0.000*	9.7 (0.2054)	7.0
ITEM0006	-0.313 0.059*	0.519 0.071*	0.604 0.141*	0.461 0.063*	0.000 0.000*	8.4 (0.4980)	9.0
ITEM0007	0.253 0.060*	0.522 0.080*	-0.485 0.117*	0.463 0.071*	0.000 0.000*	5.0 (0.6611)	7.0
ITEM0008	0.311 0.072*	0.909 0.112*	-0.342 0.071*	0.673 0.083*	0.000 0.000*	18.1 (0.0116)	7.0
ITEM0009	0.498 0.066*	0.645 0.094*	-0.772 0.119*	0.542 0.079*	0.000 0.000*	14.6 (0.0683)	8.0
ITEM0010	-0.192 0.057*	0.428 0.066*	0.449 0.157*	0.393 0.061*	0.000 0.000*	16.1 (0.0658)	9.0
ITEM0011	-0.040 0.062*	0.719 0.101*	0.056 0.089*	0.584 0.082*	0.000 0.000*	7.0 (0.4314)	7.0
ITEM0012	0.192 0.065*	0.758 0.092*	-0.253 0.081*	0.604 0.073*	0.000 0.000*	22.0 (0.0025)	7.0
ITEM0013	-0.877 0.078*	0.757 0.089*	1.159 0.142*	0.603 0.071*	0.000 0.000*	15.8 (0.0455)	8.0
ITEM0014	-0.034	0.893	0.038	0.666	0.000	14.5	7.0

	0.065*	0.109*	0.074*	0.081*	0.000*	(0.0425)	
ITEM0015	-0.401	0.371	1.081	0.348	0.000	7.5	8.0
	0.058*	0.060*	0.227*	0.056*	0.000*	(0.4884)	
ITEM0016	-0.161	0.635	0.254	0.536	0.000	11.9	8.0
	0.060*	0.086*	0.107*	0.073*	0.000*	(0.1559)	
ITEM0017	-0.098	0.611	0.161	0.521	0.000	15.2	8.0
	0.059*	0.078*	0.099*	0.066*	0.000*	(0.0547)	
ITEM0018	-0.766	0.623	1.229	0.529	0.000	10.4	8.0
	0.070*	0.078*	0.174*	0.066*	0.000*	(0.2381)	
ITEM0019	-0.326	0.763	0.427	0.607	0.000	8.5	6.0
	0.064*	0.099*	0.106*	0.078*	0.000*	(0.2041)	
ITEM0020	0.821	0.891	-0.922	0.665	0.000	11.8	6.0
	0.087*	0.122*	0.101*	0.091*	0.000*	(0.0665)	
ITEM0021	-0.069	0.737	0.094	0.593	0.000	7.1	7.0
	0.063*	0.095*	0.089*	0.076*	0.000*	(0.4146)	
ITEM0022	-0.321	0.472	0.680	0.427	0.000	8.4	8.0
	0.058*	0.072*	0.162*	0.065*	0.000*	(0.3984)	
ITEM0023	-0.415	0.485	0.855	0.437	0.000	10.8	8.0
	0.060*	0.072*	0.177*	0.065*	0.000*	(0.2108)	
ITEM0024	0.649	0.733	-0.886	0.591	0.000	7.7	7.0
	0.076*	0.106*	0.114*	0.086*	0.000*	(0.3618)	
ITEM0025	0.083	0.400	-0.207	0.372	0.000	8.9	8.0
	0.055*	0.065*	0.140*	0.061*	0.000*	(0.3498)	

EK 9. DOĞRULAYICI FAKTÖR ANALİZİ ÖLÇME MODELİ



Chi-Square=628.11, df=275, P-value=0.00000, RMSEA=0.050

EK 10. ORJİNALLİK RAPORU

https://www.turnitin.com/dv?s=1&o=551348973&u=1040241844&lang=tr&

masler master tez - TESLİM TARİHİ25-Haz-2015 Roadmap 1 / 1

Originality GradeMark PeerMark YES/NO, EBEL VE İŞARETLEME STANDART BELİRLEME YÖNTEMLERİNİN SEVDA KORKMAZ TARAFINDAN turnitin %6 BENZER 0 ÜZERİNDEN

Eşleşmeyi Gözden Geçir

1	www.edam.com.tr	Internet kaynağı	%3
2	www.slideshare.net	Internet kaynağı	%1
3	katalog.hacettepe.edu.tr	Internet kaynağı	%<1
4	egitimvebilim.ted.org.tr	Internet kaynağı	%<1
5	pegem.net	Internet kaynağı	%<1
6	www.nuraysenemoglu...	Internet kaynağı	%<1
7	ilkogretim-online.org.tr	Internet kaynağı	%<1
8	de.slideshare.net	Internet kaynağı	%<1

YES/NO, EBEL VE İŞARETLEME STANDART
BELİRLEME YÖNTEMLERİNİN KARŞILAŞTIRILMASI

COMPARING OF YES/NO, EBEL AND BOOKMARK
STANDARD SETTING METHODS

Sevda KORKMAZ

SAYFA: 1 / 58 Salt-Metin Raporu

ÖZGEÇMİŞ

Kişisel Bilgiler

<i>Adı Soyadı</i>	Sevda KORKMAZ
<i>Doğum Yeri</i>	Malatya
<i>Doğum Tarihi</i>	13.07.1988

Eğitim Durumu

<i>Lisans</i>	Gazi Üniversitesi- Fen Fakültesi	2011
<i>Yüksek Lisans</i>	Hacettepe Üniversitesi- Eğitimde Ölçme ve Değerlendirme	2015
<i>Yabancı Dil</i>	İngilizce	

İş Deneyimi

<i>Çalıştığı Kurumlar</i>	Sebit Eğitim ve Bilgi Teknolojileri A. Ş./Matematik Soru Alan Uzmanı	2013-2015
	TED Genel Merkezi/Ölçme ve Değerlendirme Uzmanı	2015-

İletişim

<i>e-Posta Adresi</i>	sevdadtli@gmail.com

<i>Jüri Tarihi</i>	03.07.2015
--------------------	------------