



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı
Eğitimde Ölçme ve Değerlendirme Programı

ÖRNEKLEM BÜYÜKLÜĞÜNÜN VE MADDE FORMATININ SEKİZİNCİ SINIF
ABİDE TESTLERİNİN EŞİTLENMESİNE ETKİSİ

Barış ÖZGÜRLÜK

Doktora Tezi

Ankara, 2019

Liderlik, arařtırma, inovasyon, kaliteli eęitim ve deęiřim ile

Daha ileriye ... En İyiyeye ...



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı
Eğitimde Ölçme ve Değerlendirme Programı

ÖRNEKLEM BÜYÜKLÜĞÜNÜN VE MADDE FORMATININ SEKİZİNCİ SINIF
ABİDE TESTLERİNİN EŞİTLENMESİNE ETKİSİ

EFFECTS OF SAMPLE SIZE AND ITEM FORMAT ON EIGHTH GRADE
ABİDE TESTS EQUATING

Barış ÖZGÜRLÜK

Doktora Tezi

Ankara, 2019

Kabul ve Onay

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼ne,
Barıř ÖZG¼RL¼K¼¼n hazırladıđı “Örneklem B¼y¼kl¼đ¼n¼n ve Madde Formatının
Sekizinci Sınıf Abide Testlerinin Eřitlenmesine Etkisi” bařlıklı bu alıřma j¼rimiz
tarafından **Eđitim Bilimleri Ana Bilim Dalı, Eđitimde Ölme ve Deđerlendirme
Bilim Dalı’nda Doktora Tezi** olarak kabul edilmiřtir.

J¼ri Bařkanı

Prof. Dr. Selahattin GELBAL

Imza


J¼ri Üyesi (Danıřman)

Prof. Dr. H¼lya KELECIOđLU

Imza


J¼ri Üyesi

Do. Dr. Hakan Yavuz ATAR

Imza


J¼ri Üyesi

Do. Dr. Dilara BAKAN
KALAYCIOđLU

Imza


J¼ri Üyesi

Dr. Öđr. Üyesi K¼bra ATALAY
KABASAKAL

Imza


Bu tez Hacettepe Üniversitesi Lisans¼st¼ Eđitim, Öđretim ve Sınav Yönetmeliđi’nin ilgili maddeleri uyarınca yukarıdaki j¼ri üyeleri tarafından 17 / 06 / 2019 tarihinde uygun gör¼lm¼ř ve Enstit¼ Yönetim Kurulunca / / tarihinde kabul edilmiřtir.

Prof. Dr. Ali Ekber řAHİN
Eđitim Bilimleri Enstitüsü M¼d¼r¼

Öz

Bu çalışmada denk olmayan gruplar ortak madde deseninde gerçek puan (Ger_Pn) ve gözlenen puan ($Göz_Pn$) eşitleme yöntemiyle yapılan eşitlemede aynı dağılım özelliğinin (ADÖ) ve eşitlik özelliğinin korunumu incelenmiştir. Araştırma 2016 yılı Sekizinci Sınıf Akademik Becerilerin İzlenmesi ve Değerlendirilmesi (ABİDE) araştırmasına katılan yaklaşık 34000 öğrenciden seçkisiz yöntemle seçilen 1000, 3000 ve 5000 kişilik örneklemeler üzerinden yapılmıştır. ABİDE araştırmasında yer alan Matematik ve Sosyal Bilgiler testinin A ve C formlarının B formuna eşitlenmesinde ADÖ ve eşitlik özelliğinin korunumu; örneklem büyüklüğü, madde formatı, testi alan alt gruplar, gerçek/gözlenen puan eşitleme ve ölçek dönüştürme yöntemleri (Ortalama-Ortalama, Ortalama-Standart Sapma, Haebara, Stocking-Lord) üzerinden değerlendirilmiştir. ABİDE testlerinin karma formatta olmasından dolayı madde formatının etkisini değerlendirebilmek için ikiden fazla kategorisi olan maddelerden alınan en büyük puana 1, kısmi puanlara 0 puan verilerek testler iki kategorili formata dönüştürülmüştür. Bu dönüştürme işleminin ardından karma formatın ve iki kategorili formatın ADÖ, birinci sıra eşitlik (BSE) ve ikinci sıra eşitlik (İSE) üzerindeki etkisi karşılaştırılmıştır. Matematik ve Sosyal Bilgiler testinde; (i) Cinsiyetin tek yönlü olarak ADÖ, BSE ve İSE üzerinde etkili olmadığı. (ii) Karakteristik eğri yöntemlerinin daha kararlı sonuçlar ürettiği, ancak en küçük BSE ve İSE değerlerinin bazı durumlarda moment yöntemlerde elde edildiği. (iii) Eşitleme öncesinde formların birikimli dağılım fonksiyonlarının (bdf) benzer ya da farklı olmasının dönüştürme sonrası ADÖ'nün korunumuna etki ettiği. (iv) Format etkisinin eşitlenen testlere göre değiştiği. (v) $Göz_Pn$ ve Ger_Pn eşitlemenin ADÖ, BSE ve İSE'nin korunumuna etkisinin testlere göre değiştiği sonucuna ulaşılmıştır.

Anahtar sözcükler: ABİDE, denk olmayan gruplar ortak madde deseni, aynı dağılım özelliği, birinci sıra eşitlik, ikinci sıra eşitlik

Abstract

In this study, preservation of same distributions property (SDP) and equity properties in non-equivalent groups common-item design was investigated for true score (*True_Sc*) and observed score (*Obs_Sc*) equating. Samples of 1000, 3000 and 5000 participants selected randomly from 34000 students having participating in 2016 Eighth Grade Monitoring and Evaluating Academic Skills (MEAS) research were used for analysis. In equating A and C forms of Math and Social Studies tests to B form in MEAS research, preservation of SDP and equity properties was evaluated according to sample size, item format, sub groups of test takers, true/observed score equating and scale transformation methods (Mean-Mean, Mean-Sigma, Haebera, Stocking-Lord). Since MEAS tests were mixed-format, to evaluate the effects of item format, tests were transformed into dichotomous format by giving 1 point for maximum score and 0 point to the others. After transformation, effects of mixed-format and dichotomous format on SDP, first-order equity (FOE) and second-order equity (SOE) values were compared. The results obtained were as follows: In Math and Social Studies tests; (i) Sex of student wasn't unidirectionally effective for preservation of SDP, FOE and SOE. (ii) Characteristic curve methods generated mostly stable results, but the smallest FOE and SOE values were estimated in moment methods in some cases. (iii) Whether cumulative distribution functions (*cdf*) of the forms were similar or different prior to equating affected preservation of SDP after equating. (iv) Format affected tests differently. (v) Effect of *Obs_Sc* and *True_Sc* equating on preservation of SDP, FOE and SOE changed according to tests.

Keywords: ABIDE, common item non-equivalent design, same distributions property, first-order equity, second-order equity

Teşekkür

Doktoraya başlamam konusunda tavsiyede bulunan Genel Müdürüm Prof. Dr. Bayram ÇETİN'e, derslere katılımım için gerekli kolaylığı sağlayan Daire Başkanım Kemal BÜLBÜL'e, değerli görüş, düşünce ve deneyimlerini paylaşan Danışman Hocam Prof. Dr. Hülya KELECİOĞLU'na, derslerini almaktan mutluluk duyduğum Hocalarıma, Tez Jürimde bulunan Prof. Dr. Selahattin GELBAL'a, Doç. Dr. Hakan YAVUZ ATAR'a, Doç. Dr. Dilara BAKAN KALAYCIOĞLU'na ve Dr. Öğr. Üyesi Kübra ATALAY KABASAKAL'a, yardım ve desteğini sürekli hissettiğim arkadaşım Araştırma Görevlisi Abdullah Faruk KILIÇ'a, her zaman yanımda olan ve beni destekleyen eşim Gülfem ÖZGÜRLÜK'e teşekkür ederim.

İçindekiler

Öz.....	ii
Abstract.....	iii
Teşekkür.....	iv
Tablolar Dizini.....	ix
Şekiller Dizini.....	xi
Simgeler ve Kısaltmalar Dizini.....	xii
Bölüm 1 Giriş.....	1
Problem Durumu.....	1
Araştırmanın Amacı ve Önemi.....	5
Araştırma Problemi.....	7
Sayıltılar.....	8
Sınırlılıklar.....	9
Tanımlar.....	9
Bölüm 2 Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar.....	10
Test Eşitleme.....	10
Eşitleme Desenleri (Equating Designs).....	11
Test Eşitleme Yöntemleri.....	13
Madde Tepki Kuramı (MTK).....	13
Madde Tepki Kuramı Modelleri.....	15
Eşitleme Sonuçlarının Değerlendirilmesi.....	21
İlgili Araştırmalar.....	23
Bölüm 3 Yöntem.....	30
Araştırma Yöntemi.....	30
Araştırmanın Evreni ve Örneklemi.....	30
Veri Toplama Süreci.....	31
Veri Toplama Araçları.....	31

Verilerin Analizi	32
Testlerin Betimsel Analizi	36
Testlerin Güvenirlik Analizi	39
Testlerin Analizi	42
Madde Parametrelerinin Kestirimi ve Ölçek Dönüştürme	50
Eşitleme	50
Bölüm 4 Bulgular ve Yorumlar	53
Birinci Alt Probleme Ait Bulgular ve Yorumlar	53
İkinci Alt Probleme Ait Bulgular ve Yorumlar	60
Üçüncü Alt Probleme Ait Bulgular ve Yorumlar	67
Dördüncü Alt Probleme Ait Bulgular ve Yorumlar	73
Bölüm 5 Sonuç, Tartışma ve Öneriler	80
Sonuçlar	80
Madde Formatına İlişkin Sonuçlar	80
Testi Alan Alt Gruplara İlişkin Sonuçlar	82
Örneklem Büyüklüğüne İlişkin Sonuçlar	83
Ölçek Dönüştürme Yöntemlerine İlişkin Sonuçlar	84
Tartışma	85
Öneriler	86
Kaynaklar	89
EK-A: <i>Mplus7</i> Syntax	94
EK-B: Matematik A Formu Yamaç-Birikinti Grafiği (Karma)	95
EK-C: Matematik B Formu Yamaç-Birikinti Grafiği (Karma)	96
EK-Ç: Matematik C Formu Yamaç-Birikinti Grafiği (Karma)	97
EK-D: Matematik A Formu Yamaç-Birikinti Grafiği (İki Kategorili)	98
EK-E: Matematik B Formu Yamaç-Birikinti Grafiği (İki Kategorili)	99
EK-F: Matematik C Formu Yamaç-Birikinti Grafiği (İki Kategorili)	100

EK-G: Sosyal Bilgiler A Formu Yamaç-Birikinti Grafiği (Karma)	101
EK-Ğ: Sosyal Bilgiler B Formu Yamaç-Birikinti Grafiği (Karma)	102
EK-H: Sosyal Bilgiler C Formu Yamaç-Birikinti Grafiği (Karma)	103
EK-I: Sosyal Bilgiler A Formu Yamaç-Birikinti Grafiği (İki Kategorili)	104
EK-İ: Sosyal Bilgiler B Formu Yamaç-Birikinti Grafiği (İki Kategorili)	105
EK-J: Sosyal Bilgiler C Formu Yamaç-Birikinti Grafiği (İki Kategorili)	106
EK-K: Matematik Testi A-B-C Formlarının Ham Puan Dağılım Grafikleri (Karma)	107
EK-L: Matematik Testi A-B-C Formlarının Ham Puan Dağılım Grafikleri (İki Kategorili)	108
EK-M: Sosyal Bilgiler Testi A-B-C Formlarının Ham Puan Dağılım Grafikleri (Karma)	109
EK-N: Sosyal Bilgiler Testi A-B-C Formlarının Ham Puan Dağılım Grafikleri (İki Kategorili)	110
EK-O: Matematik Testi A Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (Karma)	111
EK-Ö: Matematik Testi A Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (İki Kategorili) ...	112
EK-P: Matematik Testi C Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (Karma)	113
EK-R: Matematik Testi C Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (İki Kategorili) ...	114
EK-S: Sosyal Bilgiler Testi A Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (Karma)	115
EK-Ş: Sosyal Bilgiler Testi A Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (İki Kategorili) ...	116
EK-T: Sosyal Bilgiler Testi C Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (Karma)	117
EK-U: Sosyal Bilgiler Testi C Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (İki Kategorili) ...	118

EK-Ü: Verilerin Kullanımına İlişkin MEB'den Alınan İzin Yazısı.....	119
EK-V: Etik Komisyonu Onay Bildirimi	120
EK-Y: Etik Beyanı.....	121
EK-Z: Yüksek Lisans/Doktora Tez Çalışması Orijinallik Raporu.....	122
EK-AA: Thesis/Dissertation Originality Report.....	123
EK-AB: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı	124

Tablolar Dizini

Tablo 1 Çoktan Seçmeli ve Açık Uçlu Maddelerin Karşılaştırılması	3
Tablo 2 Matematik ve Sosyal Bilgiler Testinin Evreni ve Örneklem Bilgileri	31
Tablo 3 Matematik A-B Testinde Yer Alan Madde ve Ortak Madde Sayıları ve Formatları	33
Tablo 4 Matematik C-B Testinde Yer Alan Madde ve Ortak Madde Sayıları ve Formatları	33
Tablo 5 Sosyal Bilgiler A-B Testinde Yer Alan Madde ve Ortak Madde Sayıları ve Formatları	34
Tablo 6 Sosyal Bilgiler C-B Testlerinde Yer Alan Madde ve Ortak Madde Sayıları ve Formatları	35
Tablo 7 Matematik Testine İlişkin Betimsel İstatistikler (Karma Format)	36
Tablo 8 Matematik Testine İlişkin Betimsel İstatistikler (İki Kategorili Format).....	37
Tablo 9 Sosyal Bilgiler Testine İlişkin Betimsel İstatistikler (Karma Format).....	38
Tablo 10 Sosyal Bilgiler Testine İlişkin Betimsel İstatistikler (İki Kategorili Format)	39
Tablo 11 Matematik Testine İlişkin Güvenirlik Değerleri	40
Tablo 12 Sosyal Bilgiler Testine İlişkin Güvenirlik Değerleri	41
Tablo 13 Matematik Testi için Açık Uçlu ve Çoktan Seçmeli Madde Toplam Puanları Arasındaki Korelasyon Değerleri	42
Tablo 14 Sosyal Bilgiler Testi Açık Uçlu ve Çoktan Seçmeli Madde Toplam Puanları Arasındaki Korelasyon Değerleri	43
Tablo 15 Matematik Testine İlişkin Gözlem Sayıları ve Açıklanan Varyans Yüzdeleri (Karma Format)	45
Tablo 16 Matematik Testine İlişkin Gözlem Sayıları ve Açıklanan Varyans Yüzdeleri (İki Kategorili Format)	46
Tablo 17 Sosyal Bilgiler Testine İlişkin Gözlem Sayıları ve Açıklanan Varyans Yüzdeleri (Karma Format)	48
Tablo 18 Sosyal Bilgiler Testine İlişkin Gözlem Sayıları ve Açıklanan Varyans Yüzdeleri (İki Kategorili Format)	49
Tablo 19 Matematik Testi A Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi - Ham Puan ve Ölçek Puan Dağılımları için Kolmogorov T İstatistikleri (%).....	54

Tablo 20 Matematik Testinde A Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan BSE Değerleri.....	54
Tablo 21 Matematik Testinde A Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan İSE Değerleri.....	55
Tablo 22 Matematik Testi C Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi - Ham Puan ve Ölçek Puan Dağılımları için Kolmogorov T İstatistikleri (%).....	61
Tablo 23 Matematik Testinde C Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan BSE Değerleri.....	62
Tablo 24 Matematik Testinde C Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan İSE Değerleri.....	62
Tablo 25 Sosyal Bilgiler Testi A Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi - Ham Puan ve Ölçek Puan Dağılımları için Kolmogorov T İstatistikleri (%).....	68
Tablo 26 Sosyal Bilgiler Testinde A Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan BSE Değerleri.....	68
Tablo 27 Sosyal Bilgiler Testinde A Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan İSE Değerleri.....	69
Tablo 28 Sosyal Bilgiler Testi C Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi - Ham Puan ve Ölçek Puan Dağılımları için Kolmogorov T İstatistikleri (%).....	74
Tablo 29 Sosyal Bilgiler Testinde C Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan BSE Değerleri.....	75
Tablo 30 Sosyal Bilgiler Testinde C Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan İSE Değerleri.....	76

Şekiller Dizini

Şekil 1. Sekizinci sınıf ABİDE test deseni.....	6
Şekil 2. Denk olmayan gruplarda test düzeni (Kolen ve Brennan'dan (2014) uyarlanmıştır.)	12

Simgeler ve Kısaltmalar Dizini

- ABİDE** : Akademik Becerilerin İzlenmesi ve Değerlendirilmesi
- ADÖ** : Aynı Dağılım Özelliği
- bdf** : Birikimli Dağılım Fonksiyonu
- BSE** : Birinci Sıra Eşitlik
- Ger_Pn** : Gerçek Puan
- Göz_Pn** : Gözlenen Puan
- HB** : Haebara
- İSE** : İkinci Sıra Eşitlik
- KTK** : Klasik Test Kuramı
- MTK** : Madde Tepki Kuramı
- O-O** : Ortalama-Ortalama
- O-S** : Ortalama-Standart Sapma
- SL** : Stocking-Lord

Bölüm 1

Giriş

Bu bölümde; problem durumu, araştırmanın amacı ve önemi, problem cümlesi, sayılılar, sınırlılıklar ve tanımlar yer almaktadır.

Problem Durumu

Eğitim sisteminde öğrencilere aktarılan bilgilerin öğrenciler tarafından hangi düzeyde kazanıldığı genellikle yazılı sınavlar yoluyla belirlenir. Diğer ülkelerde olduğu gibi Türkiye’de de öğrenci başarıları çeşitli zamanlarda ve farklı formatlarda yapılan sınavlarla ölçülmektedir. Dünya genelinde yapılan sınavlarda kimi zaman çoktan seçmeli, kimi zaman açık uçlu (yapılandırılmış cevaplı), kimi zamansa hem açık uçlu hem de çoktan seçmeli maddeler kullanılmaktadır. Türkiye’de uygulanan sınavlarda ise çoğunlukla çoktan seçmeli maddeler tercih edilmektedir. Bu tercihin en belirgin nedeni çoktan seçmeli maddelerin hazırlanmasının ve puanlanmasının kolay olmasıdır. Okul veya sınıf temelli yapılan Türkçe ve sosyal bilgiler sınavında öğrencilerin bir konu hakkında bildiklerini sınav kâğıdına yazması, matematik ve fen bilimleri sınavında problem çözmesi, fiziksel veya kimyasal bir olayı açıklamasının öğretmenlerce değerlendirilmesinin çok zor olmayacağı aşikârdır. Ancak katılımcı sayısının çok olmasından dolayı geniş ölçekli sınavların yukarıda belirtildiği şekilde yapılıp değerlendirilmesi pek de mümkün değildir.

Türkiye’de geniş ölçekli başarı testlerine yönelik çalışmalar 1992 yılında Milli Eğitim Bakanlığı (MEB) bünyesinde kurulan Eğitimi Araştırma ve Geliştirme Dairesi Başkanlığı (EARGED) tarafından başlatılmış ve 1990’ların sonlarında somut örnekler ortaya çıkmıştır. Türkçe, sosyal bilgiler, fen bilgisi ve matematik alanlarında 5, 6, 8, 9 ve 11. sınıf düzeylerinde yapılan durum tespit sınavlarında çoktan seçmeli ve açık uçlu maddeler kullanılmıştır (Çakan, 2003).

Açık uçlu maddeler yapısal olarak çoktan seçmeli maddelerden farklıdır. Açık uçlu maddeleri içeren sınavlarda cevaplayıcı cevabı düşünür, bulur, tasarlar ve düzenler (Bağcan Büyükturan & Çıkrıkçı Demirtaşlı, 2012). Açık uçlu maddelerde, rubrik kullanılarak öğrencinin bilgisine göre kısmi puan alması sağlanabilirken çoktan seçmeli maddelerde böyle bir durum söz konusu olamamaktadır. Çoktan seçmeli maddelerin genellikle daha güvenilir puanlama sürecine sahip olduğu

düşünülür (Lissitz & Hou, 2012). Çeldiricileri çok iyi hazırlanan maddeler dışında, çoktan seçmeli maddelerin şansla doğru cevaplanma olasılığı istatistiksel olarak seçenek sayısına göre değişirken açık uçlu maddelerde doğru cevabı şansla bulmak neredeyse imkânsızdır.

Livingston (2009) testlerde çoktan seçmeli maddeler yerine neden yapılandırılmış cevaplı maddelerin kullanılması gerektiğini aşağıdaki gibi özetlemiştir:

- ✓ Okullarda öğretilen bilgiler çoktan seçmeli sorularla etkili bir biçimde ölçülemeyecek kadar karmaşıktır.
- ✓ Tarih öğrencileri için hazırlanan çoktan seçmeli bir test öğrencilerin gerçek bilgilerini test edebilir ve öğrencilerin gerçek olaylar arasındaki ilişkileri doğru ve yanlış ayırt edip etmediğini belirleyebilir. Ancak, öğrencilerin tarihle ilgili bir konu üzerinde iyi kurgulanmış bir paragraf yazıp yazamayacağını belirleyemez.
- ✓ Matematik dersi için hazırlanan çoktan seçmeli bir test öğrencilerin birçok problemi çözüp çözemeyeceğini belirleyebilir. Ancak, öğrencilerin matematikle ilgili bir kanıtı yapılandırıp yapılandıramayacağını belirleyemez.
- ✓ Yazma becerisi için hazırlanan çoktan seçmeli bir test öğrencilerin iyi veya kötü yazılan bir cümleyi ayırt edip edemeyeceğini belirleyebilir. Ancak, öğrencilerin kendi düşüncelerini mantıksal bir şekilde organize edip edemeyeceğini belirleyemez.
- ✓ Çoktan seçmeli bir testte doğru cevabı seçenekler içinden belirleyebilen bir öğrenci belki de seçenekleri görmeden doğru cevabı bulamayacaktır. Eğitim açısından bu fark önemli olabilir.
- ✓ Fen bilimleri deneyinde doğru süreci hatırlayamayan bir öğrenci sonraki adımda veya adım dizisinde doğru süreci fark edebilir.
- ✓ Bir metindeki mantıksal akışı açıklayamayan öğrenci, dört ya da beş olasılıktan biri verildiğinde akışı açıklayabilir.

Çoktan seçmeli maddelerde çeldiricileri iyi şekilde hazırlamak yukarıda anlatılan sınırlılıkların üstesinden gelemeyen ve doğru cevabı bilen öğrencilerin doğru cevaba çok yakın bir çeldiriciyi seçerek yanlış cevap vermesine neden olabilir.

Açık uçlu maddelerin, çoktan seçmeli maddelere göre avantajları yanında bazı dezavantajları da göz ardı edilmemelidir. Bu dezavantajlardan bazıları; rubriklerin dikkatsiz ve/veya özensiz hazırlanması sonucu oluşan sorunlar, puanlayıcılardan kaynaklı sorunlar ve maddelerin kapsam geçerliğini tam sağlayamaması sorunlarıdır. Puanlayıcıdan kaynaklı sorunlardan biri olan iki puanlayıcı arasındaki uyumsuzluk durumu üçüncü puanlayıcının üst değerlendirici olarak maddeyi puanlamasıyla bir noktaya kadar çözümlenebilmektedir. Çoktan seçmeli maddeler ile açık uçlu maddelerin karşılaştırmalı gösterimi, psikometrik ve bilişsel perspektifi Tablo 1’de verilmiştir (Cao, 2008).

Tablo 1

Çoktan Seçmeli ve Açık Uçlu Maddelerin Karşılaştırılması

	Çoktan Seçmeli Madde	Açık Uçlu Madde
<i><u>Psikometrik Perspektif</u></i>		
Birim zamandaki kapsam örnekleme	Geniş (+)	Dar
Güvenirlilik	Genel olarak yüksek (+)	Genel olarak düşük
Puanlama	Objektif, etkili ve ucuz (+)	Subjektif, zaman alıcı ve pahalı
Geçerlik	Genel olarak düşük	Genel olarak yüksek (+)
Doğru cevabı tahminle bulma olasılığı	Yüksek	Düşük (+)
<i><u>Bilişsel Perspektif</u></i>		
Bilişsel beceriler	Düşük veya yüksek düzey	Düşük veya yüksek düzey
Bilişsel aralık	Dar	Geniş (+)
Öğrenme ve öğretmeye etkisi	Genel olarak düşük	Genel olarak yüksek (+)

Not: (+) ile gösterilen özellikler genel olarak istenen durumlardır.

Tablo 1 incelendiğinde psikometrik olarak çoktan seçmeli maddelerin; kapsam, güvenilirlik ve puanlama, açık uçlu maddelerin ise geçerlik ve doğru cevabı tahminle bulma olasılığının düşük olması açısından etkili oldukları görülmektedir. Bilişsel olarak her iki madde formatının da düşük veya yüksek düzey bilişsel becerileri ölçebileceği, bilişsel aralık ile öğrenme ve öğretmeye etki açısından ise açık uçlu maddelerin daha etkili olduğu söylenebilir.

Bazı sınavlarda açık uçlu ve çoktan seçmeli maddeler bir arada kullanılabilir. Farklı madde formatlarını birlikte kullanan testler, tek bir format kullanan testlerden daha geniş yetenek aralıklarının ölçülmesine izin verir (Kim & Lee, 2004). Açık uçlu ve çoktan seçmeli maddeleri birlikte içeren testler, karma format (mixed-format ya da mixed-type) testler olarak adlandırılmaktadır. Karma format testlerin kullanılmasıyla yalnızca çoktan seçmeli veya yalnızca açık uçlu maddeleri kullanan testlerin bazı dezavantajları ortadan kaldırılabilir.

Kimi zaman güvenlik kimi zaman ise daha çok maddeye ilişkin parametre elde etmek için testlerin farklı formları sınavlarda uygulanmaktadır. Aynı testin yalnızca maddelerinin yerini değiştirerek farklı formlar oluşturmak ölçme açısından çok fazla sorun yaratmazken benzer özelliğe sahip olduğu düşünülen ancak kısmen farklı maddelerden oluşan testleri öğrencilere uygulamak sınav öncesi ve sonrası bazı düzenlemeler yapılmasını gerektirmektedir. Çoktan seçmeli maddeleri içeren testlerin A ve B gibi iki formu genellikle aynı sayıda maddeden oluşturulur ve bu şekilde öğrenciler farklı formları alsalar bile ortalama doğru sayılarının yaklaşık olarak aynı olması sağlanmaya çalışılır. Ancak, maddeler her ne kadar benzer ayırt edicilik ve güçlükte yazılmaya çalışılsa da formların güçlük ve standart sapmalarında farklılıklar oluşur. Karma format testlerde açık uçlu maddelerden alınabilecek en büyük puanların farklılaşması ve öğrencilerin açık uçlu maddelere karşı davranışlarının farklılaşması formların ortalama ve standart sapmalarında daha büyük değişimlere neden olabilmektedir.

Testlerin farklı formlarının uygulanması bu testlerin eşitlenmesi zorunluluğunu da beraberinde getirmektedir. Literatür incelendiğinde denk olmayan gruplar ortak madde deseninde test eşitlenmesinde kullanılması gereken ortak madde sayısına ilişkin birbirinden farklı görüşler olduğu görülmektedir. Ortak madde sayısı için temel kural, ortak madde sayısının testteki toplam madde sayısının %20-25'i olmasıdır (Hambleton, Swaminathan, & Rogers, 1991). Deneyimlere göre test

40 ve daha fazla maddeden oluşuyorsa en az %20, test çok uzun değilse 30 ortak madde yeterli olabilir (Kolen & Brennan, 2014). Budescu (1985), güvenilirlik ve daha az eşitleme hatasına neden olacağı için olabildiğince çok sayıda ortak maddenin testlerde yer almasının daha çok istenen durum olduğunu belirtmiştir.

Yeterli sayıda ortak madde ile yapılacak eşitleme öncesinde testlerin eşitleme koşullarını sağlaması gerekmektedir. Lord (1980) eşitliğin sağlanabilmesi için hangi testin alındığının önemli olmaması ve eşitliğin testlerin eşit ya da aynı güvenilirliğe sahip olması durumunda gerçekleşeceğini belirtmiştir. Testlerin eşit olması durumunda zaten eşitlemeye gerek kalmayacaktır. Bu nedenle daha az kısıtlayıcı olan yöntemlerle testlerin eşitleme sonuçlarının değerlendirilmesi gerekmektedir. Bu yöntemler, eşitleme sonrası koşullu ortalamayı dikkate alan birinci sıra eşitlik (BSE) ve koşullu standart hatayı dikkate alan ikinci sıra eşitlik (İSE) yöntemleridir. Brennan (2010) Center for Advanced Studies in Measurement and Assessment (CASMA)'nın 30 numaralı raporunda doğrusal eşitlemeyi BSE ve İSE temelinde gerçek ve gözlenen puan üzerinden, eğrisel eşitlemeyi BSE ve İSE temelinde kuadratik eşitleme üzerinden inceleyerek detaylı bilgiler sunmuştur. Lee, Lee ve Brennan (2010) CASMA 30 raporunda yer alan çalışmanın bir benzerini Tong ve Kolen (2005) tarafından oluşturulan BSE ve İSE formüllerini geliştirerek CASMA 31 raporunda yayınlanmıştır. Bu araştırmada da ABİDE Matematik ve Sosyal Bilgiler testinin eşitlik özelliği literatürde yer alan çalışmalardan esinlenilerek incelenmiştir.

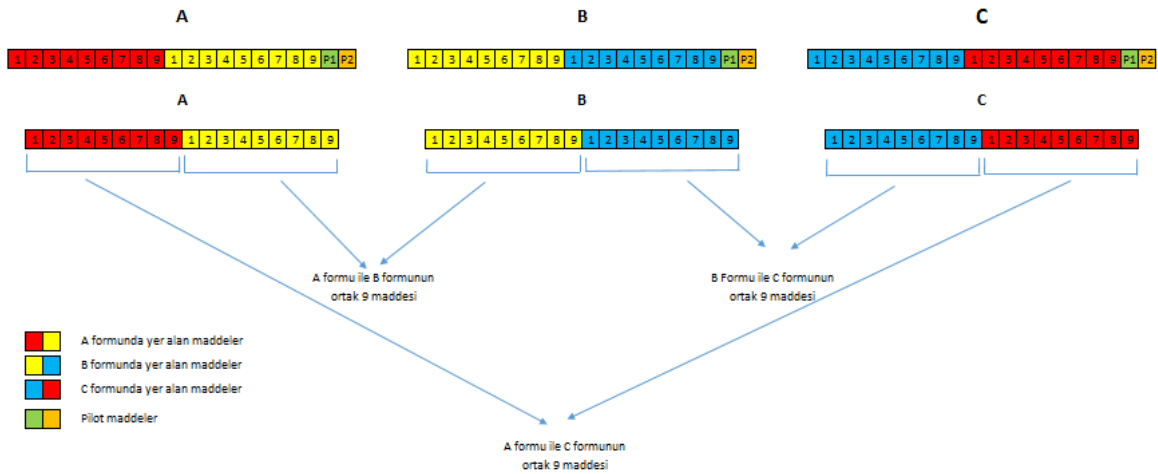
Araştırmanın Amacı ve Önemi

Bu araştırmanın amacı karma ve iki kategorili format testlerin eşitlenmesinde; örneklem büyüklüğünün, cinsiyetin, gerçek/gözlenen puan eşitleme yönteminin ve ölçek dönüştürme yönteminin aynı dağılım özelliğinin ve eşitlik özelliğinin korunumuna etkisini incelemektir.

Sınav hazırlayanlar bazen sınav güvenliği bazen de daha fazla maddeyi denemek için sınavlarda öğrencilere farklı formlar verebilmektedir. Bu formlarda benzer sorular öğrencilere yöneltirse de formların ve maddelerin güçlük ve ayırt edicilikleri farklılaşabilmektedir. Bu nedenle farklı maddeleri içeren formların uygun eşitleme yöntemleri kullanılarak eşitlenmesi gerekmektedir. Ancak, karma format maddeleri içeren testlerden elde edilecek en büyük puanların farklılaşması

testlerden elde edilen doğru sayılarına da etki edebilmektedir. Bu amaçla bu çalışmada veriler hem karma hem de iki kategori üzerinden değerlendirilerek aynı dağılım özelliği ve eşitlik özelliğinin korunumu incelenmiştir.

Açık uçlu maddelerle çoktan seçmeli maddeleri aynı test içerisinde kullanan ve çok sayıda bireyle yapılan uygulamaları ele alan çalışmalar Türkiye’de oldukça azdır. Ancak, MEB tarafından 2016 yılında birincisi yapılan ABİDE araştırması bu çeşit çalışmalar için araştırmacılara veri kaynağı sunmaktadır. ABİDE araştırmasında öğrencilere Türkçe, Matematik, Fen Bilimleri ve Sosyal Bilgiler testi uygulanmıştır. Bu çalışma kapsamında ise Matematik ve Sosyal Bilgiler testleri incelenmiştir. Biri sayısal biri sözel alanda yer alan bu iki test 81 ilde yaklaşık 34000 kişiye uygulanmıştır. Her bir test kendi içerisinde ana kitapçıklar olan A, B ve C formlarından oluşmuştur. Yine her bir ana kitapçığa bir açık uçlu bir de çoktan seçmeli pilot madde eklenerek test formuyla uyumlu olarak A testi için A1, A2, A3, A4; B testi için B1, B2, B3, B4 ve C testi için C1, C2, C3, C4 formları oluşturulmuştur. Farklı formların uygulanmasıyla her bir alt test için 24 pilot maddenin de parametre değerleri elde edilerek maddelerin sonraki döngülerde kullanılabilmesi imkânı sağlanmıştır. Formlar arasındaki ilişki Şekil 1’de yer almaktadır.



Şekil 1. Sekizinci sınıf ABİDE test deseni

Şekil 1 incelendiğinde pilot maddeler dışında ABİDE araştırmasında her bir ders için 27 maddenin uygulandığı ve her test formunda pilot maddeler dışında 18 madde yer aldığı görülmektedir. Bu test düzeninde testleri birbirine bağlayan 9 madde bulunmaktadır. Bu testleri A formu, B formu veya C formu üzerinden birbirine

bağlamak mümkündür ve testler B formu üzerinden birbirine bağlandığında B formundaki tüm bilgi kullanılmaktadır.

Bu çalışmada gerçek uygulamadan elde edilen verilerin kullanılması, madde formatı, farklı örneklem büyüklükleri, cinsiyet, ölçek dönüştürme ve gerçek/gözlenen puan eşitleme özelliklerinin birlikte ele alınarak aynı dağılım ve eşitlik özelliğinin korunumunun değerlendirilmiş olması bu çalışmayı yapılan diğer çalışmalardan farklı kılmaktadır. Literatürde ölçek dönüştürme yöntemlerinin aynı dağılım ve eşitlik özelliğinin korunumuna etkisine ilişkin çalışmalara rastlanırken örneklem büyüklüğü, cinsiyet ve madde formatlarını birlikte ele alan çalışmalara rastlanamamıştır. Bu nedenle bu değişkenlerin de aynı dağılım ve eşitliğin korunumu açısından etkileri incelenmiştir.

Araştırma Problemi

Bu çalışmanın problem cümlesi ve alt problemleri aşağıda belirtilmiştir.

Denk olmayan gruplar ortak madde deseninde gerçek/gözlenen puan ile yapılan eşitlemede aynı dağılım ve eşitlik özelliğinin korunumu; maddelerin formatına (çok kategorili puanlanan, iki kategorili puanlanan), testi alan alt gruplara (kız, erkek), örneklem büyüklüğüne (1000, 3000 ve 5000) ve ölçek dönüştürme yöntemlerine (O-O, O-S, HB ve SL) göre nasıl değişmektedir?

Alt problemler. Gerçekleştirilen bu çalışmada aşağıda yer alan alt problemlere cevap aranmıştır.

1. Matematik testi için denk olmayan gruplar ortak madde deseninde *gerçek/gözlenen* puan eşitleme yöntemi ile A formunun B formuna eşitlenmesinde aynı dağılım ve eşitlik özelliğinin korunumu,
 - a. *maddelerin formatına* (çok kategorili puanlanan, iki kategorili puanlanan),
 - b. *testi alan alt gruplara* (kız, erkek),
 - c. *örneklem büyüklüğüne* (1000, 3000 ve 5000) ve
 - d. *ölçek dönüştürme yöntemlerine* (O-O, O-S, HB, SL) göre nasıl değişmektedir?

2. Matematik testi için denk olmayan gruplar ortak madde deseninde *gerçek/gözlenen* puan eşitleme yöntemi ile C formunun B formuna eşitlenmesinde aynı dağılım ve eşitlik özelliğinin korunumu,
 - a. maddelerin formatına (çok kategorili puanlanan, iki kategorili puanlanan),
 - b. testi alan alt gruplara (kız, erkek),
 - c. örneklem büyüklüğüne (1000, 3000 ve 5000) ve
 - d. ölçek dönüştürme yöntemlerine (O-O, O-S, HB, SL) göre nasıl değişmektedir?
3. Sosyal Bilgiler testi için denk olmayan gruplar ortak madde deseninde *gerçek/gözlenen* puan eşitleme yöntemi ile A formunun B formuna eşitlenmesinde aynı dağılım ve eşitlik özelliğinin korunumu,
 - a. maddelerin formatına (çok kategorili puanlanan, iki kategorili puanlanan),
 - b. testi alan alt gruplara (kız, erkek),
 - c. örneklem büyüklüğüne (1000, 3000 ve 5000) ve
 - d. ölçek dönüştürme yöntemlerine (O-O, O-S, HB, SL) göre nasıl değişmektedir?
4. Sosyal Bilgiler testi için denk olmayan gruplar ortak madde deseninde *gerçek/gözlenen* puan eşitleme yöntemi ile C formunun B formuna eşitlenmesinde aynı dağılım ve eşitlik özelliğinin korunumu,
 - a. maddelerin formatına (çok kategorili puanlanan, iki kategorili puanlanan),
 - b. testi alan alt gruplara (kız, erkek),
 - c. örneklem büyüklüğüne (1000, 3000 ve 5000) ve
 - d. ölçek dönüştürme yöntemlerine (O-O, O-S, HB, SL) göre nasıl değişmektedir?

Sayıtlılar

ABİDE araştırmasından elde edilen veriler üzerinden yapılan bu çalışmada,

1. tüm bireylerin aynı koşullar altında testi aldıkları,
2. test için verilen sürenin yeterli olduğu ve

3. testlerin hız testi olmadığı varsayılmıştır.

Sınırlılıklar

Bu araştırma 2016 yılında sekizinci sınıf öğrencilerine uygulanan ABİDE araştırmasından elde edilen Matematik ve Sosyal Bilimler testi verileriyle sınırlıdır ve değişen madde fonksiyonuna bakılmamıştır.

Tanımlar

ABİDE: Akademik becerilerin izlenmesi ve değerlendirilmesi çalışmasıdır.

Çoktan Seçmeli Madde: Belirli bir sayıda seçenek içerisinde verilen doğru cevap için "1", yanlış cevap için "0" olarak puanlanan madde türüdür.

Açık Uçlu Madde: Bu çalışmada açık uçlu madde, cevabı yapılandırılmış madde anlamında kullanılmaktadır. Bu maddeler, belirli bir rubriğe göre puanlanan, yanlış ve ilgisiz cevaplar için "0", kısmi ve tamamen doğru cevaplara göre "0" ile madde için belirlenen maksimum puan arasında puanların alınabildiği maddelerdir. Matematik ve Sosyal Bilimler testlerinde açık uçlu maddelerin kategori sayıları Tablo 3, Tablo 4, Tablo 5 ve Tablo 6'dan görülebilir.

Bölüm 2

Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar

Test Eşitleme

Eşitleme test formlarındaki puanları uyarlamak için kullanılan istatistiksel bir yöntemdir ve bu yöntemle formlardan alınan puanlar birbirinin yerine kullanılabilir (Kolen & Brennan, 2014). Benzer içeriğe sahip testlerden farklı zamanlarda alınan puanların yaklaşık olarak aynı olması beklenir. Çoğu zaman testlerin aynı güçlükte hazırlanması amaçlansa da maddelerin güçlük ve ayırt ediciliklerinden kaynaklı testlerden elde edilen puanlarda beklenenin dışında sonuçlar elde edilebilir. Aynı kapsam için hazırlanan birden fazla testten elde edilen puanlar bazı istatistiksel tekniklerle birbirine eşitlenir. Bu eşitleme işleminden sonra her iki testten elde edilen puanların birbirinin yerine kullanılması sağlanmış olur. Örneğin; eşitleme sonrasında, farklı tarihlerde yapılan ve zor olan birinci sınavdan (Y formu) 55 alan bir öğrencinin, yaklaşık aynı içeriğe sahip olan ve daha kolay olan ikinci sınavdan (X formu) 60 alan öğrenciyle aynı başarıda olduğu söylenebilir. Eşitlemede temel alınan Y formu ise bu durumda X formundan 60 alan öğrencinin eşitleme sonrası puanı 55 olacaktır.

Eşitlemenin varsayımları. Testlerin eşitlenebilmesi için bazı özelliklerin sağlanması gerekir. Bu özellikler, Klasik Test Kuramı (KTK)'ndeki gerçek puan ve Madde Tepki Kuramı (MTK)'ndeki örtük özellikler gibi direkt olarak gözlenemeyen değişkenlere odaklanmaktadır (Kolen & Brennan, 2014). Gerçek puan ve örtük özellikler ölçme hatası içermeyen puanlardır. Gerçek puan; test teorisiyle modellenirken, eşitlemenin gözlenen puan özellikleri test teorisi modellerine dayanmaz (Kolen & Brennan, 2014).

Simetri özelliği (Symmetry property). Bu özellik; A formundaki puanı, B form ölçeğine çevirmek için kullanılan fonksiyonun, B formundaki puanı, A form ölçeğine çevirmek için kullanılan fonksiyonun tersi olmasını (Kolen & Brennan, 2014). Bu şekilde A formunda 55 olarak elde edilen puanın B formundaki karşılığı 60 ise, B formundaki 60 olan puanın karşılığı da A formunda 55 olacaktır.

Aynı spesifikasyonlar özelliği (Same specifications property). İki test formu eşitlenmek isteniyorsa, test formları aynı kapsam ve istatistiksel özelliklere

sahip olmak zorundadır (Kolen & Brennan, 2014). Aynı kapsam ve istatistiksel özelliklere sahip formların birbiri yerine kullanılabilmesinde bu özellik önem taşımaktadır.

Eşitlik özelliği (Equity property). Eşitleme yapılacak testlerin ortalama, güçlük ve güvenilirliklerinin eşit olmasıdır.

Gözlenen puan eşitleme özelliği (Observed score equating property). Puan dağılımı özelliklerinin her iki grupta da aynı olmasıdır. Bu özellik gereği, X formunu alan grupta 50 üzeri puan alan kişi oranıyla, Y formunu alan grupta 50 üzeri puan alan kişi oranının aynı olması gerekir.

Grup değişmezlik özelliği (Group invariance property). Eşitleme ilişkilerinin, testi alan kişilerin hangi grupta olduğunu etkilememesidir (Kolen & Brennan, 2014).

Eşitleme Desenleri (Equating Designs)

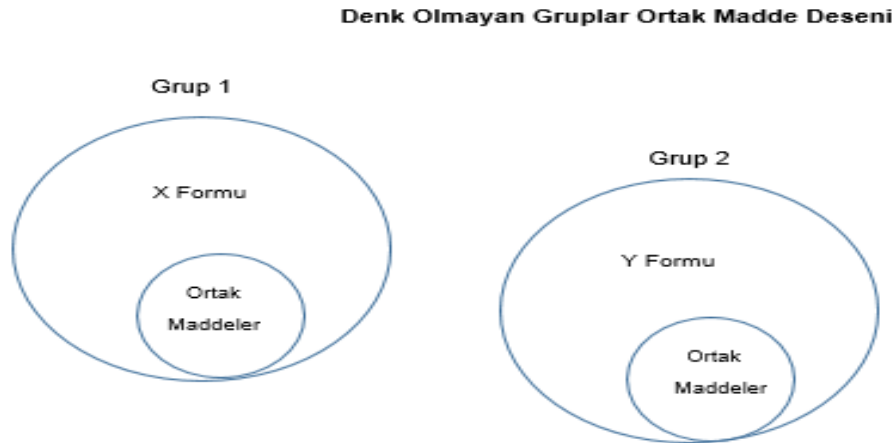
Eşitleme aşamasından önce verinin elde edilmiş durumuna göre değişik eşitleme desenleri kullanılabilir. Desenlerin seçimi hem pratik ve hem de istatistiksel konuları içermektedir (Kolen & Brennan, 2014). Random grup, tek grup ve denk olmayan gruplar ortak madde desenleri daha yaygın kullanımı olan eşitleme desenleridir. Aşağıda bu desenlere ilişkin açıklamalara yer verilmiştir.

Random grup deseni (Random group design). Bu desende testi alacak kişilere uygulanacak formlar rastgele atanır. Sarmal halde ilk sıradaki kişiye X, ikinci sıradaki kişiye Y formu verilir. Bu sarmal işlem herkes bir form aldığı anda sonlandırılır. 50 kişinin sınava girdiği durumda 25 tane X, 25 tane Y formu kullanılır. Formlar rastgele atandığı için iki grup performansları arasındaki fark testlerde yer alan maddelerin güçlüklerine atfedilebilir. Cinsiyet vb. özelliklerin etkilerini ortadan kaldırabilmek için uygulama sırasında formlar bu özelliklerle orantılı olarak dağıtılmalıdır.

Tek grup deseni (Single group design). Bu desende sınava girenlere önce X, ardından Y formu uygulanır. Ancak, yorgunluk etkisinden dolayı Y formunda düşük başarılı görülebilir ya da teste alışmadan kaynaklı Y formu daha kolay gelebilir. Bu nedenlerle pek tercih edilen bir desen değildir.

Dengelenmiş tek grup deseni (Single group design with counterbalancing). Tek grup desenindeki sıra etkisini ortadan kaldırmak için tasarlanmış bir desendir. Hazırlanan test kitapçıklarının yarısında önce X, sonra Y formu (XY), kalanında önce Y, sonra X (YX) formu yer almalıdır. Birinci sırada sınava girene XY formu, ikinci sıradan sınava girene YX formu ve üçüncü sırada sınava girene XY formu gelecek şekilde sarmal olarak en son kişiye test formu verilene kadar form dağıtımı devam eder. Hangi formun önce veya sonra alındığı dikkate alınarak nasıl bir eşitleme yöntemi uygulanacağı araştırmacı tarafından belirlenir.

Denk olmayan gruplar ortak madde deseni (Common-Item nonequivalent groups design). Ortak madde test deseni, maddelerin iyi seçilmesi durumunda eşit grup ve tek grup deseninden kaynaklanan öncelik, sonralık ve yorgunluk gibi problemleri ortadan kaldırabilmektedir (Hambleton, Swaminathan, & Rogers, 1991). Kolen ve Brennan (2014) denk olmayan gruplar ortak madde deseninin genelde her bir testin birden fazla formunun, test güvenliği ve diğer pratik nedenlerle uygulanamadığı zamanlarda kullanıldığını belirtmektedirler. Denk olmayan gruplarda test düzeni Şekil 2’de verilmiştir.



Şekil 2. Denk olmayan gruplarda test düzeni (Kolen ve Brennan’dan (2014) uyarlanmıştır.)

Şekil 2 incelendiğinde ortak maddelerin Grup 1 içerisinde X, Grup 2 içerisinde Y formuyla birlikte uygulandığı görülmektedir. Bu tarz bir yapı içerisinde kullanılan ortak madde sayısı ve oranının test eşitlemeye etkisinin de önemli olacağı gözden

kaçırılmamalıdır. Öyle ki sadece madde sayısının değil maddelerin nasıl puanlandığının (ikili ve çoklu puanlama) da eşitlemeye etkisinin farklı olacağı göz önünde bulundurulmalıdır.

Denk olmayan gruplar ortak madde deseninde ortak maddeler; testle birlikte uygulanıyor ve testi alan kişinin puanında etkili oluyorsa iç (internal), ayrı bir şekilde uygulanıyor ve testi alan kişinin puanında etkili olmuyorsa dış (external) ortak madde olarak nitelendirilir (Kolen & Brennan, 2014).

Test Eşitleme Yöntemleri

Test eşitleme KTK ve MTK'ya dayalı yapılabilmektedir. KTK'ya dayalı yöntemler; ortalama eşitleme, doğrusal eşitleme ve eşit yüzdelikli eşitleme yöntemleridir. MTK'ya dayalı eşitleme yöntemler; gözlenen ve gerçek puan eşitleme yöntemleridir. Bu çalışmada MTK'ya göre eşitleme yapıldığı için yalnızca MTK ve MTK modelleri aşağıda detaylarıyla anlatılmıştır.

Madde Tepki Kuramı (MTK)

MTK, örtük özellikler ve bu özellikler arasındaki ilişkileri tanımlayan modeller sistemidir (de Ayala, 2009). Örtük özellikler teorisi olarak da MTK; yetenek düzeyi kestirimleri, hem bireylerin cevaplarına hem de uygulanan maddelerin özelliklerine bağlı olan model tabanlı ölçmedir (Embretson & Reise, 2000). MTK, gözlenen cevapların kestiricileri olarak maddeleri ve bireylerin örtük özelliklerini kullanır (de Ayala, 2009). MTK aşağıda belirtilen iki temel varsayıma dayanır (Hambleton, Swaminathan, & Rogers, 1991).

- ✓ Testi alan bir kişinin test maddelerindeki performansı; özellik, örtük özellik ya da yetenek adı verilen faktörler kümesi tarafından kestirilebilir.
- ✓ Testi alanların madde performansları ve madde performansları altında yatan özellikler kümesi, madde karakteristik fonksiyonu veya madde karakteristik eğrisi adı verilen monoton artan bir fonksiyonla tanımlanabilir.

Testi alan bireyin yeteneği arttıkça madde karakteristik fonksiyonun yapısı gereğince maddenin doğru cevaplanma olasılığı kişi hangi gruptan gelirse gelsin artacaktır. MTK'da kişiler ve maddeler aynı ölçek üzerine yerleştirilir ve kullanışlılığı olan bir madde ölçek boyunca değişik noktalara yerleşen bireyler arasındaki farkları ortaya koyabilmelidir (de Ayala, 2009). MTK modelleri, KTK modellerinin aksine

yanlışlanabilir özelliğe sahiptir (Hambleton, Swaminathan, & Rogers, 1991). Bu durum model veri uyumuyla ilgilidir. Veri uyumu iyi olan modellerin kullanılması daha sağlıklı sonuçların elde edilmesini sağlar.

MTK'da madde ve yetenek parametreleri değişmezdir (Hambleton, Swaminathan, & Rogers, 1991). Bu nedenle, aynı yeteneğe sahip bireylerin hangi gruptan geldiklerine bakılmaksızın maddeyi doğru cevaplama olasılıkları birbirine eşittir. KTK'da testi alan tüm bireyler için tek bir standart hata değeri hesaplanırken MTK'da her birey için yetenek kestiriminin standart hatası ayrı hesaplanır.

Madde tepki kuramının varsayımları. MTK modellerinin uygulanabilmesi için bazı varsayımların sağlanması gerekir. Varsayımların sağlanamaması durumunda, MTK modellerinden elde edilen sonuçlara temkinli yaklaşmak gerekir. İstenen varsayımların tam anlamıyla sağlanması pek mümkün olmasa da yeterince sağlanması da bazı durumlarda kabul görmektedir. MTK modellerinin sağlanması gereken varsayımlar; tek boyutluluk, yerel bağımsızlık ve tanımlanan madde karakteristik fonksiyonunun, gözlenemeyen değişkenler (yetenekler) ve gözlenen değişkenler (madde cevapları) arasındaki gerçek ilişkileri yansıttığıdır (Hambleton, Swaminathan, & Rogers, 1991; Embretson, & Reise, 2000). DeMars (2010) model veri uyumunu da varsayımlar altında ele almıştır. Bu kapsamda; tek boyutluluk, yerel bağımsızlık, değişmezlik ve veri uyumuna ilişkin açıklamalar başlıklar halinde aşağıda sunulmuştur.

Tek boyutluluk (Unidimensionality). Tek boyutluluk testte yer alan maddelerin tek bir yeteneği ölçmesidir. Sağlanması kolay olan bir varsayım değildir. Bu nedenle tek bir yeteneğin baskın olarak ölçülmesi çoğu zaman yeterli olarak kabul edilir. Bazı bilişsel, kişisel faktörler yanında testin alındığı ortam ve motivasyon gibi faktörlerin ortaya çıkması bu varsayımın tam olarak karşılanamamasının nedenleri arasındadır (Hambleton, Swaminathan, & Rogers, 1991). Boyutluluğun belirlenmesinde faktör analizi tekniklerinden faydalanılmaktadır.

Yerel bağımsızlık (Local independence). Yerel bağımsızlık maddelere verilen cevapların birbirinden bağımsız olmasıdır. Bir başka ifadeyle bir maddede yer alan ifadenin ya da bilginin diğer maddeler için ipucu sağlamamasıdır. Hambleton ve diğerleri (1991) yerel bağımsızlığı; test performansını etkileyen yetenekler sabitken herhangi iki madde çiftine verilen cevapların istatistiksel olarak

bağımsız olması şeklinde yorumlamış ve matematiksel olarak Eşitlik 1'deki gibi ifade etmiştir.

$$\begin{aligned} P(U_1, U_2, \dots, U_n | \theta) &= P(U_1 | \theta) P(U_2 | \theta) \dots P(U_n | \theta) \\ &= \prod_{i=1}^n P(U_i | \theta) \end{aligned} \quad (1)$$

U_i : Rastgele seçilen bir bireyin i ($i = 1, 2, \dots, n$) maddesine verdiği cevap.

$P(U_i = 1 | \theta)$: Rastgele seçilen ve θ yeteneğine sahip bir bireyin maddeye doğru cevap verme olasılığı.

$P(U_i = 0 | \theta)$: Rastgele seçilen ve θ yeteneğine sahip bir bireyin maddeye yanlış cevap verme olasılığı.

Tek boyutlu bir modelde maddelere verilen cevaplar yerel bağımsız değilse bu durumda diğer bir boyut yerel bağımlılığa neden olur ve madde çiftleri arasındaki bağımlılığa bakmak gerekir (DeMars, 2010). Maddelerin arasında ilişki olması beklenen bir durumdur. θ kontrol altına alındığında, maddelerin yerel bağımsızlığının kontrolünü sağlamak için Yen (1984) tarafından Q_3 testinin kullanılması önerilmiştir (DeMars, 2010).

Değişmezlik (Invariance). Madde ve yetenek parametrelerinin değişmezlik özelliği MTK'yi, KTK'dan ayıran en önemli özelliğidir (Hambleton, Swaminathan, & Rogers, 1991). Bu özellik, bir maddeyi niteleyen parametrelerin testi alanların yetenek dağılımına bağlı olmadığı ve testi alanları niteleyen parametrenin madde grubuna bağlı olmadığını ifade eder (Hambleton, Swaminathan, & Rogers, 1991; Embretson & Reise, 2000).

Veri Uyumu (Fit). Veriyle uyumlu olan modelin kullanılması MTK'da en önemli hususlardan biridir. Uyumsuz modellerle yapılan çalışmalar sonuçların geçerliğini düşürür. Şans parametresi içeren bir veri kümesi için 1PL ve 2PL modelin uygulanmasından elde edilecek sonuçlar ile 3PL modelin uygulanmasından elde edilecek sonuçlar da farklı olacaktır. Uyumun incelenmesi için log-olabilirlik değerlerinden (G^2) yararlanılabilir (DeMars, 2010). Modellerden elde edilen log-olabilirlik değerlerinin arasındaki farklar incelenip χ^2 tablosuyla karşılaştırılarak veriyle en çok uyumlu olan model belirlenir.

Madde Tepki Kuramı Modelleri

İkili puanlanan maddeler için MTK modelleri. İkili puanlanan maddelerde doğru cevaplar için (1), yanlış cevaplar için (0) puan verilir. İkili puanlanan maddeler

madde güçlüğü (b), madde ayırt ediciliği (a) ve şans parametresini (c) içerecek şekilde lojistik modeller yardımıyla modellenebilir. Yalnızca madde güçlüğünün yer aldığı, madde ayırt ediciliğinin tüm bireyler için belirli bir sabite eşit olduğu ve şans parametresinin bulunmadığı modele bir parametrelili lojistik model (1PLM) adı verilir. Madde güçlüğünü, madde ayırt ediciliğini içeren ve şans parametresinin bulunmadığı model iki parametrelili lojistik model (2PLM), madde güçlüğü, madde ayırt ediciliği ve şans parametresinin bulunduğu model üç parametrelili lojistik model (3PLM) olarak adlandırılır. Bu modeller aşağıda açıklanmıştır.

Bir parametrelili lojistik model (1PLM). Bir parametrelili lojistik model (1PLM) Eşitlik 2’de yer aldığı şekilde ifade edilebilir. Bu eşitlikte θ_s bireyin yeteneğini, β_i madde güçlüğünü, α ise madde ayırt ediciliğini göstermektedir. Eşitlik 2’den de anlaşılacağı üzere bireyin yeteneği arttıkça maddeyi doğru yanıtlama olasılığı da artmaktadır. 1 PL modelde madde ayırt ediciliği ortalama bir değere (herhangi bir sabit) sahiptir.

$$P(X_{is} = 1 | \theta_s, \beta_i) = \frac{\exp(\alpha(\theta_s - \beta_i))}{1 - \exp(\alpha(\theta_s - \beta_i))} \quad (2)$$

1PL modelde $\alpha = 1$ olduğunda Rasch modeli elde edilir. Rasch modelin gösterimi Eşitlik 3’te verilmiştir.

$$P(X_{is} = 1 | \theta_s, \beta_i) = \frac{\exp(\theta_s - \beta_i)}{1 - \exp(\theta_s - \beta_i)} \quad (3)$$

İki parametrelili lojistik model (2PLM). İki parametrelili lojistik model (2PLM) Eşitlik 4’te yer aldığı şekilde ifade edilebilir. Bu eşitlikte; θ_s bireyin yeteneğini, β_i madde güçlüğünü, α_i ise madde ayırt ediciliğini göstermektedir. 1PL modelden farklı olarak α (alfa) simgesinin altında bir i indisi yer almaktadır. Bu indis, maddelerin ayırt ediciliklerinin bir sabite eşit olmadığını ve değiştiğini göstermektedir.

$$P(X_{is} = 1 | \theta_s, \beta_i, \alpha_i) = \frac{\exp(\alpha_i(\theta_s - \beta_i))}{1 - \exp(\alpha_i(\theta_s - \beta_i))} \quad (4)$$

Üç parametrelili lojistik model (3PLM). Üç parametrelili lojistik model (3PLM) Eşitlik 5’te yer aldığı şekilde ifade edilebilir. Bu eşitlikte; θ_s bireyin yeteneğini, β_i madde güçlüğünü, α_i madde ayırt ediciliğini, γ_i (düşük asimptot) maddenin şansa doğru cevaplanma olasılığını göstermektedir. 1PL modelden farklı olarak α (alfa) simgesinin altında bir i indisi yer almaktadır. Bu indis maddelerin ayırt ediciliklerinin bir sabite eşit olmadığını ve değiştiğini göstermektedir. Benzer şekilde γ (gama)

simgesinin altında yer alan i indisi de maddelere göre şansa doğru cevap verme olasılığının değiştiğini göstermektedir.

$$P(X_{is} = 1 | \theta_s, \beta_i, \alpha_i, \gamma_i) = \gamma_i + (1 - \gamma_i) \frac{\exp(\alpha_i(\theta_s - \beta_i))}{1 - \exp(\alpha_i(\theta_s - \beta_i))} \quad (5)$$

Model seçimi. Yukarıda açıklanan 3 modelden (1PLM, 2PLM, 3PLM) hangisinin veriye daha iyi uyum sağladığının belirlenmesinde -2 log Likelihood değerleri ile serbestlik derecesi değerleri kullanılır. Eşitlik 6'da iki modelin karşılaştırılmasıyla χ^2 değerinin nasıl elde edileceği gösterilmiştir (Embretson & Reise, 2000).

$$\chi^2 = -2 \log Likelihood_{1.Model} - 2 \log Likelihood_{2.Model} \quad (6)$$

Eşitlik 6'dan elde edilen χ^2 değerinin manidarlığı 2.modelin 1.modele göre veriye daha iyi uyum sağladığını gösterir. Bu durumda 2.model 1.modele tercih edilmelidir.

Çoklu puanlanan maddeler için MTK modelleri. Bir testte ya da ölçekte yer alan maddeler her zaman doğru ve yanlış (1-0) şeklinde puanlanmaz. Örneğin, cevabı yapılandırılmış testlerde öğrencilerin maddeye verdikleri cevaplar belli bir rubrik yardımıyla ikiden daha fazla kategori üzerinden puanlanabilir. Böylesi bir durumda çoklu puanlanan maddeler için bu modellere ihtiyaç duyulur. Çoklu puanlanan maddeler, daha bilgi verici ve daha detaylı bilgilerin alınabildiği maddelerdir. Teste cevap veren bireyin yetenek düzeyi ile bir kategoriye cevap verme olasılığı arasındaki doğrusal olmayan ilişkileri ortaya koyabilmek için çoklu puanlamaya uygun MTK modellerine ihtiyaç duyulmaktadır (Embretson & Reise, 2000). Çoklu puanlanan maddelerin analizi için aşamalı tepki modeli, kısmi bilgi modeli ve genelleştirilmiş kısmi bilgi modeli gibi modeller geliştirilmiş olup bu çalışmada aşamalı tepki modeli üzerinden analizler yapıldığı için aşağıda yalnızca aşamalı tepki modeline ilişkin açıklamalara yer verilmiştir.

Aşamalı tepki modeli (Graded response model). Madde cevaplarının sıralı olduğu durumlarda kullanılabilen bu model 2PL modelin genelleştirilmiş halidir (Embretson & Reise, 2000). Aşamalı tepki modelinde kategori sayısının bir eksiği (K-1) kadar eşik değer elde edilir. Eşik değerler, bir kategoriden diğer kategoriye geçme ya da geçmeme durumunun belirlendiği noktalardır. Aşamalı tepki

modelinde eşik (güçlük) değerler sıralı elde edilir. Eşitlik 7 kullanılarak hesaplamalar yapılabilir.

$$P_{ix}(\theta) = \frac{\exp[\alpha_i(\theta - \beta_{ij})]}{1 + \exp[\alpha_i(\theta - \beta_{ij})]} \quad x=j=1, \dots, K-1 \quad (7)$$

Eşitlik 7’de;

θ : yetenek düzeyini,

α_i : i maddesinin ayırt edicilik değerini,

β_{ij} : 0.5 olasılıkla j eşik değerinin üstünde cevap vermek için gerekli yetenek düzeyini,

j : kategori sayısını göstermektedir.

Madde kalibrasyon yöntemleri (Item calibration methods). Veriyle uyumlu olan modelin seçilmesinden sonra madde parametrelerinin kestirimi ayrı (seperate) ya da eş zamanlı (concurrent) kalibrasyon yöntemiyle olarak yapılabilir. Bu çalışmada madde parametreleri ayrı kalibrasyon yöntemiyle kestirildiği için aşağıda yalnızca ayrı kalibrasyon yöntemine ilişkin bilgiler verilmiştir.

Ayrı kalibrasyon yöntemleri (Seperate calibration methods). Denk olmayan gruplarda eşitleme işlemi yapılırken test formlarının aynı ölçeğe getirilmesi gerekir (Kolen & Brennan, 2014). Bu amaçla 3 parametrelili I ve J ölçeklerinin dönüşümlerinde, A ve B sabit, θ_{Ji} ve θ_{Ii} sırasıyla ölçek J ve ölçek I ’da i bireyinin θ değerlerini göstermek üzere Eşitlik 8’de verilmiştir.

$$\theta_{Ji} = A\theta_{Ii} + B \quad (8)$$

a_{Jj} , b_{Jj} ve c_{Jj} ; J ölçeğindeki j maddesi için madde parametreleri ve a_{Ij} , b_{Ij} ve c_{Ij} ; I ölçeğindeki j maddesi için madde parametrelerini göstermek üzere Eşitlik 9 ve 10’da verilmiştir.

$$a_{Jj} = \frac{a_{Ij}}{A} \quad (9)$$

$$b_{Jj} = Ab_{Ij} + B \quad (10)$$

Düşük asimptot parametresi ölçekten bağımsız olarak Eşitlik 11’de verilmiştir.

$$c_{Jj} = c_{Ij} \quad (11)$$

Ayrı kalibrasyon yöntemleri moment yöntemler ve karakteristik eğri dönüştürme yöntemleri olarak ikiye ayrılmaktadır. Aynı zamanda bu yöntemlerin ikili

ve çoklu puanlanan modeller için hesaplama yöntemleri farklılaşmaktadır. Çoklu puanlanan modellerin hesaplama işlemleri karmaşık olduğundan aşağıda yalnızca ikili puanlanan modellerin formüllerine yer verilmiştir.

Moment yöntemler. Moment yöntemler; ortalama-ortalama (mean-mean) ve ortalama-standart sapma (mean-standart deviation) yöntemlerini içerir.

Ortalama-ortalama yöntemi. Bu yöntemde iki testin yalnızca güçlük ve ayırt edicilik parametrelerinin ortalaması kullanılır. A parametresi belirlemek için Eşitlik 12, B parametresini belirlemek için Eşitlik 13'te yer alan formüller kullanılır.

$$A = \frac{\mu_{a_I}}{\mu_{a_J}} \quad (12)$$

$$B = \mu_{b_J} - A\mu_{b_I} \quad (13)$$

A: Eğim değeri,

B: Sabit,

μ_{a_I} : I ölçeğindeki maddelerin ortalama ayırt edicilik değerini,

μ_{a_J} : J ölçeğindeki maddelerin ortalama ayırt edicilik değerini,

μ_{b_I} : I ölçeğindeki maddelerin ortalama güçlük değerini,

μ_{b_J} : J ölçeğindeki maddelerin ortalama güçlük değerini göstermektedir.

Ortalama-standart sapma yöntemi. Bu yöntemde iki testteki güçlük parametrelerinin ortalama ve standart sapması yardımıyla A parametresini belirlemek için Eşitlik 14, B parametresini belirlemek için Eşitlik 15 kullanılır.

$$A = \frac{\sigma_{b_I}}{\sigma_{b_J}} \quad (14)$$

$$B = \mu_{b_J} - A\mu_{b_I} \quad (15)$$

A: Eğim değeri,

B: Sabit,

σ_{b_I} : I ölçeğindeki maddelerin güçlük değerlerinin standart sapmasını,

σ_{b_J} : J ölçeğindeki maddelerin güçlük değerlerinin standart sapmasını,

μ_{b_I} : I ölçeğindeki maddelerin ortalama güçlük değerini,

μ_{b_J} : J ölçeğindeki maddelerin ortalama güçlük değerini göstermektedir.

Karakteristik eğri dönüştürme yöntemleri (Characteristic curve transformation methods). Moment yöntemler tüm madde parametre tahminlerini

aynı anda dikkate almaz (Kolen & Brennan, 2014). Bu sınırlılığı ortadan kaldırmak için Haebara (1980) tüm madde parametrelerini aynı anda dikkate alan bir yöntem önermiştir. Ardından Stocking ve Lord (1983), Haebara'nın yöntemine benzer yeni bir yöntem önermiştir. Bu iki yöntem karakteristik eğri dönüştürme yöntemi adı verilmektedir.

Haebara yaklaşımı. Belirli bir yetenek düzeyinde sınava girenlerin her bir madde için madde karakteristik eğrileri arasındaki farkın karesinin toplamı madde karakteristik eğrileri arasındaki fark olarak ifade edilebilir (Haebara, 1980). θ_i yeteneği için Haebara yöntemi Eşitlik 16'daki gibi gösterilir.

$$Hdiff(\theta_i) = \sum_{j:V} \left[p_{ij}(\theta_{ji}; \hat{a}_{jj}, \hat{b}_{jj}, \hat{c}_{jj}) - p_{ij}(\theta_{ji}; \frac{\hat{a}_{jj}}{A}, A \hat{b}_{jj} + B, \hat{c}_{jj}) \right]^2 \quad (16)$$

Toplama işlemi ortak maddeler ($j:V$) üzerinden yapılır. *Hdiff* değerleri gözlemler üzerinden toplanır. A ve B değerlerini bulmak için yapılan kestirim Eşitlik 17'deki ölçütü minimize eder (Kolen & Brennan, 2014).

$$Hcrit = \sum_i Hdiff(\theta_i) \quad (17)$$

Stocking-Lord yaklaşımı. Stocking ve Lord (1983), Haebara yaklaşımındaki madde karakteristik eğrileri arasındaki farkın karesinin toplamı yerine, Eşitlik 18'de gösterildiği gibi toplamların farkının karesini kullanmayı önermiştir (Kolen & Brennan, 2014).

$$SLdiff(\theta_i) = \left[\sum_{j:V} p_{ij}(\theta_{ji}; \hat{a}_{jj}, \hat{b}_{jj}, \hat{c}_{jj}) - \sum_{j:V} p_{ij}(\theta_{ji}; \frac{\hat{a}_{jj}}{A}, A \hat{b}_{jj} + B, \hat{c}_{jj}) \right]^2 \quad (18)$$

Stocking-Lord yaklaşımında, eşitlemeden önce her bir parametre dizisi için maddeler üzerinde toplam alınır. Eşitlik 19'da yer alan fonksiyon MTK'da test karakteristik eğrisi olarak adlandırılır.

$$\tau(\theta_i) = \sum_j p_{ij}(\theta_i) \quad (19)$$

SLdiff(θ_i) ifadesi θ_i için karakteristik eğrilerin toplamlarının farkının karesidir. *Hdiff*(θ_i) ifadesi ise θ_i için madde karakteristik eğrileri arasındaki farkın karesinin toplamıdır. *SLdiff* değerleri gözlemler üzerinden toplanır. A ve B değerlerini bulmak için yapılan kestirim Eşitlik 20'deki ölçütü minimize eder (Kolen & Brennan, 2014).

$$SLcrit = \sum_i SLdiff(\theta_i) \quad (20)$$

Eşitleme Sonuçlarının Değerlendirilmesi

Eşitleme işlemi gerçekleştirildikten sonra sonuçların değerlendirilmesi için çeşitli yöntemler kullanılmaktadır. Bunlardan bazıları RMSD (Root Mean Square Difference), MSE (Mean Square Error), RMSE (Root Mean Square Error) ve eşitlemenin standart hatasıdır (SE).

Lord'a (1980) göre eşitliğin sağlanabilmesi için hangi testin alındığının önemli olmaması gerekir. Aynı zamanda Lord (1980), eşitliğin testlerin eşit ya da aynı güvenilirliğe sahip olması durumunda gerçekleşeceğini belirtmiştir.

Testlerin eşitlik ve aynı dağılım özellikleri CASMA raporunun 2011 yılı Aralık ayı raporunun 7. bölümünde He ve Kolen (2011) tarafından detaylı olarak incelenmiştir. Lord (1980) tarafından önerilen eşitlik özelliğinin iki form aynı olmadığı süre sağlanamayacağı ve aynı olduğu durumda ise eşitlemenin gereksiz olacağı Kolen ve Brennan (2014) tarafından belirtilmektedir. Bu nedenle Divgi (1981) ve Morris (1982) Lord'un (1980) önerdiği eşitlik özelliğinden daha az kısıtlayıcı olan zayıf eşitlik özelliğini, bir diğer adıyla birinci sıra (first-order) eşitlik ve Morris (1982) eşitleme sonrası formlara ait ölçmenin koşullu standart hatasının eşit olması durumunda sağlanan ikinci sıra (second-order) eşitlik özelliğini önermişlerdir (akt. Öztürk-Gübeş, 2014).

Birinci ve ikinci sıra eşitlik özelliği (First and second order equity properties). Birinci sıra eşitlik (BSE), eşitlemeden sonra paralel formların koşullu ortalamalarının eşit olmasını gerektirmektedir. Kolen ve Brennan ikinci sıra eşitlik (İSE) özelliğinin, eşitlemeden sonra paralel formların ölçmenin koşullu standart hatasının eşit olması durumunda gerçekleşeceğini belirtmişlerdir (He & Kolen, 2011). BSE özelliğinin matematiksel gösterimi Eşitlik 21'de, İSE özelliğinin matematiksel gösterimi Eşitlik 22'de yer almaktadır (Öztürk-Gübeş & Kelecioğlu, 2015b).

$$E[eq_Y(X) | \tau] = E(Y | \tau) \quad (21)$$

$$SEM_{eq_Y(X) | \tau} = SEM_{Y | \tau} \quad (22)$$

BSE özelliğinin değerlendirilmesi için Tong ve Kolen (2005) tarafından geliştirilen D_1 indeksi Eşitlik 23'te verilmiştir (Tong & Kolen, 2005).

$$D_1 = \frac{\sqrt{\sum[q_i(ySC|\theta_i - xSC|\theta_i)^2]}}{SD_Y} \quad (23)$$

Eşitlik 23'te;

$ySC|\theta_i$: i. quadrature noktasındaki θ_i yetenek düzeyi için Y formunun beklenen ölçek puanını,

$xSC|\theta_i$: i. quadrature noktasındaki θ_i yetenek düzeyi için X formunun beklenen ölçek puanını,

q_i : θ_i 'deki quadrature ağırlığını,

SD_Y : Y formu ölçek puanlarının standart sapmasını göstermektedir. D_1 'in küçülen değerleri BSE'nin daha iyi korunduğunu göstermektedir.

İSE özelliğinin değerlendirilmesi için Tong ve Kolen (2005) tarafından geliştirilen D_2 indeksi Eşitlik 24'te verilmiştir.

$$D_2 = \frac{\sqrt{\sum[q_i(ySEM|\theta_i - xSEM|\theta_i)^2]}}{SD_Y} \quad (24)$$

Eşitlik 24'te;

$ySEM|\theta_i$: Y formu için i. quadrature noktasında ve θ_i yetenek düzeyindeki bireylere ait koşullu ölçmenin standart hatasını,

$xSEM|\theta_i$: X formu için i. quadrature noktasında ve θ_i yetenek düzeyindeki bireylere ait koşullu ölçmenin standart hatasını,

SD_Y : Y formu ölçek puanlarının standart sapmasını göstermektedir. D_2 'nin küçülen değerleri İSE'nin daha iyi korunduğunu göstermektedir.

Aynı dağılım özelliği (Same distributions property). Eşitlik özelliğini değerlendirmek için X formundaki puanların dağılımının eşitlenen Y formundaki puanların dağılımıyla aynı olması aynı dağılım özelliğinin (ADÖ) sağlandığını gösterir. ADÖ'nün korunumunu değerlendirmek için Tong ve Kolen (2005) iki dağılım arasındaki en büyük farkı dikkate alan Kolmogorov-Smirnov T 'yi (Conover, 1999) kullanmıştır. Kim, Brennan ve Kolen (2005) ise tüm puan noktaları üzerinde toplanan dağılımların arasındaki farklar olan mutlak farklar toplamını çalışmalarında kullanmıştır.

Bu çalışmada ADÖ'nün korunumunu değerlendirmek için Eşitlik 25'te yer alan Kolmogorov-Smirnov T değeri kullanılmıştır.

$$T = \sup_{i=k} |Y(i) - X(k)| \quad (25)$$

Eşitlik 25'te yer alan $Y(i)$ ve $X(k)$; Y ve X formlarının yüzde olarak ifade edilen birikimli dağılım fonksiyonlarını (*bdf*) gösterirken, Kolmogorov T ; $i = k$ olduğu durumlarda iki *bdf* arasındaki en büyük farkı göstermektedir. Küçülen T değeri iki dağılımın daha benzer olduğunu gösterirken büyüyen T değeri dağılımların daha az benzer olduğunu göstermektedir.

İlgili Araştırmalar

Kim ve Lee (2004) MTK'ya ilişkin ortalama-ortalama, ortalama-sapma, Haebara ve Stocking-Lord yönteminin tek format (çoktan seçmeli ya da açık uçlu) testlerden karma format testlere genişletilmesine ilişkin yaptıkları çalışmada aşağıdaki bulguları elde etmiştir.

- ✓ Karakteristik eğri yöntemleri az bir istisna ile moment yöntemlerine göre daha az eşitleme hatası vermiştir.
- ✓ Haebara yöntemi diğer yöntemlere göre daha az eşitleme hatası vermiştir.
- ✓ Eş zamanlı (simultaneous) eşitleme, çoğu durumlarda tek bir madde formatıyla (çoktan seçmeli ya da açık uçlu) yapılan ilişkilendirmeden daha doğru sonuç vermiştir.
- ✓ Eşitleme doğruluğu ortak madde fazlalaştığında daha da artmıştır.
- ✓ Yalnızca ikili puanlanan ve yalnızca çoklu puanlanan maddelerle yapılan eşitleme incelendiğinde ortak madde yapısında hangi madde tipi ağırlıklı ise o maddelerle yapılan ilişkilendirmelerin daha az hata ürettiği görülmüştür.

Wang, Lee, Brennan ve Kolen (2008) denk olmayan gruplar ortak madde deseninde frekans kestirim yöntemi ile zincirleme eşit yüzdeliği yöntemini karşılaştırmışlardır. Önemli derecede grup farklılıkları olduğunda frekans kestirim yönteminin zincirleme eşit yüzdeliği yöntemden daha büyük yanlılığa sahip olduğu, ancak eşitlemenin standart hatasının daha az olduğu görülmüştür.

Hagge (2010) denk olmayan gruplarda eşitleme yöntemlerini ve karma format testlerde ortak madde yeterliliğine ilişkin çalışma yapmıştır. İleri düzey yerleştirme sınavı programında (Advanced Placement Examination Program) yer alan İngilizce, İspanyolca ve kimya testleri kullanılarak karma format testler için yapılan analizlerde elde edilen bulgulardan bazıları aşağıda verilmiştir.

- ✓ Gruplar arası yetenek farkı yükseldikçe yanlılık artma eğilimindedir.
- ✓ Belirli bir eşitleme yöntemi için ölçüt eşitleme ilişkisine göre yanlılıktaki artış frekans tahmini için en büyük, MTK eşitleme yöntemleri için en küçüktür.
- ✓ Eşitlemenin standart hatası MTK gözlenen puan eşitleme için en küçük ve zincir yüzdelli eşitleme için en büyüktür.
- ✓ Hangi ortak madde kompozisyonunun en az yanlılığa neden olduğuna ilişkin sonuçlar karışıktır.

Kilmen (2010) MTK'ya dayalı test eşitleme yöntemlerinden kestirilen eşitleme hatalarının örneklem büyüklüğü ve yetenek dağılımına göre karşılaştırılması adlı çalışmada 3 parametrelili lojistik modele uyumlu simülatif veriler kullanmıştır. Çoktan seçmeli maddelerle yapılan bu çalışmada 500 ve 1000 kişilik benzer ve farklı yetenek dağılımına sahip gruplar kullanıldığı durumlarda Stocking-Lord yöntemiyle yapılan eşitlemenin daha az hatalı olduğu sonucuna ulaşılmıştır. Ayrıca 1000 kişilik örneklem büyüklüğüne sahip gruplara ait eşitleme hatalarının 500 kişilik örneklem büyüklüğünden elde edilen eşitleme hatalarından daha az olduğu saptanmıştır.

Walker ve Kim (2010) denk olmayan gruplar deseninde hem çoktan seçmeli hem de açık uçlu madde içeren karma format testleri eşitlemek için yalnızca çoktan seçmeli maddelerin kullanılması durumunda test eşitlemenin nasıl olacağını incelemişlerdir. Yalnızca çoktan seçmeli maddelerin çoktan seçmeli ve açık uçlu maddelerin aynı yapıyı ölçmesi ya da çoktan seçmeli ortak maddelerin miktarıyla toplam test arasındaki ilişkilerin yeni ve referans gruplar boyunca sabit kalması durumunda bu iki formun etkili bir şekilde eşitlenebileceği görülmüştür. Çalışmada ayrıca yalnızca çoktan seçmeli ortak maddelerle eşitlenen karma format testlerin; çoktan seçmeli maddelerin yalnızca çoktan seçmeli maddelerle eşitlendiği ve sonra birleşik puanların çoktan seçmeli puanlara ölçeklendiği iki aşamalı stratejiden daha etkili olup olmadığı değerlendirilmiştir. Ortak madde eşitleme ve iki aşamalı eşitleme

hem lineer hem de lineer olmayan zincir eşitleme yöntemleri için neredeyse aynı sonuçları verdiği görülmüştür.

Michaelides (2010) test eşitlemede iyi çalışmayan ortak maddelerin MTK parametre tahminlerine etkisini tartışmıştır. Testin geçerliğine ilişkin çıkarımlar ve iyi çalışmayan ortak maddelerin testten çıkarılıp çıkarılmaması gerektiğine ve bu maddelerle nasıl baş edilebileceğine ilişkin önerilerde bulunmuştur.

Sinharay ve Holland (2010) denk olmayan gruplarda ortak madde deseninin tasarımı nedeniyle eksik veri içerdiğini belirterek, frekans tahmini eşit yüzdelikli eşitleme modeli, zincir eşit yüzdelikli eşitleme modeli ve MTK gözlenen puan eşitleme yöntemlerinden hangisinin daha sağlıklı sonuçlar verdiğini araştırmışlardır. Yapılan araştırma sonucunda zincir eşit yüzdelikli eşitleme yönteminin diğer yöntemlere göre daha olumlu sonuç verdiğini görmüşlerdir.

He (2011) karma format testlerin eşitleme özelliklerine ilişkin yaptığı çalışmada frekans tahmin yöntemi, zincirleme eşit yüzdelik yöntemi, MTK gerçek puan eşitleme ve MTK gözlenen puan eşitleme yöntemini kullanmıştır. Elde edilen bulgulardan bazıları aşağıda verilmiştir.

- ✓ Gerçek puan eşitleme daha iyi BSE, gözlenen puan eşitleme daha iyi İSE sağlamıştır.
- ✓ Çoktan seçmeli ve açık uçlu maddelerin daha yüksek korelasyonu, her iki MTK yöntemi için de birinci sıra eşitliğin daha iyi sağlanmasıyla ilgilidir.
- ✓ Çoktan seçmeli ve açık uçlu maddelerin daha yüksek korelasyonu, gerçek puan eşitleme yöntemi için BSE'nin daha iyi sağlanmasıyla ilgilidir.
- ✓ Çoktan seçmeli ve açık uçlu maddelerin daha yüksek korelasyonu, gözlenen puan eşitleme yöntemi için ADÖ'nün daha iyi sağlanmasıyla ilgilidir.
- ✓ Ortak madde oranı, çoktan seçmeli madde oranı ve formlar arasındaki benzerlik, eşitleme özelliklerinin korunmasıyla ilgili değildir.

Gültekin (2011) yaptığı çalışmada TIMSS 2007 sekizinci sınıf matematik testi ikinci kitapçığındaki maddelerden test uzunluğunun 15 ve 25 madde, açık uçlu madde yüzdesinin %20 ve %40 olduğu durumlar için madde ve yetenek parametrelerini, test ve madde bilgi fonksiyonlarını ve görelilik indekslerini incelemiştir. Yapılan inceleme sonucunda uzun karma testlerin, sadece çoktan seçmeli ve sadece açık uçlu maddelerden oluşan testlere göre orta güçlükte ve

daha ayırıcı testler olduğu görülmüştür. Orta yetenek düzeyinde uzun karma testlerin daha az standart hatayla yetenek kestirimi yaptığı, yanıtı sınırlı açık uçlu maddelerin çoktan seçmeli maddelerden daha fazla bilgi verdiği ve açık uçlu testin diğer testlerden etkili olduğu belirlenmiştir.

Kastner ve Stangl (2011), çoktan seçmeli ve açık uçlu maddeleri değerlendirmek için çok yüzeysel Rasch model kullanmışlar ve doğru sayısına göre puanlama yapıldığında açık uçlu maddeli testlerin çoktan seçmeli maddeli testlere eşdeğer olduğunu görmüşlerdir.

Wang (2011) denk olmayan gruplar ortak madde deseninde eşitlemenin standart hatasının kestiriminde bootstrap yöntemini incelemiştir. Kullanılan yöntemler dışında örneklem büyüklüğü ve test uzunluğunun standart hatanın kestiriminde en fazla etkiye sahip olduğu, toplam madde içerisindeki ortak madde oranının ve grup yetenek farklarının kısa testlerde çok az, uzun testlerde biraz etkiye sahip olduğu görülmüştür.

Andrews (2011) çok boyutlu MTK kullanarak denk olmayan gruplar ortak madde deseninde BSE ve İSE'yi incelemiştir. Hem gerçek hem de simülatif veri ile yürütülen çalışmada frekans tahmini, zincir eşit yüzdellikli, tek boyutlu MTK gerçek puan, tek boyutlu MTK (TBMTK) gözlenen puan ve çok boyutlu MTK (ÇBMTK) gözlenen puan eşitleme yöntemleri kullanılmıştır. Elde edilen sonuçlardan bazıları aşağıda özetlenmiştir.

- ✓ Gerçek veriler üzerinden yapılan analizler BSE ve İSE bazında eşitleme yöntemlerinin performansının hangi testin kullanıldığı ve eşitliği değerlendirmek için hangi çerçevenin (ÇBMTK/TBMTK) kullanıldığına bağlı olarak değiştiğini göstermiştir.
- ✓ Bazı testler her iki çerçeve için benzer eşitliğin korunumu gösterirken, diğerleri eşitlik değerlendirmesinde oldukça farklılaşmıştır.

He (2011) karma format testlerin eşitleme özelliklerini değerlendirdiği çalışmada ileri düzey yerleştirme sınavı programı sınavında yer alan 39 karma format teste ilişkin verilerini kullanmıştır. Frekans tahmini, zincir eşit yüzdellikli, MTK gerçek puan eşitleme ve MTK gözlenen puan eşitleme yöntemlerinin kullanıldığı çalışmada BSE, İSE ve ADÖ üzerinden sonuçlar aşağıda belirtildiği şekilde özetlenebilir.

- ✓ MTK gerçek puan eşitleme BSE'yi MTK gözlenen puan eşitleme yöntemine göre daha iyi sağlarken, MTK gözlenen puan eşitleme İSE'yi ve ADÖ'yü MTK gerçek puan eşitlemeye göre daha iyi sağlamıştır.
- ✓ Geleneksel yöntemler arasında zincir eşit yüzdelikli eşitleme BSE'yi frekans tahmini yönteminden daha iyi sağlarken, İSE ve ADÖ ele alındığında her iki yöntem benzer sonuçlar ortaya koymuştur.
- ✓ Düzgünleştirme İSE ve ADÖ'nün korunmasına yardımcı olmuştur.
- ✓ Çoktan seçmeli ve açık uçlu maddeler arasındaki yüksek korelasyon her iki MTK modeli için BSE'nin daha iyi korunumuyla ilgili bulunmuştur.
- ✓ Çoktan seçmeli ve açık uçlu maddeler arasındaki yüksek korelasyon MTK gerçek puan eşitleme modeli için İSE'nin daha iyi korunumuyla ilgili bulunmuştur.
- ✓ Çoktan seçmeli ve açık uçlu maddeler arasındaki yüksek korelasyon MTK gözlenen puan eşitleme modeli için ADÖ'nün daha iyi korunumuyla ilgili bulunmuştur.
- ✓ Ortak madde oranı, formlar arasındaki benzerlik eşitlik özelliklerinin korunumuyla ilgili bulunmamıştır.

Montgomery ve Skorupski (2012) iki örnekleme dört model ve altı madde kombinasyonu ile çalışma yapmışlardır. Madde kombinasyonları, MTK modeli ve örneklem büyüklüğü arasındaki etkileşimlerin theta (yetenek) kestirimindeki RMSE (root mean square) değerini, sınıflama doğruluğunu ve yakınsama oranını (convergence rate) etkilediğini bulmuşlardır.

Gök (2012), örneklem büyüklüğü, yetenek dağılımı, test uzunluğu ve model türüne göre türetilen test formlarını MTK'ya dayalı kestirim yöntemlerini kullanarak denk olmayan gruplar ortak madde deseninde eşitleme çalışması yapmıştır. Elde edilen bulgulara göre en iyi eşitleme sonucunu 3000 kişilik örneklem, 80 maddelik test, benzer yetenek dağılımına sahip gruplar, 2 parametrelili Lojistik Model ve Stocking-Lord yöntemi vermiştir.

Lissitz, Hou ve Slater (2012) açık uçlu ve çoktan seçmeli maddelerin değişik alanlardaki yapı farklılıklarını, test karakteristik eğrisi ve test güvenilirliğindeki değişimleri ve madde formatının ırk ve cinsiyetle etkileşimini incelemiştir. Veriler,

lise diploması almak için yüksek öneme sahip devlet sınavı olan Maryland Yüksekokulu'ndan elde edilmiştir. Açık uçlu ve çoktan seçmeli maddelerin arasında ince farklar olduğu ve bu farkların özellikle İngilizce ve siyasal bilgiler testi için boyutluluk açısından ortaya çıktığı belirlenmiştir. Ayrıca, ırk ve cinsiyet açısından da bu iki madde tipinde değişen performanslar gözlenmiştir.

Lu, Haberman, Guo ve Liu (2015) denk olmayan gruplar ortak madde deseniyle eşitlemede ortak madde seçiminin etkilerini değerlendirmek için örneklem istatistiklerinin varyansının tahmini için kullanılabilen yeniden örnekleme yöntemi olan Jackknifing yöntemini kullanmışlardır. Lu ve arkadaşları daha çok ortak madde içeren testlerin toplam grup ortalamalarında daha az standart hata ürettiklerini, hem yetenek hem de diğer özellikler bakımından farklılaşan örneklemelerin ortak madde kararlılık ölçüsünü etkileyebileceğini, ortak madde örnekleminin büyük etki gösterdiğini ve puanların raporlanmasında bu durumun göz ardı edilmemesi gerektiğini belirtmişlerdir.

Öztürk-Gübeş (2014) testlerin boyutluluğunun, ortak madde formatının, yetenek dağılımının ve ölçek dönüştürme yöntemlerinin karma testlerin eşitlenmesine etkisi araştırmışlardır. Simülatif verilerle yapılan çalışmada BSE özelliği ölçütüne göre yukarıda belirtilen tüm faktörlerin, İSE özelliğine göre ise ortak madde seti dışındaki tüm faktörlerin gerçek puan ve gözlenen puan eşitleme sonuçları üzerinde istatistiksel olarak anlamlı etkisinin olduğu sonucuna ulaşılmıştır. Elde edilen diğer sonuçlar aşağıdaki gibi özetlenebilir.

- ✓ BSE ve İSE eşitlik özellikleri en iyi tek boyutlu, en kötü ise çok boyutluluk derecesinin en yüksek olduğu koşul altında korunmuştur.
- ✓ Karma ortak madde seti kullanılarak elde edilen gerçek puan ve gözlenen puan eşitleme sonuçları BSE özelliğini çoktan seçmeli madde seti kullanılan koşullara göre daha iyi korumuştur.
- ✓ Tek boyutluluğun sağlandığı koşullarda her ne kadar moment ölçek dönüştürme yöntemleri BSE ve İSE özelliklerini korumada karakteristik eğri yöntemlerine yakın performans göstermiş olsa da özellikle çok boyutluluğun derecesinin arttığı koşullarda karakteristik eğri ölçek dönüştürme yöntemleri her iki eşitlik özelliğini korumada üstün performans göstermiştir.

Öztürk-Gübeş ve Keleciođlu (2015a) farklı test eşitleme yöntemlerinin eşitlik özelliđi ölçütüne göre karşılaştırmışlardır. 2009 Öğrenci Başarılarının Belirlenmesi Sınavı'nda yer alan A ve C kitapçıklarını çalışma kapsamında ele alınmış ve BSE özelliđinin MTK gerçek puan eşitleme, İSE özelliđinin ise MTK gözlenen puan eşitleme yönteminde en iyi şekilde sağlandığı görülmüştür.

Öztürk-Gübeş ve Keleciođlu (2015b) karma format testlerin eşitlenmesinde MTK eşitleme yöntemlerini eşitlik özelliđi korunumuna göre karşılaştırmışlardır. Öğrencilerin 2011 TIMSS sekizinci sınıf Türkiye örnekleminde yer alan matematik testine verdiği cevaplar üzerinden yapılan çalışmada MTK gerçek puan eşitleme yönteminin BSE özelliđini korumada daha iyi performans gösterdiği, MTK gözlenen puan eşitleme yönteminin İSE özelliđini korumada daha iyi olduğu sonucuna ulaşılmıştır.

Uysal ve Kilmen (2016) simülatif veri kullanarak yaptıkları çalışmada karma format testlerde MTK'ya dayalı eşitleme yöntemlerini ortak madde yüzdesinin farklılaştığı durumlar için karşılaştırmışlardır. Karşılaştırma sonucunda en az eşitleme hatası sağa çarpık, %30 ortak madde uzunluđuna sahip testlerde ve Stocking-Lord yönteminde ortaya çıkmıştır. Ayrıca benzer yetenek dağılımına sahip gruplara ilişkin eşitleme hataları, farklı yetenek dağılımına sahip gruplara ilişkin eşitleme hatalarından daha küçük elde edilmiştir.

Bölüm 3

Yöntem

Bu bölümde araştırmanın yöntemi, evreni ve örnekleme, veri toplama süreci, araştırma deseni ve verilerin analizi başlıklarına yer verilmiştir.

Araştırma Yöntemi

Bu araştırmada denk olmayan gruplar ortak madde deseninde gerçek puan ve gözlenen puan eşitleme yöntemiyle yapılan eşitlemede aynı dağılım ve eşitlik özelliğinin korunumu incelenmiştir.

Araştırmanın Evreni ve Örnekleme

Bu çalışmada 2016 yılında 81 ilde uygulanan ABİDE araştırmasında yer alan Matematik ve Sosyal Bilgiler testi verileri kullanılmıştır. ABİDE araştırmasına katılan yaklaşık 34000 öğrencinin cevap örüntüleri incelenmiş olup çalışmanın daha sağlıklı yürütülebilmesi, madde parametre kestirimleri ve testlerin güvenilirliklerinin hesaplanmasının daha tutarlı olması açısından araştırmanın evreni ve örneklemi belirlenirken aşağıdaki hususlara dikkat edilmiştir.

- ✓ Çoktan seçmeli ve açık uçlu maddelere cevap verme davranışlarının farklı olmasından dolayı testlerde yer alan 18 maddenin en az 4 tanesine ve açık uçlu maddelerin en az 2 tanesine cevap veren öğrenciler araştırmaya dâhil edilmiştir.
- ✓ A, B ve C formlarının oluşturduğu evren içerisinde rastgele seçilen 1000, 3000 ve 5000 kişilik öğrenciler ise araştırmanın örneklemi oluşturmuştur.
- ✓ Evren ve örneklerdeki erkek ve kız öğrenciler alt gruplara göre farklılıkları incelemek için dikkate alınmış olup cinsiyete göre ayrıca bir örneklem seçilmemiştir.

Araştırma kapsamında yer alan Matematik ve Sosyal Bilgiler testinin; evreninin, 1000, 3000 ve 5000 kişilik örneklerinin cinsiyete göre dağılımı Tablo 2'de verilmiştir.

Tablo 2

Matematik ve Sosyal Bilgiler Testinin Evreni ve Örneklem Bilgileri

Veri	Form	Matematik			Sosyal Bilgiler		
		Toplam	Erkek	Kız	Toplam	Erkek	Kız
Evren	A	11122	5587	5535	11327	5718	5609
	B	11027	5510	5517	11391	5737	5654
	C	10572	5346	5226	10912	5565	5347
Örneklem 1	A	5000	2506	2494	5000	2547	2453
	B	5000	2495	2505	5000	2541	2459
	C	5000	2531	2469	5000	2577	2423
Örneklem 2	A	3000	1508	1492	3000	1519	1481
	B	3000	1509	1491	3000	1501	1499
	C	3000	1530	1470	3000	1515	1485
Örneklem 3	A	1000	503	497	1000	502	498
	B	1000	509	491	1000	534	466
	C	1000	494	506	1000	539	461

Tablo 2 incelendiğinde Matematik A testinin evreninin 11122, B testinin evreninin 11027, C testinin evreninin 10572, Sosyal Bilgiler A testinin evreninin 11327, B testinin evreninin 11391, C testinin evreninin 10912 öğrenciden oluştuğu görülmektedir.

Veri Toplama Süreci

Bu çalışmada 2016 yılında uygulanan ABİDE araştırmasından elde edilen bilgiler kullanılmıştır. Bu amaçla MEB Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü'nden verilerin kullanımına ilişkin gerekli izin alınmıştır. İzin belgesi EK-Ü'de yer almaktadır.

Veri Toplama Araçları

Bu çalışmada 2016 ABİDE araştırmasında uygulanan dört testten ikisi olan Matematik ve Sosyal Bilgiler testinden elde edilen veriler kullanılmıştır.

ABİDE araştırması sekizinci sınıf öğrencilerinin zihinsel becerilerinin ortaya konulması ve öğrencilerin başarılarıyla ilişkili öğrenci, öğretmen ve okul özelliklerinin belirlenmesi amacıyla yapılmıştır. ABİDE çalışması kapsamında soru yazarlarına

akademisyenler tarafından eğitimler verilmiştir. Maddeler akademisyenlerin ve öğretmenlerin katılımıyla öğretim programından yararlanarak oluşturulan beceriler dikkate alınarak geliştirilmiştir. Alan uzmanları tarafından hazırlanan sorular, ölçme ve değerlendirme uzmanları ile dil uzmanları tarafından gözden geçirildikten sonra akademik redaksiyona tabi tutulmuş ve sorulara son şekli verilmiştir. Pilot uygulamanın ardından esas uygulamada yer alacak maddeler belirlenmiştir (MEB, 2017). MEB ABİDE raporu incelendiğinde Matematik testi A, B ve formlarının güvenilirliklerinin sırasıyla 0.85, 0.83 ve 0.79, Sosyal Bilgiler testi A, B ve C formlarının güvenilirliklerinin sırasıyla 0.73, 0,78 ve 0,73 olduğu görülmektedir.

Verilerin Analizi

Sekizinci sınıf ABİDE araştırmasında Türkçe, Fen Bilimleri, Sosyal Bilgiler ve Matematik testlerini alan öğrenciler her bir testten 20'şer madde olmak üzere 80 maddeye cevap vermiştir. Bu çalışmada kullanılan Sosyal Bilgiler ve Matematik testinde her bir öğrenciye A1, A2, A3, A4, B1, B2, B3, B4, C1, C2, C3, C4 formlarından birisi uygulanmıştır. A kitapçığının farklı formlarını oluşturan A1, A2, A3, A4 formlarını birbirinden ayıran iki adet pilot madde bulunmaktadır. Bu pilot maddelerin her ikisi çoktan seçmeli, her ikisi açık uçlu olabileceği gibi bir tanesi açık uçlu, bir tanesi çoktan seçmeli olabilmektedir. Analizler bu iki madde dışarıda bırakılarak 18 madde üzerinden yapılmıştır. A testini B testine bağlayan 9, B testini C testine bağlayan 9 ve A testini C testine bağlayan 9 ortak madde bulunmaktadır. Bu 9 ortak maddedeki açık uçlu ve çoktan seçmeli madde ağırlıkları 4-5 ya da 5-4 şeklinde oluşturulmuş olup açık uçlu maddelerin cevapları 0-1, 0-1-2, 0-1-2-3-4 şeklinde rubrikle puanlanmıştır.

Çoktan seçmeli maddeler ve açık uçlu maddelerden alınabilecek toplam puanların testlerde yer alan açık uçlu madde sayısı ve bu maddelerin kategorisine göre değişmesi nedeniyle testlerden alınabilecek maksimum puanlar değişebilmektedir. Çalışmada kullanılan madde ve ortak madde sayıları Matematik testi için Tablo 3 ve Tablo 4'te, Sosyal Bilgiler testi için Tablo 5 ve Tablo 6'da verilmiştir.

Tablo 3

Matematik A-B Testinde Yer Alan Madde ve Ortak Madde Sayıları ve Formatları

<i>Tüm Test</i>			
<u>Madde Türü</u>	<u>Kategori Sayısı</u>	<u>A</u>	<u>Madde Sayısı</u>
Çoktan Seçmeli	0-1	9	9
Açık Uçlu	0-1	3	3
	0-1-2	6	6
	<i>Toplam</i>	18	18
<i>Ortak Test</i>			
<u>Madde Türü</u>	<u>Kategori Sayısı</u>	<u>A</u>	<u>Madde Sayısı</u>
Çoktan Seçmeli	0-1	5	5
Açık Uçlu	0-1	0	0
	0-1-2	4	4
	<i>Toplam</i>	9	9
<i>Ortak Olmayan Test</i>			
<u>Madde Türü</u>	<u>Kategori Sayısı</u>	<u>A</u>	<u>Madde Sayısı</u>
Çoktan Seçmeli	0-1	4	4
Açık Uçlu	0-1	3	3
	0-1-2	2	2
	<i>Toplam</i>	9	9

Tablo 3 incelendiğinde Matematik testi için B formuna eşitlenen A formunun madde sayısı, ortak madde sayısı, ortak madde formatı açısından aynı olduğu görülmektedir.

Tablo 4

Matematik C-B Testinde Yer Alan Madde ve Ortak Madde Sayıları ve Formatları

<i>Tüm Test</i>			
<u>Madde Türü</u>	<u>Kategori Sayısı</u>	<u>C</u>	<u>Madde Sayısı</u>
Çoktan Seçmeli	0-1	8	9
Açık Uçlu	0-1	6	3
	0-1-2	4	6
	<i>Toplam</i>	18	18
<i>Ortak Test</i>			
<u>Madde Türü</u>	<u>Kategori Sayısı</u>	<u>C</u>	<u>Madde Sayısı</u>
Çoktan Seçmeli	0-1	4	4
Açık Uçlu	0-1	3	3
	0-1-2	2	2
	<i>Toplam</i>	9	9
<i>Ortak Olmayan Test</i>			
<u>Madde Türü</u>	<u>Kategori Sayısı</u>	<u>C</u>	<u>Madde Sayısı</u>
Çoktan Seçmeli	0-1	4	5
Açık Uçlu	0-1	3	0
	0-1-2	2	4
	<i>Toplam</i>	9	9

Tablo 4 incelendiğinde Matematik testi için B formuna eşitlenen C formunun madde sayısı, ortak madde sayısı, ortak madde formatı açısından aynı oldukları görülmektedir. Ortak olmayan maddelerin C testinde 4'ü çoktan seçmeli iken, B testinde 5'i çoktan seçmelidir. Ortak olmayan açık uçlu maddelerden C testinde 0-1 şeklinde puanlananların sayısı 3 iken B testinde 0-1 şekilde puanlanan madde bulunmamaktadır. Benzer şekilde ortak olmayan açık uçlu maddelerden C testinde 0-1-2 şeklinde puanlananların sayısı 2 iken B testinde bu sayı 4'tür.

Tablo 5

Sosyal Bilgiler A-B Testinde Yer Alan Madde ve Ortak Madde Sayıları ve Formatları

<i>Tüm Test</i>			
<u>Madde Türü</u>	<u>Kategori Sayısı</u>	<u>A</u>	<u>B</u>
Çoktan Seçmeli	0-1	9	9
Açık Uçlu	0-1	6	6
	0-1-2	2	3
	0-1-2-3-4	1	0
	<i>Toplam</i>	18	18
<i>Ortak Test</i>			
<u>Madde Türü</u>	<u>Kategori Sayısı</u>	<u>A</u>	<u>B</u>
Çoktan Seçmeli	0-1	4	4
Açık Uçlu	0-1	4	4
	0-1-2	1	1
	<i>Toplam</i>	9	9
<i>Ortak Olmayan Test</i>			
<u>Madde Türü</u>	<u>Kategori Sayısı</u>	<u>A</u>	<u>B</u>
Çoktan Seçmeli	0-1	5	5
Açık Uçlu	0-1	2	2
	0-1-2	1	2
	0-1-2-3-4	1	0
	<i>Toplam</i>	9	9

Tablo 5 incelendiğinde Sosyal Bilgiler testi için B formuna eşitlenen A formunun madde sayısı, ortak madde sayısı, ortak madde formatı açısından aynı olduğu görülmektedir. Ortak olmayan maddelerde A testinde bir madde 0-1-2-3-4 şeklinde puanlanırken B testinde bu şekilde puanlanan madde bulunmamaktadır.

Tablo 6

Sosyal Bilgiler C-B Testlerinde Yer Alan Madde ve Ortak Madde Sayıları ve Formatları

Tüm Test			
<u>Madde Türü</u>	<u>Kategori Sayısı</u>	<u>C</u>	<u>B</u>
Çoktan Seçmeli	0-1	10	9
Açık Uçlu	0-1	4	6
	0-1-2	3	3
	0-1-2-3-4	1	0
	<i>Toplam</i>	18	18
Ortak Test			
<u>Madde Türü</u>	<u>Kategori Sayısı</u>	<u>C</u>	<u>B</u>
Çoktan Seçmeli	0-1	5	5
Açık Uçlu	0-1	2	2
	0-1-2	2	2
	<i>Toplam</i>	9	9
Ortak Olmayan Test			
<u>Madde Türü</u>	<u>Kategori Sayısı</u>	<u>C</u>	<u>B</u>
Çoktan Seçmeli	0-1	5	4
Açık Uçlu	0-1	2	4
	0-1-2	1	1
	0-1-2-3-4	1	0
	<i>Toplam</i>	9	9

Tablo 6 incelendiğinde Sosyal Bilgiler testi için B formuna eşitlenen C formunun madde sayısı, ortak madde sayısı, ortak madde formatı açısından aynı oldukları görülmektedir. Ortak olmayan maddelerin C testinde 5'i çoktan seçmeli iken B testinde 4'ü çoktan seçmelidir. Ortak olmayan açık uçlu maddelerden C testinde 0-1 puanlananların sayısı 2 iken B testinde bu sayı 4'tür. Benzer şekilde, ortak olmayan açık uçlu maddelerden C testinde 0-1-2-3-4 şeklinde puanlananların sayısı 1 iken B testinde 0-1-2-3-4 şeklinde puanlanan madde bulunmamaktadır.

Karma format olarak hazırlanan ABİDE testlerinde çok kategorili açık uçlu maddelerin puanlamasından kaynaklı olarak ortaya çıkabilecek durumları da değerlendirebilmek için maksimum kategori puanını alan öğrencilerin puanları 1, diğerlerinin puanları 0 olacak şekilde yeniden puanlama yapılmıştır. Bu şekilde yapılan puanlama ile testlerden alınabilecek maksimum puanlar testlerde yer alan 18 maddenin karşılığı olan 18 puandır. Bu şekilde puanlanan testlere bu çalışmada iki kategorili format test adı verilmiştir. Bu şekilde karma format testler çoktan seçmeli test formatına dönüştürülmüş ve çalışmalar hem karma hem de iki kategorili format üzerinden yürütülmüştür.

Testlerin Betimsel Analizi

Matematik testine ilişkin evren, 5000, 3000 ve 1000 kişilik örneklemeler üzerinden hesaplanan betimleyici istatistikler karma format için Tablo 7’de, iki kategorili format için Tablo 8’de verilmiştir.

Tablo 7

Matematik Testine İlişkin Betimsel İstatistikler (Karma Format)

Veri	Form	Cinsiyet	N	Min. Puan	Mak. Puan	Ortalama Doğru	Standart Sapma	Başarı Oranı (%)	Çarpıklık	Basıklık	
Evren	A	Toplam	11122	0	24	8.09	5.14	33.7	0.62	-0.39	
		Erkek	5587	0	24	8.09	5.27	33.7	0.64	-0.43	
		Kız	5535	0	23	8.08	5.00	33.7	0.60	-0.36	
	B	Toplam	11027	0	24	7.20	4.57	30.0	0.73	0.03	
		Erkek	5510	0	23	7.30	4.68	30.4	0.73	-0.01	
		Kız	5517	0	24	7.09	4.45	29.6	0.72	0.05	
	C	Toplam	10572	0	21	6.66	4.23	30.3	0.65	-0.23	
		Erkek	5346	0	21	6.52	4.26	29.6	0.71	-0.13	
		Kız	5226	0	21	6.81	4.19	31.0	0.60	-0.33	
	5000	A	Toplam	5000	0	24	8.14	5.12	33.9	0.62	-0.40
			Erkek	2506	0	24	8.15	5.27	34.0	0.65	-0.41
			Kız	2494	0	23	8.13	4.96	33.9	0.57	-0.40
B		Toplam	5000	0	23	7.16	4.59	29.8	0.73	-0.04	
		Erkek	2495	0	22	7.24	4.69	30.2	0.70	-0.10	
		Kız	2505	0	23	7.07	4.49	29.5	0.74	0.03	
C		Toplam	5000	0	21	6.65	4.24	30.2	0.66	-0.17	
		Erkek	2531	0	21	6.49	4.27	29.5	0.73	-0.03	
		Kız	2469	0	20	6.81	4.19	30.9	0.60	-0.30	
3000		A	Toplam	3000	0	23	8.09	5.07	33.7	0.58	-0.43
			Erkek	1508	0	22	8.10	5.21	33.7	0.60	-0.46
			Kız	1492	0	23	8.08	4.92	33.7	0.55	-0.42
	B	Toplam	3000	0	24	7.13	4.55	29.7	0.75	0.08	
		Erkek	1509	0	23	7.24	4.66	30.2	0.76	0.02	
		Kız	1491	0	24	7.03	4.43	29.3	0.74	0.11	
	C	Toplam	3000	0	21	6.68	4.26	30.4	0.65	-0.19	
		Erkek	1530	0	21	6.53	4.30	29.7	0.70	-0.06	
		Kız	1470	0	20	6.84	4.22	31.1	0.59	-0.33	
	1000	A	Toplam	1000	0	23	8.32	5.29	34.7	0.58	-0.52
			Erkek	503	0	22	8.27	5.43	34.5	0.60	-0.58
			Kız	497	0	23	8.37	5.15	34.9	0.57	-0.46
B		Toplam	1000	0	24	7.16	4.52	29.8	0.79	0.20	
		Erkek	509	0	23	7.33	4.73	30.5	0.77	0.00	
		Kız	491	0	24	6.98	4.30	29.1	0.79	0.40	
C		Toplam	1000	0	21	6.83	4.40	31.0	0.68	-0.11	
		Erkek	494	0	20	6.78	4.43	30.8	0.67	-0.14	
		Kız	506	0	21	6.87	4.37	31.2	0.69	-0.07	

Tablo 7 incelendiğinde çarpıklık ve basıklık değerlerinin ± 1 arasında kalması nedeniyle testlerin normal dağılım koşulunu sağladığı söylenebilir (Çokluk, Şekercioğlu, & Büyüköztürk, 2014). Evrene ve seçilen örneklemelere göre bazı testlerden alınan puanların maksimum değerlere ulaşamadığı görülmektedir. Örneğin; Tablo 4 incelendiğinde C testinden alınabilecek maksimum puanın 22 olduğu görülmektedir. Ancak C testinde bu puana ulaşabilen hiçbir öğrenci

bulunmamaktadır. Maksimum puana ulaşılabilen diğer testlerde ise alt gruplara göre bazı durumlarda yine maksimum puanlara ulaşamadığı görülmektedir. Başarı oranları incelendiğinde A testinin diğer testlere göre öğrencilere daha kolay geldiği söylenebilir. Testlerin çoğunda erkek öğrencilerin kız öğrencilerden daha başarılı olduğu görülmektedir.

Tablo 8

Matematik Testine İlişkin Betimsel İstatistikler (İki Kategorili Format)

Veri	Form	Cinsiyet	N	Min. Puan	Mak. Puan	Ortalama Doğru	Standart Sapma	Başarı Oranı %	Çarpıklık	Basıklık	
Evren		<i>Toplam</i>	11122	0	18	6.04	3.67	33.6	0.61	-0.32	
	A	Erkek	5587	0	18	6.07	3.76	33.7	0.62	-0.36	
		Kız	5535	0	17	6.02	3.58	33.4	0.60	-0.28	
	B	<i>Toplam</i>	11027	0	18	5.59	3.27	31.1	0.65	0.01	
		Erkek	5517	0	18	5.53	3.19	30.7	0.63	0.00	
	C	Kız	5510	0	17	5.66	3.35	31.5	0.66	0.00	
		<i>Toplam</i>	10572	0	17	5.14	3.23	28.5	0.74	0.10	
	5000	A	Erkek	5346	0	17	5.05	3.25	28.1	0.80	0.26
			Kız	5226	0	17	5.23	3.21	29.0	0.68	-0.04
		B	<i>Toplam</i>	5000	0	18	6.09	3.66	33.8	0.61	-0.31
			Erkek	2506	0	18	6.12	3.76	34.0	0.64	-0.33
		C	Kız	2494	0	17	6.06	3.54	33.7	0.57	-0.31
<i>Toplam</i>			5000	0	17	5.57	3.29	30.9	0.64	-0.06	
3000		A	Erkek	2495	0	16	5.61	3.36	31.2	0.63	-0.08
			Kız	2505	0	17	5.52	3.23	30.7	0.65	-0.05
		B	<i>Toplam</i>	5000	0	17	5.12	3.23	28.5	0.75	0.19
			Erkek	2531	0	17	5.02	3.25	27.9	0.82	0.41
		C	Kız	2469	0	16	5.22	3.21	29.0	0.69	-0.02
			<i>Toplam</i>	3000	0	17	6.04	3.63	33.5	0.56	-0.37
	1000	A	Erkek	1508	0	17	6.05	3.74	33.6	0.58	-0.42
			Kız	1492	0	17	6.02	3.52	33.4	0.54	-0.33
		B	<i>Toplam</i>	3000	0	18	5.56	3.26	30.9	0.68	0.09
			Erkek	1509	0	17	5.64	3.33	31.3	0.69	0.06
		C	Kız	1491	0	18	5.48	3.18	30.4	0.67	0.11
			<i>Toplam</i>	3000	0	17	5.14	3.25	28.6	0.74	0.16
1000		A	Erkek	1530	0	17	5.05	3.27	28.1	0.80	0.38
			Kız	1470	0	16	5.23	3.23	29.1	0.68	-0.07
		B	<i>Toplam</i>	1000	0	17	6.21	3.80	34.5	0.57	-0.47
			Erkek	503	0	16	6.18	3.88	34.4	0.57	-0.55
		C	Kız	497	0	17	6.23	3.72	34.6	0.57	-0.37
			<i>Toplam</i>	1000	0	18	5.58	3.23	31.0	0.71	0.20
	1000	A	Erkek	509	0	17	5.71	3.37	31.7	0.68	-0.01
			Kız	491	0	18	5.45	3.08	30.3	0.73	0.44
		B	<i>Toplam</i>	1000	0	17	5.24	3.37	29.1	0.79	0.24
			Erkek	494	0	16	5.25	3.41	29.1	0.77	0.20
		C	Kız	506	0	17	5.23	3.34	29.0	0.81	0.29

Tablo 8 incelendiğinde çarpıklık ve basıklık değerlerinin ± 1 arasında kalması nedeniyle testlerin normal dağılım koşulunu sağladığı söylenebilir (Çokluk, Şekercioğlu, & Büyüköztürk, 2014).

Sosyal Bilgiler testine ilişkin evren, 5000, 3000 ve 1000 kişilik örneklemeler üzerinden hesaplanan betimleyici istatistikler karma format için Tablo 9'da, iki kategorili format için Tablo 10'da verilmiştir.

Tablo 9

Sosyal Bilgiler Testine İlişkin Betimsel İstatistikler (Karma Format)

Veri	Form	Cinsiyet	N	Min. Puan	Mak. Puan	Ortalama Doğru	Standart Sapma	Başarı Oranı (%)	Çarpıklık	Basıklık	
Evren	A	Toplam	11327	0	23	11.36	4.81	49.4	-0.07	-0.80	
		Erkek	5718	0	23	11.03	5.01	47.9	-0.05	-0.90	
		Kız	5609	0	23	11.70	4.57	50.9	-0.04	-0.72	
	B	Toplam	11391	0	21	10.10	4.27	48.1	0.01	-0.67	
		Erkek	5737	0	21	9.88	4.47	47.0	0.03	-0.78	
		Kız	5654	0	21	10.32	4.05	49.1	0.03	-0.56	
	C	Toplam	10912	0	24	12.47	5.00	51.9	-0.19	-0.71	
		Erkek	5565	0	24	12.11	5.18	50.5	-0.17	-0.82	
		Kız	5347	0	24	12.83	4.78	53.5	-0.18	-0.63	
	5000	A	Toplam	5000	0	22	11.34	4.80	49.3	-0.06	-0.81
			Erkek	2547	0	22	10.92	4.95	47.5	-0.02	-0.89
			Kız	2453	0	22	11.78	4.60	51.2	-0.07	-0.74
B		Toplam	5000	0	21	10.11	4.23	48.2	0.00	-0.62	
		Erkek	2541	0	20	9.84	4.40	46.8	0.03	-0.70	
		Kız	2459	0	21	10.40	4.02	49.5	0.00	-0.55	
C		Toplam	5000	0	24	12.48	4.98	52.0	-0.17	-0.72	
		Erkek	2577	0	24	12.06	5.13	50.2	-0.13	-0.84	
		Kız	2423	0	24	12.92	4.77	53.8	-0.17	-0.61	
3000		A	Toplam	3000	0	22	11.33	4.84	49.2	-0.05	-0.81
			Erkek	1519	0	22	10.93	5.02	47.5	-0.03	-0.91
			Kız	1481	0	22	11.73	4.62	51.0	-0.04	-0.74
	B	Toplam	3000	0	21	10.05	4.25	47.9	0.02	-0.69	
		Erkek	1501	0	21	9.71	4.43	46.2	0.07	-0.79	
		Kız	1499	0	21	10.40	4.02	49.5	0.01	-0.60	
	C	Toplam	3000	0	24	12.55	4.97	52.3	-0.20	-0.71	
		Erkek	1515	0	24	12.16	5.08	50.7	-0.14	-0.81	
		Kız	1485	0	24	12.94	4.83	53.9	-0.26	-0.59	
	1000	A	Toplam	1000	0	22	11.49	4.82	50.0	-0.03	-0.76
			Erkek	502	0	22	11.20	4.99	48.7	-0.04	-0.80
			Kız	498	1	22	11.79	4.64	51.2	0.01	-0.75
B		Toplam	1000	0	19	10.14	4.20	48.3	-0.01	-0.68	
		Erkek	534	0	19	10.08	4.37	48.0	0.06	-0.76	
		Kız	466	0	19	10.21	4.00	48.6	-0.11	-0.58	
C		Toplam	1000	1	23	12.58	5.01	52.4	-0.24	-0.69	
		Erkek	539	1	23	12.40	5.11	51.7	-0.21	-0.78	
		Kız	461	1	23	12.79	4.90	53.3	-0.28	-0.58	

Tablo 9 incelendiğinde çarpıklık ve basıklık değerlerinin ± 1 arasında kalması nedeniyle testlerin normal dağılım koşulunu sağladığı söylenebilir (Çokluk, Şekercioğlu, & Büyüköztürk, 2014). Seçilen örneklemelere göre bazı testlerden alınan puanların maksimum değerlere ulaşamadığı görülmektedir. Örneğin; Tablo 5 incelendiğinde B testinden alınabilecek maksimum puanın 21 olduğu görülmektedir. Ancak B testinde 1000 kişilik örnekleme bu puana hiçbir öğrenci ulaşamamıştır. Başarı oranları incelendiğinde C testinin öğrencilere diğer testlerden daha kolay

geldiği söylenebilir. Tüm testlerde kız öğrencilerin erkek öğrencilerden daha başarılı olduğu görülmektedir.

Tablo 10

Sosyal Bilgiler Testine İlişkin Betimsel İstatistikler (İki Kategorili Format)

Veri	Form	Cinsiyet	N	Min. Puan	Mak. Puan	Ortalama Doğru	Standart Sapma	Başarı Oranı (%)	Çarpıklık	Basıklık
Evren		<i>Toplam</i>	11327	0	18	8.87	3.60	49.3	-0.04	-0.70
	A	Erkek	5718	0	18	8.61	3.76	47.8	-0.02	-0.81
		Kız	5609	0	18	9.14	3.41	50.8	-0.02	-0.62
	B	<i>Toplam</i>	11391	0	18	8.82	3.61	49.0	-0.05	-0.63
		Erkek	5737	0	18	8.60	3.78	47.8	-0.02	-0.75
	C	Kız	5654	0	18	9.03	3.41	50.2	-0.05	-0.52
<i>Toplam</i>		10912	0	18	9.22	3.63	51.2	-0.14	-0.65	
5000	A	Erkek	5565	0	18	8.98	3.78	49.9	-0.12	-0.77
		Kız	5347	0	18	9.47	3.46	52.6	-0.12	-0.55
	B	<i>Toplam</i>	5000	0	17	8.88	3.58	49.3	-0.02	-0.72
		Erkek	2547	0	17	8.55	3.70	47.5	0.01	-0.80
	C	Kız	2453	0	17	9.21	3.42	51.2	-0.02	-0.65
		<i>Toplam</i>	5000	0	18	8.82	3.57	49.0	-0.06	-0.57
3000	A	Erkek	2541	0	17	8.55	3.72	47.5	-0.01	-0.67
		Kız	2459	0	18	9.09	3.39	50.5	-0.07	-0.48
	B	<i>Toplam</i>	5000	0	18	9.23	3.62	51.3	-0.11	-0.67
		Erkek	2577	0	18	8.93	3.74	49.6	-0.08	-0.78
	C	Kız	2423	0	18	9.54	3.47	53.0	-0.11	-0.55
		<i>Toplam</i>	3000	0	17	8.88	3.60	49.3	-0.03	-0.71
1000	A	Erkek	1519	0	17	8.56	3.74	47.6	-0.01	-0.83
		Kız	1481	0	17	9.20	3.43	51.1	0.00	-0.62
	B	<i>Toplam</i>	3000	0	18	8.79	3.60	48.8	-0.03	-0.65
		Erkek	1501	0	18	8.47	3.76	47.1	0.03	-0.77
	C	Kız	1499	0	18	9.10	3.40	50.6	-0.05	-0.53
		<i>Toplam</i>	3000	0	18	9.29	3.63	51.6	-0.15	-0.66
1000	A	Erkek	1515	0	18	9.04	3.73	50.2	-0.08	-0.75
		Kız	1485	0	18	9.54	3.50	53.0	-0.19	-0.56
	B	<i>Toplam</i>	1000	0	17	9.00	3.61	50.0	-0.03	-0.70
		Erkek	502	0	17	8.77	3.72	48.7	-0.03	-0.73
	C	Kız	498	1	17	9.22	3.49	51.2	0.00	-0.68
		<i>Toplam</i>	1000	0	17	8.84	3.58	49.1	-0.07	-0.65
1000	B	Erkek	534	0	17	8.76	3.73	48.6	0.03	-0.75
		Kız	466	0	16	8.92	3.40	49.6	-0.20	-0.51
1000	C	<i>Toplam</i>	1000	0	17	9.26	3.69	51.4	-0.17	-0.68
		Erkek	539	0	17	9.13	3.77	50.7	-0.18	-0.74
		Kız	461	1	17	9.41	3.60	52.3	-0.14	-0.63

Tablo 10 incelendiğinde çarpıklık ve basıklık değerlerinin ± 1 arasında kalması nedeniyle testlerin normal dağılım koşulunu sağladığı söylenebilir (Çokluk, Şekercioğlu, & Büyüköztürk, 2014).

Testlerin Güvenirlik Analizi

Matematik ve Sosyal Bilgiler testinin A, B ve C formlarının evren ve örneklemeler üzerinden Factor 10.8.04 programıyla hesaplanan McDonald's Omega güvenirlik değerleri karma ve iki kategorili veri formatına göre Matematik testi için

Tablo 11'de, Sosyal Bilgiler testi için Tablo 12'de verilmiştir (Lorenzo-Seva & Ferrando, 2013).

Tablo 11

Matematik Testine İlişkin Güvenirlik Değerleri

Veri	Form	Cinsiyet	N	McDonald's Omega		
				Karma	İki Kategorili	
Evren	A	<i>Toplam</i>	11122	0.91	0.91	
		Erkek	5587	0.92	0.92	
		Kız	5535	0.91	0.91	
	B	<i>Toplam</i>	11027	0.89	0.90	
		Erkek	5510	0.90	0.90	
		Kız	5517	0.89	0.90	
	C	<i>Toplam</i>	10572	0.88	0.89	
		Erkek	5346	0.89	0.90	
		Kız	5226	0.87	0.90	
	5000	A	<i>Toplam</i>	5000	0.91	0.91
			Erkek	2506	0.92	0.92
			Kız	2494	0.90	0.88
B		<i>Toplam</i>	5000	0.90	0.90	
		Erkek	2495	0.90	0.90	
		Kız	2505	0.89	0.90	
C		<i>Toplam</i>	5000	0.88	0.89	
		Erkek	2531	0.89	0.90	
		Kız	2469	0.87	0.90	
3000		A	<i>Toplam</i>	3000	0.91	0.91
			Erkek	1508	0.92	0.92
			Kız	1492	0.90	0.90
	B	<i>Toplam</i>	3000	0.89	0.90	
		Erkek	1509	0.90	0.90	
		Kız	1491	0.89	0.82	
	C	<i>Toplam</i>	3000	0.88	0.89	
		Erkek	1530	0.89	0.90	
		Kız	1470	0.88	0.88	
	1000	A	<i>Toplam</i>	1000	0.92	0.87
			Erkek	503	0.92	0.89
			Kız	497	0.83	0.81
B		<i>Toplam</i>	1000	0.89	0.90	
		Erkek	509	0.80	0.84	
		Kız	491	0.76	0.75	
C		<i>Toplam</i>	1000	0.90	0.83	
		Erkek	494	0.90	0.78	
		Kız	506	0.67	0.80	

Tablo 11 incelendiğinde 1000 kişilik örneklem dışı, karma ve iki kategorili formata dönüştürülen Matematik testlerinin güvenilirlik değerlerinin A, B ve C formları için birbirine yakın ve yüksek olduğu görülmektedir. 1000 kişilik örneklem için karma formatta A, B ve C formlarının güvenilirlik değerlerinin birbirine yakın ve yüksek, iki kategorili formatta B formunun güvenilirlik değerinin A ve C formundan daha yüksek güvenilirlik değeri aldığı görülmektedir. 1000 kişilik

örneklemelerin alt grupları olan kız ve erkek öğrenciler için A, B ve C formlarının güvenilirlik değerleri farklılaşmaktadır. Güvenirlik değerlerinin 1000 kişilik örneklerde farklılaşması eşitleme koşullarını ihlal etse de çalışmanın bütünlüğünün bozulmaması için 1000 kişilik örnekler ve alt grupları araştırma dışında bırakılmamıştır.

Tablo 12

Sosyal Bilgiler Testine İlişkin Güvenirlik Değerleri

Veri	Form	Cinsiyet	N	McDonald's Omega		
				Karma	İki Kategorili	
Evren	A	<i>Toplam</i>	11327	0.88	0.88	
		Erkek	5718	0.90	0.89	
		Kız	5609	0.87	0.87	
	B	<i>Toplam</i>	11391	0.88	0.88	
		Erkek	5737	0.89	0.89	
		Kız	5654	0.86	0.86	
	C	<i>Toplam</i>	10912	0.87	0.87	
		Erkek	5565	0.88	0.88	
		Kız	5347	0.85	0.85	
	5000	A	<i>Toplam</i>	5000	0.88	0.88
			Erkek	2547	0.89	0.89
			Kız	2453	0.87	0.86
B		<i>Toplam</i>	5000	0.88	0.87	
		Erkek	2541	0.89	0.89	
		Kız	2459	0.86	0.86	
C		<i>Toplam</i>	5000	0.87	0.87	
		Erkek	2577	0.88	0.88	
		Kız	2423	0.85	0.85	
3000		A	<i>Toplam</i>	3000	0.88	0.88
			Erkek	1519	0.90	0.89
			Kız	1481	0.87	0.86
	B	<i>Toplam</i>	3000	0.88	0.88	
		Erkek	1501	0.89	0.89	
		Kız	1499	0.86	0.86	
	C	<i>Toplam</i>	3000	0.87	0.87	
		Erkek	1515	0.88	0.88	
		Kız	1485	0.86	0.86	
	1000	A	<i>Toplam</i>	1000	0.88	0.88
			Erkek	502	0.89	0.88
			Kız	498	0.87	0.88
B		<i>Toplam</i>	1000	0.87	0.87	
		Erkek	534	0.88	0.88	
		Kız	466	0.86	0.86	
C		<i>Toplam</i>	1000	0.87	0.87	
		Erkek	539	0.88	0.88	
		Kız	461	0.87	0.86	

Tablo 12 incelendiğinde karma ve iki kategorili formata dönüştürülen Sosyal Bilgiler testlerinin güvenilirlik değerlerinin A, B ve C formları için birbirine yakın ve yüksek olduğu görülmektedir.

Testlerin Analizi

Araştırma kapsamında ele alınan Matematik ve Sosyal Bilgiler testlerinin analizi iki farklı yolla gerçekleştirilmiştir.

İlk olarak testlerde yer alan çoktan seçmeli ve açık uçlu maddelerden alınan puanlar arasındaki ilişkileri incelenmek için çoktan seçmeli ve açık uçlu maddelerden alınan toplam puanlar hesaplanmıştır. Ardından toplam puanlar arasında hesaplanan korelasyonlar Matematik testi için Tablo 13'te, Sosyal Bilgiler testi için Tablo 14'te sunulmuştur.

Tablo 13

Matematik Testi için Açık Uçlu ve Çoktan Seçmeli Madde Toplam Puanları Arasındaki Korelasyon Değerleri

Veri	Form	Toplam	Erkek	Kız
Evren	A	0.658	0.661	0.655
	B	0.611	0.618	0.605
	C	0.591	0.588	0.596
5000	A	0.661	0.653	0.656
	B	0.615	0.620	0.610
	C	0.600	0.603	0.598
3000	A	0.651	0.650	0.654
	B	0.614	0.621	0.607
	C	0.595	0.602	0.590
1000	A	0.678	0.672	0.687
	B	0.618	0.642	0.588
	C	0.647	0.666	0.632

Tablo 13 incelendiğinde açık uçlu ve çoktan seçmeli maddelerden alınan toplam puanlar arasında pozitif yönlü ve nispeten orta düzeyden daha yüksek korelasyon olduğu görülmektedir. Bu korelasyon değerleri açık uçlu ve çoktan seçmeli maddelerin aynı yapıyı ölçtüğüne dair kanıt olarak kullanılabilir.

Tablo 14

Sosyal Bilgiler Testi Açık Uçlu ve Çoktan Seçmeli Madde Toplam Puanları Arasındaki Korelasyon Değerleri

Veri	Form	Toplam	Erkek	Kız
Evren	A	0.557	0.592	0.511
	B	0.582	0.610	0.546
	C	0.536	0.564	0.501
5000	A	0.557	0.577	0.529
	B	0.581	0.604	0.550
	C	0.536	0.562	0.502
3000	A	0.556	0.583	0.519
	B	0.582	0.609	0.547
	C	0.536	0.545	0.524
1000	A	0.575	0.592	0.555
	B	0.550	0.555	0.543
	C	0.520	0.569	0.464

Tablo 14 incelendiğinde açık uçlu ve çoktan seçmeli maddelerden alınan toplam puanlar arasında pozitif yönlü ve orta düzeyde korelasyon olduğu görülmektedir. Bu korelasyon değerleri açık uçlu ve çoktan seçmeli maddelerin aynı yapıyı ölçtüğüne dair kanıt olarak kullanılabilir.

Analizler ayrıca iki kategorili olarak puanlanan maddeler için tetrakorik, ikiden fazla kategoriden oluşan maddeler için polikorik korelasyonlar kullanılarak incelenmiştir. Bu amaçla, *Mplus7* programında EK-A'da yer alan syntax çalıştırılarak kategorik veriler için açımlayıcı faktör analizi gerçekleştirilmiştir. Sonuçların değerlendirilmesinde üç hususa dikkat edilmiştir. Bunlar; yamaç-birikinti grafiği (scree plot), birinci özdeğer tarafından açıklanan varyans oranı ve özdeğerlerin sıralı olarak birbirine oranıdır.

Yamaç-birikinti grafiğinde birinci özdeğerden sonra gelen ikinci özdeğerde birinci özdeğere göre önemli miktarda azalma ve ikinci özdeğerle birlikte azalan bir eğim tek boyutluluk göstergesi olarak kabul edilmektedir (Çokluk, Şekercioğlu, & Büyüköztürk, 2014). Yamaç-birikinti grafikleri karma format Matematik testi A, B ve C formu için sırasıyla EK-B, EK-C ve EK-Ç'de, iki kategorili format Matematik testi A, B ve C formu için sırasıyla EK-D, EK-E ve EK-F'de verilmiştir. Yamaç-birikinti grafikleri karma format Sosyal Bilgiler testi A, B ve C formu için sırasıyla EK-G, EK-

Ğ ve EK-H'de, iki kategorili format Sosyal Bilgiler testi A, B ve C formu için sırasıyla EK-I, EK-İ ve EK-J'de verilmiştir.

EK-B, EK-C, EK-Ç, EK-D, EK-E ve EK-F'de yer alan yamaç-birikinti grafikleri incelendiğinde evren ve örneklem verilerine göre Matematik A, B ve C testlerinin 1000 kişilik örneklem dışında yeterince tek boyutlu olduğu görülmektedir. 1000 kişilik örneklem ve alt grubu olan erkek ve kız öğrencilerde ise birden çok boyut olma ihtimali söz konusu olmakla birlikte boyutluluğa karar vermek için varyans açıklama oranı ve art arda gelen özdeğerlerin oranlarına da dikkat etmek gerekmektedir.

EK-G, EK-Ğ, EK-H, EK-I, EK-İ ve EK-J'de yer alan yamaç-birikinti grafikleri incelendiğinde evren ve örneklem verilerine göre Sosyal Bilgiler A, B ve C testlerinin yeterince tek boyutlu olduğu görülmektedir. Ancak boyutluluğa net olarak karar verebilmek için art arda gelen özdeğerlerin birbirlerine oranlarına da dikkat etmek gerekmektedir.

Yamaç-birikinti grafiklerinde Sosyal Bilgiler testi için tüm örneklem büyüklüklerinde tek boyutlu bir yapı olduğu izlenimi varken Matematik testi için 1000 kişilik örneklem büyüklüğünde tek boyutluluğun sağlanamadığına ilişkin kuşku ortaya çıkmaktadır. Bu nedenle birinci özdeğer tarafından açıklanan varyans ve art arda gelen özdeğerlerin birbirlerine oranları da dikkate alınarak tek boyutluluğa karar verilmeye çalışılmıştır.

Birinci özdeğer tarafından açıklanan varyans oranının olabildiğince yüksek olması (%20'den büyük) tek boyutluluğun varlığı konusunda kanıt sunmaktadır (Reckase, 1979). Birinci özdeğerin ikinci özdeğere oranlanması (1/2) ile elde edilen değer diğer özdeğerlerin birbirlerine oranlanması (2/3, 3/4) sonucu elde edilen değerden büyük olması tek boyutluluk için dikkate alınabilmektedir. Bu kapsamda; tek boyutluluğu değerlendirebilmek için karma format ve iki kategorili format Matematik testlerine uygulanan açımlayıcı faktör analizi sonucu elde edilen açıklanan varyans oranları ve özdeğerlerin oranları sırasıyla Tablo 15 ve Tablo 16'da verilmiştir.

Tablo 15

Matematik Testine İlişkin Gözlem Sayıları ve Açıklanan Varyans Yüzdeleri (Karma Format)

Test Adı	Form	Veri	Gözlem Sayısı	Birinci Özdeğer Tarafından Açıklanan Varyans (%)	Art Arda Gelen Özdeğerlerin Oranı		
					1/2	2/3	3/4
MATEMATİK	A	Evren	11122	43.36	6.61	1.13	1.06
		<i>Erkek</i>	5587	45.53	6.80	1.18	1.07
		<i>Kız</i>	5535	41.39	6.14	1.13	1.03
		Örnekleme	5000	43.48	6.48	1.12	1.08
		<i>Erkek</i>	2506	45.95	6.64	1.12	1.17
		<i>Kız</i>	2494	41.30	6.15	1.05	1.11
		Örnekleme	3000	41.94	6.19	1.12	1.02
		<i>Erkek</i>	1508	44.23	6.28	1.05	1.17
		<i>Kız</i>	1492	40.02	5.30	1.24	1.05
		Örnekleme	1000	45.01	6.31	1.18	1.08
		<i>Erkek</i>	503	47.74	6.12	1.27	1.00
		<i>Kız</i>	497	43.14	5.21	1.19	1.11
	B	Evren	11027	39.52	5.19	1.31	1.04
		<i>Erkek</i>	5510	41.01	5.43	1.31	1.02
		<i>Kız</i>	5517	38.02	4.89	1.25	1.12
		Örnekleme	5000	39.90	5.27	1.31	1.05
		<i>Erkek</i>	2495	41.21	5.43	1.31	1.05
		<i>Kız</i>	2505	38.66	5.00	1.15	1.21
		Örnekleme	3000	39.56	5.20	1.35	1.02
		<i>Erkek</i>	1509	41.79	5.74	1.21	1.11
		<i>Kız</i>	1491	37.21	4.47	1.25	1.09
		Örnekleme	1000	39.36	4.63	1.26	1.20
		<i>Erkek</i>	509	40.76	4.55	1.03	1.52
		<i>Kız</i>	491	38.39	4.19	1.19	1.11
	C	Evren	10572	34.31	5.13	1.09	1.09
		<i>Erkek</i>	5346	35.42	5.20	1.16	1.04
		<i>Kız</i>	5226	33.31	4.80	1.08	1.04
		Örnekleme	5000	34.68	4.91	1.16	1.05
		<i>Erkek</i>	2531	36.08	4.88	1.25	1.02
		<i>Kız</i>	2469	33.40	4.58	1.06	1.13
		Örnekleme	3000	35.22	5.28	1.03	1.10
		<i>Erkek</i>	1530	36.73	4.92	1.21	1.03
		<i>Kız</i>	1470	33.91	4.44	1.21	1.01
		Örnekleme	1000	37.29	4.59	1.20	1.07
		<i>Erkek</i>	494	36.67	4.26	1.18	1.05
		<i>Kız</i>	506	38.11	4.50	1.08	1.15

Tablo 15 incelendiğinde karma format Matematik A, B ve C formlarından elde edilen varyans açıklama oranlarının farklılaştığı görülürken, birinci özdeğer tarafından açıklanan varyans oranı hiçbir örnekleme %33'ün altına düşmemiştir. Kız ve erkek değişkenleri için hesaplanan açıklanan varyans oranları genellikle erkekler lehine daha yüksek olmakla birlikte bu değerler evren ve örnekleme büyüklüklerine göre bir miktar değişmektedir. Birinci özdeğerin ikinci özdeğere

oranın yüksek değerler vermesi, ikinci özdeğerin üçüncü özdeğere ve üçüncü özdeğerin dördüncü özdeğere oranının 1'e yakın olması tek boyutluluk açısından önemli ipuçları vermektedir. Ayrıca özdeğerlerin oranı incelendiğinde birinci özdeğerin ikinci özdeğerden en az 4.19 kat daha büyük olduğu görülmektedir.

Tablo 16

Matematik Testine İlişkin Gözlem Sayıları ve Açıklanan Varyans Yüzdeleri (İki Kategorili Format)

Test Adı	Form	Veri	Gözlem Sayısı	Birinci Özdeğer Tarafından Açıklanan Varyans (%)	Art Arda Gelen Özdeğerlerin Oranı		
					1/2	2/3	3/4
MATEMATİK	A	Evren	11122	43.02	6.31	1.19	1.04
		<i>Erkek</i>	5587	44.97	6.53	1.22	1.07
		<i>Kız</i>	5535	41.33	5.84	1.20	1.03
		Örneklem	5000	43.36	6.49	1.07	1.12
		<i>Erkek</i>	2506	45.61	6.77	1.02	1.27
		<i>Kız</i>	2494	41.24	6.07	1.03	1.14
		Örneklem	3000	41.10	5.48	1.25	1.02
		<i>Erkek</i>	1508	43.16	5.49	1.22	1.14
		<i>Kız</i>	1492	39.17	4.68	1.33	1.07
		Örneklem	1000	43.61	5.21	1.33	1.12
		<i>Erkek</i>	503	45.68	5.09	1.44	1.01
		<i>Kız</i>	497	40.91	3.85	1.52	1.15
		Evren	11027	40.29	5.11	1.27	1.09
		<i>Erkek</i>	5510	41.52	5.30	1.26	1.10
		<i>Kız</i>	5517	38.96	4.77	1.25	1.15
		Örneklem	5000	40.19	5.05	1.30	1.11
		<i>Erkek</i>	2495	41.37	5.07	1.32	1.06
		<i>Kız</i>	2505	38.91	4.73	1.15	1.30
	B	Örneklem	3000	40.40	5.00	1.35	1.06
		<i>Erkek</i>	1509	41.88	5.21	1.26	1.19
		<i>Kız</i>	1491	38.61	4.46	1.26	1.08
		Örneklem	1000	38.89	4.10	1.31	1.23
		<i>Erkek</i>	509	39.21	3.96	1.09	1.52
		<i>Kız</i>	491	37.04	3.47	1.36	1.20
		Evren	10572	35.35	4.87	1.17	1.08
		<i>Erkek</i>	5346	36.24	5.09	1.20	1.03
		<i>Kız</i>	5226	34.64	4.66	1.10	1.10
		Örneklem	5000	35.65	4.63	1.21	1.09
		<i>Erkek</i>	2531	36.63	4.68	1.27	1.05
		<i>Kız</i>	2469	34.88	4.48	1.13	1.09
	C	Örneklem	3000	35.98	4.85	1.11	1.14
		<i>Erkek</i>	1530	37.07	4.56	1.25	1.07
		<i>Kız</i>	1470	35.16	4.48	1.13	1.13
		Örneklem	1000	38.67	4.67	1.18	1.09
		<i>Erkek</i>	494	37.48	4.17	1.14	1.05
		<i>Kız</i>	506	39.92	4.68	1.09	1.11

Tablo 16 incelendiğinde Matematik (iki kategorili format) A, B ve C formlarından elde edilen varyans açıklama oranlarının farklılaştığı görülürken, birinci boyut tarafından açıklanan varyans oranı hiçbir örnekleme %34'ün altına düşmemiştir. Kız ve erkek değişkenleri için hesaplanan açıklanan varyans oranları genellikle erkekler lehine daha yüksek olmakla birlikte bu değerler evren ve örneklem büyüklüklerine göre bir miktar değişmektedir. Birinci özdeğerin ikinci özdeğere oranının yüksek değerler vermesi, ikinci özdeğerin üçüncü özdeğere ve üçüncü özdeğerin dördüncü özdeğere oranının 1'e yakın olması tek boyutluluk açısından önemli ipuçları vermektedir. Ayrıca özdeğerlerin oranı incelendiğinde birinci özdeğerin ikinci özdeğerden en az 3.47 kat daha büyük olduğu görülmektedir.

Açıklanan varyans oranlarının her bir evren ve örneklem için %20'den daha fazla olması ve birinci özdeğerin ikinci özdeğere oranının çoğunlukla 4 kattan daha fazla olması testlerin tek boyutlu olduğu konusunda kanıtlar sunmaktadır. Yamaç-birikinti grafiklerinde özellikle 1000 kişilik örneklem büyüklüğünde tek boyutluluğun sağlanamadığı konusunda şüpheler olmakla birlikte açıklanan varyans oranların yeterince yüksek olması ve özdeğer oranlarının oldukça yüksek olması nedeniyle tüm testlerin tek boyut olarak kabul edilebileceği görülmektedir.

Sosyal Bilgiler testinin tek boyutluluğunu değerlendirebilmek için karma format ve iki kategorili format testlere uygulanan açımlayıcı faktör analizi sonucu elde edilen açıklanan varyans oranları ve özdeğerlerin oranları sırasıyla Tablo 17 ve Tablo 18'de verilmiştir.

Tablo 17 incelendiğinde karma format Sosyal Bilgiler A, B ve C formlarından elde edilen varyans açıklama oranlarının farklılaştığı görülürken, birinci özdeğer tarafından açıklanan varyans oranı hiçbir örnekleme %28'in altına düşmemiştir. Kız ve erkek değişkenleri için hesaplanan açıklanan varyans oranları genellikle erkekler lehine daha yüksek olmakla birlikte bu değerler evren ve örneklem büyüklüklerine göre bir miktar değişmektedir. Birinci özdeğerin ikinci özdeğere oranının yüksek değerler vermesi, ikinci özdeğerin üçüncü özdeğere ve üçüncü özdeğerin dördüncü özdeğere oranının 1'e yakın olması tek boyutluluk açısından önemli ipuçları vermektedir. Ayrıca özdeğerlerin oranı incelendiğinde birinci özdeğerin ikinci özdeğerden en az 4.31 kat daha büyük olduğu görülmektedir.

Tablo 17

*Sosyal Bilgiler Testine İlişkin Gözlem Sayıları ve Açıklanan Varyans Yüzdeleri
(Karma Format)*

Test Adı	Form	Veri	Gözlem Sayısı	Birinci Özdeğer Tarafından Açıklanan Varyans (%)	Art Arda Gelen Özdeğerlerin Oranı			
					1/2	2/3	3/4	
SOSYAL BİLGİLER	A	Evren	11327	35.78	5.83	1.07	1.08	
		Erkek	5718	38.11	6.21	1.07	1.08	
		Kız	5609	33.03	5.30	1.08	1.05	
		Örnekleme	5000	35.51	5.78	1.04	1.10	
		Erkek	2547	37.09	5.99	1.02	1.14	
		Kız	2453	33.47	5.22	1.07	1.03	
		Örnekleme	3000	36.05	5.94	1.05	1.04	
		Erkek	1519	38.06	6.01	1.08	1.03	
		Kız	1481	33.52	5.20	1.02	1.09	
		Örnekleme	1000	36.09	5.36	1.07	1.09	
		Erkek	502	37.92	5.27	1.03	1.09	
		Kız	498	34.24	4.58	1.11	1.15	
		Evren	11391	33.89	5.62	1.13	1.06	
		Erkek	5737	36.28	5.81	1.16	1.14	
		Kız	5654	31.41	5.33	1.07	1.05	
		Örnekleme	5000	33.31	5.39	1.11	1.08	
		Erkek	2541	35.62	5.61	1.12	1.11	
		Kız	2459	30.90	4.79	1.12	1.00	
	B	Örnekleme	3000	34.14	5.73	1.08	1.07	
		Erkek	1501	36.58	5.91	1.09	1.08	
		Kız	1499	31.39	5.00	1.09	1.03	
		Örnekleme	1000	33.27	4.78	1.12	1.10	
		Erkek	534	35.17	4.73	1.16	1.07	
		Kız	466	31.47	4.31	1.07	1.05	
		Evren	10912	31.44	4.97	1.15	1.09	
		Erkek	5565	33.87	5.46	1.14	1.08	
		Kız	5347	28.77	4.42	1.15	1.07	
		Örnekleme	5000	31.42	4.86	1.15	1.10	
		Erkek	2577	33.57	5.20	1.14	1.10	
		Kız	2423	28.91	4.30	1.14	1.08	
		C	Örnekleme	3000	31.54	4.85	1.16	1.08
			Erkek	1515	33.34	5.13	1.07	1.06
			Kız	1485	29.70	4.33	1.19	1.03
			Örnekleme	1000	31.63	4.97	1.04	1.13
			Erkek	539	31.89	4.66	1.01	1.16
			Kız	461	32.00	4.72	1.05	1.08

Tablo 18 incelendiğinde iki kategorili format Sosyal Bilgiler A, B ve C formlarından elde edilen varyans açıklama oranlarının farklılaştığı görülürken, birinci tarafından açıklanan varyans oranı hiçbir örnekleme %28'in altına düşmemiştir. Kız ve erkek değişkenleri hesaplanan açıklanan varyans oranları erkekler lehine daha yüksek olmakla birlikte bu değerler evren ve örnekleme büyüklüklerine göre bir miktar değişmektedir. Birinci özdeğerin ikinci özdeğere oranının yüksek değerler

vermesi, ikinci özdeğerin üçüncü özdeğere ve üçüncü özdeğerin dördüncü özdeğere oranının 1'e yakın olması tek boyutluluk açısından önemli ipuçları vermektedir. Ayrıca özdeğerlerin oranı incelendiğinde birinci özdeğerin ikinci özdeğerden en az 4.06 kat daha büyük olduğu görülmektedir.

Tablo 18

Sosyal Bilgiler Testine İlişkin Gözlem Sayıları ve Açıklanan Varyans Yüzdeleri (İki Kategorili Format)

Test Adı	Form	Veri	Gözlem Sayısı	Birinci Özdeğer Tarafından Açıklanan Varyans (%)	Art Arda Gelen Özdeğerlerin Oranı		
					1/2	2/3	3/4
SOSYAL BİLGİLER	A	Evren	11327	35.29	5.68	1.09	1.08
		<i>Erkek</i>	5718	37.37	6.01	1.09	1.09
		<i>Kız</i>	5609	32.81	5.19	1.10	1.04
		Örnekleme	5000	34.99	5.60	1.05	1.10
		<i>Erkek</i>	2547	36.40	5.80	1.02	1.16
		<i>Kız</i>	2453	33.10	5.03	1.10	1.02
		Örnekleme	3000	35.47	5.73	1.06	1.04
		<i>Erkek</i>	1519	37.26	5.63	1.11	1.06
		<i>Kız</i>	1481	33.13	5.11	1.02	1.07
		Örnekleme	1000	35.72	5.29	1.07	1.08
		<i>Erkek</i>	502	36.96	5.13	1.03	1.08
		<i>Kız</i>	498	34.54	4.58	1.11	1.12
	B	Evren	11391	33.56	5.49	1.13	1.04
		<i>Erkek</i>	5737	36.01	5.71	1.15	1.12
		<i>Kız</i>	5654	31.01	5.12	1.08	1.07
		Örnekleme	5000	32.85	5.22	1.11	1.04
		<i>Erkek</i>	2541	35.20	5.48	1.11	1.09
		<i>Kız</i>	2459	30.45	4.58	1.12	1.03
		Örnekleme	3000	34.02	5.63	1.06	1.07
		<i>Erkek</i>	1501	36.42	5.81	1.08	1.10
		<i>Kız</i>	1499	31.38	4.85	1.07	1.08
		Örnekleme	1000	33.17	4.60	1.07	1.16
		<i>Erkek</i>	534	34.93	4.51	1.09	1.15
		<i>Kız</i>	466	31.58	4.06	1.08	1.15
	C	Evren	10912	31.21	5.00	1.14	1.08
		<i>Erkek</i>	5565	33.62	5.46	1.13	1.04
		<i>Kız</i>	5347	28.59	4.45	1.13	1.06
		Örnekleme	5000	31.29	4.87	1.14	1.08
		<i>Erkek</i>	2577	33.38	5.22	1.11	1.06
		<i>Kız</i>	2423	28.84	4.31	1.13	1.05
		Örnekleme	3000	31.49	4.83	1.15	1.06
		<i>Erkek</i>	1515	33.62	5.08	1.09	1.03
		<i>Kız</i>	1485	29.38	4.29	1.16	1.06
		Örnekleme	1000	31.44	4.93	1.04	1.00
		<i>Erkek</i>	539	31.91	4.31	1.09	1.15
		<i>Kız</i>	461	31.35	4.19	1.09	1.06

Açıklanan varyans oranlarının her bir evren ve örneklem için %20'den daha fazla olması ve birinci özdeğerin ikinci özdeğere oranının 4 kattan daha fazla olması testlerin tek boyutlu olduğu konusunda kanıtlar sunmaktadır. Yamaç-birikinti grafiklerinde birinci özdeğerden sonra keskin bir düşüş olması, varyans oranların yeterince yüksek olması ve özdeğer oranlarının oldukça yüksek olması nedeniyle tüm testlerin tek boyut olarak kabul edilebileceği görülmektedir.

Madde Parametrelerinin Kestirimi ve Ölçek Dönüştürme

Testlerin A, B ve C formlarında yer alan maddelerin parametre kestirimlerinde çoktan seçmeli maddeler için 3 parametrelili lojistik (3PL) model, 0-1 kategorili açık uçlu maddeler için 2 parametrelili (2PL) model ve kategori sayısı 2'den fazla olan açık uçlu maddeler için aşamalı tepki (GR) modeli kullanılmıştır. Veri model uyumunda bazı maddeler modele uymamasına rağmen bu maddelerin çalışma kapsamında diğer analizleri de etkileyeceği düşünülerek analiz dışı bırakılmamasına karar verilmiştir.

Parametrelerin elde edilmesinden sonra A ve C testlerinden elde edilen parametreler temel test olarak alınan B testinin ölçeğine dönüşümlerinde R programında yer alan *plink* paketi kullanılmıştır (Jonathan, 2010). Ölçek dönüşümleri, moment yöntemleri olan ortalama-ortalama ve ortalama-standart sapma ve karakteristik eğri yöntemleri olan Haebara ve Stocking-Lord için yapılmıştır.

Eşitleme

MTK gözlenen ve gerçek puan eşitlemesi R programında yer alan *plink* paketiyle gerçekleştirilmiştir. Aynı dağılım özelliğinin (ADÖ) korunumunu değerlendirmek için parametrik olmayan bir yöntem olan ve Eşitlik 27'de yer alan Kolmogorov-Smirnov T istatistiği ve eşitlik özelliğinin korunumunun değerlendirilmesi için birinci sıra eşitlik (BSE) ve ikinci sıra eşitlik (İSE) değerleri kullanılmıştır. BSE ve İSE'yi değerlendirmek için gerekli olan ve Eşitlik 28 ve 29'da verilen D1 ve D2 indeksleri hesaplanmıştır.

$$T = \sup_{i=k} |Y(i) - X(k)| \quad (27)$$

Eşitlik 27'de yer alan $Y(i)$ ve $X(k)$; Y ve X formlarının yüzde olarak ifade edilen birikimli dağılım fonksiyonlarını (*bdf*) gösterirken, Kolmogorov T ; $i = k$ olduğu durumlarda iki *bdf* arasındaki en büyük farkı göstermektedir. Küçülen T değeri iki dağılımın daha benzer olduğunu gösterirken büyüyen T değeri dağılımların daha az benzer olduğunu göstermektedir.

$$D_1 = \frac{\sqrt{\sum[q_i(y_{SC}|\theta_i - x_{SC}|\theta_i)^2]}}{SD_Y} \quad (28)$$

Eşitlik 28'de;

$y_{SC}|\theta_i$: i. quadrature noktasındaki θ_i yetenek düzeyi için Y formunun beklenen ölçek puanını,

$x_{SC}|\theta_i$: i. quadrature noktasındaki θ_i yetenek düzeyi için X formunun beklenen ölçek puanını,

q_i : θ_i 'deki quadrature ağırlığını,

SD_Y : Y formu ölçek puanlarının standart sapmasını

göstermektedir. D_1 'in küçülen değerleri BSE'nin daha iyi korunduğunu göstermektedir.

$$D_2 = \frac{\sqrt{\sum[q_i(y_{SEM}|\theta_i - x_{SEM}|\theta_i)^2]}}{SD_Y} \quad (29)$$

Eşitlik 29'da;

$y_{SEM}|\theta_i$: Y formu için i. quadrature noktasında ve θ_i yetenek düzeyindeki bireylere ait koşullu ölçmenin standart hatasını,

$x_{SEM}|\theta_i$: X formu için i. quadrature noktasında ve θ_i yetenek düzeyindeki bireylere ait koşullu ölçmenin standart hatasını,

SD_Y : Y formu ölçek puanlarının standart sapmasını

göstermektedir. D_2 'nin küçülen değerleri İSE'nin daha iyi korunduğunu göstermektedir.

D_1 ve D_2 indekslerinin hesaplanmasında kullanılan koşullu ortalama ve koşullu standart sapmalar POLYCSEM (Kolen, 2004) programı aracılığı ile elde edilmiştir. POLYCSEM (Kolen, 2004) programının çalıştırılabilmesi için madde

parametreleri, quadrature noktalar ve bu noktaların ağırlıkları ile eşitleme sonrası oluşan X testindeki ham puanların Y testindeki karşılıklarının hesaplanmış olması gerekmektedir. Bu kapsamda, ihtiyaç duyulan madde parametreleri IRTPRO 4.2 programı ile kestirilmiştir (Cai, Thissen, & du Toit, 2017). 40 adet quadrature noktası ve bu quadrature noktaların ağırlıkları için PARSCALE (Kolen, 2004) programının çalıştırılmasıyla Phase 2'de oluşan veriler kullanılmıştır

Bölüm 4

Bulgular ve Yorumlar

Bu bölümde araştırma kapsamında elde edilen bulgular ve bulgular ışığında yapılan yorumlar yer almaktadır.

Alt problemlere ait bulgulara geçilmeden önce Matematik ve Sosyal Bilgiler testindeki ham puan dağılımlarına ilişkin grafikler örneklem büyüklüğü ve cinsiyete göre oluşturularak sırasıyla EK-K, EK-L, EK-M ve EK-N'de verilmiştir.

Ölçek dönüştürme sonrası oluşan yeni dağılımlarına ilişkin grafikler örneklem büyüklüğü ve cinsiyete göre Matematik testi A formu için EK-O, EK-Ö, C formu için EK-P, EK-R'de, Sosyal Bilgiler testi A formu için EK-S, EK-Ş, C formu için EK-T, EK-U'da verilmiştir.

Tabloların değerlendirilmesinde birçok ölçüt olmasından dolayı öncelikle alt problemlerin ana temalarına ağırlık verilmiş olup daha detaylı açıklamalar alt problemlerin son paragraflarında verilmiştir.

Birinci Alt Probleme Ait Bulgular ve Yorumlar

Matematik testi için denk olmayan gruplar ortak madde deseninde gerçek/gözlenen puan eşitleme yöntemi ile A formunun B formuna eşitlenmesinde aynı dağılım ve eşitlik özelliğinin korunumu;

- a. maddelerin formatına (çok kategorili puanlanan, iki kategorili puanlanan),*
- b. testi alan alt gruplara (kız, erkek),*
- c. örneklem büyüklüğüne (1000, 3000 ve 5000) ve*
- d. ölçek dönüştürme yöntemlerine (O-O, O-S, HB, SL) göre nasıl değişmektedir?*

Matematik testinde ham puan T değerleri ve A formunun B formuna gerçek puan (Ger_Pn) ve gözlenen puan ($Göz_Pn$) eşitleme yöntemiyle eşitlenmesi sonucu oluşan T değerleri Tablo 19'da, birinci sıra eşitlik (BSE) değerleri Tablo 20'de, ikinci sıra eşitlik (İSE) değerleri Tablo 21'de verilmiştir.

Tablo 19

Matematik Testi A Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi - Ham Puan ve Ölçek Puan Dağılımları için Kolmogorov T İstatistikleri (%)

F	Örn.	Grup	Ham Puan Dağılımları	O-O		O-S		HB		SL	
				Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn
1	1000	Toplam	9.70	10.10	8.55	3.66	3.41	5.63	4.70	3.54	3.30
		Erkek	8.69	13.99	11.48	4.37	4.12	6.34	5.72	5.05	4.84
		Kız	12.29	7.01	7.05	6.82	6.85	6.30	6.33	5.47	5.64
1	3000	Toplam	8.27	7.78	7.04	5.13	4.69	4.81	3.93	3.58	3.04
		Erkek	7.71	10.04	8.72	5.56	4.48	5.23	4.41	3.22	3.21
		Kız	9.69	5.76	5.73	3.72	3.69	5.97	5.95	3.72	3.69
1	5000	Toplam	7.90	7.59	6.84	2.85	2.60	3.58	3.23	3.56	3.20
		Erkek	7.05	8.18	6.94	2.80	2.86	2.58	2.67	3.05	2.52
		Kız	9.45	5.91	5.08	4.39	4.52	5.37	4.61	4.34	4.45
2	1000	Toplam	8.80	13.57	11.47	9.95	7.90	10.67	8.60	9.03	6.98
		Erkek	7.73	16.99	13.26	8.69	7.10	10.29	7.99	8.04	6.62
		Kız	9.94	10.98	10.21	11.39	10.64	10.79	10.14	9.29	8.80
2	3000	Toplam	6.43	12.08	11.12	10.68	9.62	8.77	7.65	8.18	7.12
		Erkek	6.38	15.00	12.12	10.75	9.27	8.95	7.44	8.40	6.97
		Kız	7.42	10.67	9.82	10.94	10.07	9.20	8.23	8.24	7.35
2	5000	Toplam	5.78	10.08	8.47	5.79	4.81	6.80	6.06	6.58	5.21
		Erkek	6.44	11.28	9.06	5.65	4.81	6.14	5.34	5.43	4.49
		Kız	6.64	9.23	8.26	7.47	6.48	8.06	7.02	8.07	7.04

F: Madde Formatı, Örn: Örneklem, 1: Karma Format Test, 2: İki Kategorili Formata Dönüştürülen Test

Tablo 20

Matematik Testinde A Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan BSE Değerleri

F	Örn.	Grup	O-O		O-S		HB		SL	
			Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn
1	1000	Toplam	0.2519	0.2408	0.2017	0.1978	0.2496	0.2418	0.2094	0.2082
		Erkek	0.2935	0.2810	0.2000	0.1956	0.2611	0.2535	0.2149	0.2130
		Kız	0.1746	0.1711	0.1738	0.1710	0.2128	0.2073	0.1932	0.2054
1	3000	Toplam	0.2042	0.1976	0.1855	0.1817	0.2260	0.2206	0.2001	0.1959
		Erkek	0.2498	0.2403	0.2033	0.1986	0.2398	0.2340	0.2083	0.2004
		Kız	0.1667	0.1630	0.1656	0.1623	0.2030	0.1990	0.1928	0.2010
1	5000	Toplam	0.1957	0.1893	0.1693	0.1684	0.1759	0.1739	0.1887	0.1826
		Erkek	0.1856	0.1793	0.1399	0.1400	0.1437	0.1434	0.1900	0.1834
		Kız	0.2033	0.1978	0.1991	0.1952	0.2176	0.2116	0.1911	0.1849
2	1000	Toplam	0.2749	0.2492	0.2027	0.2075	0.2466	0.2245	0.1905	0.1937
		Erkek	0.3120	0.2968	0.1955	0.1877	0.2426	0.2317	0.1995	0.1949
		Kız	0.1863	0.1756	0.1952	0.1837	0.2666	0.2478	0.1753	0.1937
2	3000	Toplam	0.2429	0.2365	0.2125	0.2079	0.2583	0.2470	0.1944	0.1878
		Erkek	0.2870	0.2747	0.2228	0.2158	0.2638	0.2524	0.1984	0.1896
		Kız	0.2030	0.1970	0.2085	0.2018	0.2554	0.2416	0.1897	0.1884
2	5000	Toplam	0.1679	0.1631	0.1087	0.1057	0.1214	0.1170	0.1478	0.1375
		Erkek	0.1784	0.1737	0.0812	0.0797	0.1034	0.0996	0.1495	0.1411
		Kız	0.1570	0.1527	0.1370	0.1326	0.1497	0.1446	0.1487	0.1421

F: Madde Formatı, Örn: Örneklem, 1: Karma Format Test, 2: İki Kategorili Formata Dönüştürülen Test

Tablo 21

Matematik Testinde A Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan İSE Değerleri

F	Örn.	Grup	O-O		O-S		HB		SL	
			Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn
1	1000	Toplam	0.0735	0.0651	0.0465	0.0427	0.0618	0.0567	0.0483	0.0471
		Erkek	0.0931	0.0793	0.0463	0.0431	0.0655	0.0600	0.0515	0.0508
		Kız	0.0423	0.0373	0.0410	0.0359	0.0523	0.0467	0.0432	0.0390
1	3000	Toplam	0.0596	0.0546	0.0485	0.0450	0.0575	0.0533	0.0496	0.0475
		Erkek	0.0785	0.0692	0.0553	0.0500	0.0637	0.0580	0.0557	0.0521
		Kız	0.0431	0.0407	0.0420	0.0397	0.0502	0.0472	0.0443	0.0420
1	5000	Toplam	0.0546	0.0496	0.0383	0.0354	0.0414	0.0381	0.0444	0.0421
		Erkek	0.0535	0.0474	0.0292	0.0267	0.0310	0.0285	0.0428	0.0409
		Kız	0.0521	0.0479	0.0485	0.0445	0.0548	0.0503	0.0457	0.0426
2	1000	Toplam	0.1136	0.0864	0.0825	0.0729	0.0969	0.0769	0.0763	0.0705
		Erkek	0.1308	0.1039	0.0776	0.0649	0.0935	0.0785	0.0770	0.0683
		Kız	0.0862	0.0697	0.0906	0.0728	0.1111	0.0924	0.0788	0.0729
2	3000	Toplam	0.1016	0.0846	0.0893	0.0752	0.0974	0.0845	0.0778	0.0706
		Erkek	0.1188	0.0964	0.0925	0.0767	0.0990	0.0846	0.0793	0.0702
		Kız	0.0863	0.0729	0.0886	0.0746	0.0978	0.0846	0.0760	0.0700
2	5000	Toplam	0.0725	0.0590	0.0435	0.0365	0.0500	0.0417	0.0568	0.0492
		Erkek	0.0787	0.0630	0.0367	0.0304	0.0452	0.0371	0.0553	0.0486
		Kız	0.0638	0.0545	0.0532	0.0459	0.0586	0.0505	0.0583	0.0527

F: Madde Formatı, Örn: Örneklem, 1: Karma Format Test, 2: İki Kategorili Formata Dönüştürülen Test

a) Maddelerin formatına göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 19 ve Tablo 20 ve Tablo 21'den elde edilen bulgular ve bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 19'un birinci sütunu madde formatını, ikinci sütunu örneklem büyüklüğünü, üçüncü sütunu cinsiyet bilgisini, dördüncü sütunu ham puan birikimli dağılımlar arasındaki en büyük fark değerini göstermektedir. Beşinci sütundan on ikinci sütuna kadar yer alan bilgiler eşitleme sonucu oluşan ölçek puanlarının birikimli yüzdeleri arasındaki en büyük farkı göstermektedir.

Tablo 19 incelendiğinde toplam öğrenci verisinde ham puanlara ilişkin başlangıç T değerlerinde karma formattan iki kategorili formata geçilmesiyle kısmen azalma yaşandığı görülmektedir. Bu bilgiyle iki kategorili formatın eşitleme öncesinde A ve B formlarının birikimli dağılım fonksiyonlarını (bdf) daha benzer hale getirdiği söylenebilir. Karma format testte eşitleme sonucu oluşan bdf ler arasındaki farklar incelendiğinde ölçek dönüştürme yöntemine göre değişmekle birlikte T değerlerinde belirgin azalma olduğu ve aynı dağılım özelliğinin (ADÖ) korunmasına yönelik yeterli sonuçlar üretildiği görülmektedir. Örneğin; karma formatta başlangıç ham puan T değerleri %7.05 ile %12.29 arasında değişmekte iken eşitleme

yapıldıktan sonra bu değerler ölçek dönüştürme yöntemine göre değişmekle birlikte sırasıyla %2.52 ile %5.47 düzeylerine kadar düşmüştür. İki kategorili formatta ise eşitleme sonrası T değerlerinde tek yönlü bir artış ya da azalış gerçekleşmemiştir. Ayrıca toplam öğrenci verisine göre $Göz_Pn$ eşitleme sonucu oluşan T değerlerinin Ger_Pn eşitleme sonuçlarından daha küçük olmasından dolayı $Göz_Pn$ eşitlemenin ADÖ'nün korunumuna Ger_Pn eşitlemeden daha iyi katkı sağladığı söylenebilir.

Sonuç olarak karma formattaki testlerin eşitlenmesi sonucu hesaplanan T değerlerinin iki kategorili formattaki testlerin eşitlenmesi sonucu hesaplanan T değerlerinden daha küçük elde edildiği görülmektedir. Bu bilgiyle eşitleme sonrası karma formatta ADÖ'nün korunumunun iki kategorili formata göre daha iyi sağlandığı söylenebilir.

Tablo 20 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede karma formatta en küçük BSE değerinin, 5000 kişilik örneklem büyüklüğü moment yöntemlerden olan $O-S$ yöntemi ve $Göz_Pn$ eşitlemede elde edildiği görülmektedir (0.1684). İki kategorili formatta ise en küçük BSE değerinin, 5000 kişilik örneklem büyüklüğü moment yöntemlerden olan $O-S$ yöntemi ve $Göz_Pn$ eşitlemede elde edildiği görülmektedir (0.1057). İki kategorili formata geçilmesiyle 5000 kişilik örneklem büyüklüğünden elde edilen BSE değerleri tüm ölçek dönüştürme yöntemlerine göre hem $Göz_Pn$ hem de Ger_Pn eşitlemede daha da küçülmüştür. 3000 kişilik örneklem büyüklüğünde SL , 1000 kişilik örneklem büyüklüğünde HB ve SL ölçek dönüştürme yöntemlerinde $Göz_Pn$ ve Ger_Pn eşitlemede BSE değeri küçülürken diğer durumlarda BSE değerinde artış meydana gelmiştir.

Tablo 21 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede karma formatta en küçük İSE değerinin, 5000 kişilik örneklem büyüklüğü moment yöntemlerden olan $O-S$ ölçek dönüştürme yöntemi ve $Göz_Pn$ eşitlemede elde edildiği görülmektedir (0.0354). İki kategorili formatta ise en küçük İSE değerinin, 5000 kişilik örneklem büyüklüğü moment yöntemlerden olan $O-S$ ölçek dönüştürme yöntemi ve $Göz_Pn$ eşitlemede elde edildiği görülmektedir (0.0365). İki kategorili formata geçilmesiyle tüm örneklem büyüklüklerinden elde edilen İSE değerleri tüm ölçek dönüştürme yöntemlerine göre hem Ger_Pn hem de $Göz_Pn$ eşitlemede bir miktar artmıştır.

b) Testi alan gruplara (erkek-kız) göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 19 ve Tablo 20 ve Tablo 21'den elde edilen bulgular ve bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 19 incelendiğinde ham puanlara ilişkin T değerlerinin hem karma hem de iki kategorili formatta kız öğrenciler için daha yüksek olduğu görülmektedir. Karma format testte eşitleme sonucu oluşan bdf ler arasındaki farklar incelendiğinde ölçek dönüştürme yöntemine göre değişmekle birlikte T değerlerinde belirgin azalma olduğu ve ADÖ'nün korunmasına yönelik yeterli sonuçlar üretildiği görülmektedir.

Sonuç olarak kız ve erkek öğrenciler için karma formattaki testlerin eşitlenmesi sonucu hesaplanan T değerlerinin, iki kategorili formattaki testlerin eşitlenmesi sonucu hesaplanan T değerlerinden daha küçük elde edildiği görülmektedir. Kız ve erkek öğrenciler için iki kategorili formattaki T değerlerinde tek yönlü bir artış ya da azalış gerçekleşmediği görülmektedir. Kız ve erkek öğrenciler için karma formatta ADÖ'nün korunumunun iki kategorili formata göre daha iyi sağlandığı söylenebilir.

Tablo 20 incelendiğinde kız öğrenciler için elde edilen BSE değerlerinin; 1000 ve 3000 kişilik örneklem büyüklükleri karma ve iki kategorili format testte $Ger_Pn/Göz_Pn$ eşitlemede erkek öğrencilere göre daha küçük olduğu görülmektedir. 5000 kişilik örneklem büyüklüklerinde ise iki kategorili format O-O ölçek dönüştürme yöntemi ve $Ger_Pn/Göz_Pn$ eşitleme, iki kategorili formatta SL ölçek dönüştürme yöntemi ve Ger_Pn eşitleme dışındaki tüm durumlarda erkek öğrenciler için elde edilen BSE değerleri kız öğrencilere göre daha küçüktür.

Tablo 21 incelendiğinde kız öğrenciler için elde edilen İSE değerlerinin; 1000 ve 3000 kişilik örneklem büyüklüğü karma format ve $Ger_Pn/Göz_Pn$ eşitlemede, 3000 kişilik örneklem büyüklüğü iki kategorili format ve $Ger_Pn/Göz_Pn$ eşitlemede erkek öğrencilere göre daha küçük olduğu görülmektedir. 5000 kişilik örneklem büyüklüğünde, iki kategorili format O-O ölçek dönüştürme yöntemi ve $Ger_Pn/Göz_Pn$ eşitleme, karma format Ger_Pn eşitleme, 1000 kişilik örneklem büyüklüğünde iki kategorili format O-O ölçek dönüştürme yöntemi ve

$Ger_Pn/Göz_Pn$ eşitleme dışındaki tüm durumlarda erkek öğrenciler için elde edilen İSE değerleri kız öğrencilere göre daha küçüktür.

c) Örneklem büyüklüğüne (1000, 3000 ve 5000) göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 19 ve Tablo 20 ve Tablo 21'den elde edilen bulgular ve bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 19 incelendiğinde ham puanlara ilişkin T değerleri hem karma hem de iki kategorili formatta toplam öğrenci verisi üzerinden değerlendirildiğinde örneklem büyüklüğü arttıkça T değerlerinin küçüldüğü görülmektedir. Karma format ve iki kategorili format testte eşitleme sonucu oluşan bdf ler arasındaki farklar incelendiğinde ölçek dönüştürme yöntemine göre değişmekle birlikte T değerlerinin çoğunda belirgin azalma olduğu ve ADÖ'nün korunmasına yönelik yeterli sonuçlar ürettiği görülmektedir.

Tablo 20 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede örneklem büyüklüğünün artmasıyla hem karma hem de iki kategorili formatta Ger_Pn ve $Göz_Pn$ eşitlemede BSE değerinin daha küçük elde edildiği görülmektedir.

Tablo 21 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede örneklem büyüklüğünün artmasıyla karma format Ger_Pn ve $Göz_Pn$ eşitlemede İSE değerinin daha küçük elde edildiği görülmektedir. İki kategorili formatta ise 1000 kişilik örneklem büyüklüğünden 3000 kişilik örneklem büyüklüğüne çıkıldığında İSE değerlerinin büyüdüğü, 5000 kişilik örneklem büyüklüğünde ise hem 3000 hem de 1000 kişilik örneklem büyüklüklerine göre İSE değerlerinin küçüldüğü görülmektedir.

d) Ölçek dönüştürme yöntemlerine (O-O, O-S, HB, SL) göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 19 ve Tablo 20 ve Tablo 21'den elde edilen bulgular ve bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 19 incelendiğinde toplam öğrenci verisinde karma format ve iki kategorili testte eşitleme sonucu oluşan bdf ler arasındaki farklar incelendiğinde ölçek dönüştürme yöntemine göre değişmekle birlikte T değerlerinin çoğunda

belirgin azalma olduđu ve ADÖ'nün korunmasına yönelik yeterli sonuçlar üretildiđi görölmektedir. O-O ölçek dönüştürme yönteminden elde edilen sonuçlar ADÖ'nün korunmasında diđer ölçek yöntemlerinden daha olumsuz sonuçlar üretmiştir. Karma format testlerde özellikle O-S, HB ve SL ölçek dönüştürme yöntemleri başlangıç T değerlerinden daha küçük T değerlerinin elde edilmesini sağlamıştır. İki kategorili format testlerin eşitlenmesinde ölçek dönüştürme yöntemlerinin T değerlerinde tek yönlü bir artış ya da azalışa neden olmadığı görölmektedir.

Tablo 20 incelendiđinde toplam öğrenci verisi üzerinden yapılan değerlendirmede ölçek dönüştürme yöntemlerine göre örneklem büyüklüğünün artmasıyla karma format Ger_Pn ve $Göz_Pn$ eşitlemede BSE değerinin daha küçük elde edildiđi görölmektedir. İki kategorili formatta ise 5000 kişilik örneklem büyüklüğü Ger_Pn ve $Göz_Pn$ eşitlemede BSE değerlerinin daha küçük elde edildiđi görölmektedir.

Tablo 21 incelendiđinde toplam öğrenci verisi üzerinden yapılan değerlendirmede ölçek dönüştürme yöntemlerine göre 5000 kişilik örneklem büyüklüğü karma ve iki kategorili format Ger_Pn ve $Göz_Pn$ eşitlemede İSE değerinin daha küçük elde edildiđi görölmektedir. 3000 kişilik örneklem büyüklüğünde ise 1000 kişilik örneklem büyüklüğünde O-O ve HB ölçek dönüştürme yöntemleri dışında Ger_Pn ve $Göz_Pn$ eşitlemede daha büyük İSE değerleri elde edilmiştir.

Yukarıda yer alan 4 alt problemin birlikte değerlendirilmesiyle aşağıda yer alan yorum yapılabilir.

Tablo 19 incelendiđinde karma formatta en küçük T değerinin, 5000 kişilik örneklem büyüklüğü SL ölçek dönüştürme yöntemi $Göz_Pn$ eşitleme ve erkek öğrenci verisinde elde edildiđi görölmektedir (%2.52). Bu verinin başlangıç T değeri de %7.05'tir. İki kategorili formatta en küçük T değerinin; 5000 kişilik örneklem büyüklüğü SL ölçek dönüştürme yöntemi $Göz_Pn$ eşitleme ve erkek öğrenci verisinde elde edildiđi görölmektedir (%4.49). Bu verinin başlangıç T değeri de %6.44'tür. Sonuç olarak karma format verinin eşitlenmesi sonucunda ADÖ'nün korunumunun daha iyi sağlandığı görölmektedir.

Tablo 20 incelendiđinde en küçük BSE değerinin, iki kategorili format 5000 kişilik örneklem büyüklüğü moment yöntemlerden olan O-S ölçek dönüştürme

yöntemi $Göz_Pn$ eşitleme ve kız öğrenci verisinde elde edildiği görülmektedir (0.0797). Format*Örneklem*Grup*Yöntem çaprazlamasının neredeyse tamamında (72 durumdan 70'inde) $Göz_Pn$ eşitleme yöntemi Ger_Pn eşitleme yöntemine göre daha küçük BSE değeri vermiştir. BSE değerleri küçükten büyüğe sıralandığında 144 durumdan ilk 24'ünde 5000 kişilik örneklem yer almaktadır. Yine bu 24 durumun ilk 11'ini iki kategorili format oluşturmaktadır. Bu bilgi örneklem büyüklüğünün artması ve verilerin iki kategorili olması durumunda BSE'nin daha iyi sağlandığını göstermektedir.

Tablo 21 incelendiğinde en küçük İSE değerinin, karma format test 5000 kişilik örneklem büyüklüğü moment yöntemlerden olan O-S ölçek dönüştürme yöntemi $Göz_Pn$ eşitleme ve kız öğrenci verisi için elde edildiği görülmektedir (0.0267). Format*Örneklem*Grup*Yöntem çaprazlamasının tamamında (72 durumdan 72'sinde) $Göz_Pn$ eşitleme yöntemi Ger_Pn yöntemine göre daha küçük İSE değeri vermiştir. İSE değerleri küçükten büyüğe sıralandığında 144 durumdan ilk 13'ünün 11'inde 5000 kişilik örneklem büyüklüğü yer almaktadır. Yine bu 13 durumun 9'unu karma format veri oluşturmaktadır. Bu sonuç örneklem büyüklüğünün artması ve verilerin karma formatta olması durumunda İSE'nin genellikle daha iyi sağlandığını göstermektedir.

İkinci Alt Probleme Ait Bulgular ve Yorumlar

Matematik testi için denk olmayan gruplar ortak madde deseninde gerçek/gözlenen puan eşitleme yöntemi ile C formunun B formuna eşitlenmesinde aynı dağılım ve eşitlik özelliğinin korunumu,

- maddelerin formatına (çok kategorili puanlanan, iki kategorili puanlanan),*
- testi alan alt gruplara (kız, erkek),*
- örneklem büyüklüğüne (1000, 3000 ve 5000) ve*
- ölçek dönüştürme yöntemlerine (O-O, O-S, HB, SL) göre nasıl değişmektedir?*

Matematik testinde ham puan T değerleri ve C formunun B formuna gerçek puan (Ger_Pn) ve gözlenen puan ($Göz_Pn$) eşitleme yöntemiyle eşitlenmesi sonucu oluşan T değerleri Tablo 22'de, birinci sıra eşitlik (BSE) değerleri Tablo 23'te, ikinci sıra eşitlik (İSE) değerleri Tablo 24'te verilmiştir.

a) Maddelerin formatına göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 22, Tablo 23 ve Tablo 24'ten elde edilen bulgular ve bu bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 22

Matematik Testi C Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi - Ham Puan ve Ölçek Puan Dağılımları için Kolmogorov T İstatistikleri (%)

F	Örn.	Grup	Ham Puan Dağılımları	O-O		O-S		HB		SL	
				Ger_Pn	Göz_Pn	Ger_Pn	Göz_Pn	Ger_Pn	Göz_Pn	Ger_Pn	Göz_Pn
1	1000	Toplam	3.00	3.20	2.50	3.07	2.50	2.79	2.50	3.24	2.54
		Erkek	4.52	11.56	11.06	3.47	2.99	3.70	3.11	2.60	2.72
		Kız	2.54	5.18	4.92	3.69	3.67	3.68	3.31	3.83	3.41
1	3000	Toplam	3.47	7.44	6.08	5.96	4.39	4.23	2.73	4.06	2.73
		Erkek	6.08	4.03	2.96	4.64	3.67	3.19	2.27	3.09	2.13
		Kız	2.07	6.17	4.84	6.28	4.97	5.12	3.54	4.56	3.19
1	5000	Toplam	4.06	5.06	4.83	2.08	1.71	2.12	1.53	2.12	2.03
		Erkek	6.86	3.86	3.44	2.42	2.02	2.42	1.71	2.42	2.49
		Kız	2.59	12.09	12.26	4.39	4.12	5.13	5.40	4.20	4.47
2	1000	Toplam	6.70	4.20	4.50	4.88	3.55	5.50	3.67	4.03	4.08
		Erkek	7.55	13.23	10.16	4.37	4.61	8.79	6.99	4.80	3.68
		Kız	7.64	7.64	6.78	5.00	4.03	3.92	3.53	5.34	3.74
2	3000	Toplam	6.37	6.72	5.38	4.90	3.54	4.73	2.21	4.70	2.27
		Erkek	8.94	4.86	2.24	5.32	3.11	4.78	2.33	4.62	2.07
		Kız	4.47	7.31	4.47	5.97	4.47	4.97	3.29	4.47	2.98
2	5000	Toplam	6.22	4.26	3.10	4.26	1.89	4.26	1.82	4.26	1.97
		Erkek	8.49	5.47	3.36	4.44	2.23	5.44	3.32	5.61	3.42
		Kız	4.13	4.86	2.53	7.38	4.90	8.04	5.50	5.37	2.72

F: Madde Formatı, Örn: Örneklem, 1: Karma Format Test, 2: İki Kategorili Formata Dönüştürülen Test

Tablo 22 incelendiğinde toplam öğrenci verisinde ham puanlara ilişkin başlangıç T değerlerinde karma formattan iki kategorili formata geçilmesiyle artış yaşandığı görülmektedir. İki kategorili formatın eşitleme öncesinde C ve B formlarının bdf lerini daha farklı hale getirdiği söylenebilir.

Karma format testte eşitleme sonucu oluşan birikimli dağılım fonksiyonları (bdf) arasındaki farklar incelendiğinde ölçek dönüştürme yöntemine göre değişmekle birlikte T değerlerinde tek yönlü bir azalma olduğu artış oluşmadığı görülmektedir. Ancak karma formatta $Göz_Pn$ eşitleme yönteminin genellikle daha küçük T değeri vermesinden dolayı Ger_Pn eşitlemeye göre aynı dağılım özelliğinin (ADÖ) korunmasında daha iyi sonuçlar ürettiği söylenebilir. İki kategorili formatta başlangıç ham puan T değerleri %4.13 ile %8.94 arasında değişmekte iken eşitleme yapıldıktan sonra bu değerler ölçek dönüştürme yöntemine göre değişmekle birlikte sırasıyla %2.53 ile %2.07 düzeylerine düşmüştür.

Tablo 23

Matematik Testinde C Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan BSE Değerleri

F	Örn.	Grup	O-O		O-S		HB		SL	
			Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn
1	1000	Toplam	0.1691	0.1564	0.1715	0.1596	0.1672	0.1565	0.1452	0.1209
		Erkek	0.3111	0.3139	0.1914	0.1804	0.1907	0.1841	0.1411	0.1352
		Kız	0.1234	0.1113	0.1106	0.1024	0.1116	0.1006	0.1149	0.0835
1	3000	Toplam	0.1522	0.1401	0.1508	0.1388	0.1630	0.1518	0.1515	0.1451
		Erkek	0.1468	0.1386	0.1415	0.1329	0.1614	0.1543	0.1593	0.1689
		Kız	0.1380	0.1253	0.1383	0.1256	0.1651	0.1512	0.1432	0.1311
1	5000	Toplam	0.1348	0.1242	0.1413	0.1331	0.1486	0.1399	0.1403	0.1339
		Erkek	0.1613	0.1528	0.1743	0.1676	0.1798	0.1732	0.1661	0.1625
		Kız	0.1639	0.1631	0.0952	0.0913	0.1185	0.1074	0.1077	0.1007
2	1000	Toplam	0.1029	0.0983	0.1305	0.1311	0.1328	0.1362	0.0813	0.0830
		Erkek	0.2937	0.3101	0.1463	0.1447	0.1869	0.1996	0.1145	0.1219
		Kız	0.0515	0.0435	0.0976	0.1021	0.0755	0.0765	0.0580	0.0569
2	3000	Toplam	0.0672	0.0569	0.0891	0.0843	0.1214	0.1195	0.1109	0.1125
		Erkek	0.1421	0.1490	0.0997	0.0996	0.1334	0.1374	0.1393	0.1484
		Kız	0.0836	0.0707	0.0886	0.0770	0.1309	0.1239	0.1104	0.1054
2	5000	Toplam	0.0763	0.0718	0.1113	0.1130	0.1154	0.1165	0.1022	0.1064
		Erkek	0.1601	0.1690	0.1444	0.1508	0.1576	0.1670	0.1583	0.1706
		Kız	0.1038	0.1007	0.1355	0.1375	0.1398	0.1444	0.0863	0.0892

F: Madde Formatı, Örn: Örneklem, 1: Karma Format Test, 2: İki Kategorili Formata Dönüştürülen Test

Tablo 24

Matematik Testinde C Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan İSE Değerleri

F	Örn.	Grup	O-O		O-S		HB		SL	
			Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn
1	1000	Toplam	0.0506	0.0466	0.0490	0.0451	0.0468	0.0431	0.0495	0.0414
		Erkek	0.0680	0.0538	0.0508	0.0473	0.0457	0.0410	0.0445	0.0423
		Kız	0.0497	0.0443	0.0372	0.0332	0.0394	0.0356	0.0410	0.0342
1	3000	Toplam	0.0574	0.0516	0.0512	0.0471	0.0476	0.0448	0.0449	0.0428
		Erkek	0.0478	0.0447	0.0488	0.0458	0.0483	0.0446	0.0473	0.0427
		Kız	0.0502	0.0443	0.0508	0.0447	0.0510	0.0462	0.0443	0.0418
1	5000	Toplam	0.0520	0.0479	0.0427	0.0403	0.0442	0.0417	0.0438	0.0410
		Erkek	0.0466	0.0435	0.0467	0.0422	0.0475	0.0427	0.0454	0.0409
		Kız	0.0748	0.0638	0.0479	0.0415	0.0477	0.0434	0.0451	0.0412
2	1000	Toplam	0.0370	0.0304	0.0446	0.0341	0.0473	0.0350	0.0380	0.0295
		Erkek	0.0987	0.0669	0.0482	0.0392	0.0702	0.0477	0.0520	0.0360
		Kız	0.0288	0.0266	0.0464	0.0365	0.0383	0.0319	0.0340	0.0294
2	3000	Toplam	0.0244	0.0223	0.0316	0.0256	0.0424	0.0315	0.0400	0.0273
		Erkek	0.0603	0.0397	0.0436	0.0302	0.0555	0.0375	0.0586	0.0377
		Kız	0.0312	0.0276	0.0318	0.0282	0.0446	0.0344	0.0391	0.0293
2	5000	Toplam	0.0344	0.0252	0.0477	0.0315	0.0489	0.0320	0.0456	0.0292
		Erkek	0.0745	0.0471	0.0681	0.0435	0.0739	0.0472	0.0744	0.0455
		Kız	0.0397	0.0327	0.0491	0.0371	0.0520	0.0380	0.0409	0.0307

F: Madde Formatı, Örn: Örneklem, 1: Karma Format Test, 2: İki Kategorili Formata Dönüştürülen Test

Sonuç olarak karma formattaki testlerin başlangıç T değerlerinin genellikle daha küçük olması nedeniyle eşitleme sonucunda hesaplanan T değerlerinde bazı

dönüştürme yöntemlerine göre azalış, diğerlerinde de az da olsa artışlar olduğu görülmektedir. İki kategorili formatta ise başlangıç T değerlerinin daha büyük olmasından dolayı eşitleme sonucunda hesaplanan T değerlerinde genellikle büyük ölçüde azalmalar olduğu görülmektedir.

Tablo 23 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede iki kategorili formatta en küçük BSE değerinin, 3000 kişilik örneklem moment yöntemlerden olan $O-O$ ölçek dönüştürme yöntemi ve $Göz_Pn$ eşitlemede elde edildiği görülmektedir (0.0569). Karma formatta ise en küçük BSE değerinin, 1000 kişilik örneklem büyüklüğü karakteristik eğri yöntemlerinden olan SL ölçek dönüştürme yöntemi ve $Göz_Pn$ eşitlemede elde edildiği görülmektedir (0.1209). İki kategorili formata geçilmesiyle örneklem büyüklüklerinin tamamından elde edilen BSE değerleri tüm ölçek dönüştürme yöntemlerine göre hem Ger_Pn hem de $Göz_Pn$ eşitlemede daha da küçülmüştür.

Tablo 24 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede iki kategorili formatta en küçük İSE değerinin, 3000 kişilik örneklem büyüklüğü moment yöntemlerden olan $O-O$ ölçek dönüştürme yöntemi ve $Göz_Pn$ eşitlemede elde edildiği görülmektedir (0.0223). Karma formatta ise en küçük İSE değerinin, 5000 kişilik örneklem büyüklüğü moment yöntemlerden olan $O-S$ ölçek dönüştürme yöntemi ve $Göz_Pn$ eşitlemede elde edildiği görülmektedir (0.0403). İki kategorili formata geçilmesiyle tüm örneklem büyüklüklerinden elde edilen İSE değerleri tüm ölçek dönüştürme yöntemlerine göre $Göz_Pn$ eşitlemede daha da küçülmüştür. İSE değerlerinde; 5000 kişilik örneklem büyüklüğünde $O-O$, HB ve SL , 1000 kişilik örneklem büyüklüğünde HB ölçek dönüştürme yönteminde Ger_Pn eşitlemede bir miktar artış gerçekleşmiştir.

b) Testi alan gruplara (erkek-kız) göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 22, Tablo 23 ve Tablo 24'ten elde edilen bulgular ve bu bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 22 incelendiğinde ham puanlara ilişkin T değerlerinin hem karma hem de iki kategorili formatta (1000 kişilik örneklem büyüklüğü dışında) erkek öğrenciler için daha yüksek olduğu görülmektedir. Karma format ve iki kategorili testlerde eşitleme sonucu oluşan bdp ler arasındaki farklar incelendiğinde ölçek dönüştürme

yöntemine göre değişmekle birlikte erkek öğrenciler için T değerlerinde belirgin azalmalar olduğu ve ADÖ'nün korunmasına yönelik yeterli sonuçlar üretildiği görülmektedir.

Sonuç olarak kız ve erkek öğrenciler için karma formattaki testlerin eşitlenmesi sonucu hesaplanan T değerlerinin, iki kategorili formattaki testlerin eşitlenmesi sonucu hesaplanan T değerlerinden genel olarak daha küçük elde edildiği görülmektedir.

Tablo 23 incelendiğinde kız öğrenciler için elde edilen BSE değerlerinin, (karma formatta, 5000 kişilik örneklem büyüklüğünde, O-O ölçek dönüştürme yöntemi ve $Ger_Pn/Göz_Pn$ eşitlemeyle elde edilen değerler dışında) karma ve iki kategorili format test ve $Ger_Pn/Göz_Pn$ eşitlemede erkek öğrencilere göre daha küçük olduğu görülmektedir.

Tablo 24 incelendiğinde kız öğrenciler için elde edilen İSE değerlerinin iki kategorili formatın tamamında, karma formatta 1000 kişilik örneklem büyüklüklerinin tamamında ve tüm ölçek dönüştürme yöntemlerinde erkeklere göre daha küçük olduğu görülmektedir. Karma formattaki 3000 ve 5000 kişilik örneklem büyüklüklerinde ise ölçek dönüştürme yöntemine göre İSE değerlerinin büyüklük ve küçüklüğü farklılaşmaktadır.

c) Örneklem büyüklüğüne (1000, 3000 ve 5000) göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 22, Tablo 23 ve Tablo 24'ten elde edilen bulgular ve bu bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 22 incelendiğinde ham puanlara ilişkin T değerlerinin karma formatta toplam öğrenci verisi üzerinden yapılan değerlendirmede örneklem büyüklüğü arttıkça büyüdüğü, iki kategorili formatta ise örneklem büyüklüğü arttıkça küçüldüğü görülmektedir. Karma format ve iki kategorili format testte eşitleme sonucu oluşan $bd\hat{f}$ ler arasındaki farklar incelendiğinde ölçek dönüştürme yöntemine göre değişmekle birlikte T değerlerinin çoğunda belirgin azalma olduğu ve ADÖ'nün korunmasına yönelik yeterli sonuçlar ürettiği görülmektedir.

Tablo 23 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede örneklem büyüklüğünün artmasıyla karma formatta, moment yöntemlerinde hem Ger_Pn hem de $Göz_Pn$ eşitlemede BSE değerinde azalma

görülürken karakteristik eğri yöntemlerinde kararlı artış ya da azalış gerçekleşmediği görülmektedir. İki kategorili formatta ise örneklem büyüklüğünün artmasıyla moment yöntemlerde hem Ger_{Pn} hem de $Göz_{Pn}$ eşitlemede BSE değerinde kararlı bir artış ya da azalma olmazken karakteristik eğri yöntemlerinde kararlı bir azalış gerçekleşmiştir.

Tablo 24 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede örneklem büyüklüğünün artmasıyla hem karma formatta Ger_{Pn} ve $Göz_{Pn}$ eşitlemede İSE değerinin daha küçük elde edildiği görülmektedir. İki kategorili formatta ise 3000 kişilik örneklem büyüklükleri daha küçük İSE değeri vermektedir.

d) Ölçek dönüştürme yöntemlerine (O-O, O-S, HB, SL) göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 22, Tablo 23 ve Tablo 24'ten elde edilen bulgular ve bu bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 22 incelendiğinde karma format ve iki kategorili format testte eşitleme sonucu oluşan bdf ler arasındaki farklar incelendiğinde ölçek dönüştürme yöntemine göre değişmekle birlikte T değerlerinin çoğunda belirgin azalma olduğu ve ADÖ'nün sağlanmasına yönelik yeterli sonuçlar üretildiği görülmektedir. O-O ölçek dönüştürme yönteminden elde edilen sonuçlar diğer ölçek yöntemlerinden ADÖ'nün korunmasında daha olumsuz sonuçlar üretmiştir. Karma format testlerin eşitlenmesi sonucu oluşan T değerleri özellikle HB ve SL ölçek dönüştürme yöntemleri için benzer sonuçlar üretmiştir. Hem karma hem de iki kategorili formatta $Göz_{Pn}$ eşitleme sonucu oluşan T değerleri Ger_{Pn} eşitleme sonucu oluşan T değerlerinden daha küçük elde edilmiştir.

Tablo 23 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede ölçek dönüştürme yöntemlerine göre örneklem büyüklüğünün artmasıyla karma formatta Ger_{Pn} ve $Göz_{Pn}$ eşitlemede SL ölçek dönüştürme yöntemi dışında BSE değerinin daha küçük elde edildiği görülmektedir. Karma formatta tüm ölçek dönüştürme yöntemlerine göre elde edilen BSE değerleri birbirine oldukça yakındır. İki kategorili formatta ise 3000 kişilik örneklem büyüklüğünde Ger_{Pn} ve $Göz_{Pn}$ eşitlemede moment yöntemleri, 1000 kişilik

örneklem büyüklüğünde *SL* ölçek dönüştürme yöntemi en küçük BSE değerlerini vermiştir.

Tablo 24 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede tüm ölçek dönüştürme yöntemlerine göre 5000 kişilik örneklem büyüklüğünde karma format *Ger_Pn* ve *Göz_Pn* eşitlemede İSE değerinin daha küçük elde edildiği görülmektedir. İki kategorili formatta ise 3000 kişilik örneklem büyüklüğü *SL* ölçek dönüştürme yöntemi ve *Ger_Pn* eşitlemede elde edilen değer dışında daha küçük İSE değeri vermiştir.

Yukarıda yer alan 4 alt problemin birlikte değerlendirilmesiyle aşağıda yer alan yorum yapılabilir.

Tablo 22 incelendiğinde karma formatta en küçük *T* değerinin, 5000 kişilik örneklem büyüklüğü *HB* ölçek dönüştürme yöntemi *Göz_Pn* eşitleme ve toplam öğrenci verisinde elde edildiği görülmektedir (%1.53). Bu verinin başlangıç *T* değeri de %4.06'dır. İki kategorili formatta en küçük *T* değerinin, 5000 kişilik örneklem büyüklüğü *HB* ölçek dönüştürme yöntemi *Göz_Pn* eşitleme ve erkek öğrencilerde elde edildiği görülmektedir (%1.82). Bu verinin başlangıç *T* değeri de %6.22'dir. Sonuç olarak hem karma hem de iki kategorili format verinin eşitlenmesi sonucunda ADÖ'nün korunumunun sağlandığı görülmektedir.

Tablo 23 incelendiğinde en küçük BSE değerinin, iki kategorili format 1000 kişilik örneklem büyüklüğü moment yöntemlerden olan *O-O* ölçek dönüştürme yöntemi format *Göz_Pn* eşitleme ve kız öğrenci verisi için elde edildiği görülmektedir (0.0435). Format*Örneklem*Grup*Yöntem çaprazlamasının çoğunda (72 durumdan 49'unda) *Göz_Pn* eşitleme yöntemi *Ger_Pn* eşitleme yöntemine göre daha küçük BSE değeri vermiştir. En küçük BSE değerlerinin daha çok 3000 ve 1000 kişilik örneklem büyüklüklerinde elde edildiği ve en küçük BSE değerinin elde edildiği 21 durumdan 20'sinde veriler iki kategorili formatta olduğu görülmektedir. Yine bu 21 durumun ilk 11'ini iki kategorili format oluşturmaktadır. Bu sonuç verilerin iki kategorili olması durumunda BSE'nin daha iyi sağlandığını göstermektedir.

Tablo 24 incelendiğinde en küçük İSE değerinin, karma format 5000 kişilik örneklem büyüklüğü moment yöntemlerden olan *O-S* ölçek dönüştürme yöntemi *Göz_Pn* eşitleme ve toplam öğrenci verisi için elde edildiği görülmektedir (0.0223). Format*Örneklem*Grup*Yöntem çaprazlamasının tamamında (72 durumdan

72'sinde) $Göz_Pn$ eşitleme yöntemi Ger_Pn eşitleme yöntemine göre daha küçük İSE değeri vermiştir. İSE değerleri küçükten büyüğe sıralandığında 144 durumdan ilk 24'ünü iki kategorili format oluşturmaktadır. Bu durum verilerin iki kategorili formatta olması durumunda İSE'nin genellikle daha iyi sağlandığını göstermektedir.

Üçüncü Alt Probleme Ait Bulgular ve Yorumlar

Sosyal Bilgiler testi için denk olmayan gruplarda ortak madde deseninde gerçek/gözlenen puan eşitleme yöntemi ile A formunun B formuna eşitlenmesinde aynı dağılım ve eşitlik özelliğinin korunumu,

- a. maddelerin formatına (çok kategorili puanlanan, iki kategorili puanlanan),*
- b. testi alan alt gruplara (kız, erkek),*
- c. örneklem büyüklüğüne (1000, 3000 ve 5000) ve*
- d. ölçek dönüştürme yöntemlerine (O-O, O-S, HB, SL) göre nasıl değişmektedir?*

Sosyal Bilgiler testinde ham puan T değerleri ve A formunun B formuna gerçek puan (Ger_Pn) ve gözlenen puan ($Göz_Pn$) eşitleme yöntemiyle eşitlenmesi sonucu oluşan T değerleri Tablo 25'te, birinci sıra eşitlik (BSE) değerleri Tablo 26'da, ikinci sıra eşitlik (İSE) değerleri Tablo 27'de verilmiştir.

a) Maddelerin formatına göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 25, Tablo 26 ve Tablo 27'den elde edilen bulgular ve bu bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 25 incelendiğinde ham puanlara ilişkin başlangıç T değerlerinde karma formattan iki kategorili formata geçilmesiyle önemli düzeyde azalmalar yaşandığı görülmektedir. İki kategorili formatın eşitleme öncesinde A ve B formlarının birikimli dağılım fonksiyonlarını (bdf) daha benzer hale getirdiği söylenebilir. Karma format testte eşitleme sonucu oluşan bdf ler arasındaki farklar incelendiğinde SL ve HB ölçek dönüştürme yöntemlerinde T değerlerinde önemli düzeyde azalmalar olduğu ve aynı dağılım özelliğinin (ADÖ) sağlandığı görülmektedir.

Tablo 25

Sosyal Bilgiler Testi A Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi - Ham Puan ve Ölçek Puan Dağılımları için Kolmogorov T İstatistikleri (%)

F	Örn.	Grup	Ham Puan Dağılımları	O-O		O-S		HB		SL	
				Ger_Pn	Göz_Pn	Ger_Pn	Göz_Pn	Ger_Pn	Göz_Pn	Ger_Pn	Göz_Pn
1	1000	Toplam	14.00	9.93	9.48	7.78	6.80	2.52	3.33	2.99	2.58
		Erkek	12.12	12.15	12.55	8.79	8.77	6.57	6.89	4.59	5.29
		Kız	17.15	5.60	5.49	2.92	2.44	3.17	2.83	4.08	3.93
1	3000	Toplam	12.70	18.43	17.09	14.41	14.06	3.39	3.24	2.86	2.73
		Erkek	12.31	9.36	8.77	13.37	12.85	5.59	5.84	4.39	4.59
		Kız	13.37	16.81	16.04	11.35	10.40	1.78	2.29	1.29	1.24
1	5000	Toplam	12.88	17.09	16.48	10.15	9.37	3.09	3.64	1.99	2.06
		Erkek	11.69	7.56	7.14	6.27	5.94	3.92	4.38	2.78	3.02
		Kız	14.43	20.14	18.83	10.69	9.63	2.73	2.94	1.61	1.50
2	1000	Toplam	2.40	9.47	9.84	8.19	7.61	3.20	3.15	3.66	3.54
		Erkek	3.26	12.68	13.47	7.66	7.33	5.56	6.18	4.42	5.05
		Kız	6.59	4.26	4.40	4.45	4.85	3.78	4.13	5.41	5.67
2	3000	Toplam	1.93	19.50	18.45	15.17	13.87	5.19	5.22	4.08	4.17
		Erkek	2.53	9.82	9.43	14.65	13.63	7.46	7.35	6.28	6.19
		Kız	1.61	18.66	18.06	11.93	10.66	3.51	3.56	2.75	2.65
2	5000	Toplam	1.84	17.96	17.72	11.35	10.21	4.64	4.96	3.02	3.35
		Erkek	1.48	6.89	6.34	8.16	6.36	4.88	5.31	4.05	4.21
		Kız	2.22	22.53	22.43	11.62	10.28	4.15	4.41	1.37	1.97

F: Madde Formatı, Örn: Örneklem, 1: Karma Format Test, 2: İki Kategorili Formata Dönüştürülen Test

Tablo 26

Sosyal Bilgiler Testinde A Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan BSE Değerleri

F	Örn.	Grup	O-O		O-S		HB		SL	
			Ger_Pn	Göz_Pn	Ger_Pn	Göz_Pn	Ger_Pn	Göz_Pn	Ger_Pn	Göz_Pn
1	1000	Toplam	0.1375	0.1478	0.1911	0.1978	0.2526	0.2612	0.2572	0.2764
		Erkek	0.1208	0.1359	0.1225	0.1273	0.1818	0.1922	0.2071	0.2219
		Kız	0.3757	0.3911	0.3393	0.3530	0.3542	0.3682	0.3190	0.3474
1	3000	Toplam	0.6739	0.6867	0.6888	0.6800	0.2408	0.2514	0.2524	0.2666
		Erkek	0.4711	0.4814	0.4873	0.4855	0.1998	0.2084	0.2163	0.2261
		Kız	0.6494	0.6670	0.6651	0.4865	0.2912	0.3057	0.3028	0.2779
1	5000	Toplam	0.6155	0.6326	0.6207	0.6128	0.2403	0.2494	0.2647	0.2861
		Erkek	0.3953	0.4050	0.4077	0.4073	0.2150	0.2214	0.2313	0.2402
		Kız	0.7152	0.7396	0.6449	0.6424	0.2809	0.2957	0.3064	0.3276
2	1000	Toplam	0.1492	0.1581	0.0972	0.0924	0.0227	0.0224	0.0252	0.0229
		Erkek	0.2059	0.2161	0.1338	0.1310	0.0770	0.0799	0.0535	0.0555
		Kız	0.0777	0.0783	0.0416	0.0391	0.0591	0.0589	0.0397	0.0398
2	3000	Toplam	0.4009	0.4080	0.4342	0.4082	0.0660	0.0691	0.0422	0.0463
		Erkek	0.1880	0.1942	0.2870	0.2761	0.0946	0.0975	0.0714	0.0766
		Kız	0.3897	0.3957	0.3967	0.3668	0.0451	0.0457	0.0265	0.0239
2	5000	Toplam	0.3610	0.3693	0.3630	0.3369	0.0811	0.0847	0.0413	0.0442
		Erkek	0.1256	0.1303	0.1859	0.1761	0.0979	0.1011	0.0725	0.0775
		Kız	0.4718	0.4837	0.3949	0.3636	0.0580	0.0603	0.0202	0.0158

F: Madde Formatı, Örn: Örneklem, 1: Karma Format Test, 2: İki Kategorili Formata Dönüştürülen Test

Tablo 27

Sosyal Bilgiler Testinde A Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan İSE Değerleri

F	Örn.	Grup	O-O		O-S		HB		SL	
			Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn
1	1000	Toplam	0.0482	0.0529	0.0277	0.0181	0.0434	0.0491	0.0367	0.0430
		Erkek	0.0527	0.0580	0.0299	0.0242	0.0336	0.0410	0.0309	0.0379
		Kız	0.0536	0.0533	0.0678	0.0709	0.0717	0.0742	0.0537	0.0585
1	3000	Toplam	0.1019	0.0842	0.2511	0.2422	0.0429	0.0496	0.0427	0.0503
		Erkek	0.0623	0.0537	0.1778	0.1726	0.0375	0.0442	0.0380	0.0447
		Kız	0.1140	0.0977	0.2326	0.1667	0.0582	0.0633	0.0563	0.0526
1	5000	Toplam	0.0871	0.0679	0.2078	0.2007	0.0485	0.0545	0.0449	0.0536
		Erkek	0.0472	0.0416	0.1292	0.1270	0.0478	0.0527	0.0457	0.0516
		Kız	0.1094	0.0857	0.2228	0.2167	0.0517	0.0594	0.0475	0.0557
2	1000	Toplam	0.0432	0.0343	0.0398	0.0344	0.0191	0.0140	0.0174	0.0123
		Erkek	0.0501	0.0382	0.0400	0.0371	0.0229	0.0181	0.0187	0.0141
		Kız	0.0252	0.0107	0.0289	0.0130	0.0308	0.0134	0.0213	0.0089
2	3000	Toplam	0.0735	0.0484	0.1827	0.1570	0.0209	0.0189	0.0169	0.0141
		Erkek	0.0338	0.0209	0.1228	0.1002	0.0318	0.0291	0.0268	0.0234
		Kız	0.0830	0.0541	0.1716	0.1502	0.0195	0.0136	0.0176	0.0103
2	5000	Toplam	0.0641	0.0397	0.1535	0.1308	0.0306	0.0255	0.0229	0.0176
		Erkek	0.0189	0.0119	0.0844	0.0679	0.0381	0.0312	0.0323	0.0249
		Kız	0.0882	0.0537	0.1725	0.1492	0.0274	0.0227	0.0206	0.0140

F: Madde Formatı, Örn: Örneklem, 1: Karma Format Test, 2: İki Kategorili Formata Dönüştürülen Test

Sonuç olarak karma formattaki testlerin başlangıç T değerlerinin büyük olması nedeniyle eşitleme sonucunda hesaplanan T değerlerinde karakteristik eğri yöntemlerine göre azalış yaşanmıştır. Ancak moment yöntemlere göre T değerlerinde tek yönlü bir artış ya da azalış yaşanmamıştır. İki kategorili formatta başlangıç ham puan T değerlerine göre eşitleme yapıldıktan sonra oluşan T değerlerinde özellikle moment yöntemlerde önemli düzeyde artışlar yaşanmıştır.

Tablo 26 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede karma formatta en küçük BSE değerinin, 1000 kişilik örneklem moment yöntemlerden olan O-O ölçek dönüştürme yöntemi ve Ger_Pn eşitlemede elde edildiği görülmektedir (0.1375). İki kategorili formatta ise en küçük BSE değerinin, 1000 kişilik örneklem büyüklüğü karakteristik eğri yöntemlerinden olan HB ölçek dönüştürme yöntemi ve Göz_Pn eşitlemede elde edildiği görülmektedir (0.0224). İki kategorili formata geçilmesiyle 1000 kişilik örneklem büyüklüğünde O-O ölçek dönüştürme yöntemi ile Ger_Pn ve Göz_Pn eşitlemede elde edilen değerler dışında diğer tüm örneklem büyüklükleri ve ölçek dönüştürme yöntemlerinde hem Ger_Pn hem de Göz_Pn eşitlemede BSE değerleri oldukça küçülmüştür.

Tablo 27 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede karma formatta en küçük İSE değerinin, 1000 kişilik örneklem büyüklüğü moment yöntemlerden olan O-S ölçek dönüştürme yöntemi ve *Göz_Pn* eşitlemede elde edildiği görülmektedir (0.0181). İki kategorili formatta ise en küçük İSE değerinin, 1000 kişilik örneklem büyüklüğün karakteristik eğri yöntemlerinden olan *SL* ölçek dönüştürme yöntemi ve *Göz_Pn* eşitlemede elde edildiği görülmektedir (0.0123). İki kategorili formata geçilmesiyle 1000 kişilik örneklem büyüklüğünde O-S ölçek dönüştürme yöntemi ile *Ger_Pn* ve *Göz_Pn* eşitlemede elde edilen değerler dışında diğer tüm örneklem büyüklükleri ve ölçek dönüştürme yöntemlerinde hem *Ger_Pn* hem de *Göz_Pn* eşitlemede İSE değerleri küçülmüştür.

b) Testi alan gruplara (erkek-kız) göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 25, Tablo 26 ve Tablo 27'den elde edilen bulgular ve bu bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 25 incelendiğinde ham puanlara ilişkin *T* değerlerinin karma formatta kız öğrenciler için daha yüksek olduğu görülmektedir. İki kategorili formatta ise 1000 ve 3000 kişilik örneklem büyüklüklerinde erkek, 5000 kişilik örneklem büyüklüğünde ise kız öğrenciler için başlangıç *T* değerlerinin daha büyük olduğu görülmektedir. Karma format testlerde eşitleme sonucu oluşan *bdf*ler arasındaki farklar incelendiğinde ölçek dönüştürme yöntemine göre değişmekle birlikte özellikle karakteristik eğri yöntemlerinde hem erkek hem de kız öğrenciler için *T* değerlerinde belirgin azalmalar olduğu, kız öğrenciler için *T* değerlerinin genellikle daha küçük elde edildiği ve ADÖ'nün korunmasına yönelik yeterli sonuçlar üretildiği görülmektedir.

Sonuç olarak kız ve erkek öğrenciler için karma formattaki testlerin eşitlenmesi sonucu hesaplanan *T* değerlerinin iki kategorili formattaki testlerin eşitlenmesi sonucu hesaplanan *T* değerlerinden genel olarak daha küçük elde edildiği görülmektedir.

Tablo 26 incelendiğinde erkek öğrenciler için elde edilen BSE değerlerinin karma formattaki tüm ölçek dönüştürme yöntemlerinde kız öğrencilere göre daha küçük olduğu görülmektedir. İki kategorili formatta 1000 kişilik örneklem büyüklüğünün tamamında ve 5000 kişilik örneklem büyüklüğünde karakteristik eğri

yöntemlerinde kız öğrenciler için elde edilen BSE değerleri erkek öğrencilere göre daha küçüktür.

Tablo 27 incelendiğinde erkek öğrenciler için elde edilen İSE değerlerinin karma formattaki tüm ölçek dönüştürme yöntemlerinde kız öğrencilere göre daha küçük olduğu görülmektedir. İki kategorili formatta 1000 kişilik örneklem büyüklüğünde moment yöntemlerde ve 3000 ile 5000 kişilik örneklem büyüklüklerinde karakteristik eğri yöntemlerinde kız öğrenciler için elde edilen İSE değerleri daha küçüktür.

c) Örneklem büyüklüğüne (1000, 3000 ve 5000) göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 25, Tablo 26 ve Tablo 27'den elde edilen bulgular ve bu bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 25 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede ham puanlara ilişkin T değerlerinin karma formatta örneklem büyüklüğünden çok etkilenmediği görülmektedir. İki kategorili formatta ise örneklem büyüklüğü arttıkça T değerlerinin küçüldüğü görülmektedir. Karma format ve iki kategorili format testte eşitleme sonucu oluşan bdf ler arasındaki farklar incelendiğinde ölçek dönüştürme yöntemine göre değişmekle birlikte örneklem büyüklüğünün artmasıyla T değerlerinde tek yönlü bir artış ya da azalış olmadığı görülmektedir.

Tablo 26 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede karma formatta 1000 kişilik örneklem büyüklüğünde moment yöntemlerde daha küçük BSE değerlerinin elde edildiği görülmektedir. Örneklem büyüklüğü arttıkça BSE değeri de artmıştır. Bu artış daha detaylı incelendiğinde ortak maddeler üzerinden yapılan eşitlemede bir madde için kestirilen parametre değerlerinin her iki test için oldukça farklı elde edildiği görülmektedir. Bu aşamada moment yöntemlerin madde parametrelerine karşı oldukça duyarlı olduğu unutulmamalıdır. Karakteristik eğri yöntemleri 1000, 3000 ve 5000 kişilik örneklem büyüklüklerinde yaklaşık aynı BSE değerlerini üretmiştir. İki kategorili formatta 1000 kişilik örneklem büyüklüğünde tüm ölçek dönüştürme yöntemleri en küçük BSE değerlerini vermiştir.

Tablo 27 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede karma formatta 1000 kişilik örneklem büyüklüğünde ve tüm ölçek dönüştürme yöntemlerinde, iki kategorili formatta 1000 kişilik büyüklüğünde ve O-O, O-S ve HB ölçek dönüştürme yöntemlerinde elde edilen İSE değerlerinin Ger_Pn ve Göz_Pn puan eşitlemede daha küçük elde edildiği görülmektedir.

d) Ölçek dönüştürme yöntemlerine (O-O, O-S, HB, SL) göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 25, Tablo 26 ve Tablo 27'den elde edilen bulgular ve bu bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 25 incelendiğinde karma format testte eşitleme sonucu oluşan *bdfler* arasındaki farklar incelendiğinde ölçek dönüştürme yöntemine göre değişmekle birlikte *T* değerlerinin çoğunda belirgin azalma olduğu ve ADÖ'nün sağlanmasına yönelik yeterli sonuçlar üretildiği görülmektedir. O-O ve O-S ölçek dönüştürme yönteminden elde edilen sonuçlar diğer ölçek yöntemlerinden ADÖ'nün korunmasında daha olumsuz sonuçlar üretmiştir. Karma format testlerin eşitlenmesi sonucu oluşan *T* değerleri özellikle HB ve SL ölçek dönüştürme yöntemleri için benzer sonuçlar üretmiş olup SL ölçek dönüştürme yöntemi için elde edilen *T* değerleri genel olarak daha küçüktür. İki kategorili formatta hem moment yöntem hem de karakteristik eğri yöntemlerine göre *T* değerlerinde artış yaşanmıştır. Hem karma hem de iki kategorili formatta Göz_Pn eşitleme sonucu oluşan *T* değerleri ile Ger_Pn eşitleme sonucu oluşan *T* değerleri benzerdir.

Tablo 26 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede ölçek dönüştürme yöntemlerine göre örneklem büyüklüğünün artmasıyla karma formatta Ger_Pn ve Göz_Pn eşitlemede moment yöntemlerde BSE değerinin daha büyük elde edildiği görülmektedir. Karakteristik eğri yöntemleri ise birbirine daha benzer BSE değerleri üretmiştir. İki kategorili formatta ise 1000 kişilik örneklem büyüklüğünde Ger_Pn ve Göz_Pn eşitlemede karakteristik eğri yöntemleri benzer ve en küçük BSE değerlerini üretmiştir.

Tablo 27 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede ölçek dönüştürme yöntemlerine göre 1000 kişilik örneklem büyüklüğünde karma ve iki kategorili formatta Ger_Pn ve Göz_Pn eşitlemede İSE değerlerinin çoğunlukla daha küçük elde edildiği görülmektedir.

Yukarıda yer alan 4 alt problemin birlikte değerlendirilmesiyle aşağıda yer alan yorum yapılabilir.

Tablo 25 incelendiğinde karma formatta en küçük T değerinin, 3000 kişilik örneklem büyüklüğü SL ölçek dönüştürme yöntemi $Göz_Pn$ eşitleme ve kız öğrenci verisinde elde edildiği görülmektedir (%1.24). Bu verinin başlangıç T değeri de %13.37'dir. İki kategorili formatta en küçük T değerinin, 5000 kişilik örneklem büyüklüğü SL ölçek dönüştürme yöntemi Ger_Pn eşitleme ve kız öğrencilerde elde edildiği görülmektedir (%1.37). Bu verinin başlangıç T değeri de %2.22'dir.

Tablo 26 incelendiğinde en küçük BSE değerinin iki kategorili format 5000 kişilik örneklem büyüklüğü karakteristik eğri yöntemlerinden olan SL ölçek dönüştürme yöntemi $Göz_Pn$ eşitleme ve kız öğrenciler için elde edildiği görülmektedir (0.0158). BSE değerleri küçükten büyüğe sıralandığında 144 durumdan ilk 43'ünü iki kategorili format oluşturmaktadır. Bu durum verilerin iki kategorili olması durumunda BSE'nin daha iyi sağlandığını göstermektedir. Karakteristik ölçek döndürme yöntemleri moment yöntemlere göre daha iyi performans sergilemiştir.

Tablo 27 incelendiğinde en küçük İSE değerinin iki kategorili formatta 1000 kişilik örneklem büyüklüğü karakteristik eğri yöntemlerinden olan SL ölçek dönüştürme yöntemi $Göz_Pn$ eşitleme ve kız öğrenciler için elde edildiği görülmektedir (0.0089). İSE değerleri küçükten büyüğe sıralandığında 144 durumdan ilk 31'inin 12'sini $Göz_Pn$ eşitleme sonuçları oluşturmaktadır. Yine bu 31 durumun 30'unu iki kategorili format oluşturmaktadır. Bu sonuç iki kategorili format ve $Göz_Pn$ eşitlemede İSE'nin genellikle daha iyi sağlandığını göstermektedir.

Dördüncü Alt Probleme Ait Bulgular ve Yorumlar

Sosyal Bilgiler testi için denk olmayan gruplarda ortak madde deseninde gerçek/gözlenen puan eşitleme yöntemi ile C formunun B formuna eşitlenmesinde aynı dağılım ve eşitlik özelliğinin korunumu,

- a. *maddelerin formatına (çok kategorili puanlanan, iki kategorili puanlanan),*
- b. *testi alan alt gruplara (kız, erkek),*
- c. *örneklem büyüklüğüne (1000, 3000 ve 5000) ve*

d. ölçek dönüştürme yöntemlerine (O-O, O-S, HB, SL) göre nasıl değişmektedir?

Sosyal Bilgiler testinde ham puan T değerleri ve C formunun B formuna gerçek puan (Ger_Pn) ve gözlenen puan ($Göz_Pn$) eşitleme yöntemiyle eşitlenmesi sonucu oluşan T değerleri Tablo 28’de, birinci sıra eşitlik (BSE) değerleri Tablo 29’da, ikinci sıra eşitlik (İSE) değerleri Tablo 30’da verilmiştir.

Tablo 28

Sosyal Bilgiler Testi C Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi - Ham Puan ve Ölçek Puan Dağılımları için Kolmogorov T İstatistikleri (%)

F	Örn.	Grup	Ham Puan Dağılımları	O-O		O-S		HB		SL	
				Ger_Pn	$Göz_Pn$	Ger_Pn	$Göz_Pn$	Ger_Pn	$Göz_Pn$	Ger_Pn	$Göz_Pn$
1	1000	Toplam	24.20	6.71	6.74	7.93	7.75	6.62	6.60	7.32	7.29
		Erkek	22.11	4.20	4.21	7.32	7.14	5.35	5.34	7.50	7.50
		Kız	28.86	7.10	6.91	7.07	6.88	7.43	7.17	7.66	7.40
1	3000	Toplam	23.37	4.60	4.31	5.44	5.25	3.50	3.24	4.03	3.86
		Erkek	21.98	7.48	7.09	7.42	7.02	4.23	3.99	5.10	4.80
		Kız	25.01	3.44	3.06	5.35	4.17	4.75	3.63	5.39	4.79
1	5000	Toplam	22.02	7.16	6.15	7.78	6.81	8.08	7.21	7.22	6.19
		Erkek	20.44	5.38	4.26	5.99	4.73	6.10	5.23	6.01	4.84
		Kız	24.04	9.47	8.63	10.46	8.81	10.99	9.49	9.33	8.00
2	1000	Toplam	6.30	3.96	3.63	5.38	5.51	3.19	3.35	4.70	4.48
		Erkek	6.95	1.60	2.15	4.42	4.43	2.30	2.56	4.95	4.15
		Kız	7.84	16.94	15.93	16.24	15.33	16.89	15.92	16.60	15.56
2	3000	Toplam	6.00	4.73	3.93	4.73	4.30	3.67	2.64	4.36	3.33
		Erkek	6.98	6.31	5.64	6.27	5.60	3.43	2.64	3.64	2.85
		Kız	6.57	3.22	2.19	5.65	3.98	4.65	3.45	4.95	4.26
2	5000	Toplam	5.12	4.16	3.46	5.63	4.99	4.52	3.93	5.07	4.51
		Erkek	4.80	6.91	5.72	7.46	6.58	6.65	5.55	6.24	5.28
		Kız	5.50	4.20	3.27	5.34	4.74	5.20	4.68	4.58	3.86

F: Madde Formatı, Örn: Örneklem, 1: Karma Format Test, 2: İki Kategorili Formata Dönüştürülen Test

a) Maddelerin formatına göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 28, Tablo 29 ve Tablo 30’dan elde edilen bulgular ve bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 28 incelendiğinde ham puanlara ilişkin başlangıç T değerlerinde karma formattan iki kategorili formata geçilmesiyle önemli düzeyde azalmalar yaşandığı görülmektedir. İki kategorili formatın eşitleme öncesinde A ve B formlarının birikimli dağılım fonksiyonlarını (bdf) daha benzer hale getirdiği söylenebilir. Karma format testte eşitleme sonucu oluşan bdf ’ler arasındaki farklar incelendiğinde moment ve karakteristik eğri yöntemlerinin benzer T değerleri ürettiği görülmektedir. Eşitleme sonrası iki kategorili formatta ise T değerlerinde genellikle azalış geçerken 1000

kişilik kız öğrenci örneklem büyüklüğünün tamamında ve 5000 kişilik erkek öğrenci örneklem büyüklüğünde ise kısmen T değerlerinde artış gerçekleştiği ve aynı dağılım özelliğinin (ADÖ) sağlandığı görülmektedir. Özellikle 1000 kişilik kız öğrenci örneğine ilişkin elde edilen değerler incelendiğinde iki kategorili formatta eşitleme sonrası T değerlerinin tüm ölçek dönüştürme yöntemlerinde oldukça yüksek olduğu görülmektedir. EK-N'de yer alan grafikler incelendiğinde ise iki kategorili formata dönüştürülen 1000 kişilik kız öğrenci verisinin dönüştürme sonrasında çizilen grafiklerinin diğer örneklemeler üzerinden çizilen grafiklere göre farklılaştığı görülmektedir. Seçilen örneklemden kaynaklı olarak maddelerin iki kategorili formata dönüştürülmesinden önce kısmi puan alan öğrencilerin çokluğu ve dönüştürme sonrası kısmi puan alanlara 0 puan verilmesi nedeniyle başarısız görünen öğrenci sayısının artmasının bu bozulmalara neden olduğu düşünülebilir. Daha büyük örneklemelerde bu sorunun yaşanmaması yeniden 1000 kişilik örneklem alınması durumunda bu sınırlılığın ortadan kalkabileceğini göstermektedir.

Tablo 29

Sosyal Bilgiler Testinde C Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan BSE Değerleri

F	Örn.	Grup	O-O		O-S		HB		SL	
			Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn
1	1000	Toplam	0.5708	0.5904	0.5518	0.5712	0.5476	0.5664	0.5784	0.6068
		Erkek	0.5584	0.5766	0.5236	0.5418	0.5535	0.5716	0.6033	0.6319
		Kız	0.5423	0.5628	0.5433	0.5632	0.5458	0.5660	0.5629	0.5900
1	3000	Toplam	0.5313	0.5486	0.5222	0.5399	0.4937	0.5102	0.5350	0.5643
		Erkek	0.5628	0.5832	0.5672	0.5877	0.5084	0.5265	0.5314	0.5589
		Kız	0.5269	0.5435	0.4959	0.5144	0.5052	0.5225	0.5612	0.5888
1	5000	Toplam	0.5320	0.5490	0.5159	0.5329	0.5038	0.5202	0.5310	0.5612
		Erkek	0.5425	0.5598	0.5327	0.5502	0.5155	0.5323	0.5273	0.5532
		Kız	0.5392	0.5563	0.5202	0.5377	0.5053	0.5223	0.5489	0.5765
2	1000	Toplam	0.1344	0.1388	0.1448	0.1478	0.1198	0.1225	0.1451	0.1512
		Erkek	0.0866	0.0884	0.1103	0.1085	0.1105	0.1124	0.1639	0.1719
		Kız	0.1576	0.1624	0.1530	0.1583	0.1589	0.1644	0.1507	0.1518
2	3000	Toplam	0.1554	0.1593	0.1612	0.1648	0.1325	0.1338	0.1448	0.1482
		Erkek	0.1995	0.2069	0.1983	0.2061	0.1392	0.1439	0.1438	0.1507
		Kız	0.1137	0.1144	0.1280	0.1244	0.1381	0.1375	0.1560	0.1579
2	5000	Toplam	0.1272	0.1302	0.1415	0.1435	0.1265	0.1279	0.1388	0.1415
		Erkek	0.1639	0.1683	0.1686	0.1735	0.1570	0.1607	0.1508	0.1580
		Kız	0.1082	0.1062	0.1261	0.1223	0.1087	0.1031	0.1342	0.1306

F: Madde Formatı, Örn: Örneklem, 1: Karma Format Test, 2: İki Kategorili Formata Dönüştürülen Test

Tablo 30

Sosyal Bilgiler Testinde C Formunun B Formuna Gerçek/Gözlenen Puan Eşitleme Yöntemiyle Eşitlenmesi Sonucu Oluşan İSE Değerleri

F	Örn.	Grup	O-O		O-S		HB		SL	
			Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn	Ger_ Pn	Göz_ Pn
1	1000	Toplam	0.0690	0.0627	0.0678	0.0557	0.0658	0.0588	0.0702	0.0600
		Erkek	0.0700	0.0681	0.0631	0.0528	0.0689	0.0641	0.0741	0.0648
		Kız	0.0659	0.0547	0.0660	0.0549	0.0671	0.0544	0.0692	0.0549
1	3000	Toplam	0.0617	0.0540	0.0616	0.0508	0.0564	0.0506	0.0622	0.0544
		Erkek	0.0694	0.0570	0.0687	0.0582	0.0581	0.0547	0.0618	0.0574
		Kız	0.0640	0.0594	0.0607	0.0492	0.0611	0.0516	0.0678	0.0572
1	5000	Toplam	0.0632	0.0615	0.0601	0.0535	0.0583	0.0540	0.0621	0.0545
		Erkek	0.0677	0.0637	0.0667	0.0592	0.0636	0.0598	0.0654	0.0597
		Kız	0.0593	0.0581	0.0552	0.0495	0.0536	0.0488	0.0587	0.0510
2	1000	Toplam	0.0311	0.0229	0.0430	0.0309	0.0297	0.0215	0.0354	0.0261
		Erkek	0.0283	0.0222	0.0477	0.0343	0.0348	0.0247	0.0441	0.0322
		Kız	0.0380	0.0271	0.0279	0.0200	0.0346	0.0247	0.0409	0.0297
2	3000	Toplam	0.0317	0.0225	0.0379	0.0271	0.0275	0.0195	0.0285	0.0207
		Erkek	0.0403	0.0291	0.0390	0.0283	0.0217	0.0158	0.0221	0.0158
		Kız	0.0311	0.0235	0.0472	0.0355	0.0422	0.0321	0.0429	0.0334
2	5000	Toplam	0.0241	0.0190	0.0373	0.0285	0.0300	0.0230	0.0309	0.0240
		Erkek	0.0364	0.0293	0.0407	0.0326	0.0362	0.0292	0.0358	0.0281
		Kız	0.0204	0.0135	0.0402	0.0294	0.0330	0.0236	0.0325	0.0244

F: Madde Formatı, Örn: Örneklem, 1: Karma Format Test, 2: İki Kategorili Formata Dönüştürülen Test

Tablo 29 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede iki kategorili formatta en küçük BSE değerinin, 1000 kişilik örneklem büyüklüğü karakteristik eğri yöntemlerinden olan *HB* ölçek dönüştürme yöntemi ve *Ger_Pn* eşitlemede elde edildiği görülmektedir (0.1198). Karma formatta ise en küçük BSE değerinin, 3000 kişilik örneklem büyüklüğü karakteristik eğri yöntemlerinden olan *HB* ölçek dönüştürme yöntemi ve *Ger_Pn* eşitlemede elde edildiği görülmektedir (0.4937). İki kategorili formata geçilmesiyle tüm örneklem büyüklüklerinden elde edilen BSE değerleri tüm ölçek dönüştürme yöntemlerine göre hem *Ger_Pn* hem de *Göz_Pn* eşitlemede daha da küçülmüştür.

Tablo 30 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede iki kategorili formatta en küçük İSE değerinin, 5000 kişilik örneklem büyüklüğü moment yöntemlerinden olan *O-O* ölçek dönüştürme yöntemi ve *Göz_Pn* eşitlemede elde edildiği görülmektedir (0.0190). Karma formatta ise en küçük İSE değerinin, 3000 kişilik örneklem büyüklüğü karakteristik eğri yöntemlerinden olan *HB* ölçek dönüştürme yöntemi ve *Göz_Pn* eşitlemede elde edildiği görülmektedir (0.0506). İki kategorili formata geçilmesiyle tüm örneklem büyüklüklerinden elde edilen İSE değerlerinde küçülme yaşanmıştır.

b) Testi alan gruplara (erkek-kız) göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 28, Tablo 29 ve Tablo 30'dan elde edilen bulgular ve bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 28 incelendiğinde ham puanlara ilişkin T değerlerinin karma formatta kız öğrenciler için daha yüksek olduğu görülmektedir. İki kategorili formatta ise 1000 ve 5000 kişilik örneklem büyüklüklerinde kız öğrenciler, 5000 kişilik örneklem büyüklüklerinde ise erkek öğrenciler için başlangıç T değerlerinin daha büyük olduğu görülmektedir. Karma format testlerde eşitleme sonucu oluşan bdf ler arasındaki farklar incelendiğinde ölçek dönüştürme yöntemine göre değişmekle birlikte özellikle karakteristik eğri yöntemlerinde hem erkek hem de kız öğrenciler için T değerlerinde belirgin azalmalar olduğu görülmektedir. İki kategorili format testte eşitleme sonucu oluşan bdf ler arasındaki farklar incelendiğinde T değerlerinde tek yönlü bir artış ya da azalış olmadığı görülmektedir. İki kategorili format testte 1000 kişilik örneklem büyüklüğünde kız öğrencilerde eşitleme sonrası T değerleri hem moment hem de karakteristik eğri yöntemlerine göre önemli ölçüde artmıştır. Buna neden olarak örneklem büyüklüğünün azlığından dolayı madde parametrelerindeki bozulmalar gösterilebilir.

Tablo 29 incelendiğinde kız ve erkek öğrenciler için elde edilen BSE değerleri karma formatta tüm ölçek dönüştürme yöntemlerinde birbirine oldukça yakın elde edilmiştir. İki kategorili formatta 1000 kişilik örneklem büyüklüğünde O-O, O-S ve HB ölçek dönüştürme yönteminde, 3000 kişilik örneklem büyüklüğünde SL ölçek dönüştürme yönteminde erkek öğrenciler için daha küçük elde edilirken diğer durumlarda kızlar için daha küçük elde edilmiştir.

Tablo 30 incelendiğinde kız öğrenciler için İSE değerlerinin iki kategorili ve karma formatta 5000 kişilik örneklem büyüklüğünde daha küçük elde edildiği görülmektedir.

c) Örneklem büyüklüğüne (1000, 3000 ve 5000) göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 28, Tablo 29 ve Tablo 30'dan elde edilen bulgular ve bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 28 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede ham puanlara ilişkin T değerlerinin karma ve iki kategorili formatta örneklem büyüklüğünden çok etkilenmediği, ancak bir miktar azaldığı görülmektedir. Karma ve iki kategorili format testte eşitleme sonucu oluşan bdf ler arasındaki farklar incelendiğinde ölçek dönüştürme yöntemine göre değişmekle birlikte örneklem büyüklüğünün artmasıyla T değerlerinde tek yönlü bir artış ya da azalış olmadığı görülmektedir.

Tablo 29 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede örneklem büyüklüğünün artmasıyla ölçek dönüştürme yöntemlerine göre elde edilen BSE değerlerinde kararlı artış ya da azalış gerçekleşmediği görülmektedir.

Tablo 30 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede örneklem büyüklüğünün artmasıyla ölçek dönüştürme yöntemlerine göre elde edilen İSE değerlerinde kararlı artış ya da azalış gerçekleşmediği görülmektedir.

d) Ölçek dönüştürme yöntemlerine (O-O, O-S, HB, SL) göre ADÖ, BSE ve İSE nasıl değişmektedir?

Tablo 28, Tablo 29 ve Tablo 30'dan elde edilen bulgular ve bulgulara göre oluşturulan yorumlar aşağıda verilmiştir.

Tablo 28 incelendiğinde karma format testte eşitleme sonucu oluşan bdf ler arasındaki farklar incelendiğinde ölçek dönüştürme yöntemine göre değişmekle birlikte T değerlerinin çoğunda belirgin azalma olduğu ve ADÖ'nün sağlanmasına yönelik yeterli sonuçlar üretildiği görülmektedir. Moment yöntemler ve karakteristik eğri yöntemleri benzer T değerleri vermiştir. İki kategorili formatta hem moment yöntem hem de karakteristik eğri yöntemlerine göre T değerlerinde çoğunlukla bir azalma yaşanmıştır. Hem karma hem de iki kategorili formatta $Göz_Pn$ eşitleme sonucu oluşan T değerleri Ger_Pn eşitleme sonucu oluşan T değerlerine göre çoğunlukla daha küçük elde edilmiştir.

Tablo 29 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede ölçek dönüştürme yöntemlerine göre BSE değerlerinde kararlı bir artış ya da azalış gerçekleşmediği, Ger_Pn eşitlemenin $Göz_Pn$ eşitmeye göre daha küçük BSE değeri verdiği görülmektedir.

Tablo 30 incelendiğinde toplam öğrenci verisi üzerinden yapılan değerlendirmede ölçek dönüştürme yöntemlerine göre İSE değerlerinde kararlı bir artış ya da azalış gerçekleşmediği, *Göz_Pn* eşitlemenin *Ger_Pn* eşitlemeye göre daha küçük İSE değeri verdiği görülmektedir.

Yukarıda yer alan 4 alt problemin birlikte değerlendirilmesiyle aşağıda yer alan yorum yapılabilir.

Tablo 28 incelendiğinde karma formatta en küçük *T* değerinin, 3000 kişilik örneklem büyüklüğü *O-O* ölçek dönüştürme yöntemi *Göz_Pn* eşitleme ve kız öğrenci verisinde elde edildiği görülmektedir (%3.06). Bu verinin başlangıç *T* değeri de %25.01'dir. İki kategorili formatta en küçük *T* değerinin, 1000 kişilik örneklem büyüklüğü *O-O* ölçek dönüştürme yöntemi *Ger_Pn* eşitleme ve erkek öğrenci verisinde elde edildiği görülmektedir (%1.60). Bu verinin başlangıç *T* değeri de %6.95'tir.

Tablo 29 incelendiğinde en küçük BSE değerinin, iki kategorili format 1000 kişilik örneklem büyüklüğü moment yöntemlerden olan *O-O* ölçek dönüştürme yöntemi *Ger_Pn* eşitleme ve erkek öğrenci verisinde elde edildiği görülmektedir (0.0866). *Format*Örneklem*Grup*Yöntem* çaprazlamasının hepsinde (72 durumdan 72'sinde) iki kategorili formattan elde edilen BSE değerleri karma formata göre daha küçüktür.

Tablo 30 incelendiğinde en küçük İSE değerinin, iki kategorili formatta 5000 kişilik örneklem büyüklüğü moment yöntemlerden olan *O-O* ölçek dönüştürme yöntemi *Göz_Pn* eşitleme ve kız öğrenci verisinde elde edildiği görülmektedir (0.0135). *Format*Örneklem*Grup*Yöntem* çaprazlamasının tamamında (72 durumdan 72'sinde) iki kategorili formattan elde edilen İSE değerleri karma formattan elde edilenlere göre daha küçüktür.

Bölüm 5

Sonuç, Tartışma ve Öneriler

Çalışmanın bu bölümünde elde edilen bulgulara dayalı olarak ulaşılan sonuçlara ve bu sonuçlara göre oluşturulan önerilere yer verilmiştir.

Sonuçlar

Bu çalışmadan elde edilen sonuçlar alt problem başlıklarına göre aşağıda verilmiştir.

Madde Formatına İlişkin Sonuçlar

Madde formatına ilişkin elde edilen sonuçlar aşağıda çalışma yapılan Matematik ve Sosyal Bilgiler testi başlıkları altında verilmiştir.

Matematik testleri. Formların birikimli dağılım fonksiyonları (*bdf*) arasındaki farkın az olduğu durumlarda karma formattan iki kategorili formata geçilmesinin *bdf*lerin arasındaki farkın artmasına neden olduğu ancak eşitleme sonrası hem karma hem de iki kategorili formatta *T* değerlerinde *O-S*, *HB* ve *SL* ölçek dönüştürme yönteminde azalma yaşandığı ve aynı dağılım özelliğinin (*ADÖ*) korunumunun sağlandığı görülmüştür. Formların *bdf*leri arasındaki farkların büyük olduğu durumlarda karma formattan iki kategorili formata geçilmesi *bdf*ler arasındaki farkların kısmen azalmasını sağlamıştır. Karma formatta eşitleme sonrası *T* değerlerinde; *O-S*, *HB* ve *SL* ölçek dönüştürme yönteminde keskin düşüşler yaşandığı ve *ADÖ*'nün korunumunun sağlandığı ancak iki kategorili formatta eşitleme sonrası *ADÖ*'nün korunumuna ilişkin sonuca ulaşılmadığı görülmüştür. Gözlenen Puan (*Göz_Pn*) eşitlemenin Gerçek Puan (*Ger_Pn*) eşitlemeye göre hem karma hem de iki kategorili formatta çoğunlukla daha küçük *T* değerlerinin elde edilmesini sağladığı sonucuna ulaşılmıştır.

Karma formattan iki kategorili formata dönüştürülen ve B formuyla eşitlenen A ve C formlarından BSE korunumu için farklı sonuçlar elde edilmiştir. Daha az benzer olan ve iki kategorili formata dönüştürülmesi sonucunda kısmen benzerliği artan ve eşitleme sonrası *T* değerlerinde kısmi artış meydana gelen A-B eşitlemesinde daha küçük *T* değerleri elde edilen yalnızca 5000 kişilik örneklem büyüklüklerinde BSE değerlerinin iki kategorili formatta daha küçük elde edildiği görülmüştür. Daha benzer olan C-B eşitlemesinde BSE değerleri çoğunlukla iki

kategorili format için daha küçük elde edilmiştir. Karma formatta $Göz_Pn$ için elde edilen BSE değerleri çoğunlukla Ger_Pn için elde edilen değerlerden daha küçüktür.

Karma formattan iki kategorili formata dönüştürülen ve B formuyla eşitlenen A ve C formlarına göre İSE için farklı sonuçlar elde edilmiştir. Daha az benzer olan ve iki kategorili format için T değerlerinde kısmi artış yaşanan A ve B formlarının eşitlenmesinde İSE değerleri tüm durumlarda karma format için daha küçük bulunmuştur. Daha benzer olan formlar için İSE'nin belli bir format için küçük elde edilmesi durumu oluşmamıştır.

Sosyal bilgiler testleri. Formların $bd\hat{f}$ leri arasındaki farkın çok olduğu durumlarda karma formattan iki kategorili formata geçilmesinin $bd\hat{f}$ lerin arasındaki farkın azalmasını sağladığı ancak iki kategorili formata geçilmesiyle $bd\hat{f}$ lerin arasındaki farkın oldukça azalması durumunda eşitleme sonrası T değerlerinde kısmi artışlar yaşandığı görülmüştür. Format dönüşümü sonrasında $bd\hat{f}$ lerin daha makul seviyede azaldığı durumlarda eşitleme sonrası T değerlerinde bir miktar daha azalma yaşandığı sonucuna ulaşılmıştır. Hem karma hem de iki kategorili formatta T değerlerinde meydana gelen azalışlar ADÖ'nün korunumunun sağlandığı göstermiştir. $Göz_Pn$ çoğunlukla Ger_Pn eşitlemeden daha küçük T değerlerinin elde edilmesini sağlamıştır.

Karma formattan iki kategorili formata dönüştürülen formlardan elde edilen BSE değerlerinin neredeyse tamamında oldukça büyük düşüşler yaşanmıştır (A formundaki 1000 kişilik örneklem dışında). Bu durum iki kategorili formatın BSE'nin korunumuna oldukça büyük katkı sağladığını göstermektedir. Ger_Pn eşitlemeden elde edilen BSE değerleri çoğunlukla $Göz_Pn$ eşitlemeden elde edilen değerlerden daha küçüktür.

Karma formattan iki kategorili formata dönüştürülen formlardan elde edilen İSE değerlerinin neredeyse tamamında oldukça büyük düşüşler yaşanmıştır (A formundaki 1000 kişilik örneklem dışında). Bu durum iki kategorili formatın İSE'nin korunumuna oldukça büyük katkı sağladığını göstermektedir. $Göz_Pn$ eşitlemeden elde edilen İSE değerleri çoğunlukla Ger_Pn eşitlemeden elde edilen değerlerden daha küçüktür.

Testi Alan Alt Gruplara İlişkin Sonuçlar

Test formatına ilişkin elde edilen sonuçlar aşağıda çalışma yapılan Matematik ve Sosyal Bilgiler testi başlıkları altında verilmiştir.

Matematik testleri. Cinsiyete göre formların daha az benzer olduğu durumlarda (*bdf*lerin arasındaki farkların büyük olduğu) eşitleme sonrasında *T* değerlerindeki azalışla formların benzerliğinin hem kız hem de erkek öğrenciler için oldukça arttığı ve ADÖ'nün korunumunun sağlandığı ancak formların daha benzer olduğu kız öğrenciler için (*bdf*lerin arasındaki farkların küçük olduğu) eşitleme sonrasında *T* değerlerinde kısmi artışlarla formların benzerliğinin azaldığı sonucuna ulaşılmıştır. *Göz_Pn* eşitlemenin *Ger_Pn* eşitlemeye göre hem kız hem de erkek öğrenciler için çoğunlukla daha küçük *T* değerlerinin elde edilmesini sağladığı sonucuna ulaşılmıştır.

Aynı örneklem büyüklükleri içerisinde genellikle kız öğrenciler için elde edilen BSE değerlerinin daha küçük olduğu ancak A-B eşitlemesindeki 5000 kişilik örneklem büyüklüklerinde erkek öğrenciler için kısmen BSE'nin daha küçük elde edildiği görülmüştür.

Aynı örneklem büyüklükleri içerisinde genellikle kız öğrenciler için elde edilen İSE değerlerinin daha küçük olduğu ancak formların benzerlik ve farklılığına göre kimi zaman karma kimi zamansa iki kategorili formatta kız öğrenciler için elde edilen İSE'nin genellikle daha küçük olduğu görülmüştür.

Sosyal bilgiler testleri. Cinsiyete göre formların daha az benzer olduğu durumlarda eşitleme sonrasında *T* değerlerindeki azalışla formların benzerliğinin hem kız hem de erkek öğrenciler için oldukça arttığı ve ADÖ'nün korunumunun sağlandığı görülmüştür. *Göz_Pn* eşitlemenin *Ger_Pn* eşitlemeye göre hem kız hem de erkek öğrenciler için çoğunlukla daha küçük *T* değerlerinin elde edilmesini sağladığı sonucuna ulaşılmıştır.

Aynı örneklem büyüklükleri içerisinde belli bir cinsiyet için tek yönlü olarak BSE değerlerinin daha küçük elde edilmediği, formata göre kimi zaman erkek kimi zaman kız öğrenciler için BSE'nin daha küçük elde edildiği görülmüştür.

Aynı örneklem büyüklükleri içerisinde belli bir cinsiyet için tek yönlü olarak İSE değerlerinin daha küçük elde edilmediği, formata göre kimi zaman erkek kimi zaman kız öğrenciler için İSE'nin daha küçük elde edildiği görülmüştür.

Örneklem Büyüklüğüne İlişkin Sonuçlar

Örneklem büyüklüğüne ilişkin elde edilen sonuçlar aşağıda çalışma yapılan Matematik ve Sosyal Bilgiler testi başlıkları altında verilmiştir.

Matematik testleri. Toplam öğrenci verisi üzerinden yapılan değerlendirmelere göre çoğunlukla 5000 kişilik örneklem büyüklüğünde T değerlerinin daha küçük elde edildiği ve $Göz_Pn$ eşitlemenin Ger_Pn eşitmeye göre genellikle ADÖ'nün korunumuna daha çok katkı sağladığı görülmüştür.

Toplam öğrenci verisi üzerinden yapılan değerlendirmelere göre bir A-B eşitlemesinde 5000 kişilik örneklem büyüklükleri için elde BSE değerlerinin diğer örneklem büyüklüklerine göre genellikle daha küçük elde edildiği ancak örneklem büyüklüğü artışının kararlı bir şekilde BSE değerinin küçük elde edilmesinde etkili olmadığı görülmüştür. Örneğin, 3000 kişilik örneklem büyüklüğünde elde BSE değeri 1000 kişilik örneklem büyüklüğü için elde değerden büyükken 5000 kişilik örneklem büyüklüğü için elde edilen BSE değeri kimi durumlarda her ikisinden de küçüktür. C-B eşitlemesinde ise örneklem büyüklüğü ile BSE değerleri arasında bir ilişki bulunamamıştır. $Göz_Pn$ için elde edilen BSE değerleri Ger_Pn için elde edilen BSE değerlerine göre çoğunlukla daha küçük bulunmuştur.

Toplam öğrenci verisi üzerinden yapılan değerlendirmelere göre A-B eşitlemesinde 5000 kişilik örneklem büyüklükleri için elde İSE değerlerinin diğer örneklem büyüklüklerine göre genellikle daha küçük elde edildiği ancak örneklem büyüklüğü artışının kararlı bir şekilde İSE değerinin küçük elde edilmesinde etkili olmadığı görülmüştür. Örneğin, 3000 kişilik örneklem büyüklüğünde elde İSE değeri 1000 kişilik örneklem büyüklüğü için elde değerden büyükken 5000 kişilik örneklem büyüklüğü için elde edilen İSE değeri kimi durumlarda her ikisinden de küçüktür. C-B eşitlemesinde ise örneklem büyüklüğü ile İSE değerleri arasında bir ilişki bulunamamıştır. $Göz_Pn$ için elde edilen İSE değerleri Ger_Pn için elde edilen İSE değerlerine göre çoğunlukla daha küçük bulunmuştur.

Sosyal bilgiler testleri. Toplam öğrenci verisi üzerinden yapılan değerlendirmelere belli bir örneklem büyüklüğüne göre T değerlerinin daha küçük elde edilmediği ve $Göz_Pn$ eşitlemenin ölçek dönüştürme yöntemine göre değişmekle birlikte Ger_Pn eşitlemeye göre çoğunlukla ADÖ'nün korunumuna daha çok katkı sağladığı görülmüştür.

Toplam öğrenci verisi üzerinden yapılan değerlendirmelere göre belli bir örneklem büyüklüğüne göre BSE'nin daha küçük elde edilmediği görülmüştür. $Göz_Pn$ için elde edilen BSE değerleri Ger_Pn için elde edilen BSE değerlerine göre çoğunlukla daha küçük bulunmuştur.

Toplam öğrenci verisi üzerinden yapılan değerlendirmelere göre belli bir örneklem büyüklüğüne göre İSE'nin daha küçük elde edilmediği görülmüştür. $Göz_Pn$ için elde edilen İSE değerleri Ger_Pn için elde edilen İSE değerlerine göre çoğunlukla daha küçük bulunmuştur.

Ölçek Dönüştürme Yöntemlerine İlişkin Sonuçlar

Ölçek dönüştürme yöntemlerine ilişkin elde edilen sonuçlar aşağıda çalışma yapılan Matematik ve Sosyal Bilgiler testi başlıkları altında verilmiştir.

Matematik testleri. Matematik testlerinin eşitlemesinde O-S, HB ve SL ölçek dönüştürme yöntemi daha küçük T değerlerinin elde edilmesine yardımcı olduğu, SL ölçek dönüştürme yönteminin diğer yöntemlere göre genellikle daha küçük ve kararlı T değerleri vermesi nedeniyle ADÖ'nün korunmasına daha fazla katkı sağladığı görülmektedir. $Göz_Pn$ eşitleme için elde edilen T değerlerinin Ger_Pn için elde edilen T değerlerinden genellikle daha küçük T değeri elde edilmesine katkı sağladığı ve ADÖ'nün korunmasında daha etkili olduğu sonucuna ulaşılmıştır.

Formların eşitlenmesinde karakteristik ölçek dönüştürme yöntemleri çoğunlukla daha iyi performans göstermesine rağmen en küçük BSE değerleri moment yöntemlerde elde edilmiştir. $Göz_Pn$ için elde edilen BSE değerleri Ger_Pn için elde edilen BSE değerlerine göre çoğunlukla daha küçük bulunmuştur.

Formların eşitlenmesinde karakteristik ölçek dönüştürme yöntemleri çoğunlukla daha iyi performans göstermesine rağmen en küçük İSE değerleri moment yöntemlerde elde edilmiştir. $Göz_Pn$ için elde edilen İSE değerleri Ger_Pn için elde edilen İSE değerlerine göre çoğunlukla daha küçük bulunmuştur.

Sosyal bilgiler testleri. Sosyal Bilgiler testlerinin eşitlemesinde karakteristik eğri dönüştürme yöntemlerinin (HB ve SL) daha küçük T değerlerinin elde edilmesine yardımcı olması nedeniyle ADÖ'nün korunmasına daha fazla katkı sağladığı görülmektedir. $Göz_Pn$ eşitleme için elde edilen T değerlerinin Ger_Pn için elde edilen T değerlerinden genellikle daha küçük T değeri elde edilmesine katkı sağladığı ve ADÖ'nün korunmasında daha etkili olduğu sonucuna ulaşılmıştır.

Formların eşitlenmesinde karakteristik ölçek dönüştürme yöntemleri çoğunlukla daha iyi performans göstermesine rağmen en küçük BSE değerleri bazı durumlarda moment yöntemlerde elde edilmiştir. $Göz_Pn$ için elde edilen BSE değerleri Ger_Pn için elde edilen BSE değerlerine göre çoğunlukla daha küçük bulunmuştur.

Formların eşitlenmesinde karakteristik ölçek dönüştürme yöntemleri çoğunlukla daha iyi performans göstermesine rağmen en küçük İSE değerleri bazı durumlarda moment yöntemlerde elde edilmiştir. $Göz_Pn$ için elde edilen İSE değerleri Ger_Pn için elde edilen İSE değerlerine göre çoğunlukla daha küçük bulunmuştur.

Tartışma

Bu çalışmadan elde edilen sonuçlarla ve literatürde yer alan çalışmaların sonuçları karşılaştırılarak aşağıda verilmiştir.

He (2011), Tong ve Kolen (2005), Öztürk-Gübeş (2014, 2015a, 2015b) tarafından yapılan çalışmalar MTK Ger_Pn eşitleme yöntemini daha küçük BSE değeri vermesinden dolayı BSE'nin korunumunda daha iyi olarak bulmuşlardır. Ancak bu çalışmada Matematik ve Sosyal Bilgiler testi için MTK $Göz_Pn$ ve MTK Ger_Pn eşitlemenin yaklaşık aynı sonuçları verdiği ve Matematik testinde MTK $Göz_Pn$ eşitlemenin MTK Ger_Pn eşitmeye göre genellikle daha küçük BSE değeri verdiği sonucuna ulaşılmıştır. Elde edilen bu sonuç yukarıda belirtilen çalışmalardan elde edilen sonuçlarla örtüşmemektedir. Sosyal Bilgiler testi için ise $Göz_Pn$ ve Ger_Pn eşitleme arasında BSE'nin küçük elde edilme durumunun örneklem büyüklüğü, testin formatı ve cinsiyete göre değiştiği sonucuna ulaşılmıştır. He (2011) amacın A testi yerine B testi almanın avantaj ya da dezavantajını ortadan kaldırmak olduğu durumlarda, yani BSE sağlanmak istendiğinde $Göz_Pn$ eşitleme yerine Ger_Pn eşitlemenin kullanılmasını önermiştir. Ancak bu çalışmadan elde

edilen bulgulara göre Matematik testi dikkate alındığında $Göz_Pn$ eşitlemenin daha küçük BSE değerleri verdiğini göstermektedir. Sosyal Bilgiler testi için ise kesin olarak bir yöntem (Ger_Pn ya da $Göz_Pn$) önerilememektedir.

He (2011), Tong ve Kolen (2005), Öztürk-Gübeş (2014, 2015a, 2015b) tarafından yapılan çalışmalar MTK $Göz_Pn$ eşitleme yöntemini daha küçük İSE değeri vermesinden dolayı İSE'nin korunumunda daha iyi olarak bulmuşlardır. Bu çalışmada da hem Matematik hem de Sosyal Bilgiler testi için MTK $Göz_Pn$ eşitlemenin MTK Ger_Pn eşitlemeye göre genellikle daha küçük İSE değeri verdiği sonucuna ulaşılmıştır. Elde edilen bu sonuç yukarıda belirtilen çalışmalardan elde edilen sonuçlarla örtüşmektedir. He (2011) amacın ölçmenin kesinliği ve azalan ölçme hataları olduğunda, yani İSE sağlanmak istendiğinde $Göz_Pn$ eşitlemenin daha sağlıklı sonuçlar üreteceğini belirtmiştir. Bu çalışmadan elde edilen bulgular hem Matematik hem de Sosyal Bilgiler testi dikkate alındığında $Göz_Pn$ eşitlemenin daha küçük İSE değerleri verdiğini göstermektedir.

He (2011) eğer her bir kesme puanında paralel formla aynı öğrenci yüzdesini elde etmek amaçlanıyorsa $Göz_Pn$ eşitlemenin Ger_Pn eşitlemeye tercih edilmesi gerektiğini belirtmiştir. Bu çalışmada Matematik testinde ADÖ'nün korunumunda $Göz_Pn$ eşitleme Ger_Pn eşitlemeye göre daha küçük T değerleri verirken Sosyal Bilgiler Testi için kesin olarak bir yöntem (Ger_Pn ya da $Göz_Pn$) diğerine üstünlük sağlayamamıştır.

İki kategorili format karma formata göre hem Matematik hem de Sosyal Bilgiler testinde çoğunlukla daha küçük BSE değerlerinin elde edilmesini sağlamış olup bu sonuç Öztürk-Gübeş ve Kelecioğlu (2016) tarafından elde bulgularla uyumludur. Bu çalışmada karakteristik eğri yöntemleri moment yöntemlere göre hem Matematik hem de Sosyal Bilgiler testinde değişen örneklem büyüklükleri, cinsiyet ve format değişkenleri bakımından daha sağlıklı ve kararlı sonuçlar elde edilmesini sağlamıştır. Elde edilen bu sonuç Öztürk-Gübeş ve Kelecioğlu (2016) tarafından elde bulgularla uyumludur.

Öneriler

Araştırma kapsamında elde edilen sonuçlar ışığında öneriler, araştırmaya ve yapılacak diğer araştırmalara yönelik olmak üzere iki başlık altında sunulmuştur.

Araştırmanın Bulgularına Dayalı Öneriler

- ✓ Her ne kadar moment yöntemler bazı örneklem büyüklüklerinde daha küçük İSE ve BSE değeri vermiş olsa da elde edilen sonuçların örneklem büyüklüklerine göre kararlılıklarına bakıldığında karakteristik eğri yöntemlerinin daha iyi sonuçlar ürettiği görülmektedir. Bu nedenle karakteristik eğri yöntemleri moment yöntemlere tercih edilmelidir.
- ✓ Eşitlenecek formlar değerlendirilirken formlarda yer alan madde formatlarının benzer olmasına ve formlardan alınabilecek maksimum puanların aynı olmasına dikkat edilmelidir.
- ✓ Matematik ve Sosyal Bilgiler testinden elde edilen sonuçlar farklı olmasından dolayı hangi örneklem büyüklüğünün etkisinin teste göre değiştiği söylenebilir. Ancak cinsiyet alt grubu dikkate alındığında 1000 kişilik örneklem büyüklüğündeki yaklaşık 500 kız ve 500 erkek öğrenci verisinin parametre kestirimdeki sıkıntılara neden olduğu ve ADÖ'nün korunmasında sıkıntı yarattığı söylenebilir. Bu nedenle 1000'den daha büyük örneklem büyüklükleri daha kararlı sonuçlar için kullanılmalıdır.

Sonraki Araştırmalara Yönelik Öneriler

- ✓ Bu çalışmada ABİDE araştırmasında yer alan Matematik ve Sosyal Bilgiler testi dikkate alınmıştır. Yapılacak diğer çalışmalarda da Fen Bilimleri ve Türkçe testleri kullanılarak sayısal ve sözel testlere göre benzerlikler araştırılabilir.
- ✓ Cinsiyet alt grubuna göre değişen madde fonksiyonu incelemesi yapılarak elde edilecek bulgular üzerinden değerlendirmeler çeşitlendirilebilir.
- ✓ Bu çalışmada testler içerisinde yer alan maddelerden testle çok da uyumlu olmayanları da çalışmanın bütünlüğünün bozulmaması için çalışma dışına çıkarılmamasına karar verilmiştir. Testle uyumsuz maddelerin çalışma dışında bırakılması yoluna gidilebilir.
- ✓ Bu çalışmada örneklem büyüklükleri 1000, 3000 ve 5000 olarak alınmış ve cinsiyete göre yapılan araştırmalarda yeniden örneklem seçimine gidilmemiştir. Yapılacak diğer çalışmalarda cinsiyete göre ayrı örneklem seçimine gidilebilir.
- ✓ Bu çalışma Türkiye geneli için yapılmıştır. TÜİK'in kullandığı İstatistiki Bölge Birimleri Sınıflaması Düzey 1'e göre çalışma detaylandırılabilir.

- ✓ Bu alıřmada okul trleri dikkate alınmamıřtır. Yapılacak dięer alıřmalarda okul trleri dikkate alınabilir. Bu řekilde verinin saęa ve sola arpık olması durumlarında sonuların nasıl gerekleřtięi deęerlendirilebilir.

Kaynaklar

- Andrews, B. J. (2011). *Assessing first- and second-order equity for the common-item nonequivalent groups design using multidimensional IRT*. (Unpublished doctoral dissertation), University of Iowa. Iowa City.
- Bağcan Büyükturan, E., & Çıkırcı Demirtaşlı, N. (2012). Çoktan seçmeli testler ile yapılandırılmış gridlerin psikometrik özellikleri bakımından karşılaştırılması. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 45(1), 395-415.
- Brennan, R. L. (2010). *First-order and second-order Equity in Equating*. CASMA Research Report (30).
- Budescu, D. (2015). Efficiency of linear equating as a function of the length of the anchor test. *Journal of Educational Measurement*, 22, 13-20.
- Cai, L., Thissen, D., & du Toit, S. (2017). IRTPRO 4.2 for Windows [Computer software]. Skokie, IL: Scientific Software International, Inc.
- Cao, Y. (2008). *Mixed-format test equating: effects of test dimensionality and common-item sets*. University of Maryland, College Park.
- Conover, W. J. (1999). *Practical nonparametric statistics (3rd ed.)*. New York: John Wiley.
- Çakan, M. (2003). Geniş ölçekli başarı testlerinin eğitimindeki yeri ve önemi. *Eğitim ve Bilim*, 28(128), 19-26.
- Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2014). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları*. Ankara: Pegem Akademi.
- de Ayala, R. J. (2009). *The theory and practice of item response theory*. New York: The Guilford Press.
- DeMars, C. (2010). *Item response theory*. New York: Oxford University Press.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Gök, B. (2012). *Denk olmayan gruplarda ortak madde deseni kullanılarak madde tepki kuramına dayalı eşitleme yöntemlerinin karşılaştırılması*. (Yayınlanmamış doktora tezi). Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Gültekin, S. (2011). *Çoktan seçmeli, açık uçlu ve karma testlerin psikometrik özelliklerinin Madde Tepki Kuramına dayalı olarak değerlendirilmesi*.

- (Yayımlanmamış doktora tezi). Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Haebara, T. (1980). Equating logistic ability scales by a weighted least squares method. *Japanese Psychological Research*, 22, 144-149.
- Hagge, S. L. (2010). *The impact of equating method and format representation of common items on the adequacy of mixed-format test equating using nonequivalent groups*. (Unpublished doctorate dissertation). University of Iowa, Iowa City.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, California: Sage Publications.
- Hanson, B. A., & Béguin, A. A. (2002). Obtaining a common scale for item response theory item parameters using separate versus concurrent estimation in the common-item equating design. *Applied psychological measurement*, 26(1), 3-24.
- He, Y. (2011). *Evaluating equating properties for mixed-format tests*. (Unpublished doctoral dissertation). University of Iowa, Iowa City.
- He, Y., & Kolen, M. (2011). *Equity and same distributions properties for test equating*. CASMA Monograph Number 2.1.
- Jonathan, P. W. (2010). plink: An R package for linking mixed-format tests using IRT-based methods. *Journal of Statistical Software*, 35, 1-33.
- Karasar, N. (2016). *Bilimsel araştırma yöntemi*. Ankara: Nobel Akademik Yayıncılık.
- Kastner, M., & Stangl, B. (2011). Multiple choice and constructed response tests: Do test format and scoring matter? *Procedia-Social and Behavioral Sciences*, 12, 263-273.
- Kilmen, S. (2010). *Madde tepki kuramına dayalı test eşitleme yöntemlerinden kestirilen eşitleme hatalarının örneklem büyüklüğü ve yetenek dağılımına göre karşılaştırılması*. (Yayımlanmamış doktora tezi). Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Kim, D., Brennan, R., & Kolen, M. (2005). A comparison of IRT equating and beta 4 equating. *Journal of Educational Measurement*, 42(1), 77-99.
- Kim, S., & Lee, W. C. (2004). *IRT scale linking methods for mixed-format tests*. Iowa City: ACT Research Report Series 2004-5.

- Kolen, M. J. (2004). PARSCALE [computer program]. Iowa City IA: The Center for Advanced Studies in Measurement and Assessment (CASMA), The University of Iowa.
- Kolen, M. J. (2004). POLYSEM [computer program]. Iowa City IA: The Center for Advanced Studies in Measurement and Assessment (CASMA), The University of Iowa.
- Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices (Third edition)*. New York: Springer.
- Lee, E., Lee, W., & Brennan, R. L. (2010). *Assessing equating results based on first-order and second-order equity*. CASMA Research Report (31).
- Lissitz, R., & Hou, X. (2012). *Multiple choice items and constructed response items: does it matter*. University of Maryland. <https://www.marces.org/multiple%20choice%20items%20and%20constructed%20response%20items.pdf> adresinden alındı
- Livingston, S. A. (2009). Constructed-response test questions: Why we use them; how we score them. *ETS R & D Connections 11*, 1-8.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.
- Lorenzo-Seva, U., & Ferrando, P. J. (2013). *Factor v9.20 manual of the program*. Universitat Rovira i Virgili, Departament de Psicologia, Tarragona, Spain.
- Lu, R., Haberman, S., Guo, H., & Liu, J. (2015). Use of jackknifing to evaluate effects of anchor item selection on equating with the nonequivalent groups with anchor test (NEAT) design. *ETS Research Report Series, 2015(1)*, 1-12.
- MEB. (2017). *Akademik Becerilerin İzlenmesi ve Değerlendirilmesi-8.Sınıflar Raporu*. Ankara: ÖDSGM. https://odsgm.meb.gov.tr/meb_iys_dosyalar/2017_11/30114819_iY-web-v6.pdf adresinden alındı
- Michaelides, M. P. (2010). A review of the effects on IRT item parameter estimates with a focus on misbehaving common items in test equating. *Frontiers in psychology, 1*, 1-7.
- Montgomery, M., & Skorupski, W. (2012). Investigation of IRT parameter recovery and classification accuracy in mixed format. Paper presented at the annual meeting of the Nation Council of Measurement in Education, British Columbia.

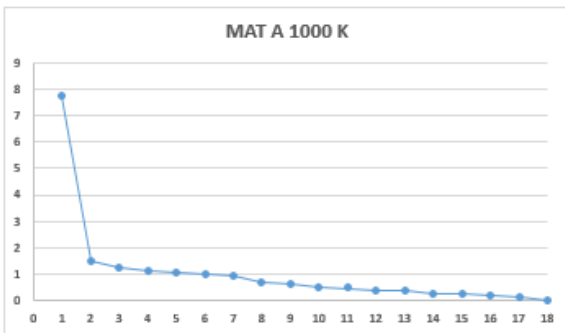
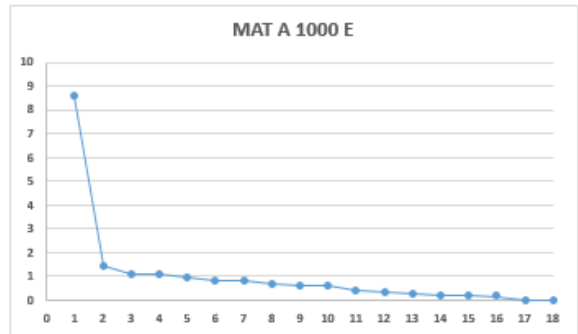
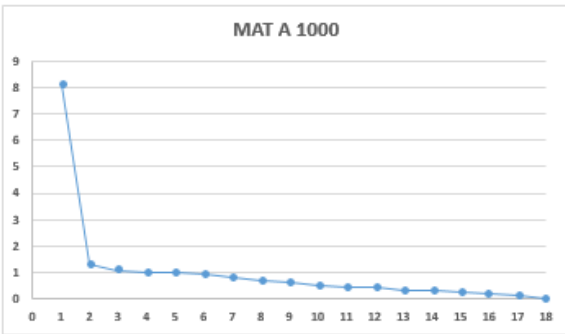
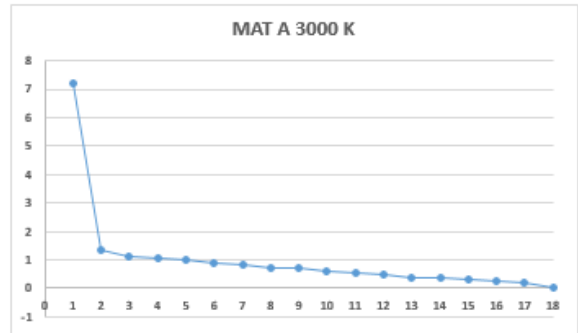
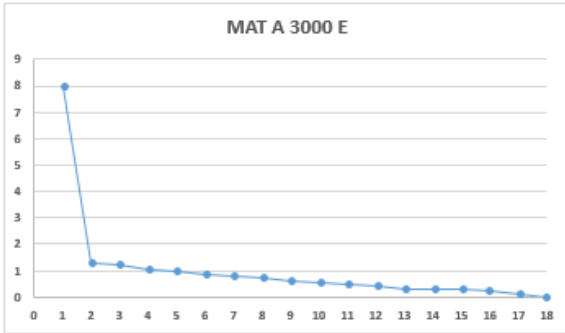
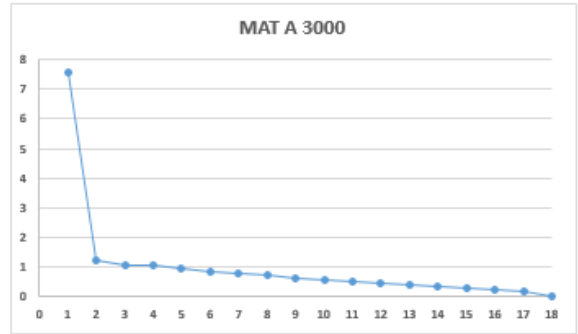
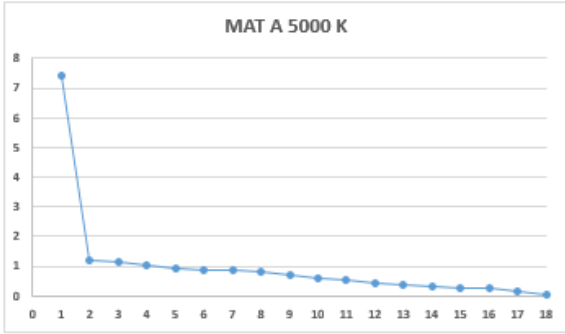
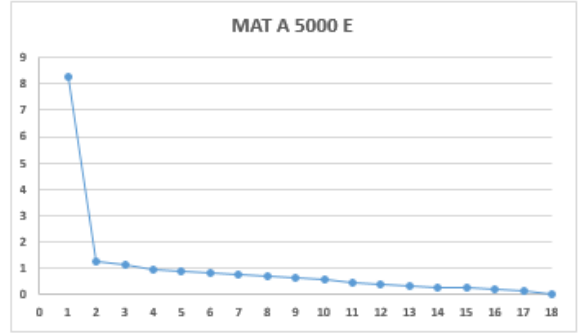
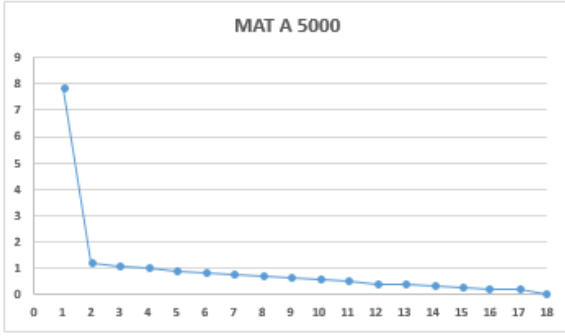
- Öztürk-Gübeş, N. (2014). *Testlerin boyutluluğunun, ortak madde formatının, yetenek dağılımının ve ölçek dönüştürme yöntemlerinin karma testlerin eşitlenmesine etkisi*. (Yayınlanmamış doktora tezi). Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Öztürk-Gübeş, N., & Kelecioğlu, H. (2015a). Farklı test eşitleme yöntemlerinin eşitlik özelliği ölçütüne göre karşılaştırılması. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 48(1), 199-214.
- Öztürk-Gübeş, N., & Kelecioğlu, H. (2015b). Karma testlerin eşitlenmesinde MTK eşitleme yöntemlerinin eşitlik özelliği korunumu ölçütüne göre karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(1), 117-125.
- Öztürk-Gübeş, N., & Kelecioğlu, H. (2016). The impact of test dimensionality, common-item set format, and scale linking methods on mixed-format test equating. *Educational Sciences: Theory & Practice*, 16, 715-734.
- Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics*, 4, 207-230.
- Ricker, K. L., & Davier, A. A. (2007). *The impact of anchor test length on equating results in a nonequivalent groups design*. ETS Research Report Series, 2007(2).
- Saen-amnuaiphon, R., Tuksino, P., & Nichanong, C. (2012). The effect of proportion of mixed-format scoring: Mixed-format achievement tests. *Procedia-Social and Behavioral Sciences*, 69, 1522-1528.
- Sinharay, S., & Holland, P. W. (2010). A new approach to comparing several equating methods in the context of the NEAT design. *Journal of Educational Measurement*, 47, 261-285.
- Tan, X., Kim, S., Paek, I., & Xiang, B. (2009). An alternative to the trend scoring method for adjusting scoring shifts in mixed-format tests. *In annual meeting of the National Council on Measurement in Education*. San Diego, CA.
- Thissen, D., Wainer, H., & Wang, X. B. (1994). Are tests comprising both multiple-choice and free-response items necessarily less unidimensional than multiple-choice tests? An analysis of two tests. *Journal of Educational Measurement*, 31(2), 113-123.
- Tong, Y., & Kolen, M. (2005). Assessing equating results on different equating criteria. *Applied Psychological Measurement*, 29, 418-432.

- Traub, R. E., & Fisher, C. W. (1977). On the equivalence of constructed-response and multiple-choice tests. *Applied Psychological Measurement*, 1(3), 355-369.
- Uysal, İ., & Kilmen, S. (2016). Comparison of item response theory test equating methods for mixed format tests. *International Online Journal of Educational Sciences*, 8(2), 1-11.
- Walker, M. E., & Kim, S. (2010). Linking mixed-format tests using multiple-choice anchors. *ETS Research Report Series*(2).
- Wang, C. (2011). *An investigation of bootstrap methods for estimating the standard error of equating under the common-item nonequivalent groups design*. (Unpublished doctoral dissertation). The University of Iowa, Iowa City.
- Wang, T., Lee, W. C., Brennan, R. L., & Kolen, M. J. (2008). A comparison of the frequency estimation and chained equipercentile methods under the common-item nonequivalent groups design. *Applied Psychological Measurement*, 32(8), 632-651.
- Yang, W. L., & Houang, R. T. (1996). *The effect of anchor length and equating method on the accuracy of test equating: comparisons of linear and IRT-based equating using an anchor-item design*. Paper presented at the annual meeting of the American Educational Research Association, New York.
- Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, 8, 125-145.

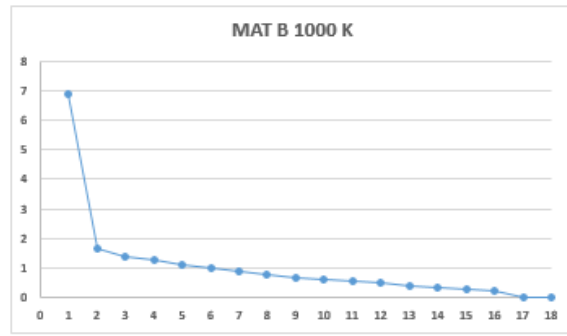
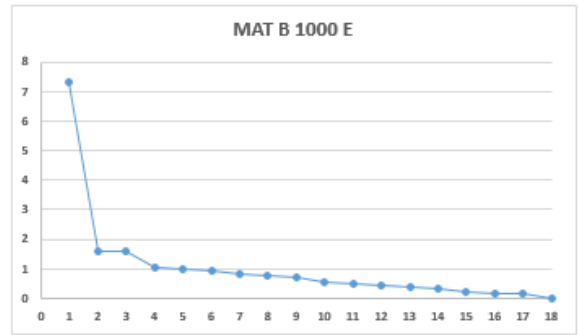
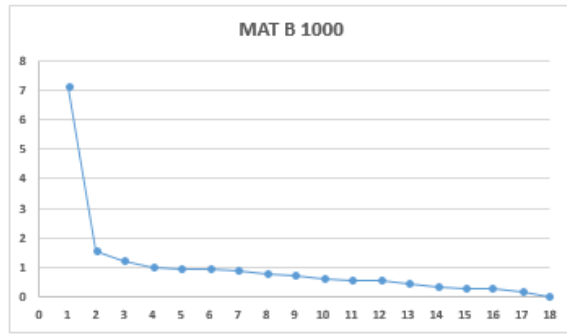
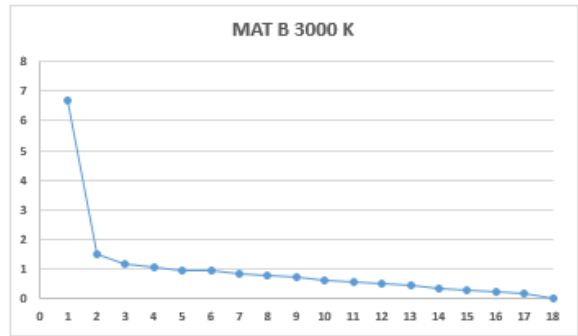
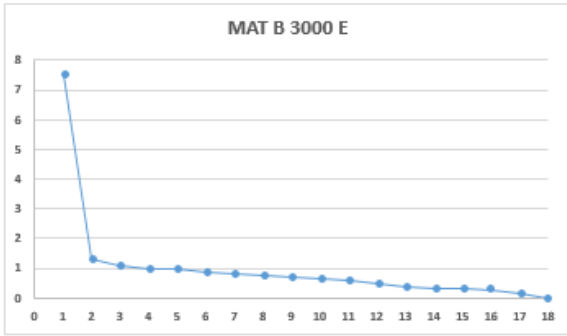
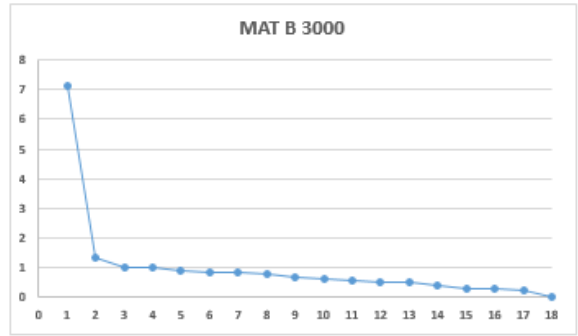
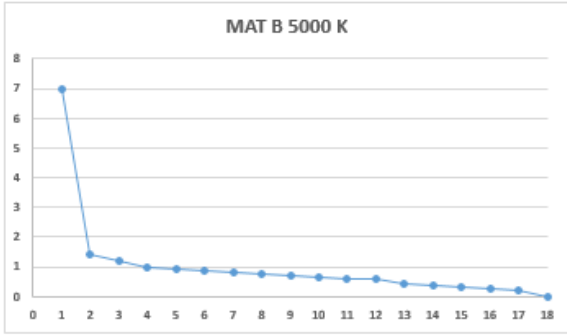
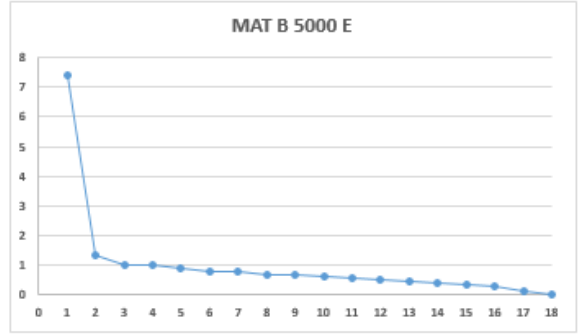
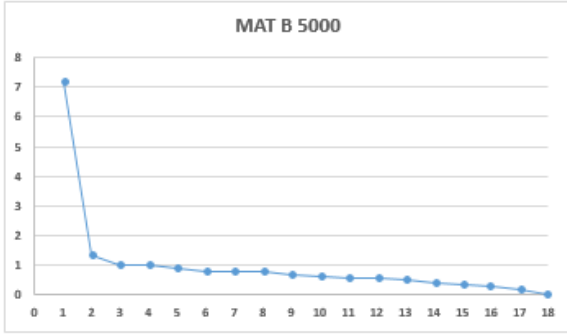
EK-A: Mplus7 Syntax

```
TITLE:      this is an example of an EFA with continuous,  
            censored, categorical, and count factor indicators  
DATA:      FILE = MATC1000.txt;  
VARIABLE:  NAMES = m1-m18;  
            CATEGORICAL = m1-m18;  
ANALYSIS: TYPE = EFA 1 2;  
PLOT:  
TYPE=PLOT2;  
SAVEDATA:  
SAMPLE IS DESKTOP;  
TYPE IS CORRELATION;
```

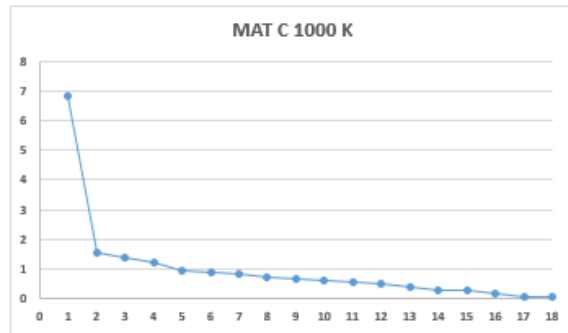
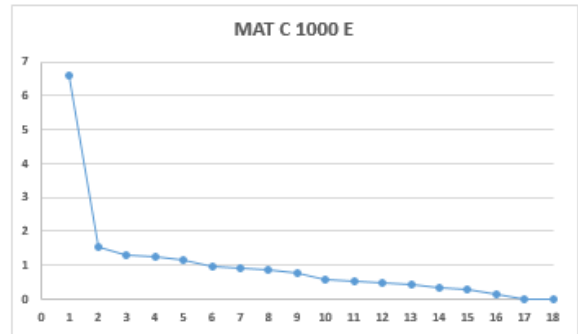
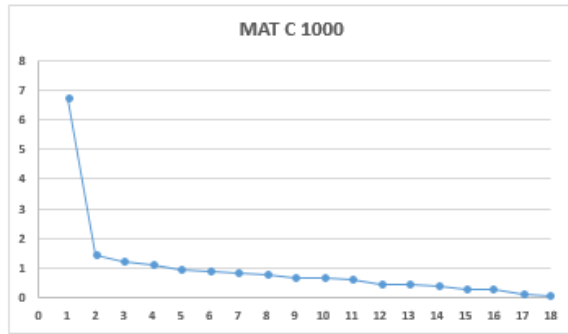
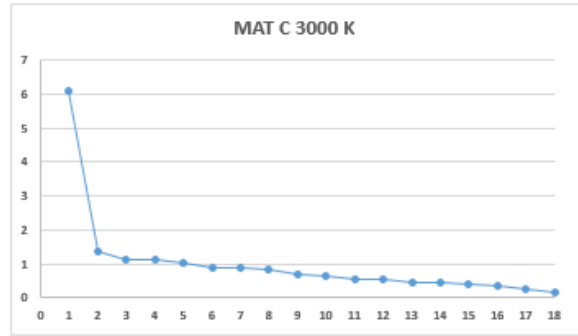
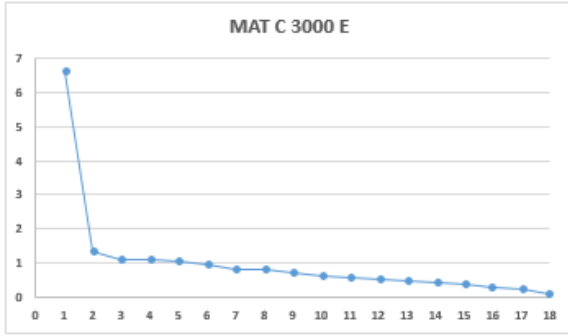
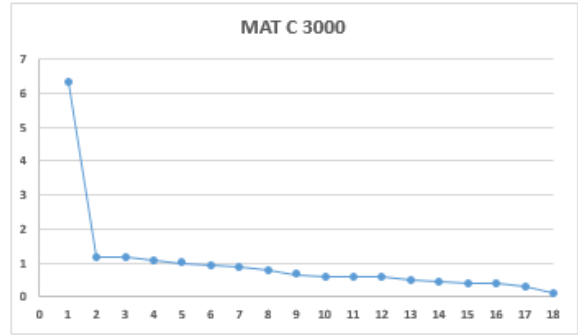
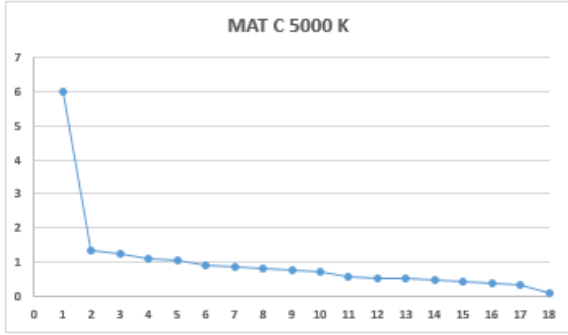
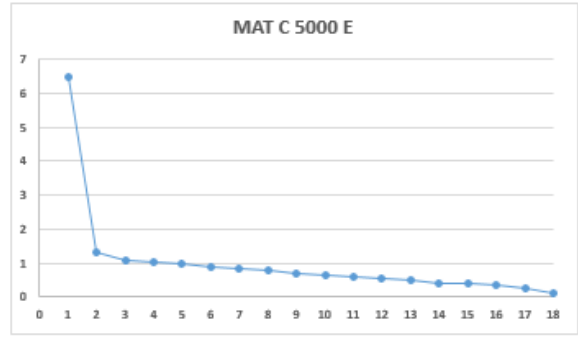
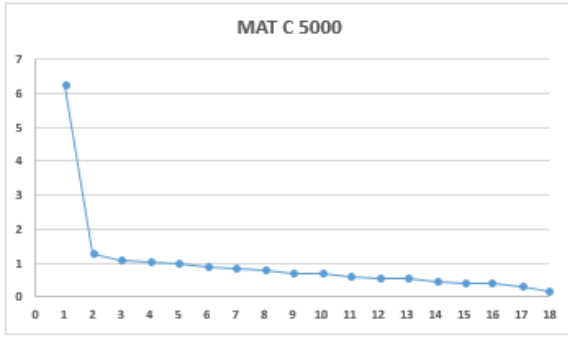
EK-B: Matematik A Formu Yamaç-Birikinti Grafiği (Karma)



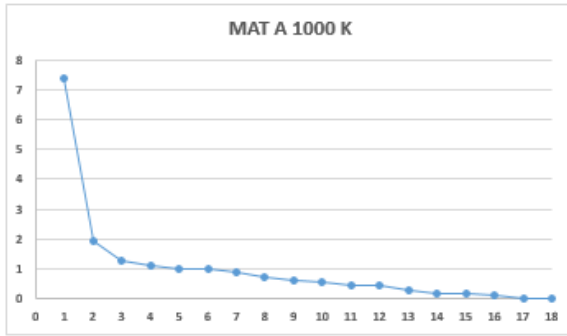
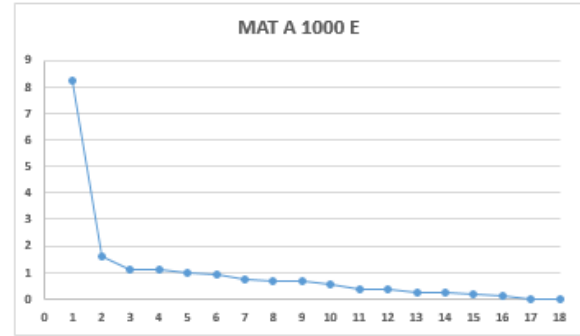
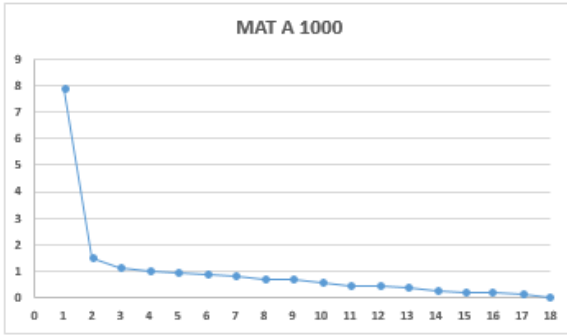
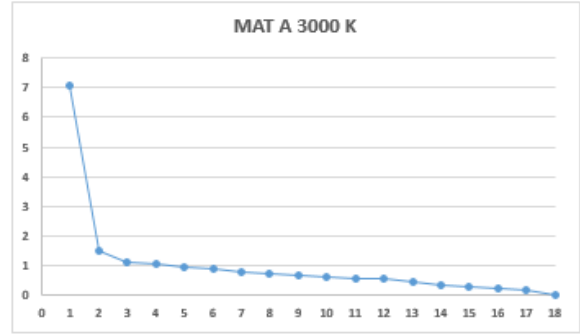
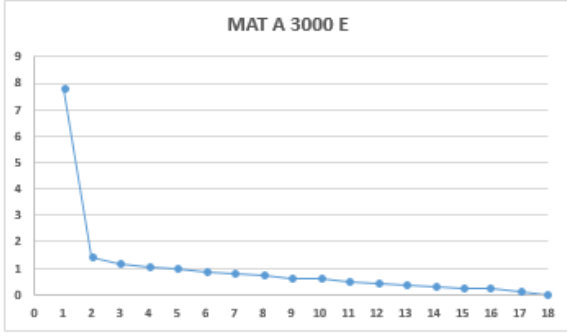
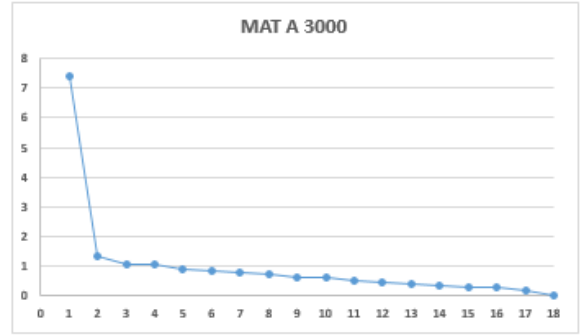
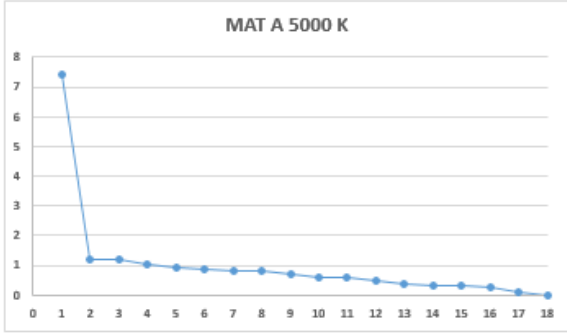
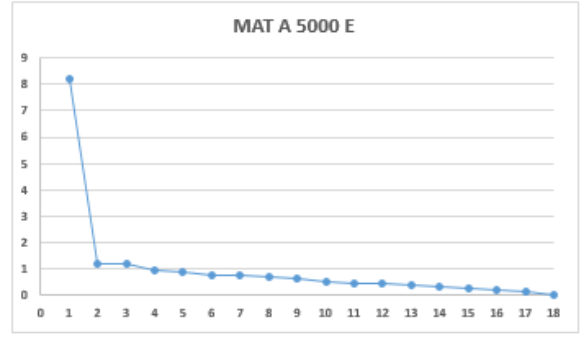
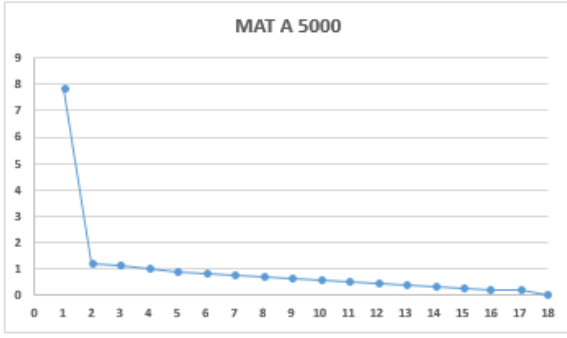
EK-C: Matematik B Formu Yamaç-Birikinti Grafiği (Karma)



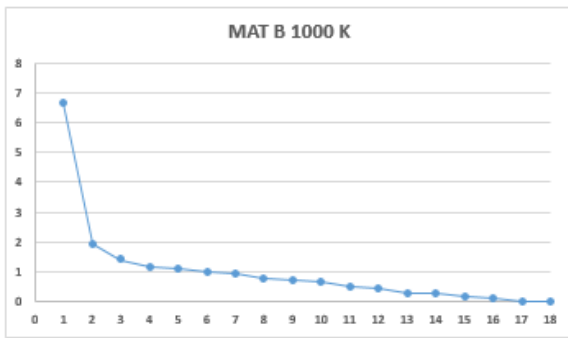
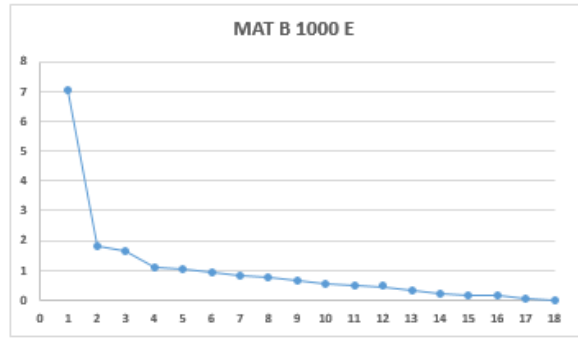
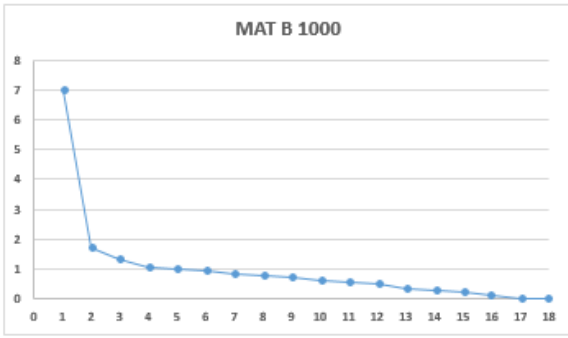
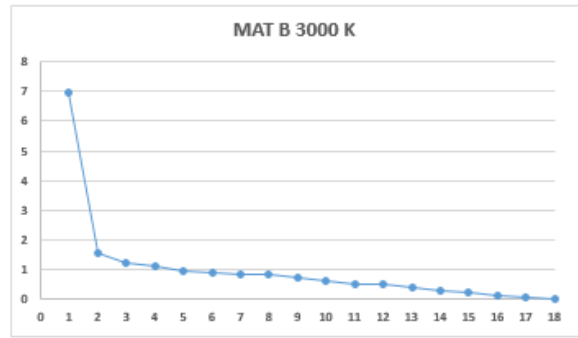
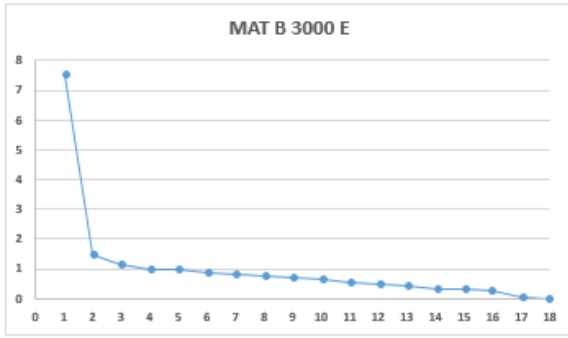
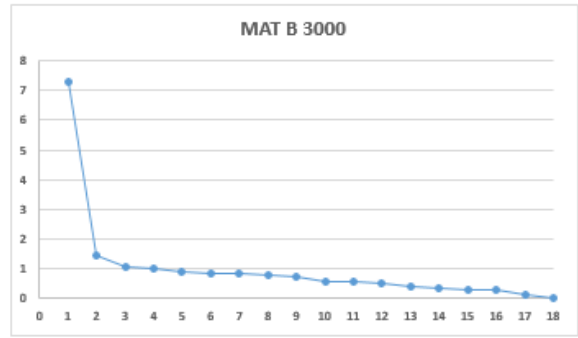
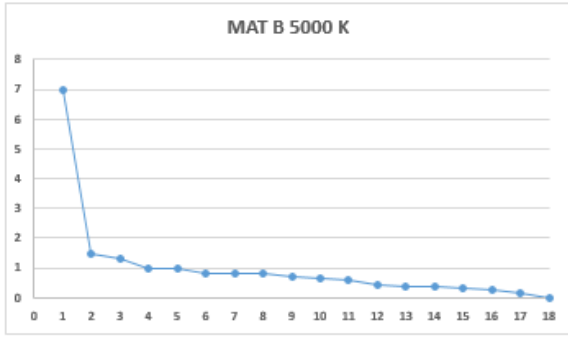
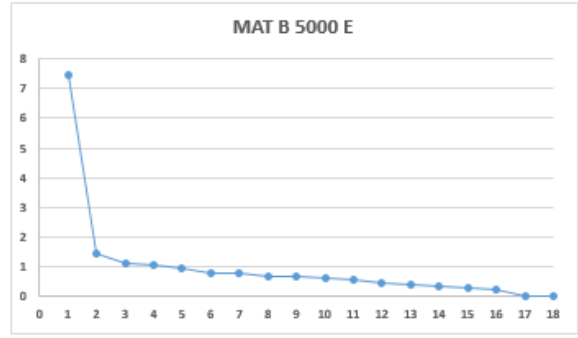
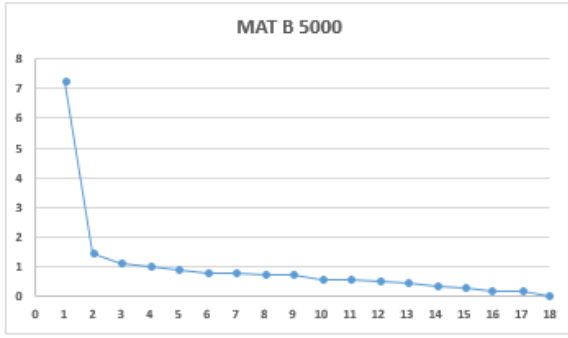
EK-Ç: Matematik C Formu Yamaç-Birikinti Grafiği (Karma)



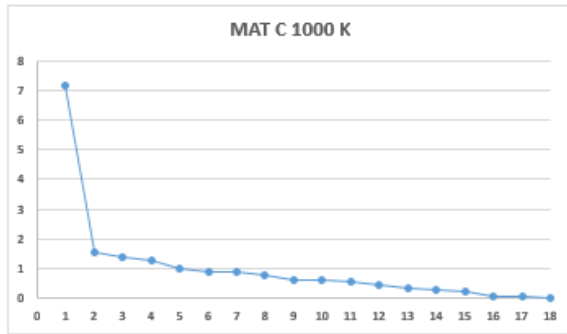
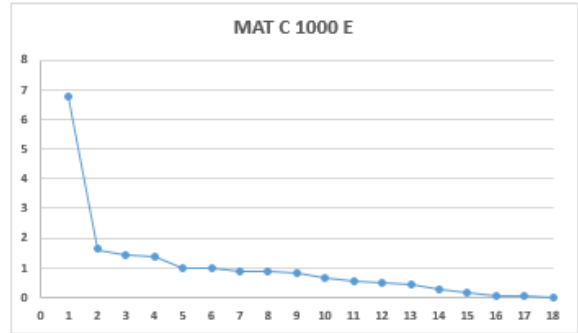
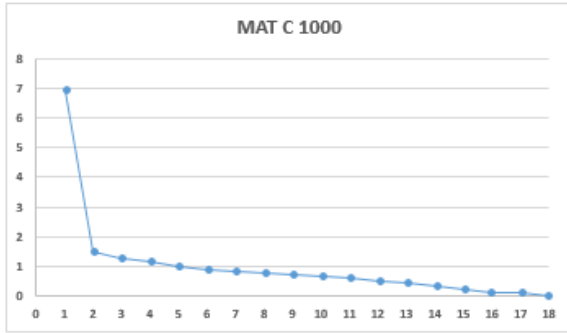
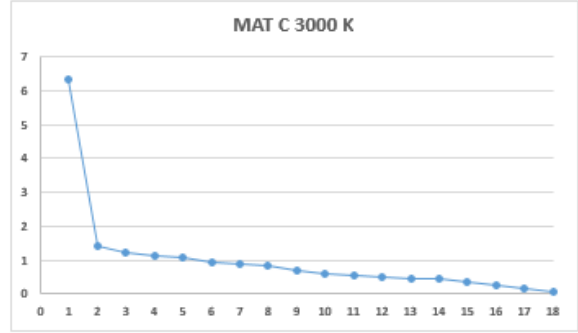
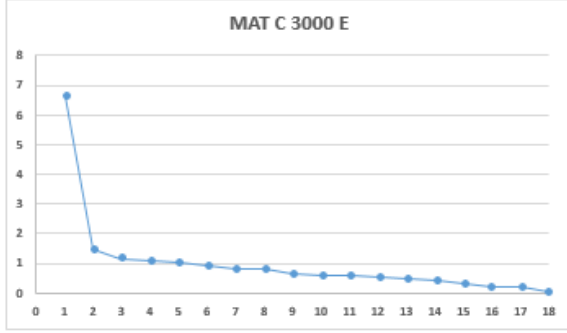
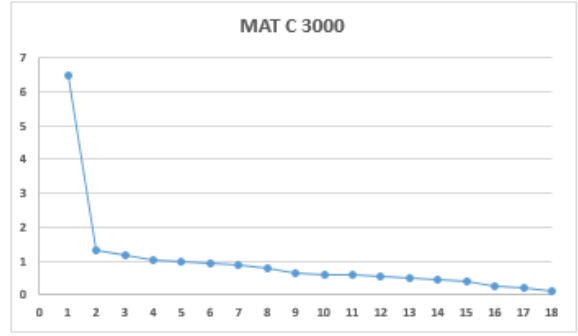
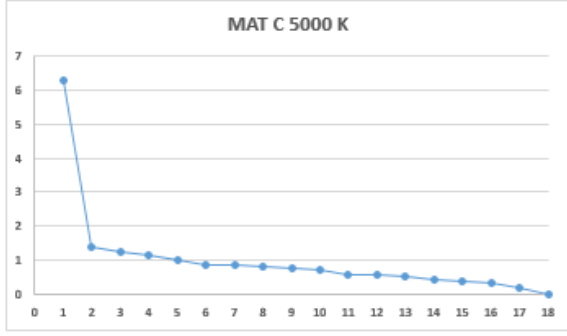
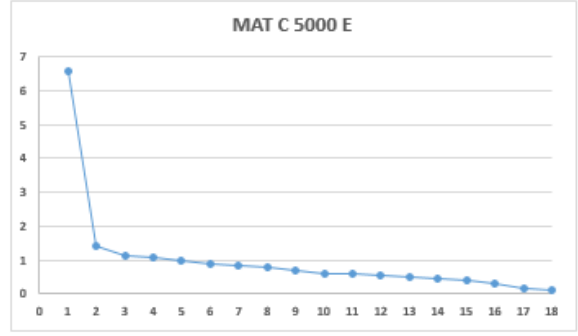
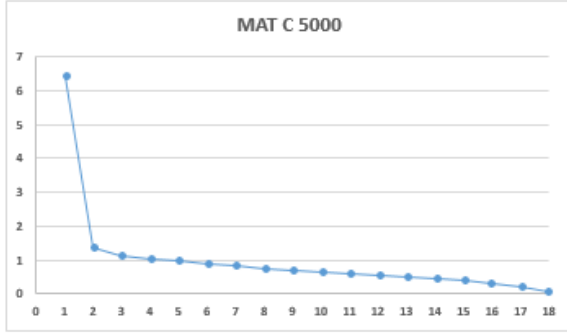
EK-D: Matematik A Formu Yamaç-Birikinti Grafiği (İki Kategorili)



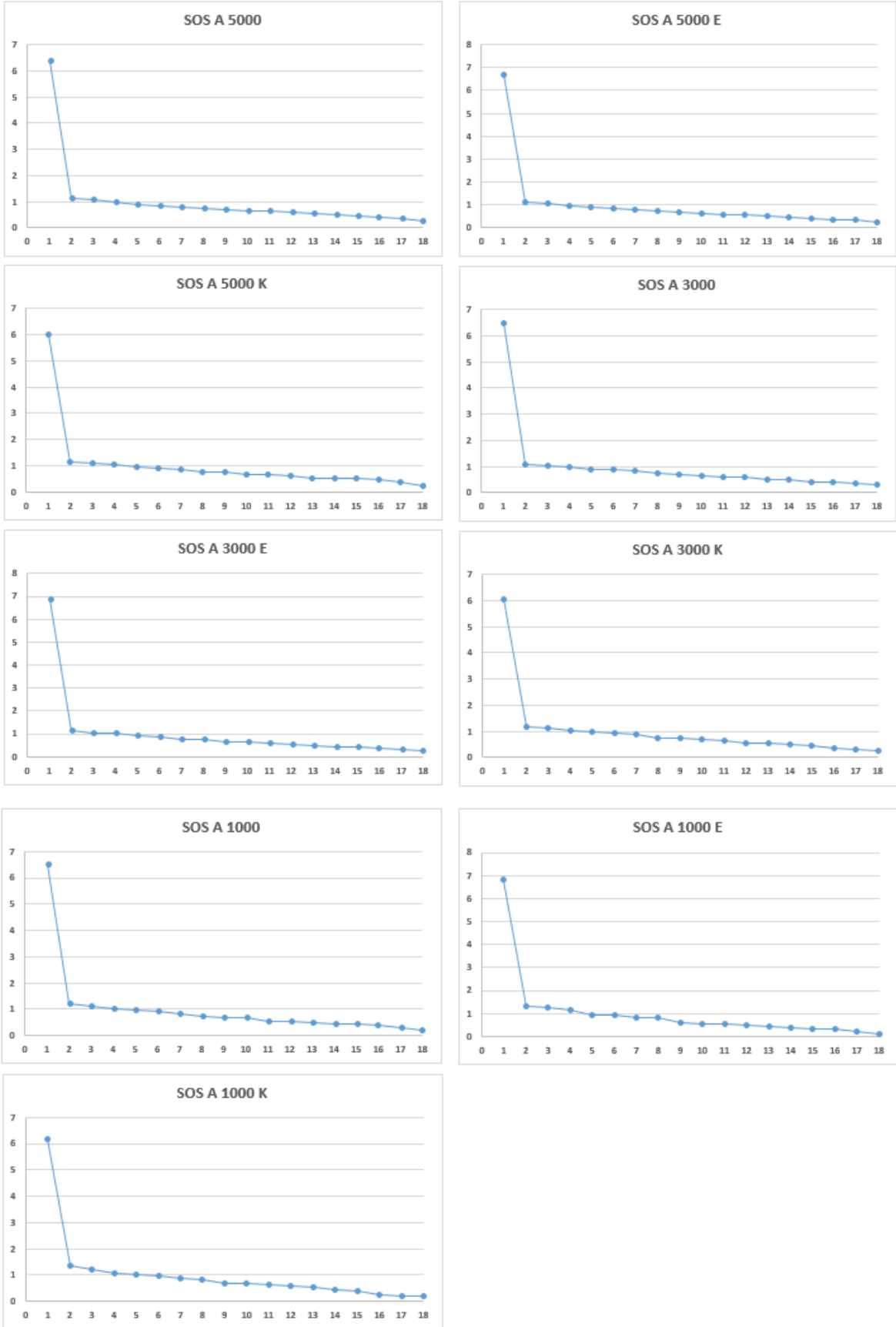
EK-E: Matematik B Formu Yamaç-Birikinti Grafiği (İki Kategorili)



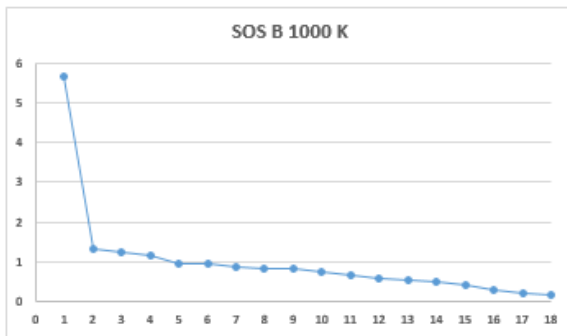
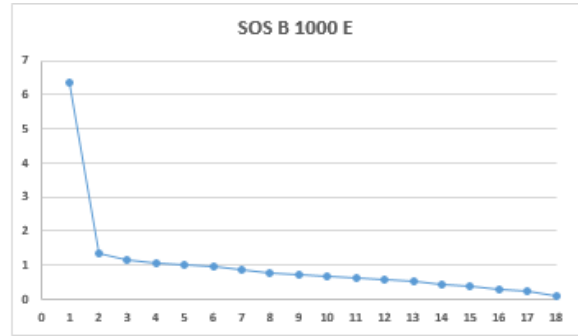
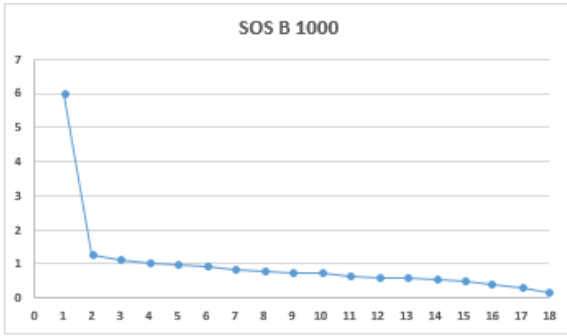
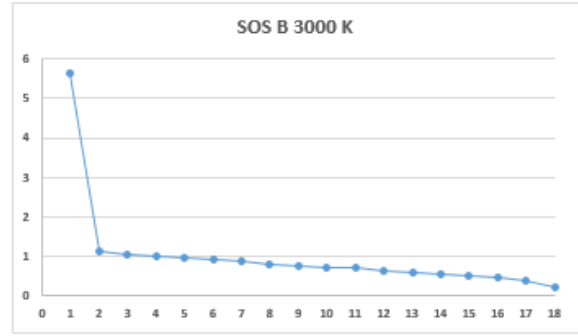
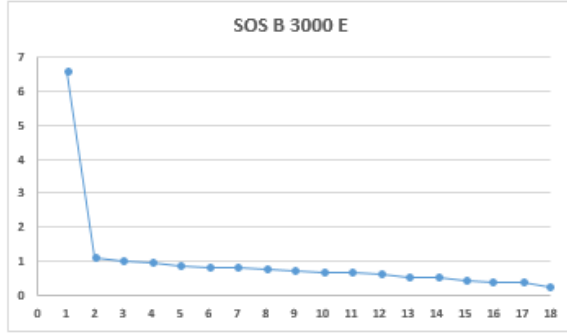
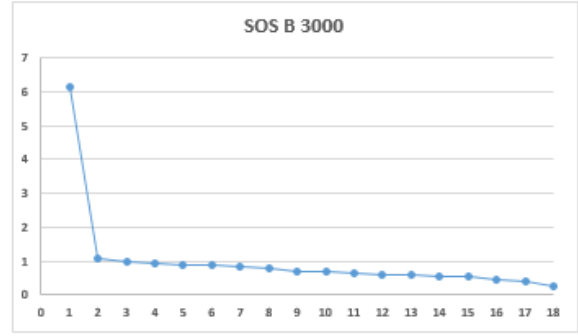
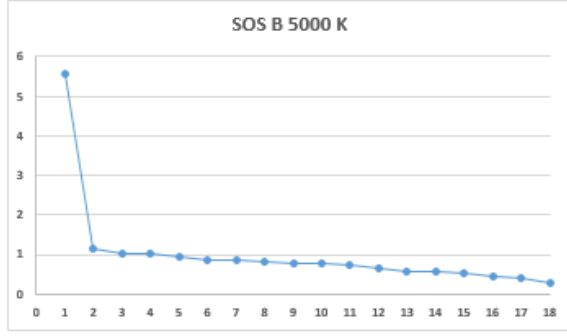
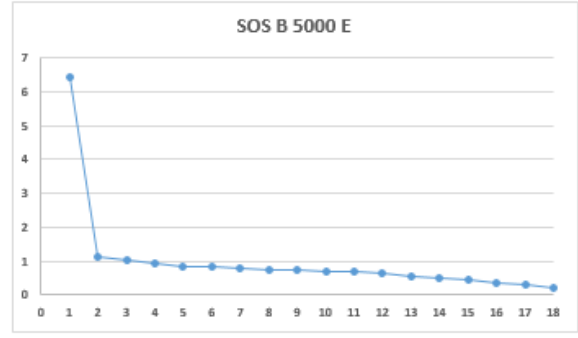
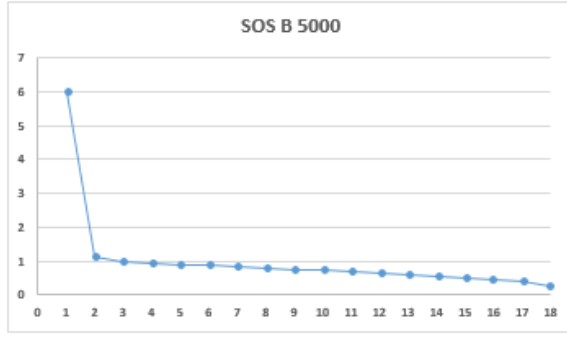
EK-F: Matematik C Formu Yamaç-Birikinti Grafiği (İki Kategorili)



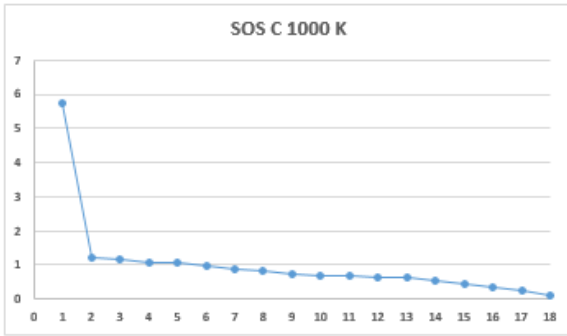
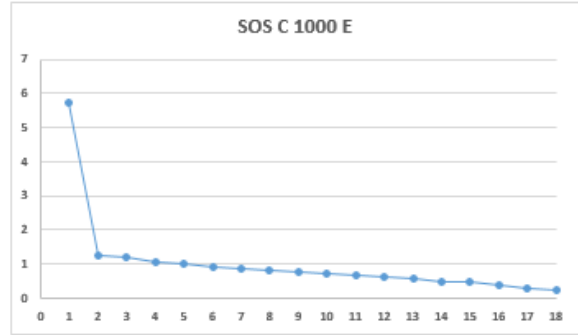
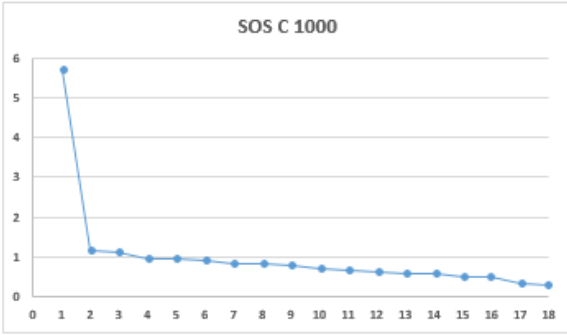
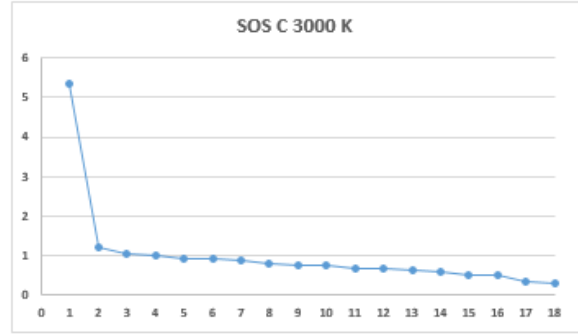
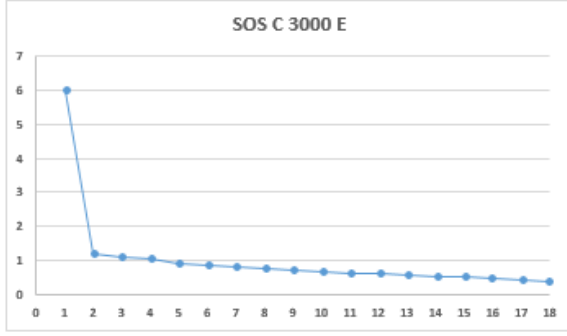
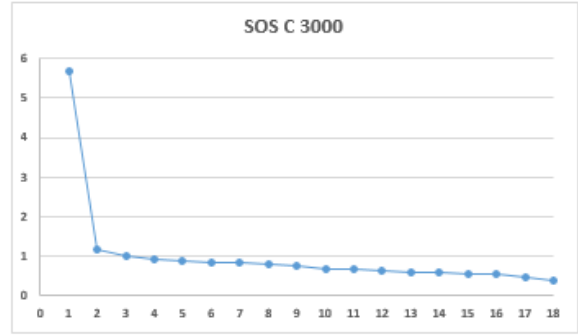
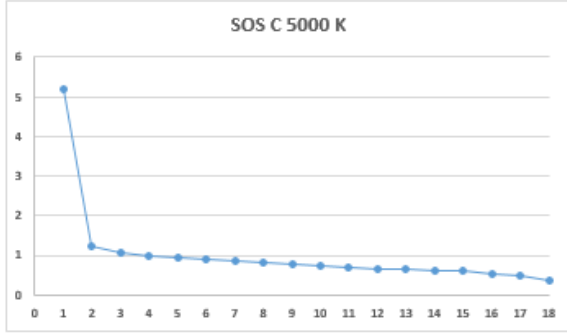
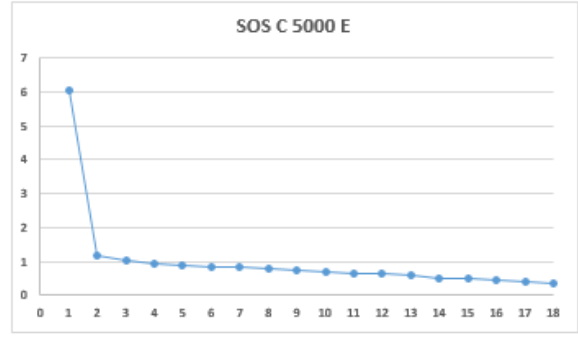
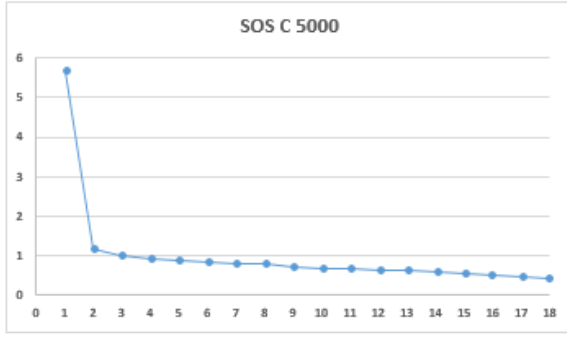
EK-G: Sosyal Bilgiler A Formu Yamaç-Birikinti Grafiği (Karma)



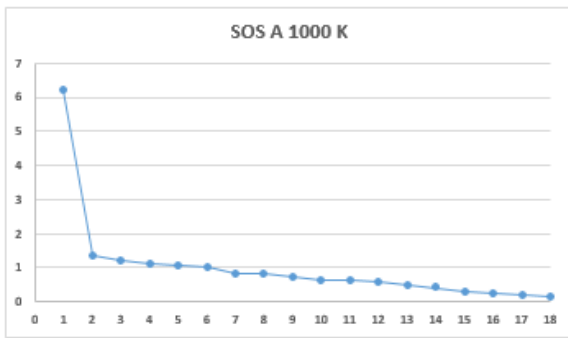
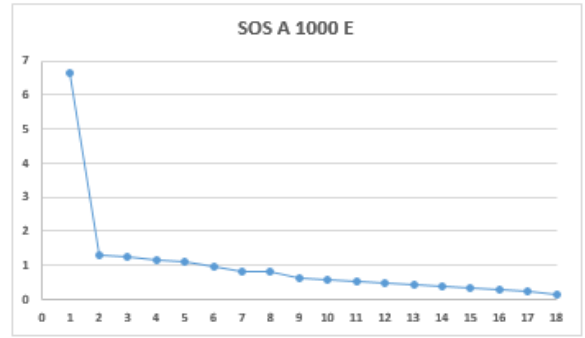
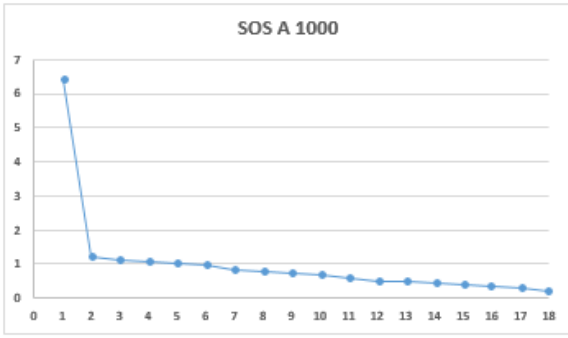
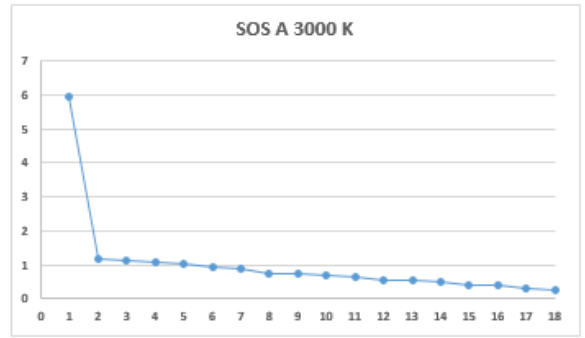
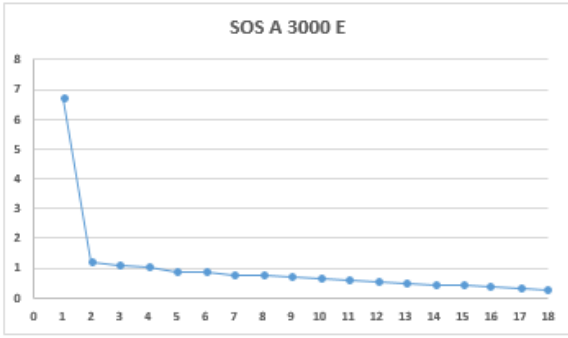
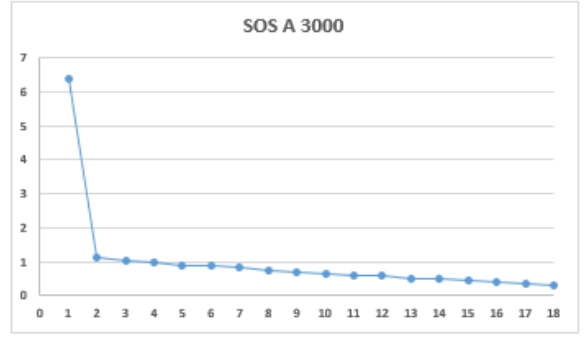
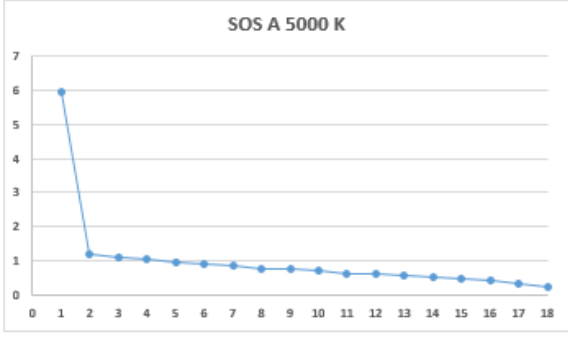
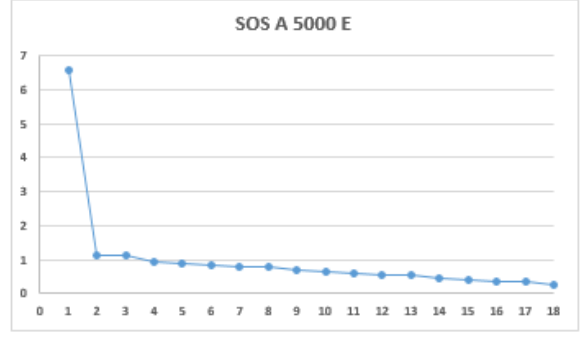
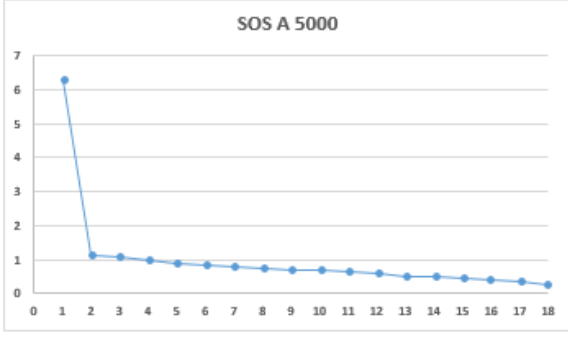
EK-Ğ: Sosyal Bilgiler B Formu Yamaç-Birikinti Grafiđi (Karma)



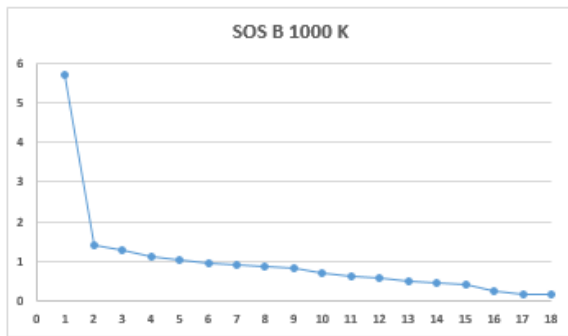
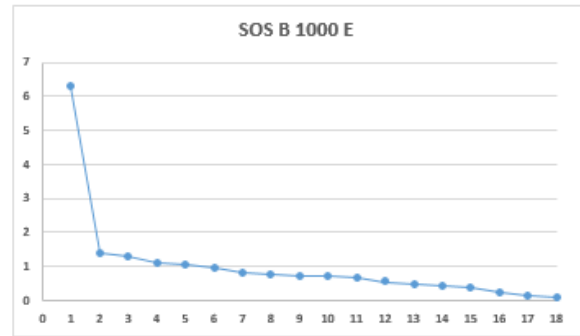
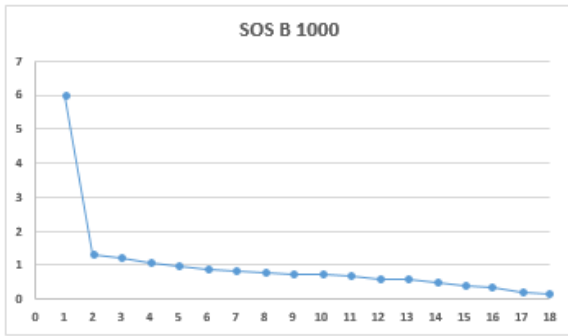
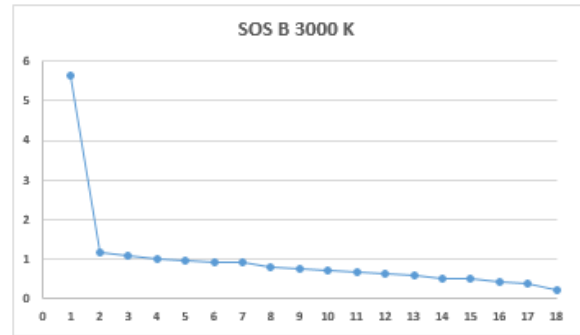
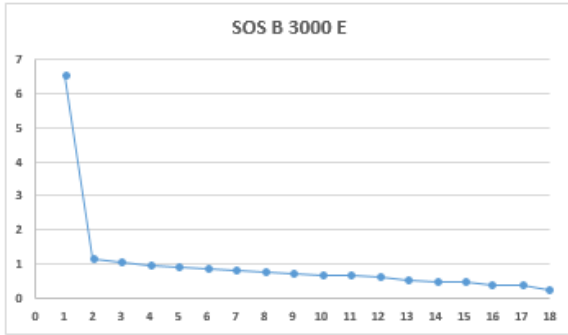
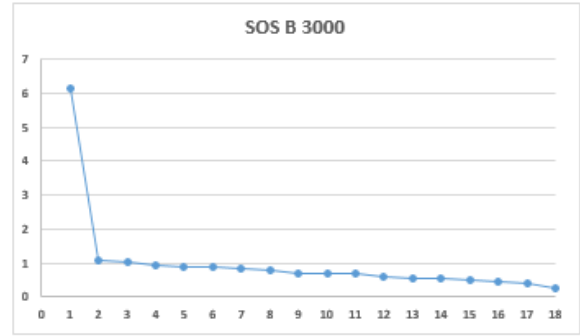
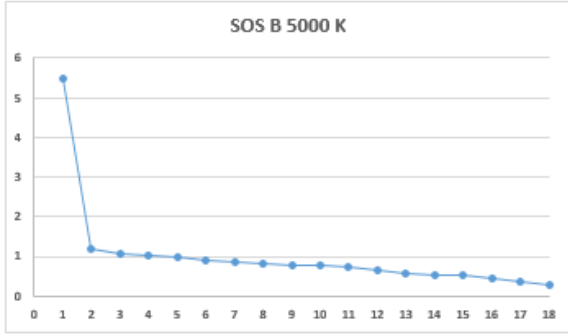
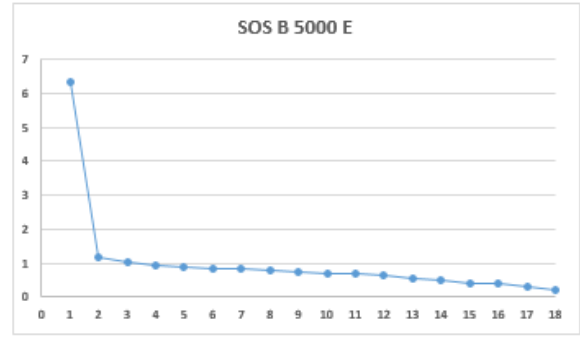
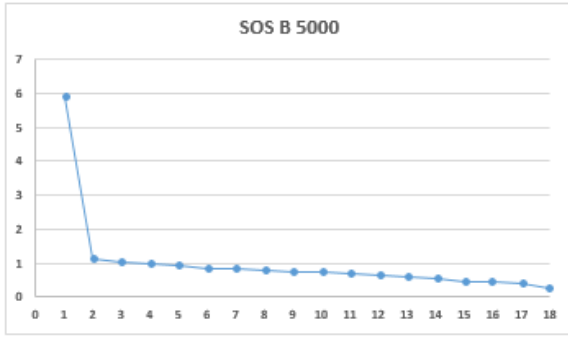
EK-H: Sosyal Bilgiler C Formu Yamaç-Birikinti Grafiği (Karma)



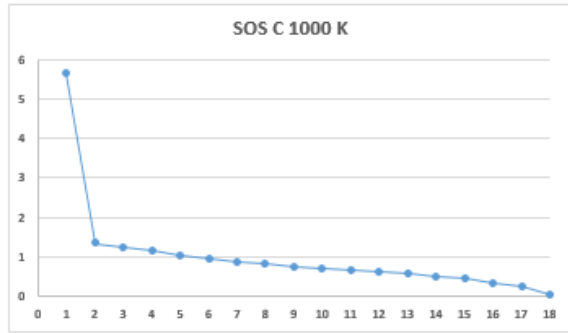
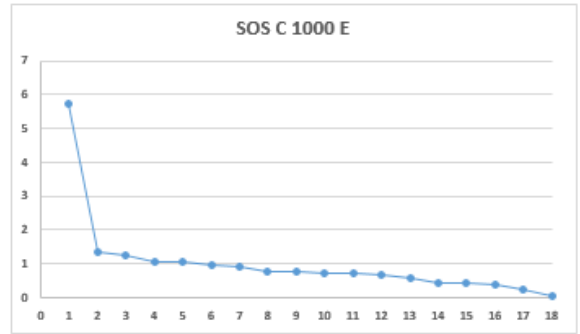
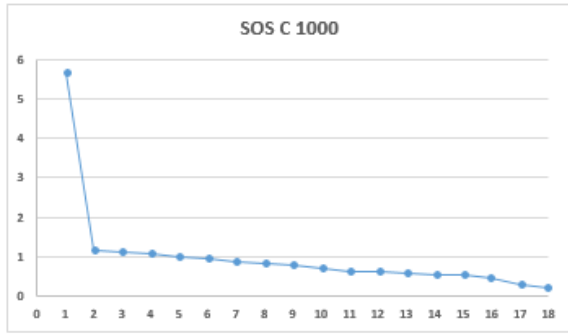
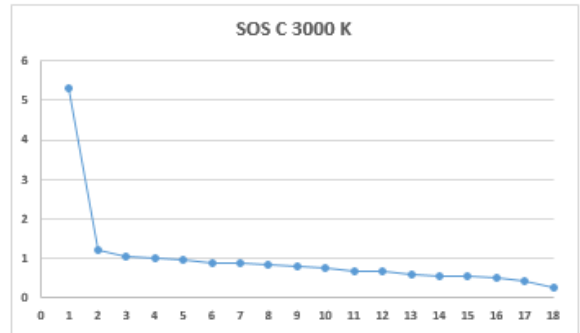
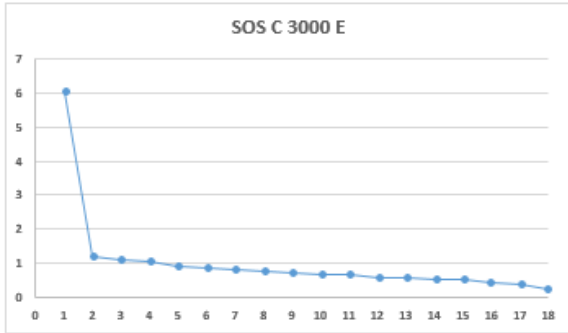
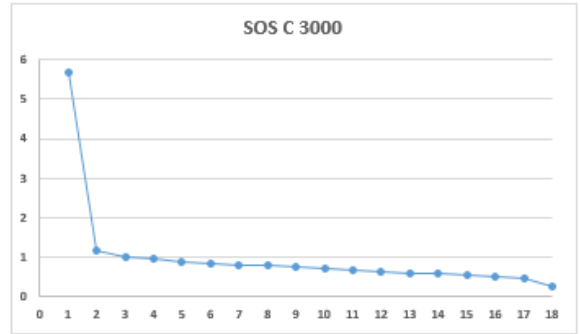
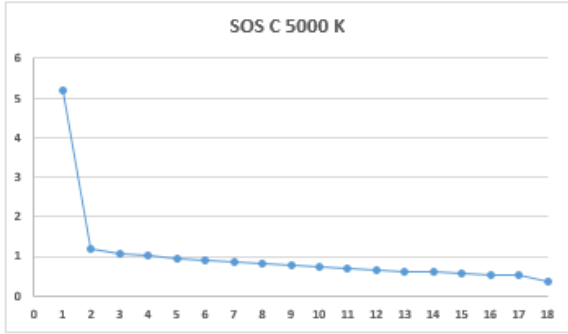
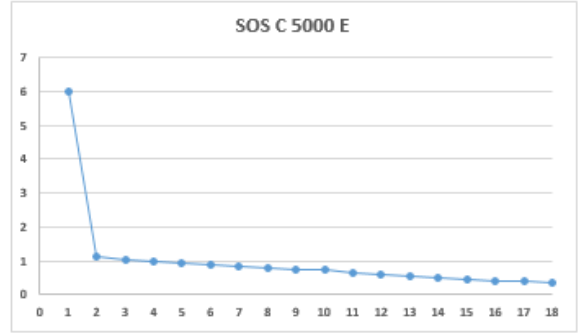
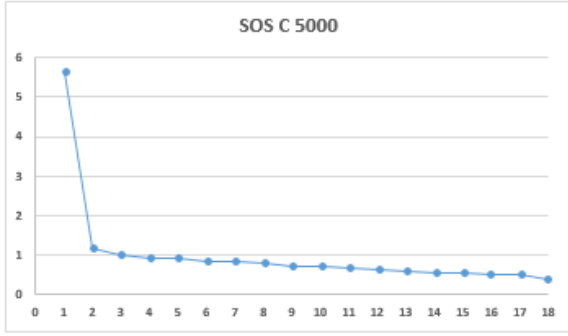
EK-I: Sosyal Bilgiler A Formu Yamaç-Birikinti Grafiği (İki Kategorili)



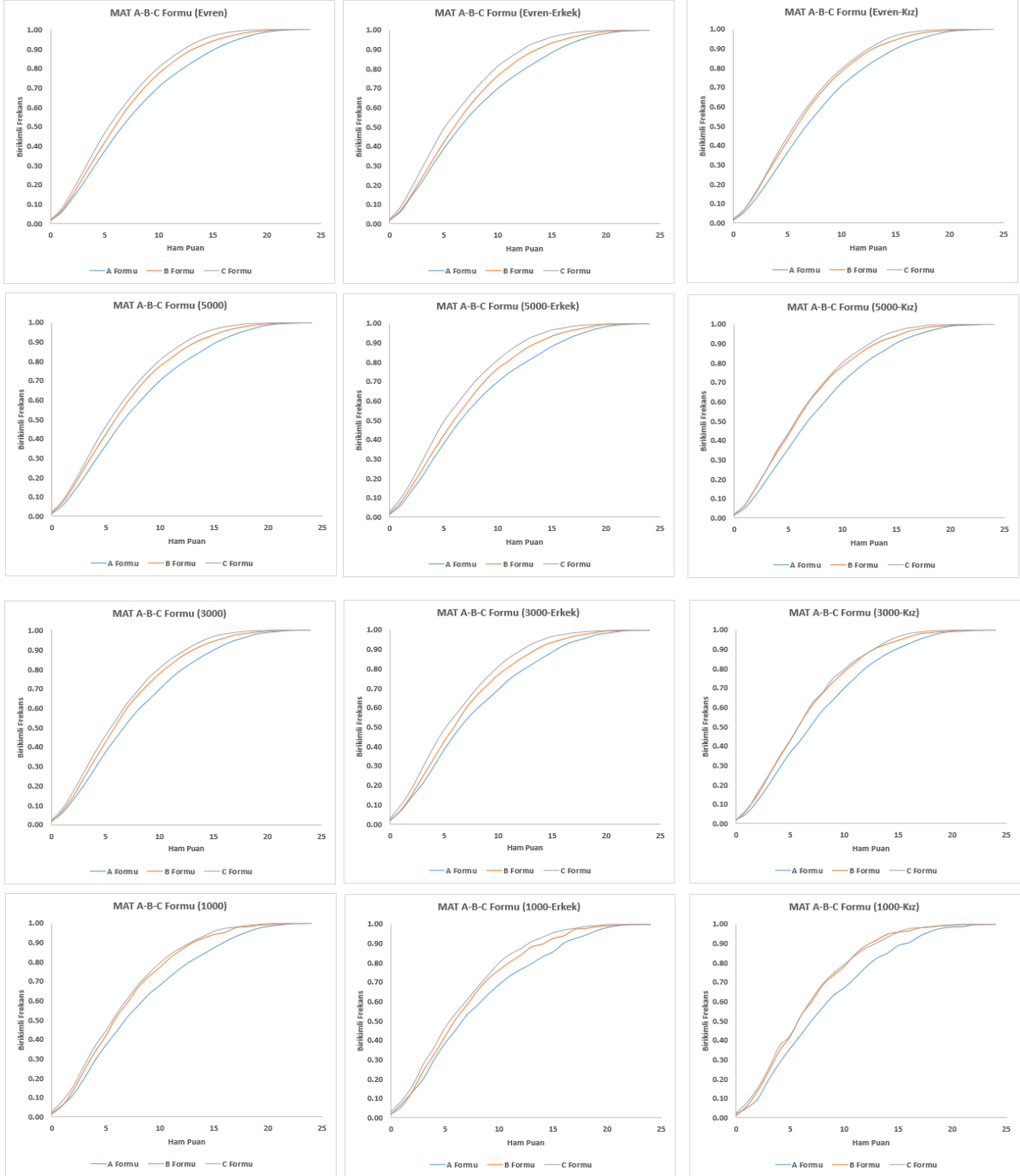
EK-İ: Sosyal Bilgiler B Formu Yamaç-Birikinti Grafiği (İki Kategorili)



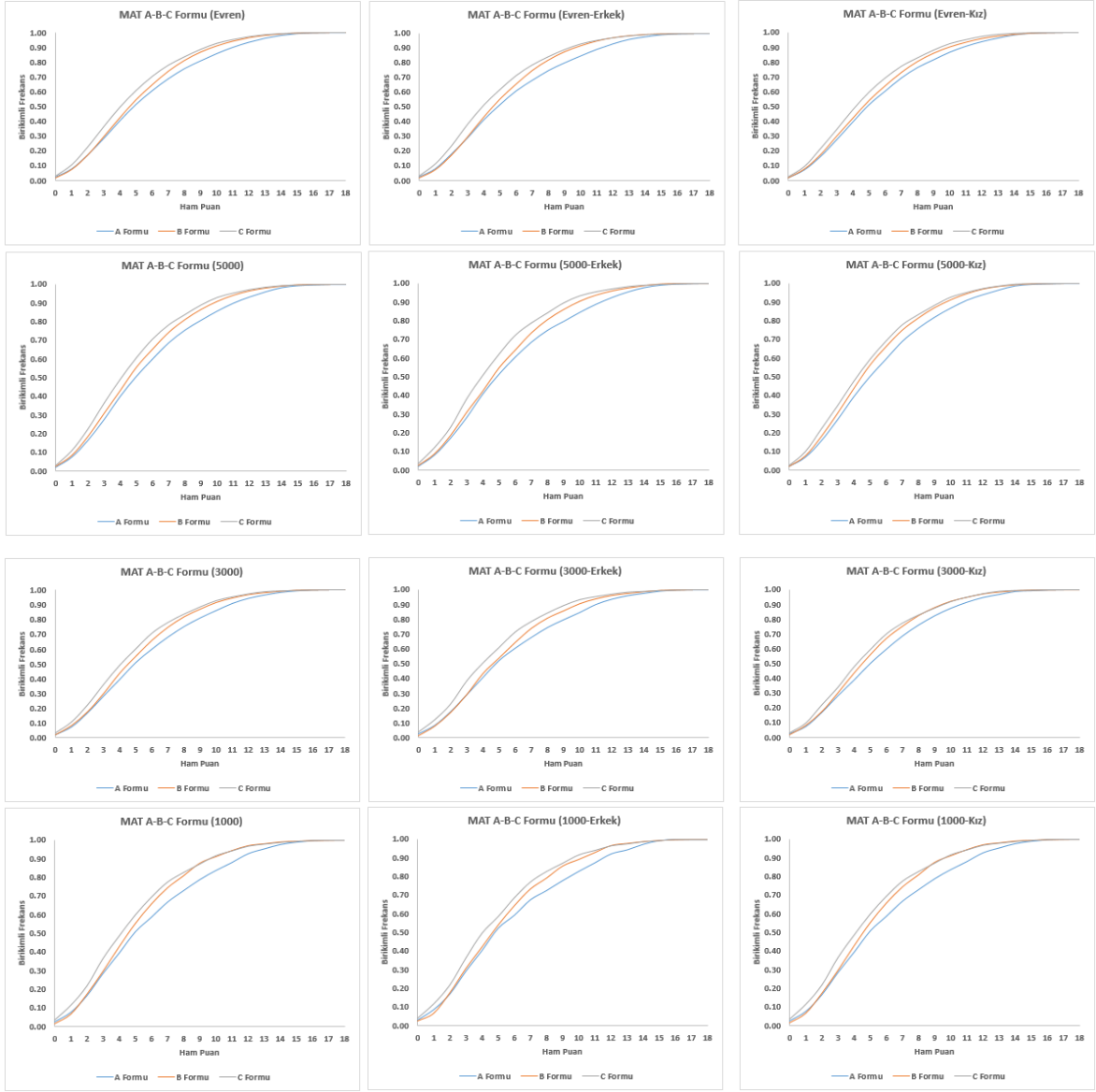
EK-J: Sosyal Bilgiler C Formu Yamaç-Birikinti Grafiği (İki Kategorili)



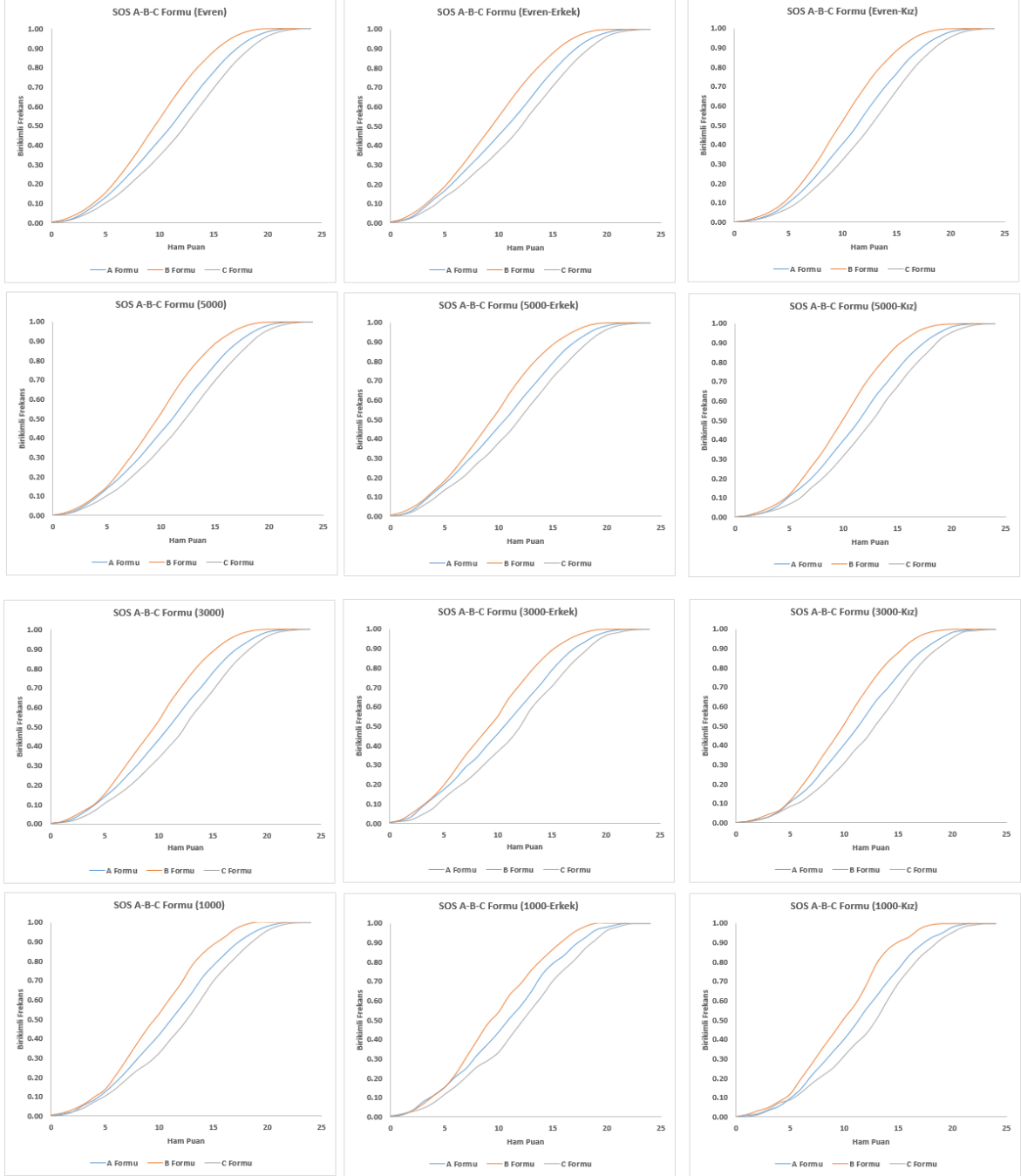
EK-K: Matematik Testi A-B-C Formlarının Ham Puan Dağılım Grafikleri (Karma)



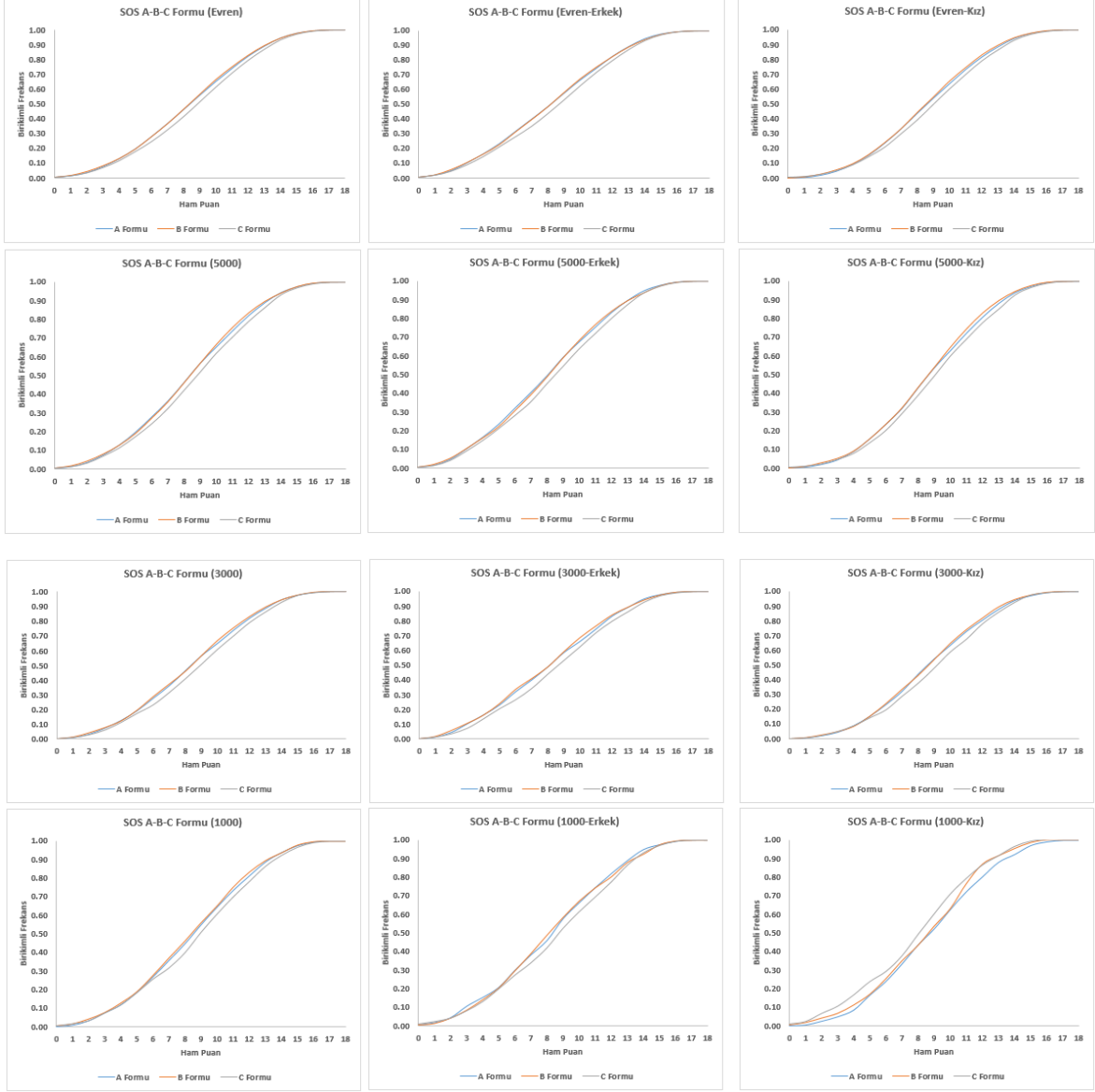
EK-L: Matematik Testi A-B-C Formlarının Ham Puan Dağılım Grafikleri (İki Kategorili)



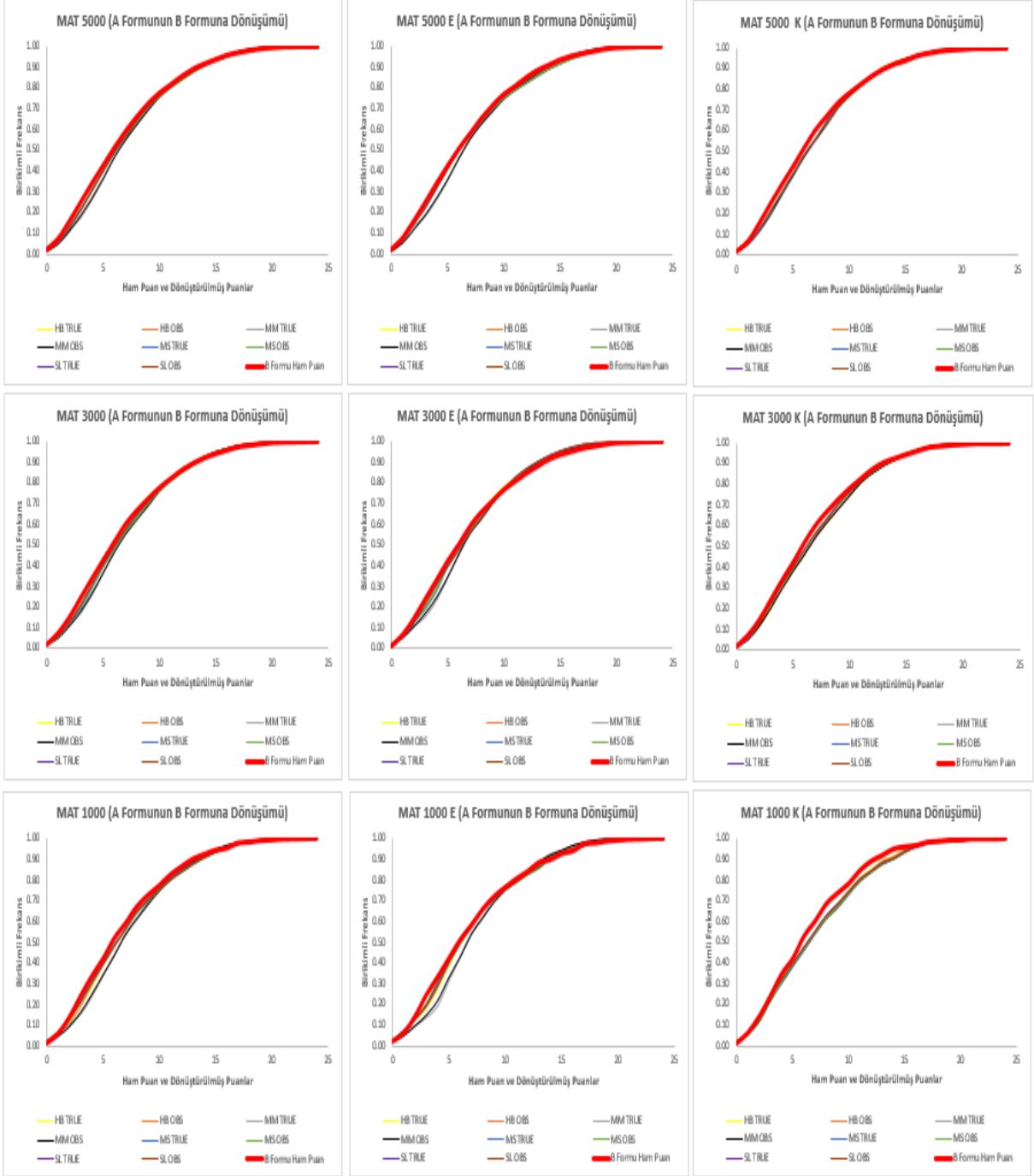
EK-M: Sosyal Bilgiler Testi A-B-C Formlarının Ham Puan Dağılım Grafikleri (Karma)



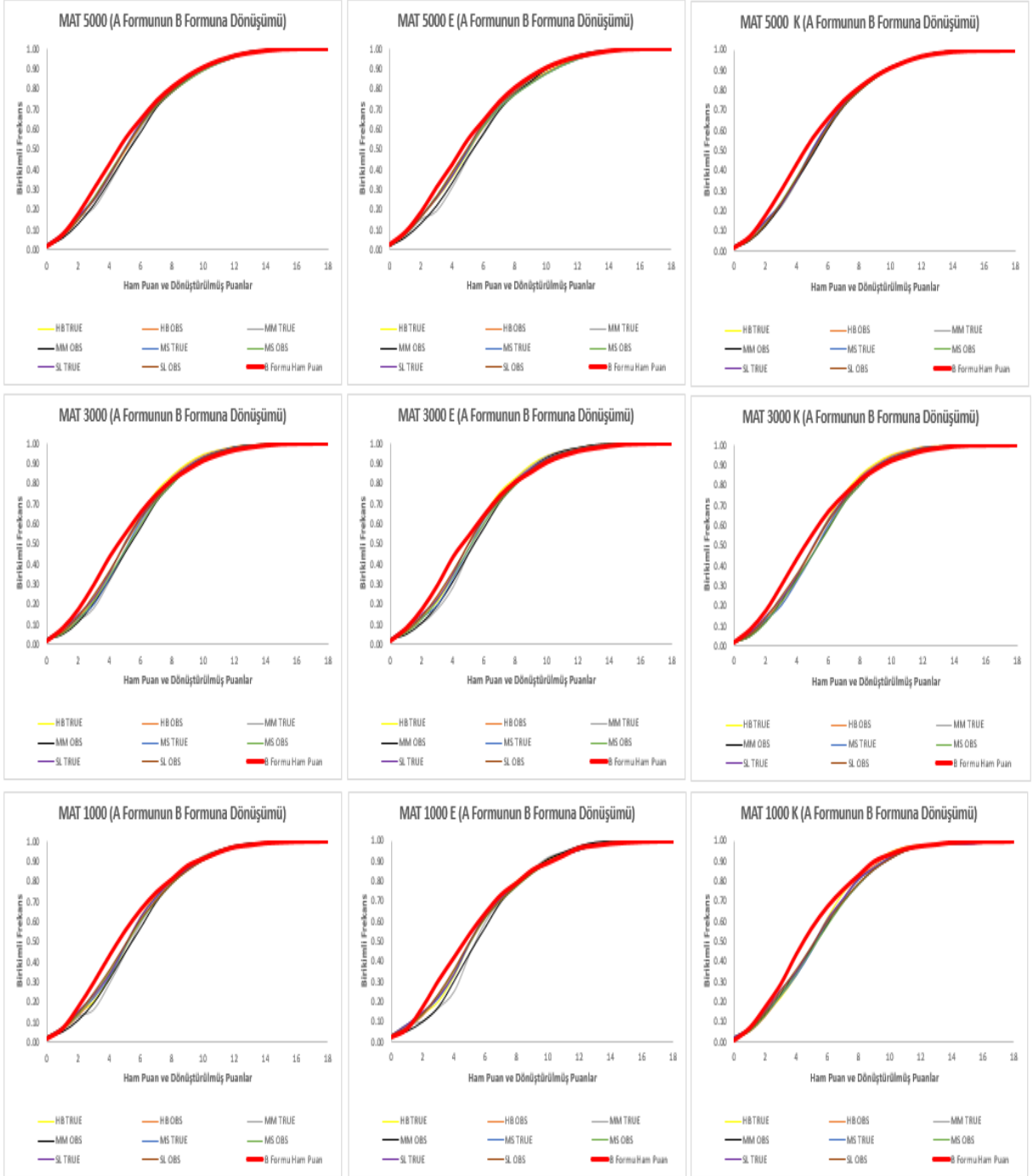
EK-N: Sosyal Bilgiler Testi A-B-C Formlarının Ham Puan Dağılım Grafikleri (İki Kategorili)



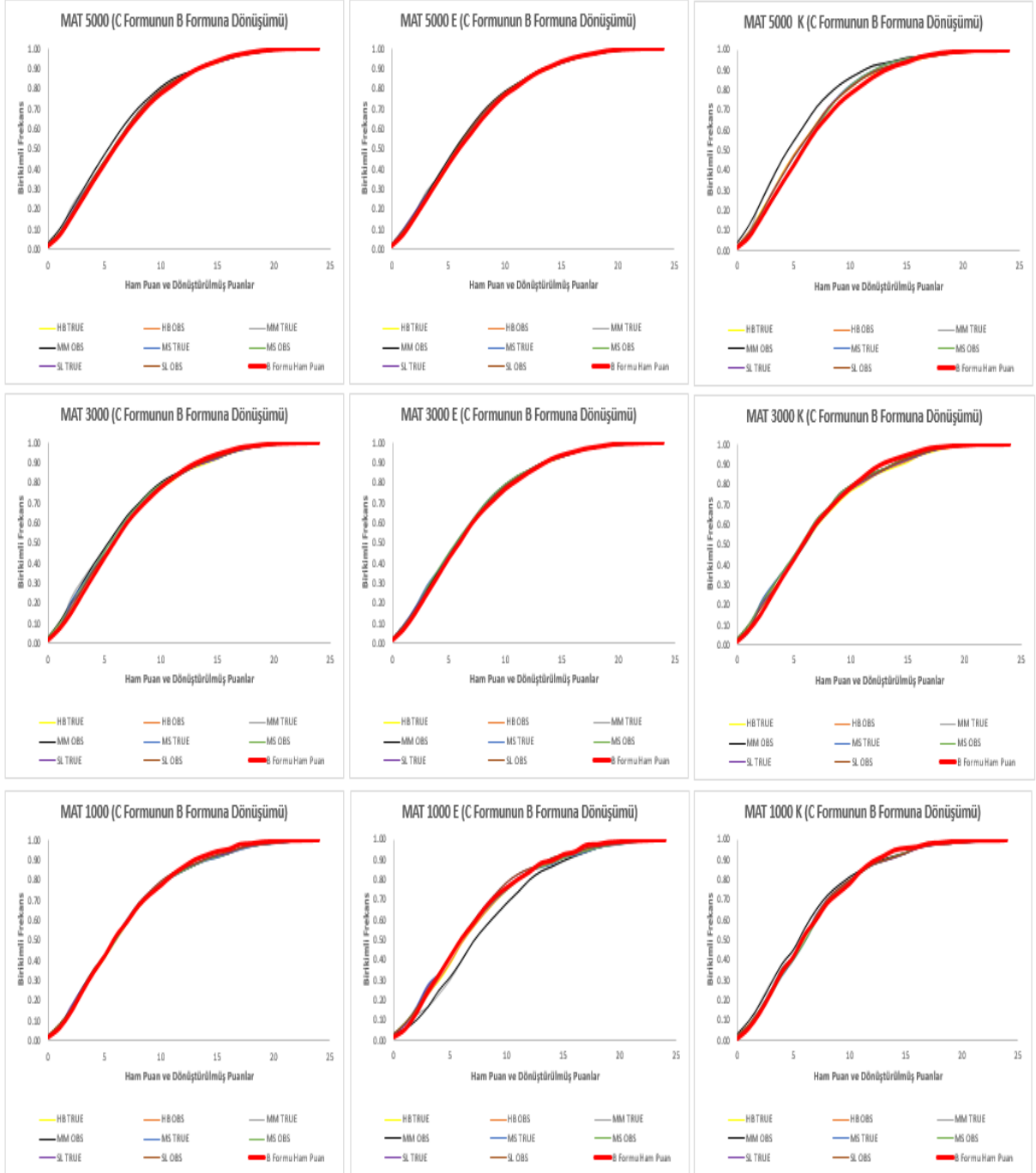
EK-O: Matematik Testi A Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (Karma)



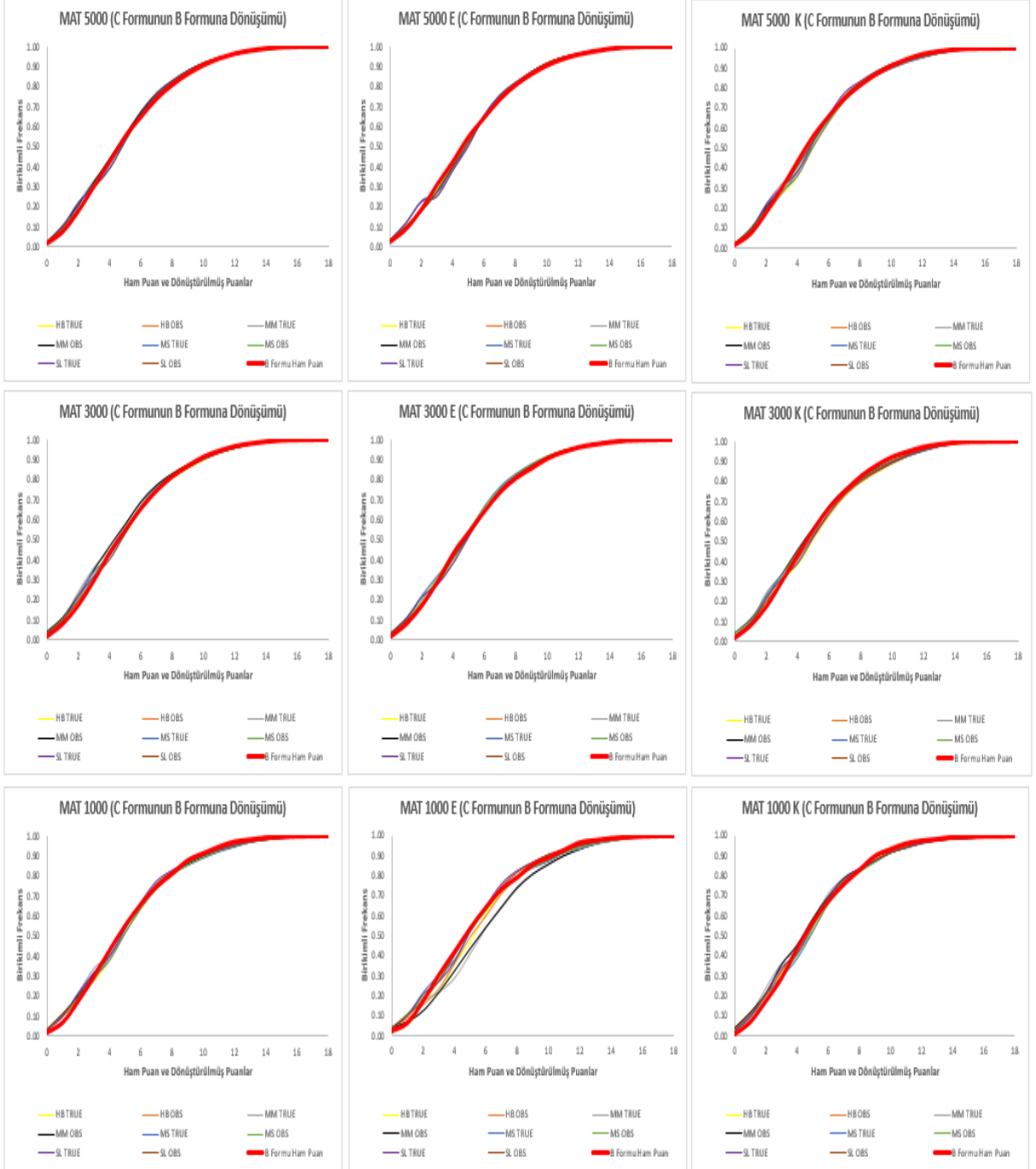
EK-Ö: Matematik Testi A Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (İki Kategorili)



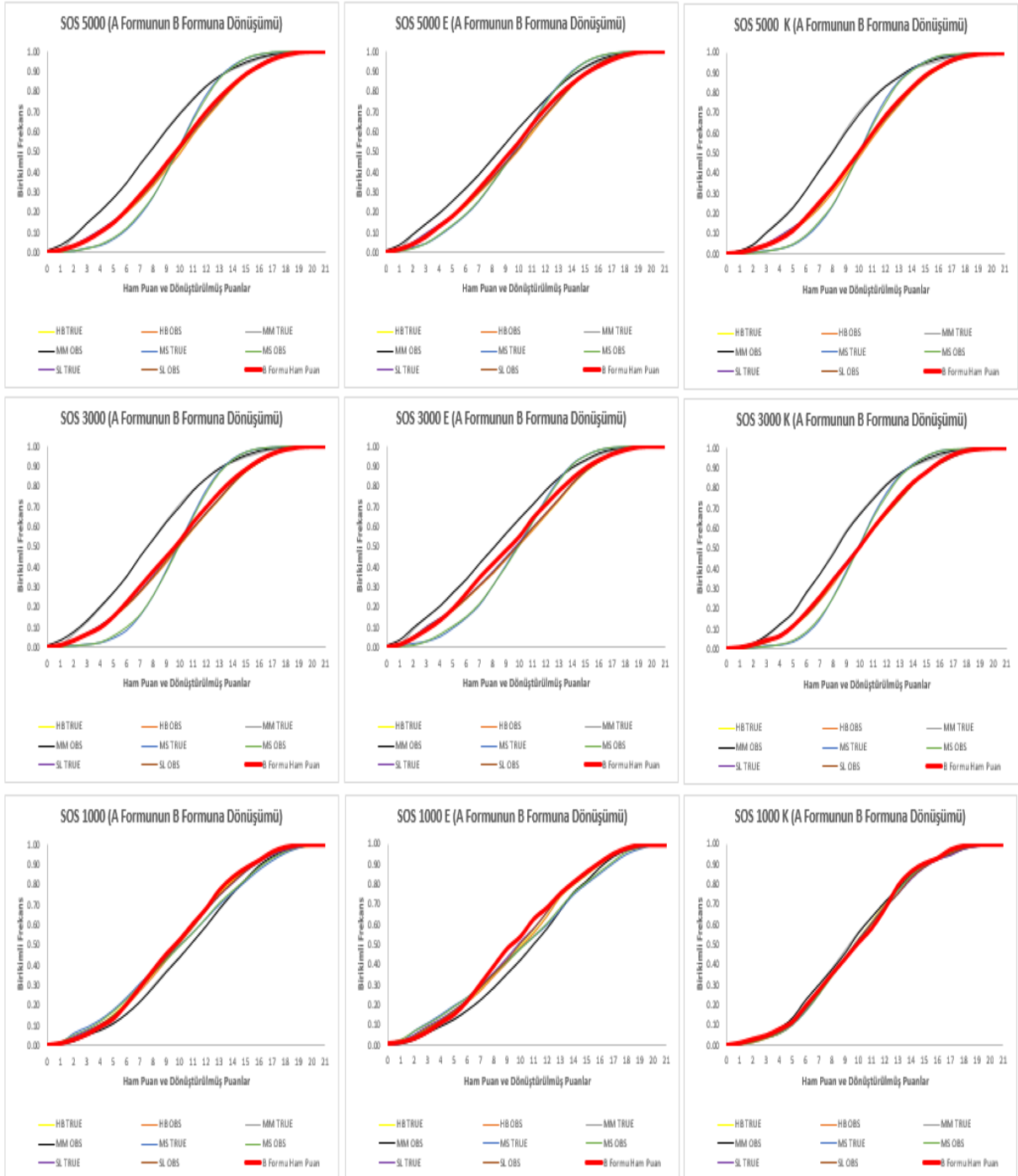
EK-P: Matematik Testi C Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (Karma)



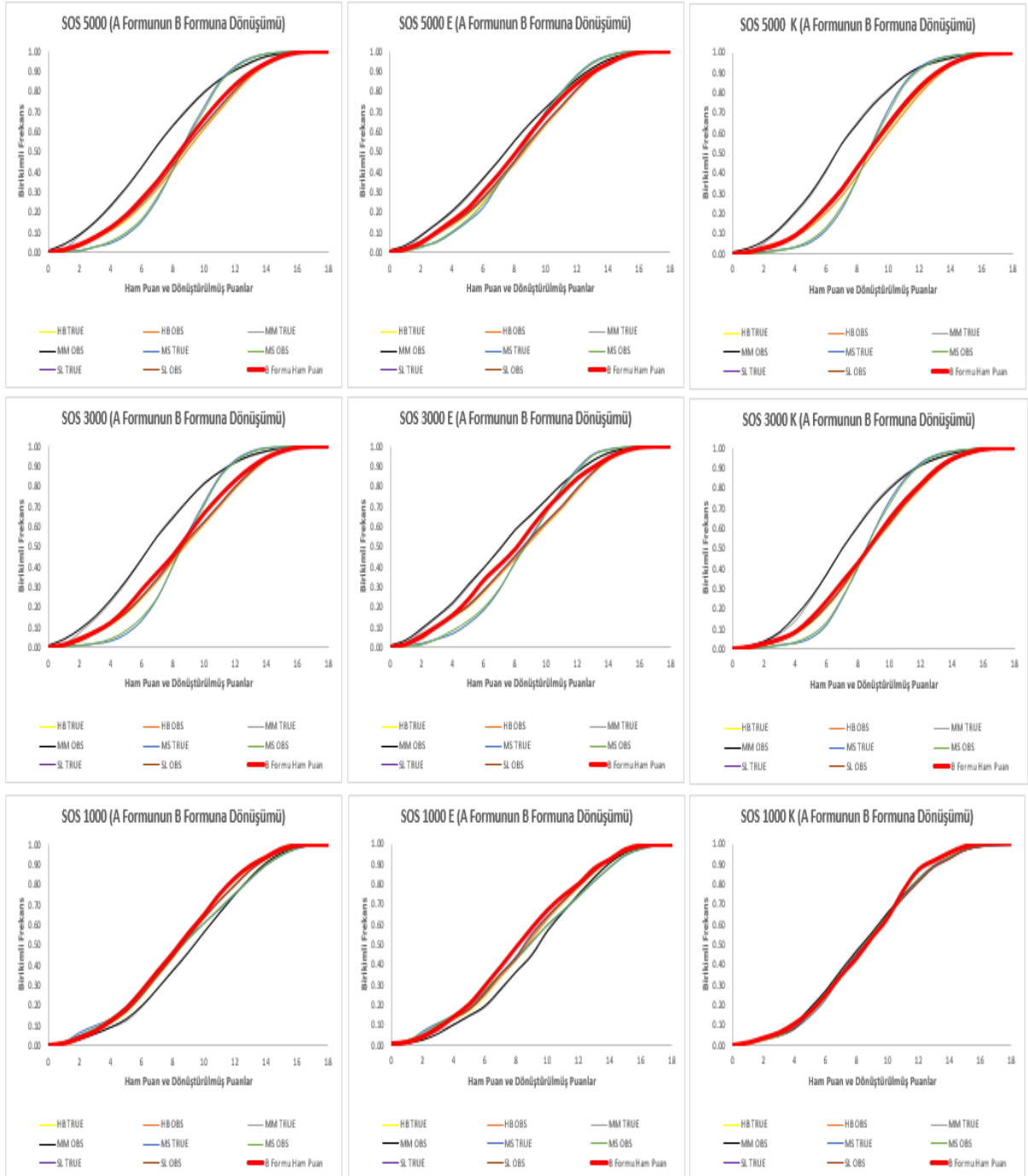
EK-R: Matematik Testi C Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (İki Kategorili)



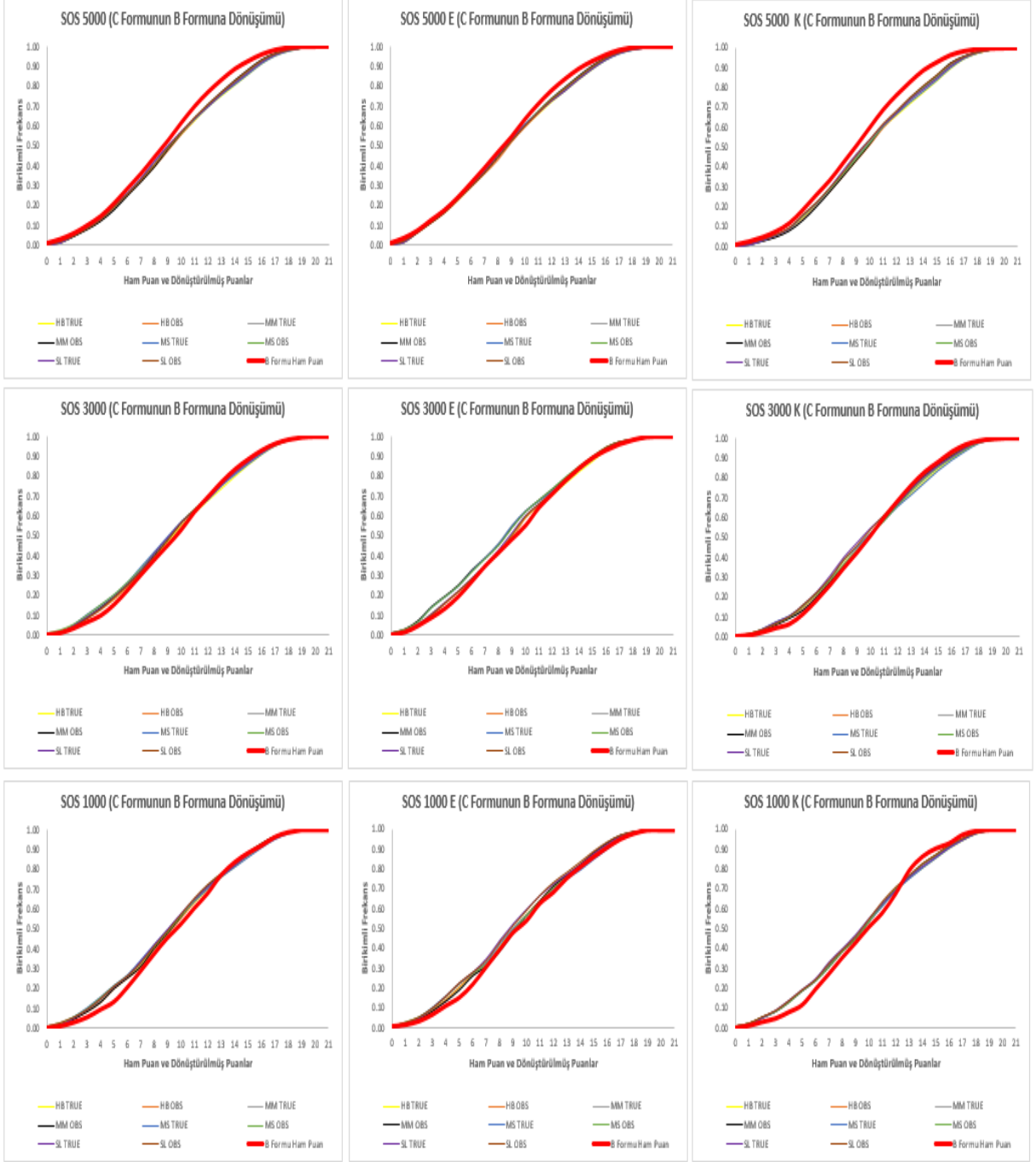
EK-S: Sosyal Bilgiler Testi A Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (Karma)



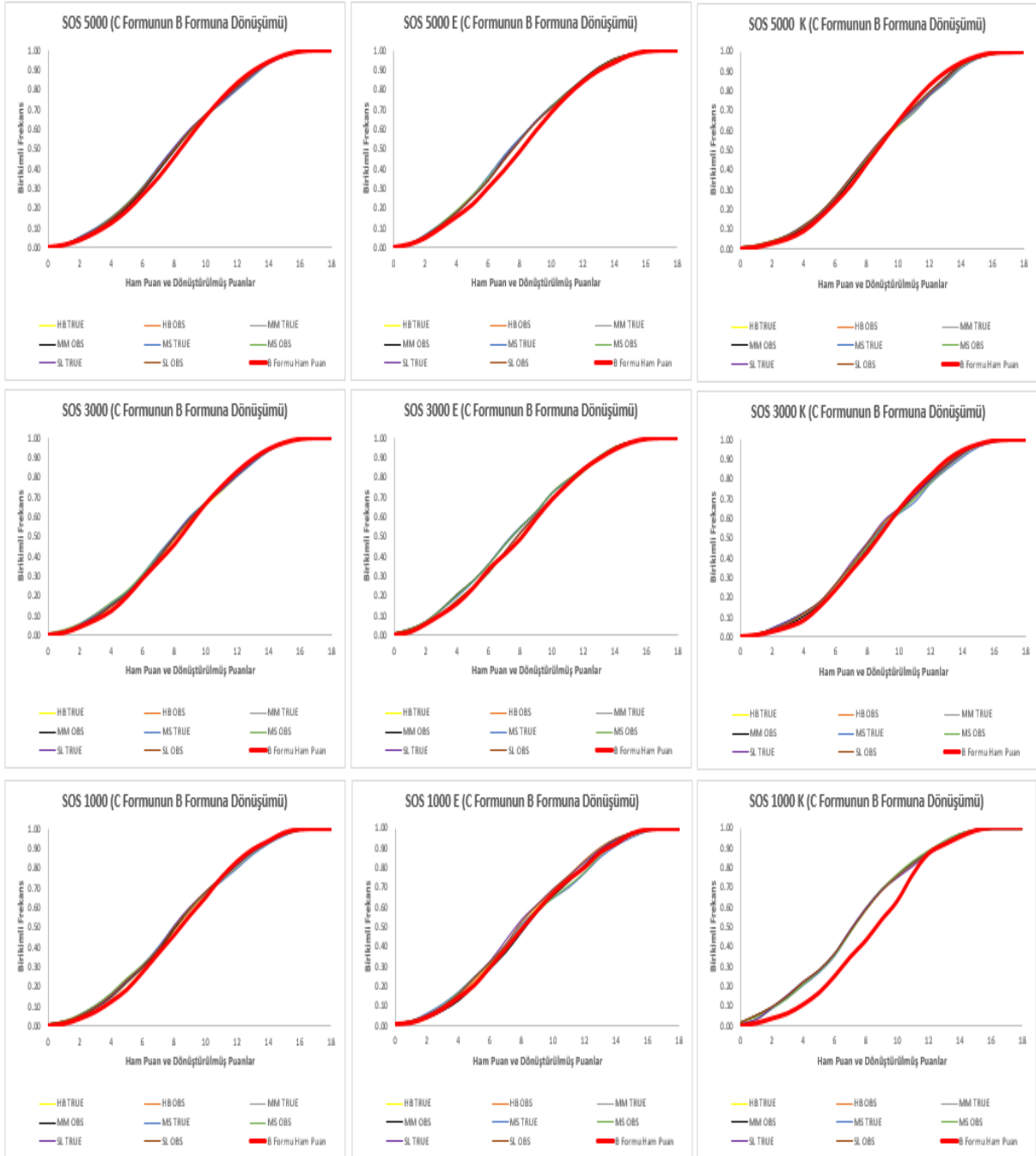
EK-Ş: Sosyal Bilgiler Testi A Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (İki Kategorili)



EK-T: Sosyal Bilgiler Testi C Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (Karma)



EK-U: Sosyal Bilgiler Testi C Formunun Ölçek Dönüştürme Sonucu Oluşan Birikimli Dağılım ile B Formunun Ham Puan Birikimli Dağılım Grafikleri (İki Kategorili)



EK-Ü: Verilerin Kullanımına İlişkin MEB'den Alınan İzin Yazısı



T.C.
MİLLÎ EĞİTİM BAKANLIĞI
Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü

Sayı : 57750415-605.01-E.12176307
Konu : Veri Talebi (Barış ÖZGÜRLÜK)

22.06.2018

HACETTEPE ÜNİVERSİTESİ REKTÖRLÜĞÜNE
(Eğitim Bilimleri Enstitüsü Müdürlüğü)

İlgi : 11.06.2018 tarihli ve 51944218-300-0092504 sayılı yazınız.

Enstitünüz Eğitim Bilimleri Ana Bilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı Doktora Programı öğrencisi Barış ÖZGÜRLÜK'ün Prof. Dr. Hülya KELECİOĞLU danışmanlığında yürüttüğü "Örneklem Büyüklüğünün, Ortak Madde Formatının ve Sayısının Sekizinci Sınıf ABİDE Testlerinin Eşitlenmesine Etkisi" başlıklı tezi kapsamında Akademik Becerilerin İzlenmesi ve Değerlendirilmesi (ABİDE) çalışmasına ait veriler ilgi yazı ile tarafımızdan talep edilmektedir.

Bu bağlamda söz konusu veri talehini içeren ilgi yazı ve eki tez önerisi Genel Müdürlüğümüze incelenmiş olup talep edilen verilerin paylaşılması uygun görülmüştür. Söz konusu veriler Barış ÖZGÜRLÜK'e teslim edilmiştir.

Bilgilerinizi ve gereğini rica ederim.

Kemal BÜLBÜL
Bakan a.
Daire Başkanı

Teknikokullar/ANKARA
Elektronik Ağ: www.meb.gov.tr
e-posta: mustafacengiz@meb.gov.tr

Ayrıntılı bilgi için: Mustafa CENGİZ (Bil. İşl.)
Tel: (0 312) 413 32 76
Faks:

Bu evrak güvenli elektronik imza ile imzalanmıştır. <https://evraksorgu.meb.gov.tr> adresinden 2c65-ca61-3af3-8780-099a koda ile teyit edilebilir.

EK-V: Etik Komisyonu Onay Bildirimi



T.C.
HACETTEPE ÜNİVERSİTESİ
Rektörlük



Sayı : 35853172-300
Konu : Barış ÖZGÜRLÜK Hk.
(Etik Komisyon İzni Hk)

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi : 11.06.2018 tarihli ve 51944218-300/00000094834 sayılı yazınız.

Enstitümüz Eğitim Bilimleri Ana Bilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı tezli doktora programı öğrencilerinden **Barış ÖZGÜRLÜK**'ün, **Prof. Dr. Hülya KELECİOĞLU** danışmanlığında yürüttüğü "**Örnekleme Büyüklüğünün, Ortak Madde Formatının ve Sayısının Sekizinci Sınıf Abide Testlerinin Eşitlenmesine Etkisi**" başlıklı tez çalışması Üniversitemiz Senatosu Etik Komisyonunun **18 Haziran 2018** tarihinde yapmış olduğu toplantıda incelenmiş olup, etik açıdan uygun bulunmuştur.

Bilgilerinizi ve gereğini saygılarımla rica ederim.

e-imzalıdır
Prof. Dr. Rahime Meral NOHUTCU
Rektör Yardımcısı

Evrakın elektronik imzalı suretine <https://belgedogrulama.hacettepe.edu.tr> adresinden **bilin@hacettepe.edu.tr** adresine yazarak ulaşabilirsiniz.
Bu belge 5070 sayılı Elektronik İmza Kanunu'na uygun olarak Güvenli Elektronik İmza ile imzalanmıştır.

Hacettepe Üniversitesi Rektörlük 06100 Sıhhiye-Ankara
Telefon:0 (312) 305 3001-3002 Faks:0 (312) 311 9992 E-posta:yazind@hacettepe.edu.tr İnternet Adresi: www.hacettepe.edu.tr



EK-Y: Etik Beyanı

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 bütün bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin bütününe kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

17/06/2019


Barış ÖZGÜRLÜK

EK-Z: Yüksek Lisans/Doktora Tez Çalışması Orijinallik Raporu

24/06/2019

HACETTEPE ÜNİVERSİTESİ
Eğitim Bilimleri Enstitüsü
Eğitim Bilimleri Ana Bilim Dalı Başkanlığına,

Tez Başlığı : Örnekleme Büyüklüğünün ve Madde Formatının Sekizinci Sınıf Abide Testlerinin Eşitlemesine Etkisi

Yukarıda başlığı verilen tez çalışmamın tamamı (kapak sayfası, özetler, ana bölümler, kaynakça) aşağıdaki filtreler kullanılarak Turnitin adlı inihal programı aracılığı ile kontrol edilmiştir. Kontrol sonucunda aşağıdaki veriler elde edilmiştir:

Rapor Tarihi	Sayfa Sayısı	Karakter Sayısı	Savunma Tarihi	Benzerlik Oranı	Gönderim Numarası
22/06 /2019	89	151194	17/06/2019	%6	1146073215

Uygulanan filtreler,

1. Kaynaklar hariç
2. Alıntılar dâhil
3. 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve çalışmamın herhangi bir inihal içermediğini aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan eder, gereğini saygılarımla arz ederim.

Ad Soyadı: Barış ÖZGÖRLÜK

Öğrenci No.: N14242604

Ana Bilim Dalı: Eğitim Bilimleri

Programı: Eğitimde Ölçme ve Değerlendirme

Statüsü: Y.Lisans Doktora Bütünleşik Dr.

İmza

DANIŞMAN ONAYI


UYGURLUR
Prof. Dr. Hülya KELECİOĞLU

EK-AA: Thesis/Dissertation Originality Report

24/06/2019

HACETTEPE UNIVERSITY
Graduate School of Educational Sciences
To The Department of Educational Sciences

Thesis Title: Effects Of Sample Size and Item Format On Eighth Grade Abide Tests Equating

The whole thesis that includes the *title page, introduction, main chapters, conclusions and bibliography section* is checked by using **Turnitin** plagiarism detection software take into the consideration requested filtering options. According to the originality report obtained data are as below.

Time Submitted	Page Count	Character Count	Date of Thesis Defense	Similarity Index	Submission ID
22/06 /2019	89	151194	17/06/2019	%6	1146073215

Filtering options applied:

1. Bibliography excluded
2. Quotes included
3. Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Educational Sciences Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility, and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.

Name Lastname: Banş ÖZGÜRLÜK
Student No.: N14242604
Department: Educational Science
Program: Educational Measurement and Evaluation
Status: Masters Ph.D. Integrated Ph.D.

Signature

ADVISOR APPROVAL


UYGUNDUR
Prof. Dr. Hülya KELECIOĞLU

EK-AB: Yayınlanma ve Fikrî Mülkiyet Hakları Beyanı

Enstitü tarafından onaylanan lisansüstü tezimin/ raporunun tamamını veya herhangi bir kısmını, basılı (kâğıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikrî mülkiyet haklarımı bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarında (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezimin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan "**Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge**" kapsamında tezimin aşağıda belirtilen koşullar hariçince YÖK Ulusal Tez Merkezi / HÜ. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- o Enstitü/Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren 2 yıl ertelenmiştir.⁽¹⁾
- o Enstitü/Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren ... ay ertelenmiştir.⁽²⁾
- o Tezimle ilgili gizlilik kararı verilmiştir.⁽³⁾

24 / 06 / 2019


Başar ÖZGÜRLÜK

"Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge"

(1) Madde 6.1. Lisansüstü teze ilişkin patent başvurusu yapılmaması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezimin erişime açılmasının ertelenmesine karar verebilir.

(2) Madde 6.2. Yeni teknik, materyal ve metodların kullanıldığı, henüz makaleye dönüştürülmemiş veya patent gibi yöntemlerle korunmamış ve internete ne yayılmış ve kurumlarca haksız kazanç imkânı oluşturabilecek bilgi ve bulguların içerdiği tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ay aşmamak üzere tezimin erişime açılması engellenebilir.

(3) Madde 7.1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat savunma ve güvenlik sağlığı vb. konularla ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tez yapıldığı kurum tarafından verilir. Kurum ve kuruluşlarla yapılan işbirliği protokolleri çerçevesinde hazırlanan lisansüstü tezlerle ilişkin gizlilik kararı ise ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.

Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir.

* Tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu tarafından karar verilir.

