

T.C.
MERSİN ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANABİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME BİLİM DALI

ARTI DEĞER ÖZELLİĞİNE SAHİP ALT TESTLERDE EŞİTLEME
YÖNTEMLERİNİN KARŞILAŞTIRILMASI

YÜKSEK LİSANS TEZİ

ARZU UÇAR

DANIŞMAN
YRD. DOÇ. DR. ÖNDER SÜNBÜL

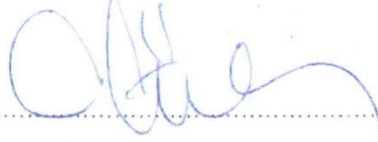
MERSİN, 2016

KABUL VE ONAY

Mersin Üniversitesi Eğitim Bilimleri Enstitüsü Müdürlüğü'ne,

Arzu UÇAR tarafından hazırlanan "Artı Değer Özelliğine Sahip Alt Testlerde Eşitleme Yöntemlerinin Karşılaştırılması" başlıklı bu çalışma, jürimiz tarafından Eğitimde Ölçme ve Değerlendirme Anabilim Dalı'nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Başkan



Doç. Dr. Cem Oktay GÜZELLER

Üye



Doç. Dr. Devrim ALICI

Üye



Yrd. Doç. Dr. Önder SÜNBÜL (Danışman)

Onay

Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylarım.



TEŐEKKÖR

Yüksek lisans eğitimin boyunca bilgilerinden faydalandığım, yanında çalışmaktan onur duyduğum ve tecrübelerinden yararlanırken göstermiş olduğu hoşgörü ve sabrından dolayı en önemli teşekkürü hak eden değerli hocam, danışmanım

Yrd. Doç. Dr. Önder Sünböl'e çok teşekkür ederim.

Yüksek lisans eğitimin sürecinde bilgi ve deneyimleriyle yol gösterici olan, tezime görüş ve eleştirileriyle katkıda bulunan Doç. Dr. Devrim ÖZDEMİR ALICI'ya, Yrd. Doç. Dr. Seçil ÖMÜR SÜNÖÖL'e ve Yrd. Doç. Dr. N. Bilge UZUN BAŐUSTA'ya ayrıca tez jürimde yer alarak görüş ve önerileriyle katkıda bulununan Doç. Dr. Cem Oktay GÖZELLER'e, tez çalışmama yardım ve katkısını esirgemeyen Uzm. Mustafa ŐAMDAN'a teşekkürlerimi sunarım.

Bütün hayatım boyunca benden maddi ve manevi desteklerini esirgemeyen, bu günlere gelmemi sağlayan aileme, çalışmam sürecinde her zaman yanımda olan ve destekleyen eşime sonsuz teşekkürlerimi sunarım.



ÖZET

ARTI DEĞER ÖZELLİĞİNE SAHİP ALT TESTLERDE EŞİTLEME YÖNTEMLERİNİN KARŞILAŞTIRILMASI

Arzu UÇAR

Yüksek Lisans Tezi

Eğitim Bilimleri Enstitüsü

Yrd. Doç. Dr. Önder SÜNBÜL (Danışman)

MERSİN, 2016

151 Sayfa

Bu çalışmada denk olmayan gruplarda ortak maddelere dayalı eşitleme deseninde artı değer özelliğine sahip olan alt testlerin alt test ve genişletilmiş alt test puanları kullanılarak eşitleme yöntemlerinin, örneklem büyüklüğüne, alt testler arasındaki ortalama güçlük düzeyi farkına ve alt test uzunluğuna göre karşılaştırılmasının yanı sıra hangi koşullar altında hangi yöntemin daha iyi sonuç verdiği araştırılmıştır.

Veri üretiminde R 3.1.1 programı kullanılmış ve iki parametrelili lojistik modele (2PLM) uygun olarak; X formu ve Y formu için 1-0 şeklinde iki kategorili veri üretilmiştir. X ve Y formları iki alt testten oluşmaktadır. Ankor test toplam test formu gibi iki alt testten oluşmaktadır. Ankor testin alt test uzunluğu, toplam testin alt test uzunluğunun %40'ı oranında madde sayısına sahiptir. 20, 25, 50, 100, 200 ve 500 örneklem büyüklükleri için alt testleri arasında 0.70, 0.80 ve 0.90 korelasyonlu; formlar arasında 0.0, 0.40, 0.70 ortalama güçlük düzeyi farkı olan; alt test uzunluğu 10, 15, 30, 50 ve 80 olan X ve Y formları oluşturulmuştur. Çalışmada birim dönüşüm, zincirlenmiş lineer eşitleme, Braun/Holland ve dairesel yay eşitleme yöntemleri kullanılarak 100 replikasyon sonucunda alt testler eşitlenmiştir. Yöntemlerin eşitleme sonuçları, eşitlemenin standart hatası (SEE), eşitleme yanlılığı (BIAS) ve eşitleme hatası (RMSE) ölçütlerine göre değerlendirilmiştir.

Genel olarak, 100, 200 ve 500 örneklem büyüklüğüne sahip, alt test uzunluğu 30 ve eşitlemenin yapılacağı test formları arasındaki ortalama güçlük düzeyi farkı 0.0 olduğunda eşitleme yapılması uygun görünürken dairesel yay eşitleme yöntemi diğer yöntemlere göre daha az hata değerleri göstermiştir. Çalışma sonucunda ortalama güçlük düzeyi farkı ve test uzunluğu arttıkça eşitlemenin yapıldığı tüm durumlarda hata değerlerinde de artış meydana gelmiştir.

Anahtar kelimeler: Test Eşitleme, Eşitleme Hatası, Artı Değer, Genişletilmiş Alt Test Puanı

ABSTRACT

COMPARING EQUATING METHODS FOR SUBTESTS WHICH HAVE ADDED VALUE

Arzu UÇAR

Master thesis

Institute of Education Sciences

Assist. Prof. Dr. Önder SÜN BÜL (Advisor)

MERSIN, 2016

151 Pages

In this study, it was aimed to compare the equating methods for nonequivalent common item group desing using subscores of subtest which had added value, using augmented subscore of subtest which had added value by the variables such as sample size, average difficulty differences between forms, test length and correlation. In addition to this which method gave better results under what conditions was also investigated.

In this study dichotomous data which was according with two parameter logistic model (2PLM) was produced for form X and form Y. R 3.1.1. programming language was used to produce data. Each test form had two subtest. Ancor test had two subtest, too. Its' subtest length was 40% of total test form X (Y). For both X and Y forms, corelation between subtests were altered in three level (0.70, 0.80 and 0.90). Moreover, average difficulty difference between subtest of form X and Y were altered in three levels (0.0, 0.4 and 0.7). Simulated forms were equated by using identity, chained linear, Braun/Holland and circle-arc methods fors ix different sample size (20, 25, 50, 100, 200 and 500) with 100 replications. The results obtained from this simulation study were evaluated based on standart error of equating (SEE), equating bias (BIAS) and equating error (RMSE) criterions.

Generally, the findings indicated in the case when sample size was 100 and more, subtest length was 30 and the level of average difficulty difference between form 0.4, it was concluded that equating forms would give better results than not equating. Furthermore circle-arc method was found to less equating error than other equating methods under most of the conditions studied. Finally, increasing the test length and average difficulty difference between forms caused to increasing equating error values of all equating methods.

Keywords: Equating, Equating Error, Added Value, Augmented Subscore

İÇİNDEKİLER

TEŞEKKÜR.....	ii
ÖZET.....	iii
ABSTRACT.....	iv
İÇİNDEKİLER.....	v
TABLolar DİZİNİ.....	x
GRAFİKLER DİZİNİ.....	xi
SİMGELER VE KISALTMALAR DİZİNİ.....	xiv
BÖLÜM I: GİRİŞ.....	1
1. TEST EŞİTLEME.....	5
1.1. Eşitlemeye Benzer Süreçler.....	5
1.1.1. Yordama.....	6
1.1.2. Ölçekleme aralığı.....	6
1.1.3. Eşitleme.....	6
1.2. Eşitlemenin Özellikleri.....	7
1.2.1. Simetri Özelliği.....	7
1.2.2. Aynı Şartlar Özelliği.....	7
1.2.3. Eşitlik Özelliği (Aynı değer).....	7
1.2.4. Gözlenen Puanın Aynı Değer Özelliği.....	8
1.2.5. Gruptan Bağımsızlık Özelliği.....	8
2. EŞİTLEME DESENLERİ.....	8
2.1. Seçkisiz Grup Deseni.....	9
2.2. Tek Grup Deseni.....	9
2.3. Karşıt Dengelenmiş Tek Grup Deseni.....	10
2.4. Aynı (Eşit) Olmayan Gruplarda Ortak Madde Deseni.....	11
3. TEST EŞİTLEME YÖNTEMLERİ.....	12
3.1. Ortalama Eşitleme Yöntemi.....	12
3.2. Lineer (Doğrusal) Eşitleme.....	13
3.3. Eşit Yüzdellikli Eşitleme.....	14
3.4. Dairesel-Yay Eşitleme.....	14
3.5. Eşit Olmayan Gruplarda Lineer Yöntemler.....	16
3.6. Tucker Yöntemi.....	17
3.6.1. Lineer Regresyon Sayıltısı.....	18
3.6.2. Şartlı Varyans Sayıltısı.....	18
3.7. Nominal Ağırlıklı Ortalama Eşitleme Yöntemi.....	18
3.8. Levine Gözlenen Puan Yöntemi.....	19
3.8.1. Korelasyon Sayıltısı.....	19
3.8.2. Lineer Regresyon Sayıltısı.....	19
3.8.3. Hata Varyansı Sayıltısı.....	20
3.9. Levine Gerçek Puan Yöntemi.....	20
3.10. Zincirleşmiş Lineer Eşitleme.....	21
3.11. Zincirleşmiş Lineer Gözlenen Puan Eşitlemesi.....	21
3.12. Zincirleşmiş Lineer Gerçek Puan Eşitlemesi.....	21
3.13. Eşit Olmayan Gruplarda Eşit Yüzdellik Yöntemi.....	21
3.14. Frekans Kestirim Yöntemi.....	22
3.14.1. Sayıltılar.....	22
3.15. Geliştirilmiş Frekans Kestirim Yöntemi.....	23

3.16. Zincirlenmiş Eşit Yüzelikli Eşitleme	23
3.17. Sentetik Fonsiyon Eşitleme	23
3.18. Braun-Holland Doğrusal Eşitleme Yöntemi.....	24
4. EŞİTLEME HATASI	24
4.1. Önyükleme (Bootstrap) Yöntemi	25
5. AMAÇ	26
6. ÖNEM	27
7. KAYNAK ARAŞTIRMASI	27
7.1. Yurt Dışında Yapılan İlgili Çalışmalar	27
7.2. Yurt İçinde Yapılan İlgili Çalışmalar	31
8. PROBLEM CÜMLESİ	34
8.1. Alt Problemler	34
9. SINIRLILIKLAR	35
BÖLÜM II: YÖNTEM	35
1. Araştırmanın Türü	35
2. Araştırmanın Değişkenleri.....	35
3. Veri Üretimi	37
4. Verilerin Analizi	39
BÖLÜM III: BULGULAR	41
1. Ortak maddelere dayalı eşitleme deseninde, artı değer özelliğine sahip olan alt testlerin, alt test /genişletilmiş alt test puanları kullanılarak yapılan eşitleme yöntemlerinin (birim eşitleme, zincirlenmiş lineer eşitleme, Braun/Holland eşitleme ve dairesel yay eşitleme yöntemleri), çeşitli faktörlere (örneklem büyüklüğü, test formları arasındaki ortalama güçlük düzeyi farkı, alt test uzunlukları) göre elde edilen eşitlemenin standart hatası (SEE), eşitlemenin yanlışlığı (BIAS) ve eşitleme hatası (RMSE) kestirimleri nasıl değişmektedir?.....	41
1.1. Çeşitli faktörlerin, ortak maddelere dayalı eşitleme deseninde, artı değer özelliğine sahip olan alt testlerin, alt test/genişletilmiş alt test puanları kullanılarak yapılan eşitleme yöntemlerinden (birim eşitleme, zincirlenmiş lineer eşitleme, Braun/Holland eşitleme ve dairesel yay eşitleme yöntemleri) elde edilen eşitlemelerin eşitleme standart hatasına (SEE), eşitleme yanlışlığına (BIAS) ve eşitleme hatasına (RMSE) temel etkisi nasıldır?	41
1.1.1.1. Örneklem büyüklüğünün, alt test puanları eşitlemelerinde, eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE) temel etkisi nasıldır?	41
1.1.1.2. Örneklem büyüklüğünün, genişletilmiş alt test puanları eşitlemelerinde, eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE) temel etkisi nasıldır?.....	46
1.1.2.1. Alt test uzunluğunun, alt test puanları eşitlemelerinde, eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE), eşitlemenin yanlışlığına (BIAS) ve eşitlemenin standart hatasına (SEE), temel etkisi nasıldır?.....	50
1.1.2.2. Alt test uzunluğunun, alt test puanları eşitlemelerinde, eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE), eşitlemenin yanlışlığına (BIAS) ve eşitlemenin standart hatasına (SEE), temel etkisi nasıldır?	54
1.1.3.1. Test formları arasındaki ortalama güçlük farkının, alt test puanları eşitlemelerinde, eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE), eşitlemenin yanlışlığına (BIAS) ve eşitlemenin standart hatasına (SEE), temel etkisi nasıldır?	58
1.1.3.2. Test formları arasındaki ortalama güçlük farkının, genişletilmiş alt test puanları eşitlemelerinde, eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE), eşitlemenin yanlışlığına (BIAS) ve eşitlemenin standart hatasına (SEE), temel etkisi nasıldır?.....	62
2. Çeşitli faktörlerin, ortak maddelere dayalı eşitleme deseninde, artı değer özelliğine sahip olan alt testlerin, alt test /genişletilmiş alt test puanları kullanılarak yapılan eşitleme yöntemlerinden (birim eşitleme, zincirlenmiş lineer eşitleme, Braun/Holland eşitleme ve dairesel yay eşitleme yöntemleri) elde edilen eşitlemelerin eşitleme standart hatasına (SEE), eşitleme yanlışlığına (BIAS) ve eşitleme hatasına (RMSE) ortak etkisi nasıldır?	66
2.1.1.1.1. Alt testleri arasında 0.70 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?	66

7. Çeşitli faktörlerin eşitleme yöntemlerinden elde edilen eşitlemenin standart hatasına (SEE) ortak etkisinin değerlendirilmesi	141
8. Çeşitli faktörlerin eşitleme yöntemlerinden elde edilen eşitleme yanlılığına (BIAS) ortak etkisinin değerlendirilmesi	143
9. Çeşitli faktörlerin eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE) ortak etkisinin değerlendirilmesi	145
ÖNERİLER	147
KAYNAKÇA	148
ÖZGEÇMİŞ.....	151



TABLÖLAR DİZİNİ

Tablo 1: Kestirilen Alt Puanların Artı Deęerini Veren İstatistikler	4
Tablo 2: Seçkisiz Grup Deseni	9
Tablo 3: Tek Grup Deseni.....	10
Tablo 4: Kaşıt Dengelenmiş Tek Grup Deseni	10
Tablo 5: Kaşıt Dengelenmiş Tek Grup Deseni Örneęi	10
Tablo 6: Aynı (Eşit) Olmayan Gruplarda Ortak Madde Deseni	12
Tablo 7: Araştırmada Deęişimlenen Faktörler.....	38



Grafik 46: 0.90 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme standart hata (SEE) değerlerine ortak etkisi.....	133
Grafik 47: 0.90 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme yanlılığı (BIAS) değerlerine ortak etkisi	135
Grafik 48: 0.90 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme hatası (RMSE) değerlerine ortak etkisi	137



SİMGELER VE KISALTMALAR DİZİNİ

NEAT	: Denk olmayan gruplarda ortak madde deseni
ID_y(x)	: Birim eşitleme fonksiyonu
m_y(x)	: Ortalama eşitleme fonksiyonu
l_y(x)	: Lineer Eşitleme fonksiyonu
e_{Lin-y}(x)	: Lineer bileşen fonksiyonu
e_{Arc-y}(x)	: Eğrisel bileşen fonksiyonu
e_{CIRC-y}	: Dairesel yay eşitleme fonksiyonu
SEE	: Eşitlemenin standart hatası
BIAS	: Eşitleme yanlılığı
RMSE	: Eşitleme hatası
a	: Madde ayırt edicilik parametresi
b	: Güçlük parametresi
2PLM	: İki parametrelili lojistik model

BÖLÜM I: GİRİŞ

İnsanların yeryüzündeki yaşama ortamına duyduğu merak, olağan gibi görünen olayları anlama çabası, aslında dünyanın bilinmeyenlerle dolu bir yer olduğunu ve bunları çözümlmek gerektiği gerçeğini doğurmuştur. Bilim fiziki ve doğal evrenin yapısının ve deney, gözlemler aracılığıyla sistematik bir şekilde incelenmesini de kapsayan pratik çalışmalar bütünüdür. Bilim kuramsal ve deneysel olmak üzere iki öğeden oluşan bir sistemdir. Bilimsel çalışmalarda öğeler arasındaki ilişkileri ortaya çıkaracak bağlantıları doğrulamak, doğrulanan bağlantıları genellemek ve genellemelerden kanunlara ulaşmak amaçlanır. Öğeler arasındaki yapıların gözlem ve deneylere dayalı olarak kurulmasında geçerli ve güvenilir veri elde edilmesi gereklidir. Bu amaçla, bilimin kendine özgü ölçme araç ve yöntemlerinden yararlanır.

Eğitim bilimlerinde, eğitim kararlarının verilmesinde de güvenilir ve geçerli bilgi sağlamada ölçmeden yararlanır. Eğitim bireylerde istenilen davranış değişikliği ya da oluşturulmasını sağlayan bir süreç olmanın yanısıra, davranış geliştiren bir sistemdir. Bu sistemin kontrolü eğitimde değerlendirme ögesi ile yapılır. Değerlendirme, “ölçme sonuçlarının bir ölçüt ile kıyaslanarak karara varılma işidir” (Akt: Baykul, 2010:98, Turgut, 1977). Değerlendirme tanımındaki ölçme ögesi, değerlendirme sonucunda verilecek olan karara esas olacak bilginin toplanmasını sağlar. Kararın isabetlilik derecesi, ölçütün uygunluğunun yanında, ölçme sonuçlarının geçerlik ve güvenilirliğine bağlıdır.

Bireylerde istenilen niteliklerin ne derece kazandırıldığını ölçmeyle belirleyebiliriz. Çünkü “ölçmenin ana konusu belli bir özellik veya nitelik” (Atılgan, 2014:2). Çoğu nitelik doğrudan gözlenemez, gözlenebilse de doğrudan ölçülemez. Doğrudan gözlenebilen ya da doğrudan gözlenemeyen özellik veya değişkenlerin miktarına ilişkin sonuçlar verilirken ölçme araçlarından yararlanır. Eğitimde de öğrencilerin ilgi, yeteneklerinin kestirilmesi, performansları, başarıları ile ilgili bilgilerin elde edilmesi için ölçme araçları kullanılır. Eğitimde kullanılan ölçme araçlarına test adı verilir (Baykul, 2010). Testlerden elde edilen bilgilerin kullanılarak karara varılması ya da çalışmaların yapılabilmesi için testlerin geçerli ve güvenilir olması gerekir. Geçerlik en genel tanımıyla “ölçme aracının ölçmek istediği özelliği başka değişkenlerle karıştırmadan ölçebilme derecesidir” (Akt: Atılgan, 2014, APA, 1974). Testlerde bulunması gereken bir diğer özellik olan güvenilirlik, test puanlarının seçkisiz hatalardan arınlık derecesidir. Başka bir ifadeyle, ölçme aracının ölçmek istediği özelliği hatasız olarak ölçebilme derecesidir.

Bir testin güvenilirliğinin yorumlanması geliştirilmiş olduğu kurama göre değişiklik gösterir. Eğitimde nitelikli ölçme aracı geliştirmek için çeşitli kuramlar ortaya çıkmıştır. Ortaya çıkan kuramlardan biri Klasik Test Teorisi’ dir. Klasik Test Teorisi istatistiksel bir kuramdır. Gözlenen bir özelliğin gerçek değeri ölçmeye karışan çeşitli hatalar nedeniyle ölçme yoluyla doğrudan elde edilemez. Gerçek puan ölçülen özelliğin gerçek değeridir. Gerçek puan Klasik Test Teorisinde ölçülen özelliğe ilişkin sonsuz sayıda ölçmelerden elde edilen puanların ortalaması olarak tanımlanır. Ölçülmeye çalışılan özellikler sabittir ve bu özellikleri ölçmek amacıyla geliştirilen ölçme araçlarından elde edilen sonuçlar ise bir birinden farklı olabilmektedir. Bu farklılık özellikleri ölçerken ölçmeye karışan hatalardan kaynaklanır. Dolayısıyla bireyin bir özelliğini ölçmeye ilişkin hazırlanmış bir testten elde ettiği puan, özelliğin bireydeki gerçek değerini ifade eden bireyin gerçek puanı ve hata puanlarından oluşur. Bu açıklama 1nolu eşitlikte ifade edilmiştir. X; bireyin testten elde ettiği gözlenen puanı, T; bireyde sabit olan ve gözlenemeyen gerçek puan, E; ölçmeye karışan hata olmak üzere;

$$X = T + E \quad (1)$$

Eşitlikteki hata ölçmeye karışan kaynağı belli olmayan seçkisiz hatadır. Eşitlikten yararlanarak bireylerin puanları arasındaki değişkenliği; S_x^2 ; bireylere ait gerçek puan varyansı, S_T^2 ; hata varyansı, S_E^2 hata varyansı olmak üzere;

$$S_x^2 = S_\tau^2 + S_e^2 \quad (2)$$

şeklinde ifade edilir. Bu eşitliğe dayanarak Klasik Test Teorisinde güvenilirlik indeksi

$$r_{tt} = \frac{S_t^2}{S_x^2} \quad (3)$$

$$r_{tt} = 1 - \frac{S_e^2}{S_x^2} \quad (4)$$

şeklinde elde edilir. Güvenirlik indeksinin elde edilmesi için gerçek puanlara ihtiyaç vardır. Oysa gerçek puanlar bir tek ölçme yoluyla elde edilemediğinden doğrudan hesaplanamaz. Hata değeri, kaynağı ve miktarı belli olmadığı için doğrudan belirlenemez. Bu nedenle güvenilirlik indeksi kuramsal bir değerdir. Güvenirlik indeksi yerine kullanılacak başka kavramların aranmasına yol açmıştır. Böylece ilgilenilen özelliği ölçen paralel testlerin geliştirilmiştir. Birbirine paralel ölçümler tanımlamak ve bu ölçümlerin gözlenen puanlar üzerindeki etkisini belirlemek söz konusudur. Böylece ölçme aracının farklı durum ve koşullarda aynı sonuçları üretip üretmediği kontrol edilmiş olacaktır. Dolayısıyla güvenilirlik, iki paralel ölçme arasındaki korelasyon katsayısıdır.

Eğitimde kullanılan testler alt içeriklerden oluşabilmektedir. Örneğin bir matematik testinin cebir ve geometriden oluşması ya da genel yetenek testinin matematik, yazma, okuma gibi alt alanlardan oluşması gibi. Testlerin alt alanlarından oluşan alt testlerden elde edilen alt test puanları kullanılarak bireyin öğrenme eksikliklerinde ya da daha başarılı olduğu konuların analiz edilmesini sağladığı gibi okulların profillerinin ortaya çıkarılmasında yararlanır. Ayrıca alt puanlar bir toplam testin yapısından daha detaylı sonuçlar çıkarmak için kullanılır. Başka bir ifadeyle ölçülmesi istenilen özellik için kullanılan maddeler başka bir özelliğin ölçümü içinde kullanılabilir. Bu durumda bu tarz maddelerin bir araya gelmesiyle yeni bir alt özellik ortaya çıkarak testin yapısıyla ilgili daha ayrıntılı bilgi elde edilebilir (Shinary, Haberman ve Puan, 2007). Testi oluşturan alt alanların sağladığı bu avantajlarla alt test puanlarına olan ilgi artmıştır. Dolayısıyla bireyin bir alt alandaki puanı, toplam testten olduğu gibi alt alanla ölçülmek istenen özelliğin gerçek değerini ifade eden bireyin alt gerçek puanı ve alt hata puanlarından oluşur. Alt puanın matematiksel olarak ifadesi aşağıdaki eşitlikte gösterilmiştir.

$$S_x = \tau_x + e_x \quad (5)$$

(5) numaralı eşitlikte S_x , τ_x ve e_x ile ifade edilen sırasıyla alt alanın gözlenen puanı, alt alanın gerçek puanı ve alt alanın hata puanıdır. Alt alandaki puanlara ilişkin değişkenlik gözlenen alt puan için σ_{S_x} , gerçek alt puan değişkenliği σ_{τ_x} ve alt hata puanı değişkenliği ise σ_{e_x} ile gösterilmek üzere;

$$\sigma_{S_x}^2 = \sigma_{\tau_x}^2 + \sigma_{e_x}^2 \quad (6)$$

(6) numaralı eşitlikle ifade edilir (Akt: Haberman, 2008:206, Lord ve Novick, 1968). Alt test puanının Klasik Test Teorisi'ne dayanarak güvenilirlik katsayısı,

$$\rho^2(S_x, \tau_x) = \frac{\sigma_{\tau_x}^2}{\sigma_{S_x}^2} \quad (7)$$

şeklinde ifade edilir.

Alt test puanlarının kullanımına olan ilginin artmasıyla birlikte toplam test puanının mı alt puanının mı kullanımı amaca daha çok hizmet edeceği araştırılmıştır. Araştırmalar sonucunda her alt test puanının kullanılmasının uygun olmadığı ortaya çıkmasıyla birlikte, alt puanın kullanılabilmesi için gereken şartlara kısmen de olsa belirtilmiştir (Sinharay, Haberman ve Puhan, 2007; Haberman, 2008; Sinharay, 2010a; Sinharay, 2010b, Sinharay ve Haberman, 2011). Alt test puanının kullanılabilir olması için alt puanın artı değer özelliğine sahip olması gerektiği vurgulanmıştır. Alt puanın artı değere sahip olması için ne gibi özellikleri olmalıdır? Başka bir ifadeyle alt puanların artı değere özelliği için alt puanların güvenilirliği farklılığı nasıl değişimlenmelidir? Bu soruların ya da artı değer özelliğinin olduğuna dair doğrudan bir cevap olmamasına karşın farklı yaklaşımları mevcuttur. Artı değer özelliği için alt puanların yüksek güvenilirliğe toplam testin görece daha düşük güvenilirliğe ve diğer alt testlerden farklı olması (ilişkinin düşük olması) gibi özelliklere bakılmaktadır (Sinharay, 2010).

Alt puanın toplam testin puanından daha çok bilgi verip vermediğinin kontrolü için alt puan üç farklı hesaplamayla ifade edilmiştir (Haberman, 2008). Gerçek alt puanın, gözlenen alt puan, gözlenen toplam test puanı, gözlenen alt puan ve toplam test puanı üzerine olan regresyonuyla üç farklı hesaplamaya dayanan alt puanları, alt puan yaklaşımı, toplam test puanı yaklaşımı, toplam test puanı ve gözlenen alt puan yaklaşımı olarak sırasıyla ifade edilen (8), (9) ve (10) numaralı eşitlikteki denklemlerde ifade edilmiştir.

$$S_x = L(\tau_x | S_x) = E(S_x) + \rho^2(S_x, \tau_x)[S_x - E(S_x)] \quad (8)$$

$$S_x = L(\tau_x | S_z) = E(\tau_x) + \beta(\tau_x | S_z)[S_z - E(S_z)] \quad (9)$$

$$S_x = L(\tau_x | S_x, S_z) = E(S_x) + \beta(\tau_x | S_x, S_z)[S_x - E(S_x)] + \beta(\tau_x | S_z, S_x)[S_z - E(S_z)] \quad (10)$$

$E(S_x)$ gözlenen alt puan ortalaması, $\rho^2(S_x, \tau_x)$ alt puanın güvenilirlik katsayısı, $\beta(\tau_x | S_z)$ toplam test puanından yararlanarak elde edilen alt puan denklemindeki regresyon katsayısını, $\beta(\tau_x | S_x, S_z)$ ve $\beta(\tau_x | S_z, S_x)$ toplam test puanı ve gözlenen alt puanının kullanılarak elde edilen alt puan denklemindeki regresyon katsayılarını göstermektedir. Eşitlik (8), (9) ve (10) kullanılarak elde edilen alt puanlar gerçek alt puan değerleriyle olan farkları yaklaşımın hatasını ifade eder (Haberman, 2008). Bu hataların varyansları ise gerçek alt puan için hata kareleri ortalaması (MSE) değerini verir. Hata kareleri ortalamasının karekökü alınarak hata kareleri ortalamasının karekök (RMSE) değeri elde edilir. Üç farklı yaklaşımla elde edilen alt puanların hangilerinin kullanımının uygun olacağı ya da toplam test puanının kullanılması gerektiği gibi sonuçlara ulaşmak için gözlenen alt puan kullanılarak, toplam test puanı kullanılarak, hem gözlenen alt puan hem de toplam test puanı kullanılarak elde edilen alt puanların RMSE değerleri karşılaştırılır. RMSE değerleri küçük olan alt puan kullanılmalıdır. Başka bir ifadeyle en az hataya sahip olan yaklaşım için alt puan artı değer özelliği vermektedir denebilir. Alt puanın artı değer özelliğine sahip olup olmadığını araştırmada yardımcı bir diğer katsayı ise alt puanın güvenilirlik katsayılarını içerisinde barındıran PRMSE değerlerine bakılır. PRMSE değeri büyük olan yaklaşım için alt puan artı değere sahiptir denir.

Kestirim Yöntemi	RMSE	PRMSE
Alt Puan	$\sigma_{\tau_x} [1 - \rho^2(S_x, \tau_x)]^{1/2}$	$\rho^2(S_x, \tau_x)$
Toplam Puan	$\sigma_{\tau_x} [1 - \rho^2(S_z, \tau_x)]^{1/2}$	$\rho^2(S_z, \tau_x)$
Toplam Puan ve Alt Puan	$\sigma_{\tau_x} [1 - \rho^2(S_x, \tau_x)]^{1/2} [1 - \rho^2(S_z, \tau_x \cdot S_x)]^{1/2}$	$1 - [1 - \rho^2(S_x, \tau_x)] [1 - \rho^2(S_z, \tau_x \cdot S_x)]$

Tablo 1: Kestirilen Alt Puanların Artı Değerini Veren İstatistikler

Tablo1’de alt puanın artı değer özelliğine sahip olup olmadığını kontrol etmek için kullanılan RMSE ve PRMSE ifadeleri verilmiştir. Gözlenen alt puan kullanılarak elde edilen alt puan S_x , toplam test puanı kullanılarak elde edilen alt puan S_z , hem gözlenen alt puan hem de toplam test puanı kullanılarak elde edilen alt puan S_{xz} ile ifade edilsin. Alt puanların PRMSE değerleri ise sırasıyla $PRMSE_x$, $PRMSE_z$ ve $PRMSE_{xz}$ ile gösterilmek üzere; $PRMSE_x > PRMSE_z$ olduğunda alt puan artı değere sahiptir. Dolayısıyla toplam test puanının ifade ettiği anlamdan daha çok anlama sahiptir. Başka bir ifadeyle alt puanın kullanılabilir olduğunu, alt puanla profil oluşturulabileceğini gösterir. Aksi durumda toplam test puanı kullanılmalıdır. $PRMSE_{xz} > PRMSE_x$ ve $PRMSE_{xz} > PRMSE_z$ olduğunda toplam test puanı ve gözlenen puandan elde edilen puan artı değer özelliği taşımaktadır. Bu puana özel olarak genişletilmiş puan adı verilir (Haberman, 2008). Genişletilmiş puan alt puandan daha çok bilgi ifade etmektedir. Ayrıca genişletilmiş puanlar test formları arasındaki karşılaştırmayı kolaylaştırılır (Sinharay, 2011).

Eğitimde testler, bireylerin ya da okulların profillerinin oluşturulmasında, bireylerin öğrenim düzeylerinin takip edilmesinde, bireylerin bir kuruma seçilmesi ve yerleştirilmesinde, bireylerin işe alımında kullanılmaktadır. Kullanılan testlerden elde edilen puanlar doğrultusunda bireyler ile ilgili kararlar verilmektedir. Verilen kararlar bireylerin hayatlarının şekillenmesinde son derece önemli rol oynamaktadır. Doğru kararların verilebilmesi için testlerin yanlı kararlara yol açmaması, geçerli ve güvenilir olması gerekir. Bu nedenle ölçmek istenilen özellikle ilgili testlerin bir defadan fazla kullanılmaması gerekir. Başka bir ifadeyle sınav güvenliği nedeniyle farklı tarihlerde ya da eş zamanda testlerin farklı paralel formları uygulanabilir. Klasik Test Teorisi’ne göre farklı test formlarından elde edilen puanların birbirleriyle karşılaştırılabilmesi için test formlarının paralel olması gerekir. Ancak paralel test formu hazırlansa da testi alan bireylerin yetenek düzeyleri farklı olacağından her birey için farklı hata kaynakları söz konusu olur. Dolayısıyla testten elde edilen puanlarda farklı miktarda hatalara sahip olur. Klasik Test Teorisi’nde madde güçlüğü ve maddenin ayırt ediciliği gruplara özgüdür. Bu nedenle farklı testleri alan kişileri paralel formlar ile karşılaştırmak güçtür.

Uygun koşullar altında uygun kararların verilebilmesini sağlamak için test puanlarının karşılaştırılması gerekmektedir. Bu karşılaştırma testlerin eşitlenmesiyle gerçekleştirilebilmektedir.

1. TEST EŞİTLEME

Farklı sınav günlerinde farklı testlerin uygulanması farklı test formu zorluklarıyla karşılaşılmasına sebep olmuştur. Örneğin; ikinci kez alınan bir testten elde edilen puanlar birinci alıma göre daha yüksek olabilir. Bu durumun iki nedeni vardır.

İlk olarak ikinci teste kadar geçen sürede testi alan kişi öğrenmelerine devam etmiştir. Dolayısıyla başarısı artmıştır. İkinci olaak ikinci kez testi alan öğrencinin birinci kez aldığı testteki maddeleri hatırlama etkisinden dolayı yüksek puan almıştır.

Yukarıdaki iki durumunda ayırt edilmesi zordur. Dolayısıyla aynı testin birden fazla defa kullanılması istenen bir durum değildir ve farklı test formlarının kullanımı zorunlulaşmıştır. Farklı test formlarının kullanılması da başka problemlere yol açmıştır. Örneğin; aynı okuldaki iki öğrenciyle iki farklı test formlarından biri uygulanmıştır. Birinci öğrenci ikinci öğrenciden yüksek puan almıştır. Bu sonuçlar karşısında birinci öğrenci ikinci öğrenciden daha yüksek başarıya sahiptir denebilir. Ancak test formlarının zorlukları farklıysa? İkinci öğrencinin aldığı test formu birinci öğrencinin aldığı test formundan zor olabilir. Bu durumda sadece formlardan elde edilen puanları ele alarak yorumda bulunmak yanlış kararlar alınmasına neden olur.

Test puanlarının kullanımı ya da açıklanması için test puanlarının güvenilir, geçerli ve karşılaştırılabilir olması gerekmektedir. Test puanlarının karşılaştırılabilir olması ise bireyin hangi test formunu almış olduğuna bakılmaksızın tüm puanların aynı doğru üzerine yerleştirilebilmesini sağlamaktadır. Dolayısıyla puanların anlamlı olaması için puanlar ölçeklenmiş puanlara çevrilmelidir. Ölçeklenmiş puanları elde etmek için test formları arasında ilişkilendirme yapılarak test formlarından elde edilen puanlar bağlantılandırılmalıdır.

Puanların bağlantılandırılması ile puanların eşitlemesi aynı durumu ifade etmemekle birlikte puanların eşitlenmesi puanların bağlantılandırılmasının özel bir halidir. Puan bağlantılandırılması bir test puanının bir diğer test puanına dönüştürülmesi anlamındadır. Ancak test eşitlemede ise bu bağlantılandırma bir testin diğerine dönüştürülmesi dışında özelliklerin de sağlanmasıyla gerçekleşmektedir. Belli bir özelliği ölçen testin ya da belli özelliği ölçen test formlarının kullanılacak olan fonksiyona uygun olarak verilerin elde edilmesiyle eşitlenmesidir. Dolayısıyla eşitleme yapabilmek için aynı özelliğin ölçüldüğü test formları, testin uygulanacağı grup, testte kullanılacak yöntem gibi durumların uygun hale getirilip sonrasında puanların birbirine dönüştürülmesi işlemidir.

Eşitleme istatistiksel bir süreçtir ve aynı özelliği ölçen farklı test formlarından elde edilen puanların birbirleri yerine kullanılmasını sağlar. Eşitleme ile test formları eşitlenir. Test formları başarılı bir şekilde eşitlenirse test formlarından elde edilen eşitlenmiş puanlar kullanılarak yorum yapılabilir. Eşitleme yapılabilmesi için farklı içerikli test formları olmamalıdır. Diğer ifadeyle test formları farklı zorlukta olabilir ancak farklı içerikte olmamalıdır. Aynı özelliği ölçen farklı testler olmalıdır.

1.1. Eşitlemeye Benzer Süreçler

Bağlantılandırmanın farklı türleri vardır. Hollan ve Dorans (2006) farklı türlere vurgu yapmaktadır(Akt:Sinharay ve Rao, 2006:170, Hollan ve Dorans, 2006). Hollan ve Dorans bağlantılandırmayı iki test puanı bir birine dönüştürme olarak tanımlamıştır ve bağlantılandırmayı üçe ayırmıştır.

- Yordama
- Ölçekleme aralığı (ölçek hizalama) (scale alignig)
- Eşitleme (equating)

1.1.1. Yordama

Puan bağlamanın en eski yöntemidir. Psikometrinin ilk dönemlerinde eşitlemeyle karıştırılmıştır. Yordama, test formlarından birinin diğerinin üzerindeki regresyonuyla puanların elde edilmesini sağlar. Ancak bu işlemde test puanlarını etkileyen çok değişken yer almaktadır. Örneğin demografik bilgiler, bilişsel bilgiler gibi. Bilişsel olmayan bilgiler (motivasyon), bilişsel bilgiler gibi bilgileri yordar (Sinharay ve Rao, 2006). Dolayısıyla bu yöntem asimetri oluşturur. (Yordanan ile yordayıcı arasında bir asimetri oluşturur.) Bu şekilde yapılan bir bağlantılandırma test puanlarının eşitlenmesinde ya da test puanlarının karşılaştırılmasında kullanılamaz.

1.1.2. Ölçekleme aralığı

Farklı test formlarından elde edilen puanları aynı ölçek üzerine yerleştirmeyi sağlayacak bir dönüşümdür. Bu yöntem sayesinde puanların karşılaştırılabilirliği ortaya çıkmıştır (Sinharay ve Rao, 2006).

1.1.3. Eşitleme

Eşitleme, bağlantılandırma türlerinden güçlü olanıdır. Eşitleme aynı ölçek üzerinde test puanlarının yerleştirilmesini sağlamanın yanında bunun için gerekli verileri daha kaliteli olmasını sağlayan özelliklerin de barındırmasında hasastır. Örneğin eşitlemenin yapılabilmesi için aynı özelliği ölçen test formları olması, aynı güvenilirliğe ve aynı zorluk düzeyinde olması gibi şartları barındıran test formlarındaki puanları aynı ölçek üzerine yerleştirilmesini sağlamaktadır.

y, yeni test formu olan Y testinden elde edilen puan; $s(y)$, Y puanlarının aynı ölçek üzerine yerleştirilmesini sağlayan ölçekleme fonksiyonu; x, eski test formu olan X testinden elde edilen puan olmak üzere bu iki formun birbiriyle ilişkilendirecek fonksiyon $e(x)$ fonksiyonudur. X ve Y test formlarından elde edilen puanları aynı ölçek üzerinde göstermek için x puanını Y testindeki bir puana eşitledikten sonra $s(y)$ fonksiyonu yardımıyla aynı ölçek üzerine yerleştirilir.

İki test formundan elde edilen puanların eşitlenmesi, iki testten elde edilen puanlar aynı testten elde edilmiş puanlar gibi karşılaştırmayı mümkün kılar. Bunun gerçekleştirebilmek için testlerin aynı yapıyı ölçmesi ve aynı zorluk da olması gerekir (Sinharay ve Rao, 2006).

Crocker ve Algina (1986) test eşitlemeyi iki ölçme aracından eşdeğer puanlar oluşturma işlemi olarak tanımlamıştır. Angoff test eşitlemeyi bir test formunun birim sistemini diğer formun birim sistemine dönüştürme olarak tanımlamıştır (Angoff, 1984:7).

Ham Puan (raw score): Çoktan seçmeli testlerde cevaplayıcının maddeleri doğru olarak cevapladığı madde sayısıdır.

Ölçek Puanları: Ham puanların dönüştürülmüş değerleridir.

1.2. Eşitlemenin Özellikleri

Eşitleme için gerekli olan beş özellik vardır. Bu beş özelliğin gerçekte sağlanabilirliği zor olmanın dışında bazılarının sağlanması diğerlerinde varlığını aşikar olarak gösterdiği araştırmacılar tarafından ifade edilmiştir. (Holland and Dorans, 2006, Lord, 1980; Dorans and Holland, 2000).

Bazı özellikler kişisel puanlara odaklanırken bazıları puan dağılımına odaklanmıştır. Kişisel düzeyde katılımcının aldığı test formundaki puanı aynı formlara etki etmeksizin aynıdır. Puan dağılımında da benzer bir yaklaşım mevcuttur. Bir grup katılımcının elde ettiği puanların diğer testteki oranı da aynı olmalıdır. Bu özellikler eşitleme sürecinin gelişmesinde temel alınarak kullanılmıştır (Kolen ve Breannan, 2014).

Eşitleme diğer bağlantılandırma türlerinden güçlü olanıdır. Gücünü sağlaması gereken özelliklerden almaktadır. Bu özellikler, simetri, aynı şartlar özelliği (eşitlenecek test formları aynı yapı ölçmelidir), eşitlik özelliği, gözlenen puanın aynı değer özelliği, gruptan bağımsızlık özellikleridir.

1.2.1. Simetri Özelliği

Lord (1980) eşitlemenin simetri özelliğinden bahsetmiştir. X ve Y test formları olmak üzere, X formundan elde edilen puanlar Y formuna dönüştürülebiliyorsa bu dönüşümün tersi de vardır. Y formundaki puanların ters görüntüsü X formundaki puanlara denk gelmektedir. Örneğin 26 alan bir kişinin Y formundaki eşitlenmiş puanı 27 puana denk geliyorsa bu dönüşümün tersi kullanılarak Y formundan 27 alan kişinin X formundaki eşitlenmiş puanı 26 puana denk gelecektir (Kolen ve Breannan, 2014).

$y \in Y$ ve $x \in X$ ve f ve g dönüşüm fonksiyonları olmak üzere;

$$y = f(x) \text{ ise } g(y) = x \text{ dir.}$$

1.2.2. Aynı Şartlar Özelliği

Eşitlemenin yapılabilmesi için testler aynı yapıyı ölçmelidir. Testlerin ölçtüğü içerik aynı olmalıdır. Testler aynı istatistiksel özelliklere sahip olmalıdır. Aynı istatistiksel özelliklere sahip olmayan testler kullanılarak elde edilmiş eşitlemedeki puanlar birbirinin yerine kullanılamazlar. Bu özellik alterne formların birbirlerinin yerine kullanıldığı zaman gereklidir (Kolen ve Breannan, 2014).

1.2.3. Eşitlik Özelliği (Aynı değer)

Lord (1980) eşitlemenin eşitlik özelliğine göre, eşitlenecek testler eşitlik özelliği taşıyorsa, cevaplayıcının hangi testi aldığı önem taşımamaktadır. Çünkü testler aynı yapıyı ölçmektedir ve cevaplayıcının bu yapıdaki gerçek tek bir puanı vardır. Cevaplayıcının gerçek puan dağılımı her test formu için aynıdır. Dolayısıyla τ gerçek puan, x cevaplayıcının X formundan aldığı puan, y ; cevaplayıcının Y formundan aldığı puanı, G ; Y formunu alan grup için puanların kümülatif dağılımını,

eq_y ; X formundaki puanın Y formuna dönüştürülmüş ölçek puanını ifade etmektedir. G^* ise eq_y puanlarının yığılmalı frekansı değerini göstermek üzere;

$$G^*[eq_y(x|\tau)] = G(y|\tau) \quad (11)$$

Eşitliğe göre verile gerçek puan için, gözlenen puan ortalaması, standart sapması ve X formundan Y formuna dönüştürülmüş puanların dağılımı aynıdır. Bu özellik hipotetiktir. Gerçek hayatta sağlanması zor bir özelliktir.

Morris' in aynı değerler özelliği Lord'un aynı değerler özelliğine göre daha az kısıtlayıcıdır. Aynı değer özelliğine göre verilen bir gerçek puan değeri için X formundan Y formuna dönüştürülebilir puanların ortalaması ile Y formundan X formuna dönüştürülen puanların ortalaması aynıdır.

E beklenen değer operatörü olmak üzere ve her τ değeri için

$$E[eq_y(x|\tau)] = E(y|\tau) \quad (12)$$

(12) numaralı denklem şeklinde ifade edilmiştir (Kolen ve Breannan, 2014).

1.2.4. Gözlenen Puanın Aynı Değer Özelliği

Angoff (1971), gözlenen puan eşitlemesinin puan dağılımı eşitlenmiş olan formun dağılımıyla aynıdır. Dolayısıyla;

$$G^*[eq_y(x)] = G(y) \quad (13)$$

bu özellik eşit yüzdelli eşitleme olarak bilinir. Eşitlenen formlar eşit yüzdelli eşitleme özelliğini sağlamışsa, X formunda ölçek puanı 26'nın altında olanların oranı ile Y formunda ölçek puanının 26'nın altında olanların aynıdır (Kolen ve Breannan, 2014).

1.2.5. Gruptan Bağımsızlık Özelliği

X ve Y puanlarını birbiriyle ilişkilendiren eşitleme fonksiyonu evren-alt evren seçiminden bağımsız olarak aynıdır (Sinharay ve Rao, 2006). Başka bir ifadeyle gruptan bağımsızlık özelliği sağlanmışsa aynı grupta kadın ya da erkek içinde aynı eşitleme sonuçları bulunacaktır (Kolen ve Brennan, 2004).

2. EŞİTLEME DESENLERİ

Eşitleme yapmak için veri kümesinin belli kurullarla elde edilmesine eşitleme deseni denir (von Davier, 2010). Verileri üç farklı şekilde toplayabiliriz.

1- Aynı gruba eşitlemenin yapılacağı referans ve yeni test birlikte uygulanarak.

2-Aynı özelliklere sahip gruplara iki form uygulanarak.

3-Farklı test formlarından biri bir gruba uygulanarak.

Verilerin toplanması farklılık göstereceği için kullanılacak istatistiksel yöntemler de farklılık gösterecektir. Aşağıda eşitlemede en çok kullanılan verilerin elde edildiği eşitleme desenlerinden bahsedilecektir.

2.1. Seçkisiz Grup Deseni

Tek grup deseni ve karşıt dengelenmiş grup desenlerinde eşitlenecek olan test formlarının ikisinde uygulanarak veriler elde edilmektedir (Kolen ve Brennan, 2014). Ancak aynı özelliği ölçen farklı test formlarını aynı bireyler tarafından cevaplandırılması kolay gerçekleştirilememektedir. O zaman bireylere birer test formu uygulayarak gerekli sonuçlar elde edilerek eşitleme yapmak mümkün müdür? Denk grup deseniyle birbirine eşit yetenek düzeyindeki gruplardan elde edilen verilerle eşitleme yapmak mümkündür. Bunun için bir birine eşit yetenek düzeyinde olan gruplara seçkisiz olarak X ya da Y testleri verilebileceği gibi aynı salonda bulunan eşit yetenek düzeyine sahip bireylere eşitlenecek testlerden bir tanesini uygulayarak iki grup elde edilmiş olur. Testlerin uygulanacağı bireylerin sayısı arttıkça grupların özellikleri bir birine daha yaklaşır. Dolayısıyla büyük örnekleme çalışıldığında denk grup deseniyle elde edilen veriler daha sağlıklı sonuçlar verir. Bu özelliği aynı zamanda eşitleme deseninin sınırlılığı da olabilmektedir. Eşitlenecek test formlarının aynı anda uygulanmadığı durumlarda da sınırlılık olabilmektedir. Daha önce uygulanmış test maddeleri bir şekilde yeni testin uygulanacağı bireylerce elde edilebilir. Dolayısıyla aynı anda uygulanıldığında daha doğru sonuçlar elde edilebilir.

Yukarıda bahsedildiği gibi sınava katılacakların seçkisiz olarak formlara atandığı desendir. Örneğin formlar katılımcılara, biri X diğeri Y formu olacak şekilde spiral olarak verilebilir. Bu durumda katılımcılar formlara atanmış olur. Seçkisiz grup deseninin kullanıldığı durumda, gruplar arasındaki performans farklılığı, farklı formların zorluklarından kaynaklı olan farklıdır. Örneğin X formunun ham puan ortalaması 72, Y formunun ortalaması 77 puan ise X formu Y formundan 5 puan daha zordur şeklinde yorum yapılabilir. Dolayısıyla ardaki bu farkı dengeleyecek şekilde test formları gruplara verilmelidir.

Seçkisiz grup deseninde katılımcılar tek bir test formu almaktadırlar. Dolayısıyla ikiden fazla birçok test formu uygulanarak zaman kaybetmeksizin veri toplanabilir.

Tablo 2: Seçkisiz Grup Deseni

Evren	Örneklem	Form1	Form2
P	O1	×	
P	O2		×

2.2. Tek Grup Deseni

Eşitlemenin yapılacağı iki test aynı bireylere uygulanır. Testlerin uygulandığı örneklemin yetenek düzeyi evrenden düşük ya da evrenden büyük olabilir. Bilgiler aynı bireylerden toplandı için elde edilen veriler istatistiksel açıdan daha güçlüdür. Ayrıca desenden elde edilen veriler diğer desenlerden elde edilen verilere göre ilişkilendirme de en doğru bilgilerin elde

edilmesini sağlar. Dezavantajı ise aynı bireylere iki test formunun uygulanması sonucu ikinci test formundaki performans birinci uygulanan test formundan etkilenmektedir (von Davier, 2010).

Tek grup deseninde sıralama ve alışma etkisi görünmektedir. Sıralama etkisi olarak ikinci testte bireyler yorulmuş olduğu için performanslarında düşüklük yaşanabiliyorken, test formlarının paralel olmasına özen gösterilerek formlar hazırlandığı için hatırlama etkisi bireylerde görülmektedir. Dolayısıyla test maddelerine cevap verirken bir önceki test formunda gösterdiği performanstan daha az performans göstermeyebilir.

Katılımcılar iki test formunu da alırlar. Önce X formunu alan katılımcı ardından Y formunu alarak sınavını tamamlar. Katılımcıya iki form birden uygulandığı için X formunda gösterdiği performansı Y formunda göstermeyebilir. Dolayısıyla Y formu X formuna göre zor gelebilir. Diğer taraftan birbirine benzeyen formları alan katılımcılar cevaplama sürecinde X formundaki performanslarını hatırlayarak Y formunda kullanır. Böylece katılımcıya Y formu X formuna göre daha kolay gelebilir. Bu duruma sıra etkisi denir.

Tablo 3: Tek Grup Deseni

Evren	Örneklem	Form1	Form2
P	O1	×	×

2.3. Karşıt Dengelenmiş Tek Grup Deseni

Tek grup deseninde sıralama etkisini ortadan kaldırmak için oluşturulan bir desendir. Birbirinden çok farklı olmayan gruplardan her birine eşitlenecek olan test formlarından (X ve Y test formlarından) bir tanesi verilir. Yeterli bir sürenin ardından gruplara almadıkları ikinci test uygulanır. Bu desen kullanılarak ilk formu alan gruplardan elde edilen veriler ile eşitleme yapılır. Böylece tek grup desenindeki sıra etkisi eşitlemede görülmeyecektir. İkinci formda ise yorulmadan kaynaklı performans düşmesi X ve Y formları için dengelenmiş olacaktır. Karşıt dengelenmiş tek grup deseninde farklı sıra etkisi ortaya çıkacaktır. İlk formların kullanıldığı eşitlemeyle ikinci formların kullanıldığında elde edilecek eşitleme farklı olacaktır.

Tablo 4: Karşıt Dengelenmiş Tek Grup Deseni

Evren	Örneklem	Form1 ₁	Form2 ₂	Form2 ₁	Form1 ₂
P	O1	×	×		
P	O2			×	×

Tablo 5: Karşıt Dengelenmiş Tek Grup Deseni Örneği

	Grup 1	Grup 2
İlk Test Formu	X	Y
	72	77
İkinci Test Formu	Y	X
	75	71

Tabloya göre ilk test formunun uygulanması sonucu elde edilen ortalama X formu için 72 Y formu için 77' dir. Elde edilen sonuçlara göre X formu Y formundan 5 puan daha zordur. İkinci test formları alındığında elde edilen sonuçlara bakılacak olursa Y formu X formundan 4 puan daha kolaydır. X ve Y testlerinin önce ve sonra uygulanmasıyla elde edilen 4-5 puanlık farklar farklı sıralama etkisi göstermektedir (Kolen ve Brennan, 2010).

2.4. Aynı (Eşit) Olmayan Gruplarda Ortak Madde Deseni

Eşitlenecek yada eşitleme yapılacak testler tamamen paralel olması günlük hayatta gerçekleştirilmesi ya da ulaşılması zor bir durumdur. Test formları arasında zorluk farkı meydana gelmektedir. Ayrıca test formlarının uygulandığı gruplarda yetenek düzeyleri de bir birinden farklı olduğunda test formlarının eşitlenmesi için ortak maddelerin kullanıldığı ortak maddelere dayalı eşitleme deseni kullanılır (von Davier, 2010). Bu desende ortak maddelerden oluşan ankor test eşitlenecek iki testle yüksek korelasyon göstermelidir. Ayrıca ankor test eşitlenecek iki testin içeriğini yansıtacak özellikte olmasının yanı sıra istatistiksel olarak da benzer özelliklere sahip olmalıdır. Böylece yapılacak olan eşitleme doğru sonuçlar verecektir.

Farklı evrenlerden seçilen iki örneklemden ilk örnekleme X, diğer örnekleme Y formu uygulanarak verilerin toplandığı desendir. Bu desende eşitlemenin yapılabilmesi için X ve Y formuyla aynı özelliklere sahip ortak maddelerden oluşan bir test daha uygulanır. Bu test ile X ve Y test formları eşitlenir. Ortak maddelerden oluşan test formu X ve Y test formunun içerisinde yer alıyorsa iç ankor denir. X ve Y formunun yanında ayrıca verilen bir form şeklindeyse dış ankor adını alır. İç ankor test formunun içerisinde yer aldığı için tüm cevaplayıcılar tarafından cevaplanmaktadır. İç ankordan elde edilen puan toplam test formundan elde edilen puanı etkilemektedir. Dış ankor ise isteyen katılımcılar tarafından cevaplanır. Dış ankor test formundan ayrı verildiği için toplam test formundan elde edilen puana etkisi yoktur. Ankor testin kullanımına bağlı olarak toplam test formundan elde edilen puanların ifadesi

iç ankor

$$x_{P_1} = x_x + x_v \quad (14)$$

$$x_{P_2} = x_y + x_v \quad (15)$$

dış ankor;

$$x_{P_1} = x_x \quad (16)$$

$$x_{P_2} = x_y \quad (17)$$

şeklinde dir. V ankor test, X test formu, Y test formu sırasıyla; x_v, x_x, x_y formlarından alınan puanlar ve x_p sınavlardan alınan puanlar, V, X formu P_1 popülasyonu tarafından alınmış ve V, Y formu ise P_2 popülasyonu tarafından alınmıştır.

İç ankor test formlarında aynı sırada yer almalıdır. Form içerisinde aynı sırada yer almazsa sınav süresince olabilecek yorulma, zor maddelerden sonra gelmişse motivasyon düşüklüğünden kaynaklı performansın azalmasına neden olur. Ankor testinde elde edilecek veriler sıralama etkisiyle toplanacaktır.

Tablo 6: Aynı (Eşit) Olmayan Gruplarda Ortak Madde Deseni

Evren	Örneklem	Form1	Form2	V
P	O1	×		×
Q	R1		×	×

Ankor testte dayalı desen kullanılması diğer desenlere göre daha avantajlıdır. Çünkü ankor test eşitlenecek olan testlerin istatistiksel özelliklerini taşıyan, içerik özelliği aynı olan ayrıca testlerin mini versiyonudur. Dolayısıyla ankor test deseni tek grup desenine göre daha avantajlıdır. Çünkü ankor Y testine de benzediğinden aynı anda hem X hem Y testi uygulanmış gibi düşünülebilir. Test madde sayısı tek testin uygulanmasından dolayı hatırlama ya da yorulma etkisini ortadan kaldıracaktır. Ankor testte dayalı desen denk grup desenine göre de avantajlıdır. Çünkü X ve Y formunun ikisi her iki popülasyonda da uygulanmıştır.

Eşitleme desenlerinin gerektirdiklerine bakıldığında tek grup deseni için küçük örneklerim kullanılması uygundur. Denk grup deseni için ise örneklem büyüklüğü fazla olmalıdır. Ankor teste dayalı desende ise her ne kadar ankor testlerin testlerle olan korelasyonunun yüksek olması durumunu gerçekleştirilmesi gerekse bile örneklem büyüklüğü tek grup deseni için gerekli olan büyüklük ile denk grup deseni için gerekli örneklem büyüklüğü arasında olmalıdır (Doran, Moses ve Eignor, 2010).

3. TEST EŞİTLEME YÖNTEMLERİ

Test formlarının eşitlenebilmesi için eşitlemenin yapılacağı formların aynı yapıyı ölçmesi, aynı içeriğe sahip olması ve aynı istatistiksel özelliklere sahip olması gerekir. Bu şartlar altında eşitleme yapılacak Klasik Test Teorisi' ne dayalı eşitleme yöntemleri; ortalama eşitleme, lineer eşitleme, yüzdelik eşitleme, dairesel yay eşitleme. Braun/Holland eşitleme, nominal ağırlıklı ortalama eşitleme, Tucker eşitleme ve Levine eşitlemedir.

3.1. Ortalama Eşitleme Yöntemi

Eşitlenecek iki test formunun aynı özelliğe sahip gruplardaki dağılımların bilindiğini varsayalım. Eşitlenecek olan test formlarındaki madde özelliklerine bakıldığında yeni formda zor madde sayısı eşitlenecek olan referans test formuna göre daha fazla olsun. Yetenek düzeyi yüksek olan kişilerin aldıkları test formlarına tepkilerinde, kolay ve orta zorluktaki maddelerle ilgili bir sorun yaşanmayacakken zor maddelerde ise sorun yaşanacaktır. Test puanlarını karşılaştırabilir yapmak için zor formun test puanına gerekli puan eklemesi yapılarak eşitleme sağlanabilir. Yetenek düzeyi düşük olan bireylerin bu test formlarını aldığını düşünelim. Yetenek düzeyi düşük olan bireyler zor sorularda yeterli performansı göstermeyecekler. Test formlarındaki kolay test maddelerinin sayısı yeni test formunda fazlayken referans test formunda sayı azdır. Dolayısıyla yeni form eski forma göre daha kolay olacaktır. Bu yüzden eşitleme yapıldığında puanlardan belirli miktar düşülecektir. Test puanlarının eşitlenmesi için test formlarının dağılımlarından yararlanarak testlerin arasındaki fark hesaplandıktan sonra eşitlenecek olan test puanlarına uygun olarak puan eklenir ya da eksiltir. Bu şekilde yapılan eşitleme ortalama eşitlemedir.

Puanların eklenmesi ya da düşürülmesi uygun bir yöntem değildir. Sadece ortalama eşitlemenin basit yapısı için anlatılan bir örnektir. Eğer test dağılımları aynı standart sapmaya sahipse test formları birbirine denktir. Dolayısıyla test formlarıyla eşitleme gerçekleştirilebilir.

Aşağıda ortalama eşitleme yönteminde kullanılan eşitleme fonksiyonu ifade edilmiştir.

Formlar arasındaki zorluk değeri tüm ölçek puanlarında sabittir. Başka bir ifadeyle X formu iki puan daha kolaysa Y formundan, bu fark düşük puanlı katılımcılar arasında da aynıdır.

x ; X formundan alınan puan, y ; Y formundan alınan puan, $\mu(x)$; X formunun ortalaması, $\mu(y)$; Y formunun ortalaması olmak üzere;

$$\begin{aligned} x - \mu(x) &= y - \mu(y) \\ y = m(x) &= x - \mu(x) + \mu(y) \end{aligned} \quad (18)$$

(18) nolu denklem ortalama eşitleme yönteminin fonksiyonudur.

3.2. Lineer (Doğrusal) Eşitleme

X ve Y formlarından elde edilen puanların karşılaştırılabilir olması için birim puanlara dönüştürülmesi gerekir. İki form eşit ise;

$$\frac{x - \mu(x)}{\sigma(x)} = \frac{y - \mu(y)}{\sigma(y)} \quad (19)$$

(19) nolu eşitlik şeklinde ifade edilir. $\sigma(x)$, $\sigma(y)$ sırasıyla X ve Y formlarının standart sapması olmak üzere, y denklemden çekilirse;

$$y = l_y(x) = \frac{\sigma(y)}{\sigma(x)}x + \left[\mu(y) - \frac{\sigma(y)}{\sigma(x)}\mu(x) \right] \quad (20)$$

(20) nolu eşitliğinde verilen lineer eşitleme yöntemi denklemini elde edilir. $\frac{\sigma(y)}{\sigma(x)}$ denklemin eğimi, $\mu(y) - \frac{\sigma(y)}{\sigma(x)}\mu(x)$ denklemin sabit terimidir.

(20) nolu denklem de formların standart sapmaları bir birine eşit olursa $\frac{\sigma(y)}{\sigma(x)} = 1$ olur ve lineer eşitleme yöntemi fonksiyonu $y = l_y(x) = x + [\mu(y) - \mu(x)]$ ortalama eşitleme yönteminde kullanılan fonksiyonu ifade eder.

Eşitleme yapılan test formlarından eşitlenecek test formu referans test formundan zor olduğunda ya da bazen test formlarındaki en yüksek ve en düşük test puanları referans test formunun puan ranjı dışına çıkabilmektedir. Bu hatalı ya da yanlış bir durum değildir. Bu lineer eşitleme sürecinde olabilen beklenen bir durumdur.

Lineer eşitleme grubun yetenek düzeyinden oldukça etkilenmektedir. Zorluk yönünden birbirinden farklı olan test formlarıyla yapılan eşitlemede test formlarının yetenek düzeyi yüksek grupta yapılan eşitleme ile yetenek düzeyi düşük grupta yapılan eşitleme bir birinden farklı olmaktadır. Örneğin, zorluk düzeyleri farklı olan iki test formu yetenek düzeyi yüksek ve yetenek düzeyi düşük olan iki gruba da uygulansın. Bu iki testin her gruptaki eşitlemesi farklı olur çünkü yetenek düzeyi yüksek olan gruba uygulanan kolay formun puan dağılımı yüksek puanlarda olmanın yanı sıra ranjı dar olacaktır. Aynı zaman da zor test formundaki puan dağılımındaki ranjı daha büyük olacaktır. Düşük yetenek gruptaki bireyler için ise zor test formundan elde edilen bilgiler hem düşük puanlarda yığılmaya neden olacak ve ranj dar olacakken kolay test formundaki puan dağılımı ranjı ise daha büyük olacaktır. Dolayısıyla yapılacak olan eşitleme sonrasında elde edilecek eşitleme sonuçları birbirinden farklı olacaktır.

3.3. Eşit Yüzdellikli Eşitleme

Eşit yüzdellikli eşitleme Braun ve Holland tarafından 1982 yılında tanımlanmıştır. X ve Y test formları olmak üzere, X formu Y formuna eşitlendiğinde eşitlemenin dağılımı Y formunun dağılımıyla aynıdır. Eşit yüzdellikli eşitleme birim dönüşüm üzerine geliştirilmiştir. Dolayısıyla X formundaki yüzdellik sıra ile Y formundaki yüzdellik sıra aynıdır.

F; X formunu alanların yığılma fonksiyonu

G; Y formunu alanların yığılma fonksiyonları

e_Y ; X formunun, Y formuna eşitleyen simetrik eşitleme fonksiyonu

G*; e_Y eşitlemesiyle elde edilmiş popülasyonun yığılmalı dağılım fonksiyonu olmak üzere;

“G*=G ise e_Y eşitleme fonksiyonudur.” şeklinde tanımlanmıştır (Braun ve Holland, 1982). Dolayısıyla;

$$e_Y(x) = G^{-1}(F(x)) \quad (21)$$

(21) nolu eşitliğiyle dönüşüm değerleri elde edilir.

Eğer formdaki puanlar ayrık ise, yani X formundaki yüzdellik sırasına karşılık Y formundaki puanların dağılımından elde edilen yüzdellik sırası karşılık gelmiyorsa puanları sürekli olarak düşünüp, örneğin 28 puan (27.5, 28.5) aralığında dağılıma sahiptir, Y formundaki yüzdellik sırası bulunur.

Lineer eşitleme ve eşit yüzdellikli eşitleme sonuçları test formlarındaki dağılımlar aynı olduğunda aynı sonuçları verirler. Dağılımların aynı olması yüzdellik dilimlerinin aynı olması anlamına gelir. Test formlarının dağılımları farklı olduğunda yapılan lineer eşitleme sağlıklı sonuçlar vermez.

3.4. Dairesel-Yay Eşitleme

Lineer eşitleme yönteminin yetenek dağılımı aynı olan durumlarda kullanılması gereklidir. Oysa bu çoğunlukla test formları arasındaki zorluk farkı ya da grup farklılığından dolayı karşılanması mümkün olmayan bir hal alır. Örneğin yetenek

düzeyi yüksek olan bir gruba kolay bir form verildiğinde yüksek puandaki yoğunluk daha çok olacakken düşük puandaki yoğunluk az olacaktır. Bu durumlarda gerçek eşitleme lineer olmadığından lineer eşitleme etkili olmayacaktır. Dairesel-yay eşitlemede gerçek eşitleme lineer olmadığından üç puan baz alınır. En düşük anlamlı puan, testten alınabilecek en yüksek puan ve bu puanların birleşiminden oluşan orta puan değeri mevcuttur. Eşitleme bu puanlarla gerçekleştirilir. Dairesel-yay eşitleme varsayımları daha az ve karşılanabilir olduğundan uygulaması daha rahattır. Ayrıca ortalama eşitleme yöntemine göre daha doğru ve daha az yanlılık içerir (Babcock, Albano ve Raymond, 2012).

Divigi tarafından önerilen bu yöntem eşitleme eğrisinin test formunun sınır puanlarından geçecek şekilde düzenler. Divigi'nin yönteminde test formundan alınabilecek maksimum değer ile minimum puan değerleri olan sınır puanlar kullanılarak eşitleme gerçekleştirilmektedir. Orta puan ortalama puan yardımıyla hesaplanmaktadır.

Eğrinin üst sınırı maksimum puan olarak alınabilecek puan olmak üzere alt sınırı ise test formundan alınabilecek anlamlı en düşük puandır (Çoktan seçmeli maddelerden oluşan testlerde şans puanı en düşük anlamlı puan olmaktadır). Orta puan ise verilerden hesaplanmaktadır (puan dağılımının orta noktasında eşitlenerek elde edilir). Eğer bu üç nokta aynı doğru üzerinde olduğunda bu doğru, eşitleme eğrisi olarak adlandırılabilir. Eğer üç nokta aynı doğru üzerinde yer almıyorlarsa daire yayı olarak bir eğri hesaplanır. Hesaplanan bu yay eşitleme eğrisi olarak kullanılır (Livingston ve Kim, 2009). Bu yöntem alternatif üstelik yüzdeler eşitlemeye yakın sonuçlar veren bir yöntem daha mevcuttur. Alternatif yöntemde eşitleme fonksiyonu lineer ve lineer olmayan iki bileşene ayrılmaktadır (Livingston ve Kim, 2010). Lineer olmayan bileşen dairesel yay eğrisinin kendisidir.

Eşitlenecek test formlarından alınacak anlamlı en az puan alt sınır (x_1, y_1) , en çok üst puan üst sınırı (x_3, y_3) belirtmek üzere bu noktalardan geçen doğru denklemi olan $L(x)$ eşitlik (22) de belirtildiği gibi hesaplanır (Livingston ve Kim, 2009).

$$L(x) = y_1 + \frac{y_3 - y_1}{x_3 - x_1}(x - x_1) \quad (22)$$

$$y^* = y - L(x) \quad (23)$$

Eğer orta nokta (puan dağılımının orta noktası) (22) nolu eşitlikteki doğru üzerindeyse kestirilen eşitleme doğrusu bu doğrudur. Eğer orta nokta bu doğru üzerinde değilse bu doğrunun x_2 değeri için y değeri bulunur $y_2 - L(x_2) = y_2^*$ değeri dönüştürülmüş puanın yüksekliğini verir. Elde edilen bu değer test formları arasındaki farkın yorumlanmasında kullanılır. Eğer yeni form eski formdan daha zorsa orta nokta (x_2, y_2) , $L(x)$ doğrusunun üzerinde kalır ve y_2^* değeri pozitif değerini alır. Eğer yeni test eski testten kolay ise orta nokta $L(x)$ doğrusunun altında kalır ve y_2^* değeri negatif olur.

Dönüştürülmüş üç nokta dairesel yayın hesaplanmasında kullanılır. r dairenin yarı çapını göstermek üzere, (x_c, y_c) dairenin merkez noktasıdır. Eşitlik (24) ve (25) kullanılarak dairenin merkez koordinatları elde edilirken eşitlik (26) da dairenin yarı çapı bulunur (Livingston ve Kim, 2009). Eşitliklerden yararlanarak alternatif yöntemin lineer olmayan bileşeni eşitlik (27)'de ifade edilmiştir (Livingston ve Kim, 2009). Eğrisel bileşen formülünde, y_2^* değeri pozitif ise “+”, negative ise “-“ değerini alır.

$$x_c = \frac{x_3^2 - x_1^2}{2(x_3 - x_1)}$$

(24)

$$y_c = \frac{x_1^2(x_3 - x_2) - (x_2^2 + (y_2^*)^2)(x_3 - x_1) + x_3^2(x_2 - x_1)}{2y_2^*(x_1 - x_3)}$$

(25)

$$r^2 = (x - x_c)^2 + (y - y_c)^2$$

(26)

$$y^* = y_c \pm \sqrt{r^2 - (x - x_c)^2}$$

(27)

$$e_{circle-Y}(x) = y_c \pm \sqrt{r^2 - (x - x_c)^2} + y_1 + \frac{y_3 - y_1}{x_3 - x_1}(x - x_1)$$

(28)

Eşitleme deseni tek grup deseni, karşıt dengelenmiş desen ve denk grup deseni ise yeni ve eski test formlarının ortalamaları orta noktadır. Ortak madde deseninde ise zincirlenmiş lineer eşitleme formülünden yararlanarak orta nokta bulununur.

$$y = m_{YB} + \frac{S_{YB}}{S_{VB}}(m_{VA} - m_{VB}) + \frac{S_{YB}S_{VA}}{S_{VB}S_{XA}}(x - m_{XA})$$

(29)

A ve B sırasıyla X ve Y formunu alan grupları göstermek üzere V ankor testini temsil etmektedir. m ve s sırasıyla ortalama ve standart sapmayı belirtmektedir. (29) nolu denklem yardımıyla denk olmayan gruplarda orta noktanın y_2 değeri hesaplanır (Livingston ve Kim, 2009). Ayrıca denk olmayan gruplarda test formların uygulandığı örnekleme sayısı eşit olabilir. Bu durumda yeterince büyük örneklem büyüklüğüne sahip grubun orta noktası kullanılarak diğer grubun orta noktası elde edilir.

3.5. Eşit Olmayan Gruplarda Lineer Yöntemler

Örneklem hatası vermeyecek kadar büyük olan farklı yetenek düzeyindeki gruplara farklı zorlukta testler uygulandığını düşünelim. Böyle bir durumda iki faktör ortaya çıkmaktadır. Biri test formlarının zorluk farkıyla oluşmaktadır, diğeri ise yeteneklerin farklılığıyla ortaya çıkmaktadır. Test formlarının zorluk farkının olmasından dolayı testlerden elde edilen puanlar eşitlenerek anlamlı hale getirilip karşılaştırılmaktadır. Gruplar arasındaki yetenek farklılıkları için ise eşitleme sürecine geçmeden önce gerekli düzenleme yapılır. Örneğin iki test formu tüm yetenek düzeyine uygulanarak eşitleme yapılır. Böylece farklı yetenek düzeylerine her iki form uygulanarak yetenek düzeyleri ayrı popülasyon alınır ve bu popülasyonlar da ayrı ayrı eşitleme yapılır. Başka bir yöntem ise test formlarının eşitlenmesi için ankor test kullanılır ve ankor test kullanımıyla yapılacak eşitlemede ortak bir popülasyon oluşturulur.

Eşit olmayan grup desenine göre test formları her iki popülasyona da uygulanması gerekir. Uygulama sonra yapılacak eşitlemede ki farklı popülasyondan elde edilecek veriler tek bir popülasyondan alınmış gibi işlenir. Bu tek bir

popülasyondan alınmış gibi görünen popülasyona, sentetik popülasyon adı verilir. Sentetik popülasyon popülasyon 1'den ω_1 ağırlığında, popülasyon 2'den ω_2 ağırlığında popülasyonlar çekilerek sentetik popülasyon

$$\omega_1 + \omega_2 = 1 \text{ ve } \omega_2, \omega_1 \geq 0$$

şeklinde eşitliği sağlayacak popülasyondan oluşur (Braun ve Holland, 1982).

Sentetik popülasyonda lineer eşitleme fonksiyonu, X ve Y formunun aritmetik ortalaması aşağıdaki eşitliklerde sunulmuştur.

$$l_{Y_s}(x) = \frac{\sigma_s(y)}{\sigma_s(x)}(x - \mu_s(x)) + \mu_s(y) \quad (30)$$

$$\mu_s(x) = \omega_1\mu_1(x) + \omega_2\mu_2(x) \quad (31)$$

$$\mu_s(y) = \omega_1\mu_1(y) + \omega_2\mu_2(y) \quad (32)$$

$$\sigma_s(x) = \omega_1\sigma_1^2(x) + \omega_2\sigma_2^2(x) + \omega_1\omega_2(\mu_1(x) - \mu_2(x))^2 \quad (33)$$

$$\sigma_s(y) = \omega_1\sigma_1^2(y) + \omega_2\sigma_2^2(y) + \omega_1\omega_2(\mu_1(y) - \mu_2(y))^2 \quad (34)$$

Eşit olmayan gruplarda lineer eşitleme yöntemlerinden gözlenen puanlara dayalı üç eşitleme yöntemi, gerçek puanlara dayalı ise bir eşileme yöntemi vardır.

$\mu_2(x), \sigma_2^2(x), \mu_1(y), \sigma_2^2(y)$ değerleri doğrudan elde edilmektedir. Tucker ve Levine gözlenen puan yöntemleriyle doğrudan elde edilemeyen $\mu_2(x), \sigma_2^2(x), \mu_1(y), \sigma_2^2(y)$ parametrelerini sayılıtlar ile kestirmeyi sağlamışlardır.

Form X		Form X
Popülasyon 1	\mapsto	Popülasyon 2
$\mu_1(x)$ ve $\sigma_1^2(x)$		$\mu_2(x)$ ve $\sigma_2^2(x)$
Form Y		Form Y
Popülasyon 2	\mapsto	Popülasyon 1
$\mu_2(y)$ ve $\sigma_2^2(y)$		$\mu_1(y)$ ve $\sigma_1^2(y)$

3.6. Tucker Yöntemi

Gulliksen tarafından temelleri atılmıştır (Kolen ve Brennan, 2014). Parametrelerin kestirilmesi için iki sayılıtlı vardır.

1. Toplam puanların ortak madde puanları üzerindeki regresyonu
2. Şartlı puan varyansı

3.6.1. Lineer Regresyon Sayılısı

α regresyon doğrusunun eğimini ve β regresyon doğrusunun sabitini göstermek üzere;

$$\alpha_1(x|v) = \frac{\sigma_1(x, v)}{\sigma_1^2(v)} \quad \beta_1(x|v) = \mu_1(x) - \alpha_1(x|v)\mu_1(v)$$

$$\alpha_1(y|v) = \frac{\sigma_2(y, v)}{\sigma_2^2(v)} \quad \beta_2(y|v) = \mu_2(y) - \alpha_2(y|v)\mu_2(v)$$

X puanlarının V puanları üzerindeki regresyonu popülasyon 1 ve popülasyon 2 için aynı lineer fonksiyondur.

$$\alpha_1(x|v) = \alpha_2(x|v) \quad \beta_1(x|v) = \beta_2(x|v)$$

Y puanlarının V puanları üzerindeki regresyonu popülasyon 1 ve popülasyon 2 için aynı lineer fonksiyondur.

$$\alpha_1(y|v) = \alpha_2(y|v) \quad \beta_1(y|v) = \beta_2(y|v)$$

3.6.2. Şartlı Varyans Sayılısı

X puanlarının V puanları üzerindeki regresyonu popülasyon 1 ve popülasyon 2 için varyanslarının olduğunu varsayar.

$$\sigma_2^2(x)(1 - \rho_2^2(x, v)) = \sigma_1^2(x)(1 - \rho_1^2(x, v))$$

Y puanlarının V puanları üzerindeki regresyonu popülasyon 1 ve popülasyon 2 için varyanslarının olduğunu varsayar.

$$\sigma_1^2(y)(1 - \rho_1^2(y, v)) = \sigma_2^2(y)(1 - \rho_2^2(y, v))$$

Lineer regresyon sayılısı ve şartlı varyans sayılısı kullanılarak $\mu_2(x)$, $\mu_1(y)$, $\sigma_2^2(x)$ ve $\sigma_1^2(y)$ parametreleri kestirilir. Elde edilen parametreler yardımıyla sentetik popülasyonun ortalaması ve standar sapması hesaplanarak,

$$l_{Y_s}(x) = \frac{\sigma_s(y)}{\sigma_s(x)}(x - \mu_s(x)) + \mu_s(y) \quad (35)$$

lineer eşitleme fonksiyonu elde edilir. Ankor testin ortalama ve standart sapmaları her iki popülasyon için eşit olduğunda sentetik popülasyon parametreleri gözlenen ortalama ve varyansa eşit olur.

3.7. Nominal Ağırlıklı Ortalama Eşitleme Yöntemi

Tucker yönteminin daha basitleştirilmiş halidir. Tucker eşitleme fonksiyonunda yer alan test varyansları yerine test formunda yer alan madde sayılarıyla elde edilen ağırlıklı oranlara yer verilir. Küçük örneklem için uygun bir eşitleme yöntemidir. Çünkü küçük örneklemde varyans kestirimleri sağlıklı olmamaktadır. K test maddesi sayısı olmak üzere $N(X)$, $N(Y)$ ise X ve Y testlerini alan birey sayısını göstermektedir. (36) numaralı eşitlik nominal ağırlıklı eşitleme fonksiyonunu göstermektedir (Albano ve Bablock, 2012)

$$l_Y(x) = x - \mu_2(X) + \mu_1(Y) + \left[\frac{N(X)K(Y) + N(Y)K(X)}{N(Y) + N(X)K(V)} \right] [\mu_2(V_X) + \mu_1(V_Y)] \quad (36)$$

3.8. Levine Gözlenen Puan Yöntemi

Tucker yöntemi gözlenebilir verileri referans olarak yürütülen bir yöntemdir. Gerçek puanlara dayalı bir istatistiksel süreci içinde doğrudan barındırmamaktadır. Eşitleme işleminde gerçek puanları baz alan çalışmalar Levine (1995) tarafından geliştirilmiştir. Levine gözlenen puan yöntemiyle X formunda gözlenen bir puanın Y formundaki gözlenen ölçek puanıyla ilişkilendirmiştir.

Parametreleri yöntemin korelasyon sayıltısı, lineer regresyon sayıltısı ve hata varyansı sayıltılarını gerçekleştirmesiyle sentetik popülasyon parametrelerine ulaşılır.

3.8.1. Korelasyon Sayıltısı

X, Y, V testlerinin aynı davranışı ölçtüğü için T_x ve T_v popülasyon 1 ve popülasyon 2' de mükemmel korelasyon gösterir. Aynı şekilde T_y ve T_v mükemmel korelasyon gösterir. Dolayısıyla matematiksel ifadesi

$$\begin{aligned} \rho_1(T_x, T_v) &= \rho_2(T_x, T_v) \\ \rho_1(T_y, T_v) &= \rho_2(T_y, T_v) \end{aligned}$$

şeklindedir.

3.8.2. Lineer Regresyon Sayıltısı

T_x değişkeninin T_v değişkeni üzerindeki regresyonu popülasyon 1 ve popülasyon 2 için aynı doğruyla ifade edilir. Benzer şekilde, T_y değişkeninin T_v değişkeni üzerindeki regresyonu popülasyon 1 ve popülasyon 2 için aynı doğruyla ifade edilir.

T_x ' in T_v üzerindeki regresyon doğrusunun eğimi;

$$\alpha_1(T_x|T_v) = \frac{\sigma_1(T_x, T_v)}{\sigma_1^2(T_v)} = \frac{\rho_1(T_x, T_v)\sigma_1(T_x)\sigma_1(T_v)}{\sigma_1^2(T_v)} = \frac{\rho_1(T_x, T_v)\sigma_1(T_x)}{\sigma_1(T_v)} \quad (37)$$

$$\alpha_2(T_y|T_v) = \frac{\rho_2(T_y, T_v)\sigma_2(T_y)}{\sigma_2(T_v)} \quad (38)$$

Regresyon doğrusunun sabiti;

$$\mu_2(x) - \alpha_2(T_x|T_v)\mu_2(v) = \mu_1(x) - \alpha_1(T_x|T_v)\mu_1(v) \quad (39)$$

$$\mu_1(y) - \alpha_1(T_y|T_v)\mu_1(v) = \mu_2(y) - \alpha_2(T_y|T_v)\mu_2(v)$$

(40)

(39) ve (40) nolu denklemlerdir. Popülasyon 1 için Y testinin standart sapma ve aritmetik ortalama parametreleri regresyon doğrularının eşitliği sayılısıyla kestirilir. Benzer şekilde popülasyon 2 için X testinin standart sapma ve ortalama parametreleri kestirilir.

3.8.3. Hata Varyansı Sayılısı

Levine yöntemine göre X formunun ölçme hatası varyansı popülasyon 1 ve popülasyon 2 için aynıdır. Benzer şekilde Y formunun ölçme hatası her iki popülasyon için aynıdır. Matematiksel olarak ifade edilecek olursa;

$$\sigma_1^2(x) - \sigma_1^2(T_x) = \sigma_2^2(x) - \sigma_2^2(T_x) \quad (41)$$

$$\sigma_1^2(y) - \sigma_1^2(T_y) = \sigma_2^2(y) - \sigma_2^2(T_y) \quad (42)$$

$$\sigma_1^2(v) - \sigma_1^2(T_v) = \sigma_2^2(v) - \sigma_2^2(T_v) \quad (43)$$

Levine gözlenen puanlar yöntemi gerçek puanları baz aldığından hipotetiktir. Hipotetik haliyle kullanılamaz. Klasik Test Teorisi sayılılarından yola çıkarak regresyon doğrularının eğimi;

$$\gamma_1 = \frac{\sigma_1(x)\sqrt{\rho(x, x')}}{\sigma_1(v)\sqrt{\rho(v, v')}} \quad (44)$$

$$\gamma_2 = \frac{\sigma_2(y)\sqrt{\rho(y, y')}}{\sigma_2(v)\sqrt{\rho(v, v')}} \quad (45)$$

şeklinde ifade edilir.

3.9. Levine Gerçek Puan Yöntemi

Levine (1955) gözlenen puanların yanı sıra gerçek puanlar yöntemine dayalı eşitleme çalışmaları yapmıştır. Gerçek puan yöntemi sayılıları gözlenen puan yöntemi sayılılarıyla aynıdır. Levine gerçek puan yönteminin Levine gözlenen yöntemden farkı gerçek puanlara dayalı olarak ifade edilmesidir.

$$l_{y_s}(t_x) = \frac{\sigma_s(T_y)}{\sigma_s(T_x)}(t_x - \mu_s(T_x) + \mu_s(T_y)) \quad (46)$$

Gerçek puan yöntemi popülasyon ağırlığından etkilenmez. Diğer bir ifadeyle eşitleme popülasyon ağırlığından bağımsızdır.

3.10. Zincirlenmiş Linear Eşitleme

Angoff (1971), Holland ve Dorans (2006) yılında üzerinde çalıştığı bir yöntemdir. Bu yöntemde X formu V formuna, V formu Y formuna eşitlenmiş olur. Dolayısıyla X formundaki herhangi bir puan V formu aracılığıyla Y formuna eşitlenmiş olur. Matematiksel olarak ifade edilecek olunursa;

$$l_y(x) = l_y(l_v(x)) \quad (47)$$

şeklinindedir. Yöntemde uzun içerikli ya da madde sayısı ankor teste göre fazla olan bir testten ankor testine ilişkilendirme yapılmaktadır. Ankor testinden madde sayısı fazla olan testte eşitleme yapılmaktadır. Bu durum içerik problemiyle karşılaşılmasına neden olabilmektedir.

Diğer problem ise X formu yalnız popülasyon 1 tarafından alınırken Y formu da yalnız popülasyon 2 tarafından alınmaktadır. Holland ve Dorans (2006) çalışmalarında iki popülasyonunda ağırlıklarından bağımsız olarak çalışmışlardır.

3.11. Zincirlenmiş Linear Gözlenen Puan Eşitlemesi

Brennan (2006) γ – terimiyle yeniden eşitleme fonksiyonunu ifade etmiştir. $\gamma_1 = \frac{\sigma_1(x)}{\sigma_1(v)}$ ve $\gamma_2 = \frac{\sigma_2(y)}{\sigma_2(v)}$

γ – terimleri olmak üzere;

$$l_y(x) = \frac{\gamma_2}{\gamma_1} x + \mu_2(y) + \gamma_2(\mu_1(v) - \mu_2(v)) - \frac{\gamma_2}{\gamma_1} \mu_1(x) \quad (48)$$

(48) nolu denklem ile ifade edilir.

3.12. Zincirlenmiş Linear Gerçek Puan Eşitlemesi

$\gamma_1 = \frac{\sigma_1(T_x)}{\sigma_1(T_v)}$ ve $\gamma_2 = \frac{\sigma_2(T_y)}{\sigma_2(T_v)}$ γ – terimleri olmak üzere;

$$l_y(x) = \frac{\gamma_2}{\gamma_1} x + \mu_2(y) + \gamma_2(\mu_1(v) - \mu_2(v)) - \frac{\gamma_2}{\gamma_1} \mu_1(x) \quad (49)$$

(49) nolu denkleminle ifade edilir.

3.13. Eşit Olmayan Gruplarda Eşit Yüzdellik Yöntemi

Yöntem gözlenen puanları veri olarak kullanır. Toplam puan dağılımı ve ortak maddelerden elde edilen puan dağılımı üzerinden işlemler yapılır. Frekans kestim yöntemi, geliştirilmiş kestirim yöntem, zincirlenmiş eşit yüzdellikli eşitleme yöntemi olmak üzere üç farklı tekniği vardır.

3.14. Frekans Kestirim Yöntemi

Angoff (1971) tarafından tanımlanmış, Braun ve Holland (1982) tarafından ise ortak maddelere dayalı eşit olmayan grup deseni kullanılarak sentetik popülasyondan elde edilen veriler ile X ve Y formlarındaki puanların yığılma dağılımını tahmin etmede kullanılır.

Şartlı dağılım; X ve V formlarında sırasıyla x ve v puanlarını alanların dağılımı, $f(x.v) = f(x|v)h(v)$ şeklinde ifade edilir. $h(v)$; dağılımda v puanı alanların oranını gösterir. $f(x|v)$; X formundaki şartlı puan dağılımını ifade eder.

3.14.1. Sayıtlar

X formundan elde edilen bir x puanının dağılımdaki frekansı ya da oranı $f(x)$ 'in bulunması oldukça kolaydır. Ancak eşit olmayan gruplarda iki farklı popülasyondan elde edilen veriler $f(x)$ fonksiyonunun sentetik fonksiyon için yeniden tanımlanmasını gerektirmiştir. Sentetik popülasyon için $f(x)$ fonksiyonu;

$$f(x) = f_s(x) = w_1 f_1(x) + w_2 f_2(x) \quad (50)$$

$$g(y) = g_s(y) = w_1 g_1(y) + w_2 g_2(y) \quad (51)$$

(50) ve (51) nolu eşitliklerdeki $f_1(x)$, $g_1(y)$ popülasyon 1 deki değerleri gösterir. $f_2(x)$, $g_2(y)$ ise popülasyon2' deki değerleri gösterir. Ancak eşit olmayan grup deseninde popülasyonlardaki bireylere yalnızca bir tek form uygulanmaktadır. Dolayısıyla X formundan elde edilen bir x puanı için $f_1(x)$ doğrudan gözlenebiliyorken, $f_2(x)$ doğrudan gözlenememektedir. Benzer şekilde Y formundan elde edilen bir y puanı için $g_2(y)$ doğrudan gözlenebiliyorken $g_1(y)$ doğrudan gözlenememektedir. Şartlı dağılım fonksiyonu kullanarak X ve Y formu için $f_1(x)$, $g_1(y)$ değerleri hesaplanır.

$$f_2(x) = \sum_v f_2(x, v) = \sum_v f_1(x|v)h_2(v) \quad (52)$$

$$g_1(y) = \sum_v g_1(y, v) = \sum_v g_2(y|v)h_1(v) \quad (53)$$

$F_s(x)$; sentetik popülasyonda x 'in yığılmalı frekansını $G_s(y)$; ise y 'nin yığılmalı frekansını verir. X formunda x puanı eşit yüzdellikli eşitleme ile Y formunda $e_y(x)$ puanına eşlenir. Bu eşleme $e_y(x) = Q_s^{-1}(P_s(x))$ denklemiyle ifade edilir. Denklemden ki $P_s(x)$; X formunun sıralı yüzdellik fonksiyonudur. Q_s ise Y formunun sıralı sıralı yüzdellik fonksiyonudur.

Yöntem popülasyonların bir birlerine benzerlikleri çok yakın ise tercih edilmelidir. Popülasyonlar bir birlerinin aynısı ise $f_1(x|v) = f_2(x|v)$ ve $g_1(y|v) = g_2(y|v)$ gözlemlenebilir.

Popülasyonlar bir birinden farklı ise frekans kestirim yöntemi gerçek puanları baz alarak yürütülebilir. Gerçek puanlar doğrultusunda uygulanan yöntem ise geliştirilmiş frekans kestirim yöntemi adı verilir.

3.15. Geliştirilmiş Frekans Kestirim Yöntemi

Frekans kestirim yöntemi sayılı olan şartlı dağılım frekanslarında X ve Y formu için şartlı frekanslar iki popülasyon içinde eşittir. Geliştirilmiş frekans yönteminde ise ortak maddelerin gerçek puanı tek ve değişmezdir. Dolayısıyla şartlı dağılım gerçek puanlar yardımıyla değişme göstermeksizin hesaplanır.

3.16. Zincirlenmiş Eşit Yüzdelikli Eşitleme

X formundaki puanlar ortak maddelerin oluşturduğu ortak testteki puanlara eşit eşitleme yöntemiyle eşitlenir. Ortak test formu Y formuna eşit yüzdelikli eşitle yöntemiyle eşitlenerek X formundaki bir x puanı zincirlenmiş eşit yüzdelikli eşitleme yöntemiye Y formuna eşitlenir. Yöntemin matematiksel ifadesi (54) nolu denklemdeki gibidir.

$$e_{y(Chain)}(x) = e_{y_2}(e_{y_1}(x)) \quad (54)$$

3.17. Sentetik Fonksiyon Eşitleme

Hemen hemen paralel test formlarının elde edilmesi pratikte oldukça zor bir durumdur. Küçük örneklemelerde iki test içerik ve zorluk bakımından hemen hemen paralel olduğu durumlarda ve testlerin oluşumu titizlikle gerçekleştirildiklerinde herhangi bir dönüşüm kullanılmaması ya da birim dönüşümün kullanımı tercih edilmektedir. İki form bir birine oldukça denk olduğunda birim eşitlemenin kullanımı eşitlemede daha az hata vermektedir. Fakat örneklem büyüklüğü 50'den daha az olduğunda ve formların zorlukları farklı olduğunda da birim eşitleme doğru bilgi verir edilmektedir (Babcock, Albano ve Raymond, 2012). Sentetik fonksiyon ile yapılan eşitlemede örneklem büyüklüğü 100 ve daha büyük olduğunda daha az yanlışlık hata değeri ve daha az eşitleme hatası elde edilmektedir (Babcock, Albano ve Raymond, 2012). Bununla birlikte eşitlemenin sistematik hatası da diğer eşitleme yöntemlerine göre daha fazla değere sahiptir.

Küçük örneklemelere sahip sentetik popülasyonlarda hem seçkisiz hem de sistematik hatanın daha az olduğu bir eşitleme Kim, von Davier ve Haberman (2008) tarafından geliştirilmiştir. Sentetik eşitleme fonksiyonu birim dönüşüm ile başka bir dönüşümün belirli ağırlıklar doğrultusundaki eşitliklerle olan lineer kombinasyon olarak ifade edilen eşitleme fonksiyonudur.

$$Synthetic_Y(x) = w \cdot e_Y(x) + (1 - w) \cdot ID_Y(x) \quad (55)$$

$$\mu(Synthetic_Y(x)) = w \cdot Mean(e_Y(x)) + (1 - w) \cdot Mean(ID_Y(x)) \quad (56)$$

$$Var(Synthetic_Y(x)) = w^2 \cdot e_Y(x) \quad (57)$$

$$SEE(Synthetic_Y(x)) = w \cdot SEE(e_Y(x)) \quad (58)$$

(55) nolu eşitlikte sentetik fonksiyonun matematiksel ifadesi verilmiştir. Denklemden $e_Y(x)$, Tucker ve Levine eşitleme fonksiyonları, $ID_Y(x)$ birim dönüşüm fonksiyonu ve W populasyonun ağırlığını ifade etmektedir (Kim, von Davier ve Haberman, 2008). (56), (57) ve (58) nolu denklemler de sırasıyla sentetik populasyonun ortalamasının, varyansını ve eşitlemenin standart hatasını göstermektedir (Kim, von Davier ve Haberman, 2008).

Test puanlarının rapor edilebilmesi için test formlarının güvenilirliği, geçerliği ve karşılaştırılabilirliği kontrol edilmesi gerekmektedir. Test formlarının karşılaştırılabilir olması için elde edilen puanların aynı ölçekleme cetvelinde konumlandırılmalıdır. Bu durumun gerçekleştirilebilmesi için testlerden elde edilen bir birine eşitleme fonksiyonuyla dönüştürülerek aynı doğrultu üzerine yerleştirilmektedir. Peki yapılan eşitlemenin doğruluğundan nasıl emin olunabilir? Eşitlemenin doğruluğunu test etmek için eşitlemede yapılan hata değeri kullanılır. Aşağıda eşitlemenin ne kadar hatayla yapıldığını hesaplamak için eşitleme hatası hakkında bilgi verilecektir.

3.18. Braun-Holland Doğrusal Eşitleme Yöntemi

Frekans tahmin yöntemindeki varsayımları, doğrusal eşitleme yöntemine adapte edilmiştir (Albano, 2016). Bu eşitleme işlemi yapılırken ortalama ve standart sapma değerlerinden yararlanılmaktadır. Ayrıca Braun ve Holland eşitleme işlemini yapabilmek için Tucker doğrusal eşitleme yönteminde bulunan varsayımları kabul etmişlerdir (Kolen ve Brennan, 2004). Braun-Holland eşitleme yöntemi, Tucker doğrusal eşitleme yönteminin genelleştirilmiş bir halidir. X ve Y formlarından elde edilen puanlar ile ortak madde formundan elde edilen puanlar arasındaki grafiğin doğrusal olduğu varsayımının kabul edilmediği durumlarda da kullanılabilir. Ancak Braun-Holland eşitleme yöntemi Tucker doğrusal eşitleme yönteminden daha karmaşık hesaplama sistemine sahiptir. Bu yüzden pratikte daha az kullanılan bir yöntemdir (Kolen ve Brennan, 2004).

4. EŞİTLEME HATASI

Farklı test formlarını alan bireyler arasında yanlılığı önlemek, farklı formlardan alınan puanları aynı ölçek üzerinde rapor etmek ve rapor edilen puanların anlamlarının korunması (Akt: Kan, 2010:3, Barnard, 1996). eşitlemenin amaçları arasındadır. Eşitlemenin amaçlarına ulaşılması eşitlemenin hatasız yapıldığı anlamına gelir. Eşitleme hatası, eşitleme desenlerinin doğru kullanılmaması, uygun eşitleme yönteminin uygulanmaması, yeterince büyük örneklemin olmaması, farklı test formlarının bir istatistiksel açıdan birbirine benzememesi durumlarında ortaya çıkar. Kolen ve Brennan (2004), eşitleme hatasını seçkisiz ve sistematik eşitleme hatası olarak ikiye ayırmıştır. Seçkisiz hata, örneklemden ve evrenden kestirilen eşitleme ilişkisi arasındaki fark olarak da adlandırılabilir. Kolen (1988) örneklem büyüklüğü artırılarak ve uygun eşitleme deseni seçilerek seçkisiz örneklem hatasının azaltılabileceğini belirtmiştir. Örneklem yerine tüm evrenle çalışıldığında bu hata ortadan kalkacaktır (Kolen ve Brennan, 2004). Sistematik hata, kullanılan eşitleme yönteminin varsayımları veya koşulları karşılanmadığı durumlarda oluşur. Örneğin, tek grup desenlerde yorgunluk etkisi yüzünden başarısız olmak veya pratik etkisi yüzünden yüksek puan almak birer sistematik hatadır. Sistematik hatalar özellikle eşdeğer olmayan gruplar deseninde (NEAT) daha fazla sorun oluşturmaktadır. Çünkü lineer eşitleme yöntemleri olan Tucker ve Levine eşitleme yöntemlerinin sayıltılarını karşılamak zor olduğu için sistematik hata yapma olasılığı yüksektir. Bu koşullar karşılanmadan yapılan eşitlemeler sonucunda sistematik hata oluşmaktadır. Örneğin ankor maddelerin tüm test formunun içeriğiyle aynı olması veya istatistiklerin benzer özellikte olması gerçekleştirmesi zor bir sayıltıdır. Ayrıca NEAT desenlerde, ortak maddenin X formunda farklı Y formunda farklı davranması da sistematik hataya yol açacaktır. Sonuç olarak, herhangi bir

eşitleme deseninde aynı testin X ve Y formları arasında içerik, güçlük ve güvenilirlik farkı varsa sistematik hatanın olduğu söylenebilmektedir (Kolen,1988).

Livingston (2004), eşitlemenin standart hatasını (SEE), eşitlenen puanların örneklem dağılımlarının standart sapması şeklinde tanımlamıştır. Wang (2006) eşitlemenin standart hatasını, seçkisiz örneklem hatasının indeksi olduğunu belirtmiş ve tekrarlı örneklemelerden elde edilen eşitlenmiş puanların standart sapması olarak tanımlamıştır. Her tekrarlı yapılan eşitleme işleminde, örneklem evren veya evrenlerden seçkisiz olarak seçilir. Daha sonra, X formundan elde edilen ham puanlarının Y formunda eşitleme yöntemleri kullanılarak farklı düzeydeki eşdeğerlikleri bulunur. Her puan düzeyinde eşitlemenin standart hatası, tekrarlı eşitlemelerin standart sapmasına eşittir.

4.1. Önyükleme (Bootstrap) Yöntemi

Efron ve Tibshirani (1993) tarafından ortaya atılan bir yöntemdir (Cui veKolen, 2008). Bu yöntemde evrenden bir örneklem seçkisiz olarak seçilir. Seçilen örneklem verilerinin yerlerini değiştirilerek, seçkisiz ve farklı birçok örneklem (bootstrap) çekilir ve her örneklem için eşitleme fonksiyonu tahmin edilir ve ilgili diğer istatistikler hesaplanır. Hesaplanan değerlerin standart sapmaları eşitlemenin standart hatasını verir.

Aşağıdaki matematiksel ifadelerde sırasıyla eşitlemenin standart hatası, eşitleme yanlılığı ve eşitleme hatasını göstermektedir. i , puan sayısını, R , replikasyon sayısını, $\hat{e}_{y_r}(x_i)$; r örneklemindeki hücrelerin birinden hesaplanmış ham puan eşdeğerini, $\bar{s}d_i$; eşitlemenin standart hatasını, \hat{d}_i ; eşitleme yanlılığı, $\hat{e}_y(x_i)$; ölçüt eşitleme fonksiyonundan elde edilen ham puan eşdeğerini göstermek üzere eşitlemenin standart hatası (Kolen ve Breannan, 2004);

$$\begin{aligned}\hat{e}_y(x_i) &= \frac{\sum \hat{e}_{y_r}(x_i)}{R} \\ \bar{s}d_i = SEE_i &= \sqrt{\frac{\sum [\hat{e}_{y_r}(x_i) - \hat{e}_y(x_i)]^2}{r}} \\ &= \sqrt{\frac{\sum [\hat{e}_{y_r}(x_i) - \hat{e}_y(x_i)]^2}{R-1}}\end{aligned}\quad (59)$$

(59) nolu denklem şeklinde ifade edilir. Kim ve diğerleri (2006) yanlılığı (60) nolu

$$\hat{d}_i = \frac{\sum_{r=1}^r [\hat{e}_{y_r}(x_i) - \hat{e}_y(x_i)]}{R}\quad (60)$$

denklemde ki gibi hesaplamıştır. Eşitleme yanlılığı ve eşitleme hatasının kareleri toplamı eşitleme hatasını göstermektedir (Skaggs, 2005, Kim ve diğerleri, 2006). Eşitleme hatası (61) nolu eşitlikte ifade edilmiştir.

$$RMSE_i = \sqrt{\hat{d}_i^2 + \bar{s}d_i^2}$$

(61)

5. AMAÇ

Bu arařtırmada, küçük örnekleme sahip denk olmayan gruplarda ortak maddelere dayalı eşitleme desenlerinde test formlarının alt testlerinde uygulanan eşitleme yöntemlerinin, örneklem büyüklüğü, alt test uzunluğu, farklı güçlük düzeyleri ve ankor test uzunluğunun değişimlenmesine bağlı olarak karşılaştırılması amaçlanmaktadır. Bireylerin belli bir özelliğe dair durumlarıyla ilgili bilgi edinebilmek için test formundan yararlanılır. Test formundaki maddelerin hatırlanması ya da test maddelerinin paylaşılması gibi durumlardan dolayı ilgilenilen özellikle ilgili başka test formları kullanılması uygun görülmüştür. Aynı özelliği ölçen farklı test formları aynı bireylerce yanıtlanabileceği gibi farklı test formlarından yalnızca biri bireylere uygulandığı durumlar da olabilmektedir. Dolayısıyla aynı özelliği ölçen test formlarından elde edilen bilgiler birbirinin yerine kullanılabilir ve birbiriyle karşılaştırılabilir olmalıdır (Akt: Sinharay, 2010b:150, APA, 1999). Aynı özelliği ölçen farklı test formlarının karşılaştırılabilir ve birbirinin yerine kullanılabilir olması için testlerden elde edilen test puanları aynı doğru üzerinde konumlandırılması testlerin eşitlenmesiyle sağlanır. Alt test içeren test formlarında da test formlarının ya da alt test formlarının birbiriyle karşılaştırılması için alt test formları ölçülen özellik bakımından birbiriyle eşitlenmektedir. Alt test formlarından elde edilen bilgiler alt test puanlarının bir biriyle karşılaştırılmasının yanı sıra bireylerin ilgilenilen özellik bakımından profilleri hakkında doğru yargılara varmamızı sağlayacak sonuçlar vermektedir (Sinharay 2010a, Sinharay, 2010b). Bireylerin ilgili özellik bakımından zayıf ya da güçlü olduğu durumları gözlemlememizi sağlayacaktır. Bu bilgileri edinebilmek için test formlarının alt test formları artı değere sahip olmalıdır. (Sinharay, Haberman ve Puhan, 2007a.). Testlerin eşitlenmesinde elde edildiği gibi artı değere sahip alt testlerin eşitlenmesi sonucunda da eşitleme hatası elde edilmektedir. Eşitleme yönteminin sayıtlarının karşılanamaması ya da kontrol edilemeyen başka durumların sistematik hataya sebep vermesi gibi örneklem büyüklüğünde eşitlemede hataya (seçkisiz hata) sebebiyet vermektedir. Eşitlemede büyük örneklem ya da örneklem büyüklüğünün artırılarak değişimlendiği çalışmalar sonucunda eşitlemenin standart hatasında azalma gözlenmiştir. Büyük örneklemelerle çalışıldığında örneklemin evreni temsil etme gücü arttığı ve eşitlemedeki hata durumu azaldığı için yapılan eşitlemenin kalitesi artmaktadır. Aynı özelliği ölçen farklı test formlarından elde edilen bilgilerle işlem yapabilmek için büyük örneklemere ulaşmak her zaman mümkün olmamaktadır. Öğretmen yapımı sınavlarda veya bazı kurs merkezlerinin yaptıkları sınavlarda küçük örneklemelerle çalışılmak zorunda kalılabilmektedir. Örneklem sayısı küçük olduğu durumlarda, örneklemin evreni temsil etmemesi ve kullanılacak yöntemlerin sayıtlarını karşılanamama problemleri ortaya çıkmaktadır. Bu problemler eşitleme hatasının fazla olmasına neden olacağı için Kolen ve Brennan (2004) tarafından küçük örneklemelerde eşitleme yapılmaması gerektiğini belirtilmesine rağmen Livingston (2009) küçük örneklemelerin neden olacağı problemlerin üstesinden gelinebilecek modellerin kullanılabileceğini ifade etmiştir. Küçük örneklemelerde eşitleme yapıldığında hangi eşitleme deseninin ve hangi eşitleme yönteminin kullanılmasının uygun olunacağına karar verilmelidir.

Testlerin eşitlemesini örneklem büyüklüğü, kullanılacak eşitleme deseni, test uzunluğu, test formları arasındaki güçlük düzeyi farklılığı ve eşitlemede kullanılacak olan yöntemler gibi etmenler etkilemektedir. Test düzeyleri paralel olduğu zaman test formları arasındaki güçlük düzeyi farkı da az olacaktır. Ancak pratikte paralel test formlarının elde edilmesi oldukça zordur. Bu durumun yanı sıra küçük örneklemelerde çalışılmak zorunda kalındığında paralel testin elde edilmesi daha da zorlaştıracaktır. Çünkü elde edilen verilerden madde parametrelerine dair yapılacak kestirimler

güçleşecektir. Küçük örneklerde ve test formlarının güçlük düzeylerinin farklı olduğu durumlarda farklı eşitleme yöntemlerinin kullanılarak alt testlerin eşitlenmesine dair yeterli bilgi yoktur. Bu çalışmada örneklem büyüklüğü, test formları arasındaki güçlük farkı düzeyleri, alt test uzunlukları, alt test formları arasındaki koorelasyon miktarı şeklindeki faktörlerin değişimlenen farklı düzeylerinin hangilerinde alt test formlarının artı değere sahip olduğunu, artı değer özelliğine sahip alt testlerin alt test ve genişletilmiş alt test puanları kullanılarak çeşitli eşitleme yöntemlerinden elde edilecek olan eşitleme hatalarının hangi yöntemlerde en küçük hata değeri vereceğini belirlemek amaçlanmıştır.

6. ÖNEM

Eğitim alanında bireylerin istedik davranışları kazanıp kazanmadığını anlamak için ülkemizde kağıt kalem testlerine sıkça başvurulmaktadır. Her yıl bireyler birçok sınava girmektedir. Bazı aynı içeriğe sahip sınavlar her yıl, yıl içerisinde bir ve birden fazla defa yapılmaktadır. Yapılan sınavlardan elde edilen sonuçlar doğrultusunda bireylere gerekli yönlendirme yapılmakta ya da bireyler hakkında kararlar alınmaktadır. Son yıllarda bireylere uygulanan testlerden elde edilen sonuçların yanı sıra testlerin alt alanları baz alınarak da bireylerin alt alanlardaki profilleri oluşturulmaktadır. Başka bir ifadeyle bireylerin testlerin alt alanlarındaki konuların hangisinde daha başarılı hangisinde daha çok eksikliği ve yeterliliğinin olmadığını anlamak için testlerin alt puanlarından yararlanılmaktadır. Alt alanlardan elde edilen bilgiler doğrultusunda daha sağlıklı yönlendirme ve kararlara varılabilmektedir. Dolayısıyla bireylerle ilgili karar vermek, bireylerin farklı testlerden aldıkları puanların karşılaştırılabilir olması test puanlarının bir birleriyle ilişkilendirilmeleri gerekmektedir. İlişkilendirme ise testlerin birimlerin birbirine dönüştürülmesiyle mümkün olmaktadır. Dönüştürme sürecinde eşitleme hataları meydana gelmektedir. Eşitleme hatasının azaltmak için eşitleme yöntemleri ve desenlerinin dikkatli seçilmesi gerekir. Doğru eşitleme yapıldığında ölçme ve değerlendirmenin de kalitesi artacaktır. Ülkemizde yapılan çalışmalar toplam test puanı bazındadır. Ülkemizde alt konu alanları baz alınarak test eşitleme çalışması yapılmamıştır. Bu çalışmada denk olmayan gruplar deseniyle elde edilen veriler Klasik Test Teoresine dayalı eşitleme yöntemlerinden zincirleşmiş lineer eşitleme, dairesel yay eşitleme, Braun/Holland eşitleme ve birim eşitleme yöntemleri kullanılarak alt konu alanları bazında oluşturulan alt testler eşitlenerek elde edilen eşitlemeleri eşitleme hataları karşılaştırılacaktır. Alt test bazında yapılacak olan eşitleme çalışmalarını daha mikro düzeyde inceleme sağladığı için eşitlemenin niteliğini arttıracak düşünülmemektedir. Buna yönelik herhangi bir çalışmaya rastlanmamıştır. Bu çalışmada belirli sınıflar altında öne sürülen görüş test edilecektir. Görüşü destekleyici bulguya rastlanması durumunda bulgunun özgün değer taşıyacağı ve ölçme değerlendirme alanına katkı sağlayacağı düşünülmektedir.

7. KAYNAK ARAŞTIRMASI

Bu bölümde ülkemizde ve yurt dışında test eşitleme ve artı değer özelliğine sahip alt testler le yapılan çalışmalara yer verilmektedir.

7.1. Yurt Dışında Yapılan İlgili Çalışmalar

Livingston (1993) çalışmasında ortak maddelere dayalı desenle elde edilmiş 25, 50, 100 ve 200 örneklem büyüklüğündeki verilerde log-lineer ön düzgünleştirmenin etkisini incelemiştir. Elde edilen sonuçlara göre ön düzgünleştirme işlemi küçük örneklerde eşitleme hatasının azlmasında daha etkili olarak bulunmuştur (Akt: Skaggs, 2005:310).

Hanson ve diğerleri (1994), çalışmalarında, örneklem büyüklüğü 100 ve altı olduğu durumlarda özdeş eşitlemenin, doğrusal eşitleme ve eşit yüzdelli eşitleme yöntemlerine göre daha az eşitleme hatası verdiğini belirtmişlerdir (Akt: Skaggs, 2005:310).

Hanson, Zeng ve Colton (1994) çalışmasında 100'den 3000'e aralığında örneklem büyüklüğü değişimleyerek birim eşitleme, lineer eşitleme, düzgünleştirilmemiş, ön düzgünleştirilmiş, son düzgünleştirilmiş eşit yüzdelli eşitleme yöntemlerini hangisinin etkili olduğuna yönelik çalışma gerçekleştirmişlerdir. Çalışmada küçük örneklemde düzgünleştirme yöntemi uygun görülürken düzgünleştirme yöntemlerinin hangisinin daha etkili olduğu belirtilmemiştir.

Parshall, Houghton ve Kromery (1995) çalışmasında ortak maddelere dayalı desende lineer eşitleme yöntemini kullanmıştır. Test formları arasındaki farklılık 0.00- 0.40 arasında değişimleyerek eşitlemeyi 15, 25, 50 ve 100 örneklem büyüklüğündeki verilerde gerçekleştirmiştir. Elde edilen sonuca göre örneklem büyüklüğünün azalması eşitlemenin standart hatasının artmasına neden olmaktadır. Ayrıca ankor testin toplam testle olan korelasyonu yüksek olmakla birlikte ankor testin toplam testteki oranı yüksek olmalıdır.

Skaggs (2005) çalışmasında seçkisiz desene dayalı çok küçük örnekleme dayalı gruplarda eşitlemenin etkililiğini incelemek için yapmıştır. Çalışmada ortalama eşitleme, lineer eşitleme, düzgünleştirme uygulanmamış eşit yüzdelli eşitleme yöntemi, ve 6 dereceye kadar değişimlenen ön düzgünleştirilmiş eşit yüzdelli eşitleme yöntemi ve örneklem büyüklüğü 25, 50, 75, 100, 150 ve 200 değişimlenmiştir. Elde edilen sonuçlara göre 25 örneklem büyüklüğü olan en küçük örneklem büyüklüğüne sahip örneklemlerde birim dönüşüm ile yapılan eşitleme en uygun sonuçları vermişken 50 ve üzeri örneklem büyüklüğündeki örneklemlerde ise elde edilen verilerde eşitleme yönteminin uygunluğu performansa dayalı olarak farklılık göstermiştir. Geçme puanının ortalamaya olan uzaklığına bağlı olarak bazı eşitleme yöntemleri daha iyi sonuçlar üretmiştir. Geçme puanı ortalamanın üzerinde iken 2 ve 3 üncü dereceden ön düzgünleştirilmiş eşit yüzdelli eşitleme yöntemleri en iyi sonucu vermiştir. Geçme puanı ortalamaya yakın olduğunda çalışmada kullanılan eşitleme yöntemleri lineer eşitleme gibi sonuçlar üretmiştir. Geçme puanı ortalamanın altında olduğunda ortalama eşitleme küçük bir yüzdeyi temsil etmek üzere uygun sonuç vermiştir.

Kim, Davier ve Haberman (2008), gerçekleştirdikleri çalışmada küçük örnekleme sahip sentetik populasyon için alternatif bir sentetik eşitleme fonksiyonu tanımlanıp eşitleme gerçekleştirilmiştir. $S(x)=w.Ey(x)+(1-w).IDy(x)$. $Ey(x)$; lineer zincirleme fonksiyonu, Tucker ya da Levine fonksiyonudur. Çalışmada 10, 25, 50, 100, 200 örneklem büyüklüğüne sahip sentetik populasyonlar için tanımlanan sentetik fonksiyon yardımıyla yapılan eşitlemenin eşitleme hatası ve bağlanmanın yanlılığı, SEE, RMSE değerleri baz alınarak karşılaştırılmıştır. Çalışmada gerçekleştirilen zincirleme eşitleme, birim eşitleme ve tanımlanan sentetik fonksiyonla yapılan eşitlemelerin yanlılık, standart hata ve RMSE değerleri karşılaştırılmıştır. Çalışma sonunda birim dönüşüm ile yapılan eşitlemede hatalar daha düşük olmasına karşın diğer eşitleme fonksiyonlarıyla yapılacak eşitlemelerden gelecek olan eşitleme hataları daha fazla olacağından tanımlanan sentetik fonksiyon fonksiyonuyla yapılan eşitlemede daha az hata elde edilmektedir.

Bazı durumlarda örneğin daha özel ilgi gerektiren durumlarda büyük örneklem gruplarına ulaşma imkanı olamamaktadır. Bu gibi kurs ya da eğitimlerde öğrenci sayısı az olmaktadır. Hatta yeni kursiyerlerinde bir sonraki kursta yeterli sayıda olup olmayacağı belli olmamaktadır. Bu grupların performanslarının izlenmesinde ya da başarı durumu ile ilgili bilgi edinmek için her ne kadar ölçülmek istenen özellikle ilgili ayrıntılı bir içeriğe uygun test formu hazırlansa dahi bu test formunun istatistiksel özellikleri konusunda ön uygulama yapılamadığı durumlar olmaktadır. Sonuç olarak uygulanan test

formu daha önceki test formundan kolay ya da zor olmaktadır ve yeni test formu eski test formuyla eşitlenmektedir. Bu gibi durumlar söz konusu olduğu için küçük örneklemelerde eşitlemenin daha iyi yapılması için Livingston ve Kim (2009) Divigi'nin önermiş olduğu, eşitlemede ilişkinin doğrusal olmadığı durumlarda daha iyi sonuçlar veren dairesel-yay eşitleme fonksiyonunun yanı sıra zincirlenmiş eşityüzdelikli eşitleme yöntemi, Levine, Tucker, zincirlenmiş lineer ve ortalama eşitleme yöntemi kullanarak küçük örneklemelerde eşitleme gerçekleştirilmiştir. Geliştirilen dairesel-yay eşitleme yönteminde test formundan elde edilen puanların yanı sıra test formunun sınır puanları ve türetilen bir orta puandan yararlanarak eşitleme yapılmıştır. Elde edilen sonuçlar eşitleme için kullanılan diğer yöntemlere göre 10'uncu ve 90'ıncı ranklarda dairesel-yay eşitleme yöntemi diğer yöntemlere göre daha az hatalı değerler vermesinin yanı sıra genel olarak diğer yöntemlere göre daha az hata vermiştir.

Küçük örneklemelerde örneklem popülasyonu temsil edemediğinden eşitlemeler sonucu elde edilen eşitleme hataları yüksek olmaktadır. Kim ve Livingston (2010) çalışmasında küçük örneklemelerde ortak maddelere dayalı eşitleme deseninde eşitleme yöntemlerinin karşılaştırılması yapılmıştır. En az 110 madde içeren dört test formunun verilerinden yararlanarak 10, 25, 50, 100 örneklem büyüklüğüne sahip örneklemeler elde edilerek zincirlenmiş eşit yüzdelikli eşitleme, zincirlenmiş lineer eşitleme, zincirlenmiş ortalama eşitleme, simetrik dairesel-yay eşitleme ve dairesel-ay eşitleme yöntemleri kullanılarak bu eşitlemelerden elde edilen RMSE değerleri karşılaştırılmıştır. Elde edilen sonuçlarda düşük puanlarda zincirlenmiş ortalama eşitleme yöntemiyle elde edilen hata değeri diğer yöntemlere göre daha azdır. Performansın yüksek olduğu puan dağılımlarında dairesel-yay eşitleme yöntemlerinin RMSE değerleri diğer yöntemlere göre daha az çıkmıştır. Çalışma sonuçlarına göre örneklem büyüklüğünün fazla olduğu durumlarda dairesel-yay eşitleme yöntemi eşit yüzdelikli eşitleme yöntemine göre daha iyi sonuçlar vermiştir.

Eşitlenecek test formları birbirine tamamen paralel ve örneklem büyüklüğü çok küçük ise birim eşitleme fonksiyonu uygulanmamalıdır ya da birim dönüşümün yapılması uygundur. Eşitlemeden elde edilecek hata da sıfır olur. Dolayısıyla bazı uzmanlarca küçük örneklemelerde eşitleme çalışması yapıldığında birim dönüşümün yapılması gerektiği savunulmuştur. Sentetik fonksiyon birim dönüşümün belirlenen bir ağırlıkla elde ettiği eşitleme fonksiyonudur. Kim, von Daver ve Haberman (2011) çalışmalarında sentetik fonksiyon birim dönüşüm fonksiyonuyla, zincirlenmiş lineer eşitleme fonksiyonu, Tucker, Levine eşitleme yöntemlerinin ikili kombinasyonu şeklinde oluşturulmuş, popülasyon ağırlıkları birim dönüşüm 0.5 ve 0.3 olmak üzere diğer eşitleme fonksiyonlarının ağırlıkları 0.5 ve 0.7 olacak şekilde değişimlenmiştir. Çalışmada küçük örneklemelerde sentetik fonksiyonun yanı sıra Tucker, zincirlenmiş lineer fonksiyon, Levine eşitleme fonksiyonu ve ortalama eşitleme fonksiyonları da kullanılmıştır. Çalışmada eşitlenecek test formları arasında zorluk farkı mevcuttur. Dolayısıyla sentetik fonksiyon ile yapılan eşitleme diğer eşitleme yöntemleri kadar performans göstermemiştir. Farklı test formu güçlük farkları sentetik fonksiyonda eşitleme yanlılığının fazla olmasına neden olmuştur. Dolayısıyla sentetik fonksiyon zorluk farkı olmadığı durumlarda iyi sonuçlar verdiği sonucuna ulaşılmıştır. Çalışmalarında küçük örneklemelerde sentetik fonksiyon ile yapılan eşitleme bir alternatif olarak görülebileceğini ancak küçük örneklemelerde gözlenen sorunların çözümünü veren bir yöntem olmadığını ayrıca ortak madde sayısı fazla, toplam testle yüksek korelasyon gösterilmesi durumlarında kullanıldığında eşitleme hatası ve yanlılık değerlerinde azalma gözlendiğini belirtmişlerdir.

Sinharay ve Haberman (2011), alt test puanlarına olan ilginin artmasıyla birlikte elde edilen alt test puanlarının eşitlemesi konusunda bir çalışma yapmışlardır. Çoktan seçmeli test maddelerinden oluşan testler kullanılmıştır. Çalışmada gözlenen alt test puanlarının yanı sıra genişletilmiş (augmented) alt test puanları kullanılmıştır. Gerçek verilerin yanı sıra gerçek veriler yardımıyla simülasyon verileri üretilerek ortak maddelere dayalı eşitleme deseninde gözlenen alt puan ve genişletilmiş alt puanlar elde edilerek eşitleme zincirlenmiş eşitleme gerçekleştirilmiştir. Ortak maddelerle yapılan eşitleme

için toplamda yedi farklı eşitleme deseni oluşturulmuştur. Yedi farklı desenin üçü gözlenen alt puanların eşitleme deseni olarak tasarlanırken dördü ise genişletilmiş alt puanlar için oluşturulmuştur. Gözlenen alt puanların ankor testle eşitlenmesinin yapılması için ilk desende X formundaki alt testler ankor testteki alt testiyle eşitlenir. Ankor testteki alt form Y testinin uygulandığı grubun ankor testinin alt testine eşitlenir. Ankor testteki alt testler ise Y formundaki alt testlere eşitlenerek X formundaki alt testler Y formundaki alt testlere eşitlenmiş olur. İkinci desende X ve Y testleri birbirine eşitlenir. Eşitlendikten sonra X formundaki alt test puanları eşitlenmiş olan test puanlarına dönüştürülür. Elde edilen puanlar Y testindeki puanlara dönüştürülür. Y testindeki test puanlarına dönüştürülen puanlar tekrar Y formundaki alt testlere dönüştürülerek eşitleme gerçekleştirilmiş olur. Üçüncü desende alt test puanları toplam ankor testi puanlarına dönüştürülerek eşitleme gerçekleştirilir. Genişletilmiş alt puanlar için oluşturulan desenlerin ilkinde, X testindeki alt testlerin genişletilmiş alt puanları ankor testin alt test puanına dönüştürülür. Dönüştürülen alt puanlar Y testini alan grubun alt ankor test puanlarına dönüştürülür. Son olarak elde edilen dönüştürülmüş alt puanlar Y testindeki genişletilmiş alt puanlara dönüştürülerek eşitleme gerçekleştirilmiş olur. İkinci desende gözlenen alt testlerin ikinci eşitleme desenine benzer olarak, X testinin genişletilmiş alt test puanları önceden eşitlenmiş X ve Y testinin eşitlenmiş toplam test puanlarına dönüştürülerek elde edilen dönüşüm puanları tekrar Y formundaki toplam test puanlarına dönüştürülür. Elde edilen dönüşüm puanları Y formundaki genişletilmiş alt test puanlarına eşitlenir. Üçüncü desende ise X formundaki genişletilmiş alt puanlar ankor testindeki genişletilmiş alt puanlara dönüştürülür. Elde edilen puanlar Y testinin uygulandığı grubun ankor formundaki genişletilmiş alt test puanlarına bağlanır. Son olarak elde edilen alt puanlar Y formundaki genişletilmiş alt test puanlarına dönüştürülerek eşitleme tamamlanmış olur. Son desende ise, X formundaki genişletilmiş alt test puanları toplam ankor testi puanına dönüştürülür. Dönüştürülmüş puanlar ankor testler aracılığıyla bağlandıktan sonra elde edilen puanlar Y testindeki genişletilmiş alt test puanlarına eşitlenir. Çalışmada genişletilmiş puanlar gözlenen puandan daha çok bilgi ifade ettiği için kullanılmıştır. Çalışmada genişletilmiş puanların kullanılarak eşitleme yapıldığında eşitleme hatasının değeri daha az olduğu sonucuna ulaşılmıştır.

Puhan ve Liang (2011) çalışmasında alt testlerin eşitlenmesinde iki farklı yöntem kullanmıştır. Yöntemlerinden biri eşitlenmiş toplam test puanlarının kullanarak alt test eşitlemesi gerçekleştirirken diğerinde ise alt test puanlarını ankor testin alt test puanlarını kullanarak eşitlemiştir. Çalışmada kullanılan X ve Y test formları çoktan seçmeli test maddelerinden oluşmaktadır. X formu X1 ve X2 olmak üzere iki test formuna ayrılmıştır. Y formu da Y1 ve Y2 test formuna ayrılmıştır. Test formları oluşturulurken toplam testin (X ve Y) içerik özelliği yansıtılacak şekilde yapılandırılmıştır. X1 ve X2 test formunun birbirine eşitlenmesi ankor test aracılığıyla gerçekleştirilmiştir. Benzer olarak Y testi için gerçekleştirilmiştir. Çalışmada alt alan sayıları değişimlenmiştir. Örneğin; X testi dört alt alandan oluşurken Y testi üç alt alandan oluşmuştur. Ankor test madde sayıları eşitlenecek olan testlerin %50'si ve %25'i şeklinde değişimlenmiştir. Ankor madde sayısı %50 olduğu durumlarda eşitleme yöntemlerinden ikincisinden elde edilen eşitleme hataları daha azdır. Dolayısıyla ankor madde sayısı testin %50'si kadar olduğunda alt test puanları kullanılarak eşitleme yapılması tercih edilirken ankor madde sayısı %25 olduğunda ise toplam testin eşitlenmesi ya da toplam testlerin eşitlenmiş puanlarının ankor olarak kullanılması uygun görülmüştür.

Babcock, Albano ve Raymond (2012) küçük örneklemeler kullanarak eşitleme yapmıştır. Araştırmada eşitleme yedi farklı eşitleme yöntemi (Zincirilmiş eşit yüzdelli eşitleme, Tucker lineer eşitleme, ortalama eşitleme, nominal ağırlıklandırılmış ortalama, birim dönüşüm, sentetik eşitleme (birim dönüşüm+Tucker) (.25 ve .75 ağırlığında olacak popülasyonlar) ve Dairesel-yay yöntemi), üç farklı örneklem büyüklüğü (20, 50, 80), farklı zorluklarda testler (0.00, 0.35, 0.70) farklı yetenek düzeylerinde ($N(0,1)$, $N(1,1)$) incelenerek karşılaştırılmıştır. Araştırmada üç farklı veri kümesi oluşturulmuştur. Desen1 de eşit yetenek düzeyindeki bilgiler doğrultusunda karşılaştırma yapılırken, Desen2 de yetenek düzeyi yüksek grubun basit formu almasıyla elde edilen bilgiler ile karşılaştırma ve Desen3 de yetenek düzeyi düşük grup

kolay testi aldığıında elde edilen bilgiler doğrultusunda yapılan eşitlemeler karşılaştırılmıştır. Desen1 düzeneğinde elde edilen sonuçlar doğrultusunda sentetik eşitleme fonksiyonuyla elde edilen RMSE değeri Tucker ve birim dönüşümle elde edilmiş RMSE değerine göre daha etkili sonuç vermiştir. Desen2’de testlerin arasında güçlük farkı yokken en iyi sonucu birim dönüşüm vermiştir. Orta ve yüksek düzeyde güçlük farkı olan durumlarda ise en iyi sonucu yani en az hata değerini nominal ağırlıklı ortalama eşitleme yöntemi vermiştir. Dairesel-yay ve sentetik eşitleme sonuçlarında da RMSE değerleri düşük gözlenmiştir. Bunun yanı sıra nominal ağırlıklı ortalama ve sentetik eşitlemede yanlılık değeri diğer yöntemlere göre daha düşük çıkmıştır. Çalışmada yetenek düzeylerindeki farkla birlikte eşitleme sonuçları ele alındığında en iyi sonuç veren nominal ağırlıklı ortalama eşitleme ve dairesele-yay eşitleme yöntemleridir. Desen2’deki gibi test formları arasında herhangi bir fark olmadığında en az RMSE değerini veren birim eşitleme iken, test formları arasında güçlük farkı varken ve yetenek düzeyleri farklı olduğunda en az RMSE değerini veren nominal ağırlıklı ortalama eşitleme ile elde edilmiş olmasına rağmen dairesele-yay eşitleme yöntemiyle elde edilen RMSE değeri diğer geleneksel yöntemlere göre daha azdır. En fazla RMSE değeri ise düzgünleştirilmiş eşit yüzdelliğe eşitleme yöntemiyle elde edilmiştir. Eşit yetenek düzeyinde ve zorluk bakımından arasında fark olmayan testler kullanıldığında birim dönüşümden elde edilen sonuçlar hatadan daha arınırken, test formları arasındaki zorluk farkıyla birlikte elde edilen eşitlemelerde ortalama ve nominal ağırlıklı ortalama yöntemiyle yapılan eşitlemeleri hatadan daha arınık hale gelmiştir. Test formlarının güçlüklerinin de farklı olmasının yanı sıra yetenek düzeyleri de farklıyken en az değeri nominal ağırlıklı ortalama ile yapılan eşitleme vermiştir. Ardından dairesele-yay ile yapılan eşitlemede de elde edilen hata değeri az olarak gözlenmiştir. Çalışmada lineer eşitleme yöntemlerinin RMSE değerlerine bakıldığında lineer yöntemlerin hata değerleri eşit yüzdelliğe eşitleme yöntemine göre daha az sonuçlar elde edilmiştir. Grupların yetenek düzeylerinde farklılık olduğu durumlarda dairesele-yay yöntemi de etkili sonuçlar vermiştir.

7.2. Yurt İçinde Yapılan İlgili Çalışmalar

Şahhüseyinoğlu (2006) çalışmasında, İngilizce yeterlik sınavında kullanılan testlerin birbirine eş değer olup olmadığı, eş değer değilse, test eşitleme yöntemleriyle farklı sınavlara ait alınan puanların birbirine dönüştürülmesi ve testlerin birimlerinin aynı olmasının sağlanmasını hedeflemiştir. İngilizce yeterlik sınavında kullanılan testlerin eşitlenmesinde Rasch yöntemi ve eşit yüzdelliğe eşitleme yönteminin uygun yöntemler olduğuna ulaşmıştır.

Çetin (2009) tarafından yapılan çalışmada, 2005 ÖBBS sınavına giren 6., 7. ve 8. sınıf öğrencilerine ait veriler üzerinde dikey ölçeklemeler yapılmıştır. Bu araştırma sonucunda, 250, 500 ve 1000 kişilik örneklem büyüklüklerinden, en az hata veren örneklem büyüklüğünün 1000 kişilik olduğu saptanmıştır.

Bozdağ ve Kan (2010a) çalışmalarında şans başarısının test eşitlemeye etkisi olup olmadığını incelemiştir. Araştırmada, tek grup düzeneği, doğrusal eşitleme ve eşit yüzdelliğe eşitleme yöntemleri kullanılmışlardır. Araştırmanın sonucunda, şanstın arındırılmamış test puanlarının eşitlenmesinde doğrusal eşitlemenin, şanstın arındırılmış test puanlarının eşitlenmesinde ise eşit yüzdelliğe eşitlemenin uygun yöntemler olduğunu ve tüm hatalar göz önünde bulundurulduğunda, şanstın arındırılmış eşit yüzdelliğe eşitleme yöntemine ait hatanın en düşük olduğuna ulaşmışlardır. Kan (2010b) çalışmasında aynı davranışları ölçen farklı madde formlarına sahip matematik testlerinin istatistiksel eşitliğini sınamak ve bu yolla madde formunun öğrencilerin performansı üzerinde etkili olup olmadığını araştırmıştır. Aynı davranışları ölçen sayısal köklü matematik testi ile sözel köklü matematik test versiyonlarının istatistiksel olarak birbirine eşit olmadığı ve birbirinin yerine kullanılamayacağını saptamıştır. Kan (2011) çalışmasında farklı yıllarda uygulanan OKS test formlarının istatistiksel eşitliğini sınamayı amaçlamıştır. Bu amaç doğrultusunda klasik eşitleme metotlarından lineer eşitleme ve tek grup düzeneği

kullanmıştır. 2003 ve 2005 yıllarında uygulanan OKS test versiyonlarının istatistiksel olarak birbirine eşit olmadığı ve birbirinin yerine kullanılamayacağı saptamıştır.

Öztürk ve Anıl (2012) yaptığı çalışmada 2008 ilkbahar ve sonbahar sayısal-1 testi puanlarını doğrusal ve eşit yüzdelliği eşitleme yöntemleri kullanarak eşitlemek ve bu yöntemlerden uygun olanını önermeyi amaçlamışlardır. Araştırma sonunda testlerin eşitlenmesi için yüzdelliği eşitleme olduğu sonucuna ulaşmışlardır.

Gök (2012), yaptığı çalışmada, farklı koşullara (örneklem büyüklüğü, yetenek dağılımı, test uzunluğu ve model türü) göre türetilen test formlarını, madde tepki kuramına dayalı kestirim yöntemlerini (ortalama-ortalama, ortalama-standart sapma ve Stocking-Lord) kullanarak eşitlemek ve bu yöntemlerden elde edilen sonuçları karşılaştırmayı amaçlamıştır. Araştırmada, eşitleme deseni olarak “denk olmayan gruplarda ortak madde/test (NEAT) deseni” kullanılmıştır. Araştırmada kullanılan veriler, simülasyon veridir. Araştırmada elde edilen sonuçlar, eşitleme yanlılığı (BIAS) ve eşitleme hatası (RMSE) ölçütlerine göre değerlendirilmiştir. Ele alınan koşullar doğrultusunda, büyük örneklemeler ile daha uzun testler kullanıldığında ve benzer yetenek dağılımına sahip gruplarda, yöntemlerin daha az hatalı olduğu sonucuna ulaşılmıştır.

Kelecioğlu ve Gübeş (2013), 2009 yılında yapılan Öğrenci Başarılarının Belirlenmesi Sınavı (ÖBBS)’nin sosyal bilimler alt testi C kitapçığından elde edilen ham puanlar doğrusal eşitleme ve üç farklı eşit yüzdelliği eşitleme (düzgünleştirilmemiş, ön-düzdünleştirilmiş ve son-düzdünleştirilmiş eşit yüzdelliği eşitleme) yöntemleri ile A kitapçığından elde edilen puanlara eşitmişlerdir. Araştırmanın sonucunda doğrusal eşitleme ile elde edilen A formu eşdeğer puanları ile C formu ham puanları arasında doğrusal bir ilişki olduğu, puan ölçeği boyunca C formunun A formundan daha kolay olduğu görülmüştür. Üç farklı eşit yüzdelliği eşitleme yönteminin birbiriyle tutarlı sonuçlar ürettiği, A formu eşdeğer puanları ile C formu ham puanları arasında eğrisel bir ilişki olduğu ve testin güçlük düzeyinin puan ölçeği boyunca farklılık gösterdiği tespit edilmiştir. Araştırmanın sonucunda, en küçük BSHO ve RMSD katsayısına doğrusal eşitleme yöntemi ile elde edilen eşdeğer puanların sahip olduğu görülmüştür. Ancak, ham puanlar ile eşitlenmiş puanlar arasında eğrisel bir ilişki olmasına dayalı olarak puanların eşitlenmesi için göreceli olarak daha az random hataya sahip olan son-düzdünleştirilmiş eşit yüzdelliği eşitlemenin en uygun eşitleme yöntemi olduğu kararına varılmıştır.

Mutluer (2013), 2011 ilkbahar ve sonbahar dönemlerinde ÖSYM tarafından merkezi olarak yapılan Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı (ALES) Puanlarını, doğrusal ve eşit yüzdelliği eşitleme yöntemleri ile eşitleyerek, bu iki yöntemden en uygun olanını belirlemeyi ve bu yöntemlerden en uygununu önermeyi amaçlamıştır. Araştırmada, tek grup deseni kullanılmıştır. Araştırmanın sonucunda, en uygun yöntemin ağırlıklandırılmış hata kareleri ortalamaları katsayısı daha küçük olan eşit yüzdelliği eşitleme yöntemi olduğu görülmüştür.

Çetin ve Gelbal (2014), ülkemizde yapılmış olan ÖBBS verilerinden yararlanarak bir dikey ölçekleme çalışması gerçekleştirmişlerdir. Öğrencilerin akademik başarılarının yıldan yıla incelemek ve öğrenci başarısındaki gelişimin ne kadar olduğunu ortaya koyabilmek amacıyla geliştirilen yöntemlerden biri dikey ölçeklemedir. Araştırmanın verilerini, 2005 yılında Türkiye genelinde yapılan İlköğretim Öğrencilerinin Başarılarının Belirlenmesi Sınavı (ÖBBS)’na ait 6., 7. ve 8. sınıf öğrencilerinin Türkçe ve Matematik testlerine ait puanları oluşturmaktadır. Matematik ve Türkçe testinde 6 ve 7. sınıflarda 20, 8. sınıfta ise 25 soru bulunmaktadır. Her testte ortak madde sayısı 4’tür. Bu sonuçlar, ölçekleme çalışması yapmak için gereken asgari sayı ölçütünü karşılamıştır. KTT’ye dayalı Thurstone Ölçekleme ve MTK’ya dayalı dikey ölçekleme tekniği kullanılmıştır. Thurstone ölçeklemede, ilk olarak tüm örneklem için 6, 7 ve 8. sınıf seviyelerine ait ölçekleme işlemi ham puanlar ve frekanslara dayalı olarak normalleştirilmiş z puanlarına dönüştürülmüştür. Aynı işlemler, her bir testteki ortak maddeler için de tekrarlanmıştır. Elde edilen bulgular çerçevesinde, Thurstone ölçeklemede, matematik testinde ortalamalar

sınıf seviyesi ile birlikte düzenli artış göstermezken, standart sapmalar tutarlı artma eğilimindedir. Türkçe testinde ise ölçeklenmiş ortalamalar, sınıf seviyesi arttıkça düşerken, standart sapmalar yine artış göstermiştir. Bu durumda, ortalama örüntüsünden bağımsız olarak Thurstone ölçekleme her iki derste de sınıf seviyesi ile birlikte standart sapmalarda artış göstermiştir. MTK'ya dayalı ölçekleme sonunda, matematik dersinde, sınıf seviyesi ile ortalamalar tutarlı bir artış ya da azalma göstermez iken, standart sapmalar seviyelere bağlı olarak artış göstermiştir. Fakat Türkçe testine ait, MTK ölçeklemesinde gerek ortalama gerekse standart sapmalarda düzenli bir örüntü meydana gelmemiştir.

Gök ve Kelecioğlu (2014) yaptıkları çalışmada, farklı koşullara göre türetilen test formlarını madde tepki kuramına dayalı kestirim yöntemlerini kullanarak eşitlemek ve bu yöntemlerden elde edilen sonuçları karşılaştırmasını amaçlamışlardır. Çalışmada eşitleme hatasını etkileyen faktörlerden test uzunluğu, örneklem büyüklüğü, yetenek dağılımı ve model türü ele alınarak eşitleme hatasının hangi durumlarda daha az olacağı belirlenmeye çalışılmıştır. Çalışmada, değişkenler üzerinde kontrol sağlamak amacıyla simülasyon verileri kullanılmıştır. Araştırmada ele alınan faktörlerden test uzunluğu, ülkemizde gerçekleştirilen test uzunluklarının ortalama sayıları göz önünde bulundurularak her bir formda 30, 60 ve 80 maddenin yer aldığı üç koşul belirlenmiştir. Her üç koşul için maddelerin en az %20'si ortak madde olarak alınmıştır. Bu nedenle n=30 koşulunda 10 madde, n=60 koşulunda 20 madde, n=80 koşulunda ise 30 madde ortak madde olarak değişimlenmiştir. Araştırmada ele alınan ikinci faktör, her bir gruptaki cevaplayıcı sayısıdır. Her bir grup için 1000 ve 3000 cevaplayıcı olmak üzere iki koşul belirlenmiştir. Üçüncü faktör olan yetenek dağılımları gruplar için üç farklı durum değişimlenmiştir. İlk yetenek dağılımı koşulunda Grup 1, ortalaması 0 standart sapması 1 olan standart normal dağılıma sahipken Grup 2'de ise yetenek dağılımı grup 1'den 0.5 daha yüksektir ve grupların standart sapmaları eşit özellikte olacak şekilde düzenlenmiştir. İkinci yetenek dağılımı koşulunda, Grup 1'in dağılımı pozitif çarpık dağılıma; Grup 2 standart normal dağılım özelliğine sahip olacak şekilde oluşturulmuş. Üçüncü yetenek dağılımı koşulunda ise Grup 1 negatif çarpık dağılıma; Grup 2 standart normal dağılım özelliğine sahip olacak şekilde yapılmıştır. Çarpıklık katsayısı testlerde genellikle 0.30–0.40 arasında değiştiğinden, araştırmada da ortalama çarpıklık değeri söz konusu değerler arasında farklılık göstermiştir. Dördüncü faktör, veriyi türetmede kullanılan modelin türü olmak üzere, maddeler iki ve üç parametrelili lojistik modele (2PLM ve 3PLM) göre türetilmiştir. Gerçekleştirilen çalışmada madde tepki kuramına dayalı eşitleme yöntemlerinden OO (Ortalama-ortalama (mean-mean) yöntemi), OS (Ortalama-standart sapma (mean-sigma) yöntemi) ve SL (Karakteristik eğri dönüştürme (characteristic curve transformation) yöntemleri) yöntemlerinin test uzunluğu, örneklem büyüklüğü, yetenek dağılımı ve model türüne göre karşılaştırılması amaçlanmıştır. Araştırmadan elde edilen sonuçlar doğrultusunda; örneklem büyüklüğü ve test uzunluğu arttıkça daha kararlı madde parametre kestirimleri elde edilmiştir. Bu doğrultuda, en kararlı parametre kestirimlerine örneklem büyüklüğü 1000'in üzerinde olduğunda ve test uzunluğu da 60'ın üzerinde olduğu durumlarda ulaşılabileceği belirtmişlerdir. Benzer yetenek dağılımına sahip gruplarda ise farklı yetenek dağılımına sahip gruplara göre yöntemlerin hata değerlerinin daha düşük olduğu bulunmuştur. Madde parametreleri açısından yöntemler genel olarak incelendiğinde a ve b parametresi için en az hatalı yöntemin OO yöntemi olduğu görülürken, en fazla hataya sahip yöntemin ise a parametresi için OS, b parametresi için ise SL yöntemi olduğu görülmüştür. Araştırmadan elde edilen sonuçlar genel olarak değerlendirildiğinde, en iyi eşitlemelerin 3000 kişilik örneklem, 80 maddelik testler, benzer yetenek dağılımına sahip gruplar ve OO yöntemi kullanılarak elde edilebileceği ifade etmişlerdir. Ayrıca çalışmada ele alınan koşullar doğrultusunda, büyük örneklem ile daha uzun testler kullanıldığında ve benzer yetenek dağılımına sahip gruplarda yöntemlerin daha az hatalı olduğu sonucuna da ulaşılmıştır.

Gübeş ve Kelecioğlu (2015) çalışmalarında karma testlerin eşitlenmesinde madde tepki kuramına (MTK) dayalı gerçek-puan eşitleme ve gözlenen-puan eşitleme yöntemleri test eşitlemenin eşitlik özelliği korunumu ölçütüne dayalı olarak karşılaştırmışlardır. Çalışmada eşitleme deseni olarak, eşdeğer olmayan gruplar ortak test deseni kullanılmıştır. Çalışma

gerçek verilere dayalı olarak yürütülmüştür. Verileri, TIMSS 2011 8. Sınıf Türkiye örnekleminde yer alan öğrencilerin 5. ve 6. kitapçıkta yer alan matematik testlerine verdiği cevaplar oluşturmaktadır. Araştırmanın çalışma grubunu, TIMSS 2011 Türkiye örnekleminde yer alan Kitapçık 5'e cevap veren 490 ve Kitapçık-6'ya cevap veren 494 öğrenci oluşturmuştur. Kitapçık-6'da yer alan matematik testi (X Formu) puanları MTK gerçek-puan eşitleme ve MTK gözlenen-puan eşitleme yöntemleri kullanılarak Kitapçık-5'teki matematik testi (Y Formuna) puanlarına eşitlenmiştir. Araştırma, Kitapçık-5 ve Kitapçık-6'da yer alan 10 çoktan seçmeli ve 4 açık uçlu madde ile ortak olan 8 çoktan seçmeli ve 9 açık uçlu madde kullanılarak gerçekleştirilmiştir. Çalışma kapsamında eşitlenen her bir test toplam 31 maddeden oluşmuştur. Bir test formundan elde edilebilecek en yüksek puan 32'dir. Elde edilen bulgulara dayalı olarak MTK gerçek puan eşitlemenin birinci-sıra eşitlik özelliğini korumada MTK gözlenen-puan eşitlemeden daha iyi performans gösterdiği; MTK gözlenen-puan eşitlemenin ise ikinci-sıra eşitlik özelliğini korumada MTK gerçek-puan eşitlemeden daha iyi performans gösterdiği sonucuna ulaşılmıştır.

8. PROBLEM CÜMLESİ

Ortak maddelere dayalı eşitleme deseninde, artı değer özelliğine sahip olan alt testlerin, alt test /genişletilmiş alt test puanları kullanılarak yapılan eşitleme yöntemlerinin (birim eşitleme, zincirlenmiş lineer eşitleme, Braun/Holland eşitleme ve dairesel yay eşitleme yöntemleri), çeşitli faktörlere (örneklem büyüklüğü, test formları arasındaki ortalama güçlük düzeyi farkı, alt test uzunlukları) göre elde edilen eşitlemenin standart hatası (SEE), eşitlemenin yanlılığı (BIAS) ve eşitleme hatası (RMSE) kestirimleri nasıl değişmektedir?

8.1. Alt Problemler

- 1- Çeşitli faktörlerin, ortak maddelere dayalı eşitleme deseninde, artı değer özelliğine sahip olan alt testlerin, alt test /genişletilmiş alt test puanları kullanılarak yapılan eşitleme yöntemlerinden (birim eşitleme, zincirlenmiş lineer eşitleme, Braun/Holland eşitleme ve dairesel yay eşitleme yöntemleri) elde edilen eşitlemelerin eşitleme standart hatasına (SEE), eşitleme yanlılığına (BIAS) ve eşitleme hatasına (RMSE) temel etkisi nasıldır?
 - a. Örneklem büyüklüğünün,
 - b. Alt test uzunluğunun,
 - c. Test formları arasındaki güçlük farkının,
 eşitleme yöntemlerinden elde edilen eşitlemenin standart hatasına (SEE), eşitlemenin yanlılığına (BIAS), eşitleme hatasına (RMSE) temel etkisi nasıldır?
- 2- Çeşitli faktörlerin, ortak maddelere dayalı eşitleme deseninde, artı değer özelliğine sahip olan alt testlerin, alt test /genişletilmiş alt test puanları kullanılarak yapılan eşitleme yöntemlerinden (birim eşitleme, zincirlenmiş lineer eşitleme, Braun/Holland eşitleme ve dairesel yay eşitleme yöntemleri) elde edilen eşitlemelerin eşitleme standart hatasına (SEE), eşitleme yanlılığına (BIAS) ve eşitleme hatasına (RMSE) ortak etkisi nasıldır?
 - a. Örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük farkı faktörlerinin altında alt test puanları eşitlenmesiyle elde edilen
 - i. eşitlemenin standart hatasına (SEE)
 - ii. eşitlemenin yanlılığına (BIAS)
 - iii. eşitleme hatasına (RMSE)

ortak etkisi nasıldır?

- b. Örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük farkı faktörlerinin altında genişletilmiş alt test puanları eşitlenmesiyle elde edilen
- i. eşitlemenin standart hatasına (SEE)
 - ii. eşitlemenin yanlılığına (BIAS)
 - iii. eşitleme hatasına (RMSE)

ortak etkisi nasıldır?

9. SINIRLILIKLAR

1. Araştırma ortak maddelere dayalı eşitleme deseniyle sınırlıdır.
2. Araştırmada sadece Klasik Test Teorisi' ne dayalı eşitleme yöntemlerinden zincirleşmiş lineer eşitleme, dairesel yay eşitleme, Braun/Holland ve birim dönüşüm eşitleme yöntemleri kullanılmıştır.
4. Araştırma, simülasyon düzeneğinde yer alan faktörlerle örneklem büyüklüğü, test uzunluğu, ankor madde sayısı, test formları arasındaki güçlük düzeyi farklılıkları ve replikasyon düzeyleriyle sınırlıdır.



BÖLÜM II: YÖNTEM

1. Araştırmanın Türü

Bu araştırmada alt test formu içeren testlerde alt testler ortak maddelere dayalı eşitleme yöntemleriyle eşitlenerek elde edilen eşitleme hatalarının karşılaştırılması amaçlanmıştır. Böylece farklı örneklem büyüklüğü, farklı test uzunluğu ve farklı güçlüğü sahip test formlarının alt test/genişletilmiş alt test puanlarının eşitlenmesiyle elde edilen eşitlemenin standart hataları, eşitlemenin yanlılık değerleri saptanarak bu konuda yapılan kuramsal çalışmalara katkıda bulunulması amaçlanmaktadır. Araştırma bu yönüyle temel araştırma niteliğindedir.

2. Araştırmanın Değişkenleri

Bu araştırmada incelenen bağımlı değişkenler; alt testlerin artı değer gösterme oranı, eşitlemenin standart hatası eşitlemenin yanlılık ve eşitleme hatası değerleridir. Araştırmanın bağımsız değişkenleri eşitleme yöntemleri, örneklem büyüklükleri, alt test uzunluğu ve test formları arasındaki güçlük düzeyleridir. Eşitleme yöntemleri altı düzeyli olarak; Birim Dönüşüm, Lineer Zincirilmiş Eşitleme Yöntemi, Dairesel-yay Eşitleme Yöntemi, Braun/Holland Eşitleme Yöntemi şeklinde belirlenmiştir. Örneklem büyüklüğü sayısı 500, 200, 100, 50, 25, 20 kişi olarak değişimlenmiştir. Test formundaki alt testlerin uzunlukları aynı olmak koşuluyla 10, 15, 30, 50 ve 80 olarak belirlenmiştir. Test formları arasındaki güçlük düzeyleri 0.00, 0.40 ve 0.70 olarak belirlenmiş olup ankor test uzunluğu da toplam testin %40'ı uzunluğunda değişimlenmiştir. Araştırma kapsamındaki faktörler ve düzeyleri ilgili literatürde yer alan benzer çalışmalar ve bu çalışmalardan elde edilen bulgular dikkate alınarak belirlenmiştir.

Örnelem büyüklüğü; İlgili literatür araştırıldığında, örneklem büyüklüğünün eşitleme hatasını etkileyen önemli bir faktör olduğu görülmektedir. Parshall ve diğerleri (1995), araştırmalarında 15, 25, 50 ve 100 örneklem büyüklüğüne sahip örneklemle çalışmasını gerçekleştirmişler. Skaggs'ın (2005) değişimlediği örneklem büyüklükleri 25, 50, 75, 100, 150 ve 200 şeklindeyken, Kim ve diğerleri (2008) örneklem büyüklüğünü 15, 25, 50, 100 ve 200 olarak değişimlemiştir. Kim ve Livingstone (2010), gerçekleştirdikleri çalışmada örneklem büyüklüklerini 10, 25, 50 ve 100 olarak belirtmişlerdir. Babcock ve diğerleri (2012) çalışmalarında örneklem büyüklüğünü 20, 50 ve 80 olmak üzere üç düzeyde değişimlemiştir.

Bu araştırma da eşitlemenin standart hatasına, eşitlemenin yanlılığına ve eşitleme hatasına örneklem büyüklüklerinin etkisini belirlemek amacıyla, 20, 25, 50, 100, 200 ve 500 olacak şekilde örneklem büyüklüğü altı farklı düzey olarak değişimlenmiştir.

Test formları arasındaki güçlük düzeyi; Test eşitlemede test formlarının paralel olması istenmektedir. Ancak paralel testlerin oluşturulması oldukça zordur. Test formları arasında güçlük düzeyi farkı da test eşitlemeyi etkileyen faktörlerden biridir. Güçlük düzeyleri farklarının değişimlendiği test eşitleme çalışmalarına bakıldığında Parshall ve diğerleri (1995), ortalama güçlük farkı 0.02, 0.17, 0.26, 0.33, 0.44 SMD olan beş çift form kullanmışlardır. Skaggs (2005) çalışmasında 0.10 SMD, Kim, von Davier ve Haberman (2008), formlar ve puan dağılımları arasındaki farklılığın önemli olduğu durumlarda, özdeş eşitlemenin uygun olmayacağını belirtmişlerdir. Livingstone ve Kim (2009), 0.36 SMD güçlük farkına sahip bir çift test kullanmışlardır. Bunlarla birlikte, Livingstone ve Kim (2010), güçlük düzeyleri 0.17-0.30 SMD arası değişen dört çift test, Babcock (2012) 0.0, 0.35 ve 0.75 SMD ortalama güçlük düzeyine sahip test formlarını kullanmışlardır.

Yapılacak çalışma da eşitlemenin standart hatasına, eşitleme yanlılığına ve eşitleme hatasına örneklem büyüklüklerinin etkisini belirlemek amacıyla, 0.00, 0.40 ve 0.70 olacak şekilde test formları arasındaki güçlük düzeyi farkı değişimlenmiştir.

Test uzunluğu, Test formlarının uzunluğu arttıkça, testin güvenilirliğinin arttığı görülmektedir. Düşük güvenilirliğe sahip testlerde eşitlemenin standart hatası yüksek, yüksek güvenilirliğe sahip testlerde ise, standart hata küçük olacaktır. Dolayısıyla yapılan eşitlemeleri etkileyen faktörlerden bir diğeri de testteki madde sayısıdır. Test uzunluğunun değiştiği çalışmalarda Skaggs (2005), 50 maddelik; Livingston ve Kim (2009 ve 2010), 110 maddelik; Babcock (2012), 100 maddelik maddelik test formlarını kullanmışlardır. Ayrıca Gök (2012), 30, 60 ve 80 maddeden oluşan test formlarını kullanmış ve test uzunluğunun, test eşitlemenin doğruluğuna etkisini gözlemlemiştir.

Bu çalışma kapsamında, oluşturulacak testin alt test uzunlukları her alt test aynı sayıda test maddesine sahip olacak şekilde 10, 15, 30, 50 ve 80 maddeden oluşturulmuştur. Ayrıca iki alt test kullanılarak alt testler oluşturulacağından toplam test uzunluğu, 20, 30, 60, 100 ve 180 maddeden oluşmuştur.

Ankor Test Uzunluğu; Ortak maddelere dayalı desenlerde test eşitlemesi ortak maddeler yardımıyla gerçekleşmektedir. Ortak maddelerden oluşan ankor test eşitlenecek testlerin içeriğini kapsayacak şekilde olmalı ve eşitlenecek testlerle olabildiğince yüksek korelasyon sağlanmalıdır. Ortak maddelere dayalı test eşitleme için ankor madde sayısı eşitlemeye etkisi vardır. Angof (1984) ankor madde sayısının eşitlenecek testlerin %20 si olması gerektiğini belirtmiştir. Sinharay ve Holland (2007) ankor testin minyatür, mini ve midi versiyonları üzerinde yaptıkları çalışmalarında ankor madde sayısını %40, %44 ve %45 şeklinde değiştirmişlerdir. Puan ve Liang (2011) ankor test kullanarak yaptıkları alt testlerin eşitlemesi çalışmalarında ise ortak madde sayısını eşitlenecek testlerin %50'si ve %25'i olacak şekilde belirtmişler ve bu değişimlenen düzeyler alt testlerde de %50 ve %25 olacak şekilde düzenlenmiştir. Sinharay ve Haberman (2012) çalışmalarında ankor testin uzunluğunu %50 ve %55 oranında değiştirmişlerdir.

Çalışmada ankor test uzunluğu ilgili literatürde referans alınarak eşitlenecek testlerin %40 oranında olacak şekilde tek düzeyde değiştirilmiştir.

Korelasyon Düzeyleri; Alt test formlarının artı değer özelliğine sahip olması konusunda gerekli şartların belirgin olarak sınırları çizilmemiştir. Ancak korelasyon düzeyinin artı değer özelliği üzerinde etkiye sahip olduğu araştırmacılar tarafından savunulmaktadır (Sinharay, 2010a; Sinharay, 2010b, Sinharay ve diğerleri, 2007a). Sinharay (2010a) çalışmada 0.70, 0.75, 0.80, 0.85, 0.90 ve 0.95 korelasyon değerlerini kullanarak alt test formlarının artı değer özelliğine sahip olma durumlarını araştırmıştır. Yüksek korelasyona sahip olan alt testlerin artı değer özelliği göstermediğini gözlemlemiştir.

Çalışmada artı değer özelliğinin incelendiği literatür referans alınarak alt test formları arasında 0.70, 0.80 ve 0.90 korelasyon düzeyleri değiştirilmiştir.

Eşitleme yöntemleri; Kullanılacak eşitleme yöntemleri, kullanılan kurama, veri toplama desenlerine ve örneklem büyüklüklerine göre değişmektedir. İlgili literatür incelendiğinde Skaggs (2005) düzgülendirme yapılmamış ve farklı düzeylerde düzgülendirme yapılmış zincirilmiş eşit yüzdelli, özdeş, ortalama ve doğrusal eşitleme yöntemlerini, Kim ve diğerleri (2008), zincirilmiş doğrusal, özdeş ve sentetik fonksiyon yöntemlerini; Livingston ve Kim (2009), zincirilmiş eşit yüzdelli ve doğrusal, ortalama, Tucker, Levine ve Dairesel-yay yöntemlerini kullanmışlardır. Benzer şekilde, Livingston ve Kim (2010), düzgülendirme yapılmış eşit yüzdelli, doğrusal, ortalama simetrik ve basitleştirilmiş Dairesel yay yöntemlerini, bir başka çalışmada ise zincirilmiş eşit yüzdelli, doğrusal, ortalama, simetrik ve basitleştirilmiş Dairesel yay ve özdeş eşitleme yöntemlerini kullanmıştır. Babcock ve diğerleri (2012) zincirilmiş eşit yüzdelli eşitleme, Tucker lineer

eşitleme, ortalama eşitleme, nominal ağırlıklandırılmış ortalama, birim dönüşüm, sentetik eşitleme ve Dairesel-yay eşitleme yöntemini kullanarak eşitlemeyi gerçekleştirmişlerdir.

Çalışmada NEAT desen kullanılarak alt test verileri elde edilecektir. Alt testler Lineer Zincirlenmiş Eşitleme Yöntemi, Dairesel Yay Eşitleme Yöntemi, Braun/Holland Eşitleme Yöntemi ve Birim Eşitleme Yöntemi kullanılarak eşitleme yapılacaktır. Eşitleme yöntemlerinden elde edilen eşitlemenin standart hatası, eşitleme yanlılığı ve eşitleme hataları diğer faktörlerde göz önünde bulundurularak hangi faktörlerce daha iyi sonuçlar verileceği incelenecektir.

3. Veri Üretimi

Bu çalışmada değişimlenen faktörler ve bu faktörlerin düzeylerine göre veri üretimi 2 parametrelili lojistik Madde Tepki Kuramı modeli kullanılarak gerçekleştirilmiştir. Üretilen veriler 1-0 olacak şekilde iki kategoriktir. Bu amaç doğrultusunda 4 farklı birey yetenek dağılımı kullanılmıştır. X formu X1 ve X2 alt testlerinden oluşurken, Y formu Y1 ve Y2 alt testlerinden oluşmaktadır. X1 ve Y1 için bireylerin yetenek dağılımı ortalaması 0 standart sapması 1 olan standart normal dağılımdan elde edilmiştir. X2 ve Y2 alt testleri için kullanılacak olan yetenek dağılımlarının daha önce X1 ve Y1 için üretilen yetenek dağılımlarıyla, istenilen korelasyonları gösterecek şekilde üretilmesi sağlanmıştır.

X ve Y test formunun alt testleri ve ankor testlerinde yer alan maddelerin ayırt edicilik parametreleri (a) ortalaması -0.15 standart sapması 0.14 olan log-normal dağılımdan elde edilmiştir.

X formunun iki alt testinin ve ankor testin güçlük (b) parametreleri ortalaması -2, standart sapması 2 olan uniform dağılımdan elde edilmiştir. Y formunun alt testlerinin güçlük (b) parametresi ise X formundaki güçlük (b) parametresine, aşağıdaki tabloda da belirtildiği gibi 0.00, 0.40 ve 0.70 değerleri eklenerek elde edilmiştir.

Tablo 7'de sunulmuş olan faktörler ve düzeylerinin çaprazlanmasından elde edilen bütün koşullar için 500 replikasyon yapılarak veri üretimi yapılmıştır.

Tablo 7'de belirtildiği gibi değişimlenen faktörelere uygun toplamda örneklem büyüklüğü düzeyi (6)x alt test uzunluğu düzeyi (5)x test formları arasındaki güçlük düzeyi (3)x alt testler arasındaki korelasyon düzeyi (3) olmak üzere 270 adet X formu ve 270 adet Y formu üretilmiştir.

Tablo 7: Araştırmada Değişimlenen Faktörler	
Değişimlenen Faktör	Düzy
Örneklem Büyüklüğü	20
	25
	50
	100
	200
	500
Alt Test Uzunluğu	10
	15
	30
	50
	80
Test Formları Arasındaki Güçlük Düzeyi	0.00
	0.40
	0.70
Ankor Test Uzunluğu	%40
Alt Testler Arasındaki Korelasyon Düzeyi	0.70
	0.80
	0.90
Eşitleme Yöntemleri	<p>Birim Eşitleme Yöntemi</p> <p>Braun/Holland Eşitleme Yöntemi</p> <p>Dairesel Yay Eşitleme Yöntemi</p> <p>Zincirlenmiş Lineer Eşitleme Yöntemi</p>

4. Verilerin Analizi

Faktör düzeylerine uygun olarak 270 adet X ve 270 adet Y formuna ilişkin 1-0 şeklinde iki kategorili veri oluşturulmuştur. Çalışmada denk olmayan gruplarda ortak maddelere dayalı eşitleme deseni kullanılmıştır. Veri matrisinin satırları toplanarak X ve Y formunun alt test puanları, ankor testin alt test ham puanları elde edilmiştir.

Çalışmada alt testleri arasında 0.70, 0.80 ve 0.90 korelasyona sahip test formları oluşturulmuştur. Oluşturulan testlerin her iki alt testi de artı değer özelliğine sahip X ve Y test formları seçilerek kullanılmıştır.

Dört farklı eşitleme düzeneği oluşturulup aynı özelliği ölçen alt testlerin birbiriyle eşitlenmesi gerçekleştirilmiştir. İlk eşitleme düzeneğinde; aynı faktör düzeylerinde üretilen X ve Y formlarının artı değer özelliğine sahip her iki alt testinden elde edilen gözlenen puanlar kullanılarak eşitleme yapılmıştır. İkinci eşitleme düzeneğinde ise, aynı faktör düzeylerinde üretilen X ve Y formlarının artı değer özelliği gösteren her iki alt testlerinin, genişletilmiş alt test puanları kullanılarak eşitleme gerçekleştirilmiştir. Benzer şekilde üçüncü eşitleme düzeneğinde; aynı faktör düzeylerinde üretilen X ve Y formlarının artı değer özelliği gösteren her iki alt testinin gözlenen puanları kullanılarak eşitleme yapılmıştır. Son eşitleme düzeneğinde ise aynı faktör düzeylerinde üretilen X ve Y formlarının artı değer özelliği göstermeyen her iki alt testlerinin genişletilmiş alt test puanları kullanılarak eşitleme gerçekleştirilmiştir. Tüm eşitlemeler R.3.1.1. programının “equate” paketi kullanılarak yapılmıştır.

Ankor test formu, alt test uzunluklarının %40 oranında değişilerek her bir alt test eşitlemesi için ankor testin ilgili alt testi kullanılarak gerçekleştirilmiştir. Tüm örneklem büyüklükleri için 100 replikasyon yapılarak örneklem çekilmiştir. Her replikasyon için X formu, Y formu ölçeğine eşitlenerek, ham puan ve genişletilmiş puan düzeyinde eşitlenmiş puanlar elde edilmiştir. Eşitleme aynı özelliklere sahip test formları ve test puanlarıyla gerçekleştirilmiştir.

Değişimlenen faktörlerin tüm düzeylerinde eşitlenmiş alt test formlarının, her bir eşitleme yöntemi için eşitlemenin standart hatası (SEE), eşitlemenin yanlılığı (BIAS) ve eşitleme hataları (RMSE) değerleri hesaplanmıştır.

Çalışmada yer alan örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı şeklindeki faktörlerin eşitlemenin standart hatasına, yanlılığına ve eşitleme hatasına temel etkisini incelemek için tüm faktörlerin her bir yöntem için ortalama standart hataları, ortalama yanlılıkları ve ortalama eşitleme hataları hesaplanmış ve temel etki grafikleri oluşturulmuştur. Ayrıca temel etkileri incelenen faktörlerin eşitlemenin standart hatasına, eşitleme yanlılığına ve eşitleme hatasına olan ortak etkileri incelenerek ortak etki grafiği oluşturulmuştur.

Değerlendirme ölçütü olarak kullanılan eşitlemenin standart hatası, eşitleme yanlılığı ve eşitleme hatası değerleri hesaplanmıştır. Eşitleme hata değerlerinin hesaplanmasına ilişkin formüller aşağıda verilmiştir

Eşitlemenin standart hatası (SEE_i) ve eşitleme yanlılığı sırasıyla;

$$\hat{e}_y(x_i) = \frac{\sum_r \hat{e}_{y_r}(x_i)}{R}$$

$$\begin{aligned} \bar{s}d_i &= SEE_i \\ &= \sqrt{\frac{\sum_r [\hat{e}_{y_r}(x_i) - \hat{e}_{y.}(x_i)]^2}{R-1}} \end{aligned} \quad (62)$$

$$\begin{aligned} \hat{d}_i &= \frac{\sum_{r=1}^r [\hat{e}_{y_r}(x_i) - \hat{e}_{y.}(x_i)]}{R} \end{aligned} \quad (63)$$

eşitlikleri kullanılarak hesaplanmıştır. Eşitliklerdeki i , puan sayısını, R , replikasyon sayısını, $\hat{e}_{Y_r}(x_i)$; r örneklemindeki hücrelerin birinden hesaplanmış ham puan eşdeğerini göstermektedir.





BÖLÜM III: BULGULAR

Bu bölümde artı değer özelliğine sahip alt testlerde çeşitli faktörlerin (örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük farkı) test eşitleme yöntemlerinden elde edilen eşitlemenin standart hatasına (SEE), eşitlemenin yanlılığına (BIAS) ve eşitlemenin hatasına (RMSE) olan temel etkisi ve ortak etkilerine ait bulgulara yer verilmiştir. Temel etki grafikleri tek boyutluyken ortak etki grafikleri çok boyutludur.

1. Ortak maddelere dayalı eşitleme deseninde, artı değer özelliğine sahip olan alt testlerin, alt test /genişletilmiş alt test puanları kullanılarak yapılan eşitleme yöntemlerinin (birim eşitleme, zincirlenmiş lineer eşitleme, Braun/Holland eşitleme ve dairesel yay eşitleme yöntemleri), çeşitli faktörlere (örneklem büyüklüğü, test formları arasındaki ortalama güçlük düzeyi farkı, alt test uzunlukları) göre elde edilen eşitlemenin standart hatası (SEE), eşitlemenin yanlılığı (BIAS) ve eşitleme hatası (RMSE) kestirimleri nasıl değişmektedir?

Artı değer özelliğine sahip iki alt test elde edilerek aynı özelliğe sahip alt test formları arasında ortak maddelere dayalı eşitleme deseni kullanılarak birim eşitleme, zincirlenmiş lineer eşitleme, Braun/Holland yöntemi ve dairesel yay eşitleme yöntemleriyle eşitleme gerçekleştirilmiştir. Aşağıda çeşitli faktörler altında eşitleme hatalarının yöntemlere göre incelendiği temel etki ve ortak etki grafikleri, elde edilen bulgulara yer verilmiştir.

Temel etki grafikleri tek satır ve üç sütun olmak üzere toplamda üç grafikten oluşmaktadır. Grafikteki sütunlar sırasıyla eşitleme hatası (RMSE), eşitleme yanlılığı (BIAS) ve eşitlemenin standart hatasına (SEE) dair elde edilen sonuçları göstermektedir. Grafiklerde yer alan eşitleme yöntemlerinden IDEN (birim eşitleme), LIN (zincirlenmiş lineer eşitleme), MEAN (Braun/Holland eşitleme) ve CIRC (daireysel yay eşitleme yöntemini) göstermektedir. Ortak etki grafiklerinde ise örneklem büyüklüğü, alt test uzunluğu ve güçlük düzeyi faktörleri altında; 0,70, 080 ve 0.90 koreelayona sahip alt testlerin eşitlemesi sonucu eşitleme yöntemlerin SEE, BIAS ve RMSE değerlerine ait bulgular yer almaktadır. Grafiklerin sağ tarafında yer alan renk cetvelinde MEAN, LIN, CIRC ve IDEN etiketleriyle yöntemlerin hataları temsil edilmektedir.

1.1. Çeşitli faktörlerin, ortak maddelere dayalı eşitleme deseninde, artı değer özelliğine sahip olan alt testlerin, alt test/genişletilmiş alt test puanları kullanılarak yapılan eşitleme yöntemlerinden (birim eşitleme, zincirlenmiş lineer eşitleme, Braun/Holland eşitleme ve dairesel yay eşitleme yöntemleri) elde edilen eşitlemelerin eşitleme standart hatasına (SEE), eşitleme yanlılığına (BIAS) ve eşitleme hatasına (RMSE) temel etkisi nasıldır?

Aşağıda artı değer özelliğine sahip alt testlerin, alt test / genişletilmiş alt test puanları kullanılarak yapılan eşitleme yöntemlerinden elde edilen eşitleme hatalarının çeşitli faktörler altındaki davranışlarıyla ilgili bulgulara yer verilmiştir.

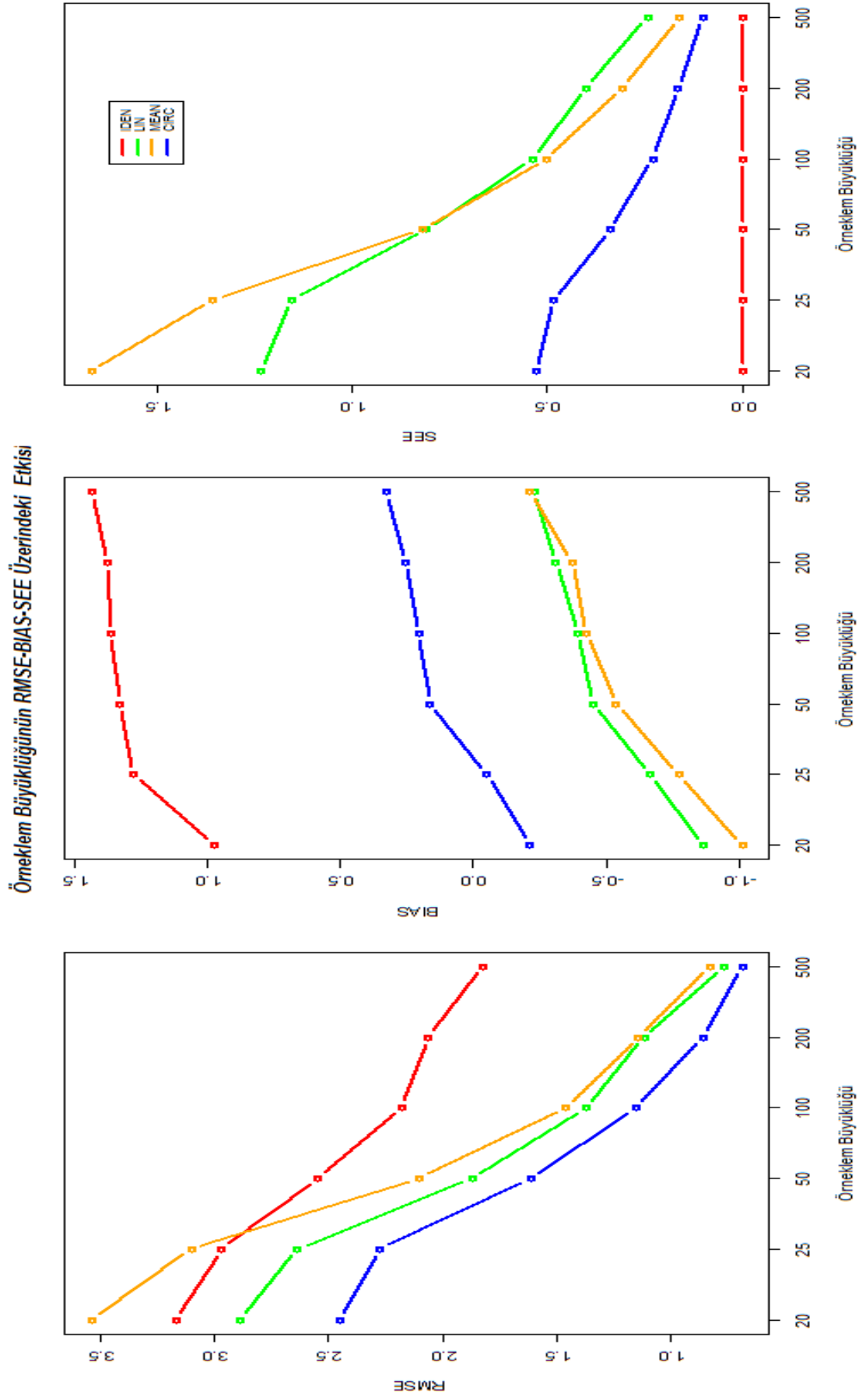
1.1.1.1. Örneklem büyüklüğünün, alt test puanları eşitlemelerinde, eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE) temel etkisi nasıldır?

Grafik1'te RMSE değerlerinin örneklem büyüklüğüne göre davranışlarını gösteren grafik incelendiğinde; örneklem büyüklüğü değiştiğinde eşitleme yöntemlerinden elde edilen RMSE değerlerinin de farklılaştığı görülmektedir. Örneklem büyüklüğü arttıkça birim eşitleme yönteminin RMSE değerlerinde azalma gözlenmektedir. Tüm örneklem büyüklüklerinde en düşük RMSE değerine dairesel yay eşitleme yönteminin sahip olduğu görülmektedir. 20 ve 25 örneklem büyüklüklerinde Braun/Holland eşitleme yöntemi en yüksek RMSE değerine sahipken; 50, 100, 200 ve 500 örneklem büyüklüklerinde birim eşitlemeden sonra en yüksek RMSE değerlerine sahip eşitleme yönteminin Braun/Holland eşitleme olduğu gözlenmektedir. Genel olarak örneklem büyüklüğü arttıkça yöntemlerin RMSE değerlerinde azalma görülmektedir.

BIAS deęerlerinin deęişiminin yer aldığı Grafik1 incelendięinde; birim eşitleme yönteminin BIAS deęerleri 20 örneklem büyüklüęü dışındaki dięer deęerlerde farklılaşmamaktadır. 20 örneklem büyüklüęü dışındaki dięer deęerlerde en yüksek BIAS deęerlerini birim eşitleme yönteminin verdięi görülmektedir. Tüm örneklem büyüklüklerinde en düşük BIAS deęerini dairesel yay eşitleme yöntemi verirken; örneklem büyüklüęü arttıkça zincirlenmiş lineer eşitleme ve Braun/Holland eşitleme yöntemlerinin BIAS deęerlerinde azalma gözlenmektedir.

Grafik1'te SEE deęerlerinin davranışları incelendięinde; birim eşitleme yönteminin tüm örneklem büyüklüklerinde SEE deęerlerinin sıfırı olduęu gözlenmektedir. 20, 25 ve 50 örneklem büyüklüklerinde en yüksek SEE deęerlerini Braun/Holland eşitleme yöntemi verirken; 100, 200 ve 500 örneklem büyüklüklerinde en yüksek SEE deęerlerine zincirlenmiş lineer eşitleme yönteminin sahip olduęu görülmektedir. Genel olarak örneklem büyüklüęü arttıkça yöntemlerin SEE deęerlerinde azalmanın meydana geldięi gözlenmektedir.





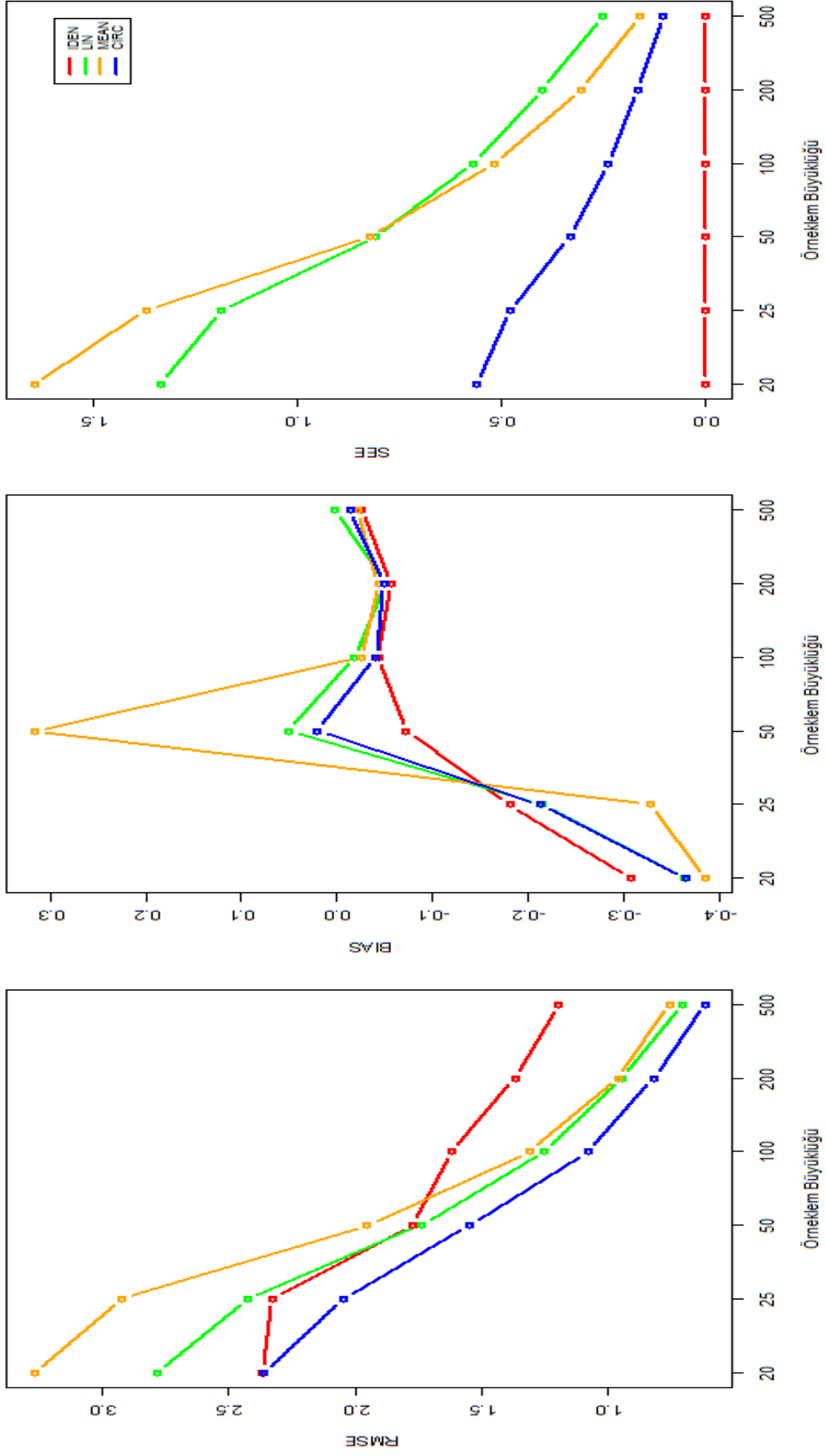
Grafik 1: Çeşitli örneklem büyüklüklerine göre, artı değer özelliğine sahip birinci alt testin alt test puanları kullanılarak gerçekleştirilen eşitlemenin eşitleme yöntemlerinden elde edilen RMSE-BIAS ve SEE değerleri

Grafik2'deki RMSE değerlerine ait olan grafik incelendiğinde; 20, 25 ve 50 örneklem büyüklüklerinde en yüksek RMSE değerlerine Braun/Holland eşitleme yönteminin sahip olduğu gözlenirken, 100, 200 ve 500 örneklem büyüklüklerinde ise en yüksek RMSE değerlerine birim eşitleme yöntemi sahiptir. Tüm örneklem büyüklüklerinde en düşük RMSE değerlerine dairesel yay eşitleme yönteminin sahip olduğu görülmektedir.

BIAS değeri davranışlarının gözlemlendiği Grafik2'te; en düşük BIAS değerini 50 örneklem büyüklüğünde zincirlenmiş lineer eşitleme yönteminden elde edildiği görülürken; 100, 200 ve 500 örneklem büyüklüklerinde yöntemlerin BIAS değerlerinin birbirlerine oldukça yakın olduğu gözlenmektedir. 20, 25 ve 50 örneklem büyüklüklerinde en yüksek BIAS değerine sahip olan yöntem Braun/Holland eşitlemedir.

Grafik2'nin standart hata (SEE) değerleri incelendiğinde, birim eşitleme yöntemi tüm örneklem büyüklüklerinde en düşük SEE değerine sahiptir. 20 ve 25 örneklem büyüklüğünde en yüksek SEE değerine Braun/Hollan eşitleme yöntemi sahipken; 50 örneklem büyüklüğünde en yüksek SEE değerini Braun/Holland ve zincirlenmiş lineer eşitleme yöntemi vermektedir. 100, 200 ve 500 örneklem büyüklüklerinde en yüksek SEE değerine zincirlenmiş lineer eşitleme yöntemi sahiptir. Birim eşitleme yönteminden sonra tüm örneklem büyüklüklerinde en düşük SEE değerlerine dairesel yay eşitleme yöntemi sahiptir. Genel olarak örneklem büyüklüğü arttıkça birim eşitleme yöntemi dışındaki diğer yöntemlerin SEE değerlerinde azalmanın meydana geldiği görülmektedir.

Örneklem Büyüklüğünün RMSE-BIAS-SEE Üzerindeki Etkisi



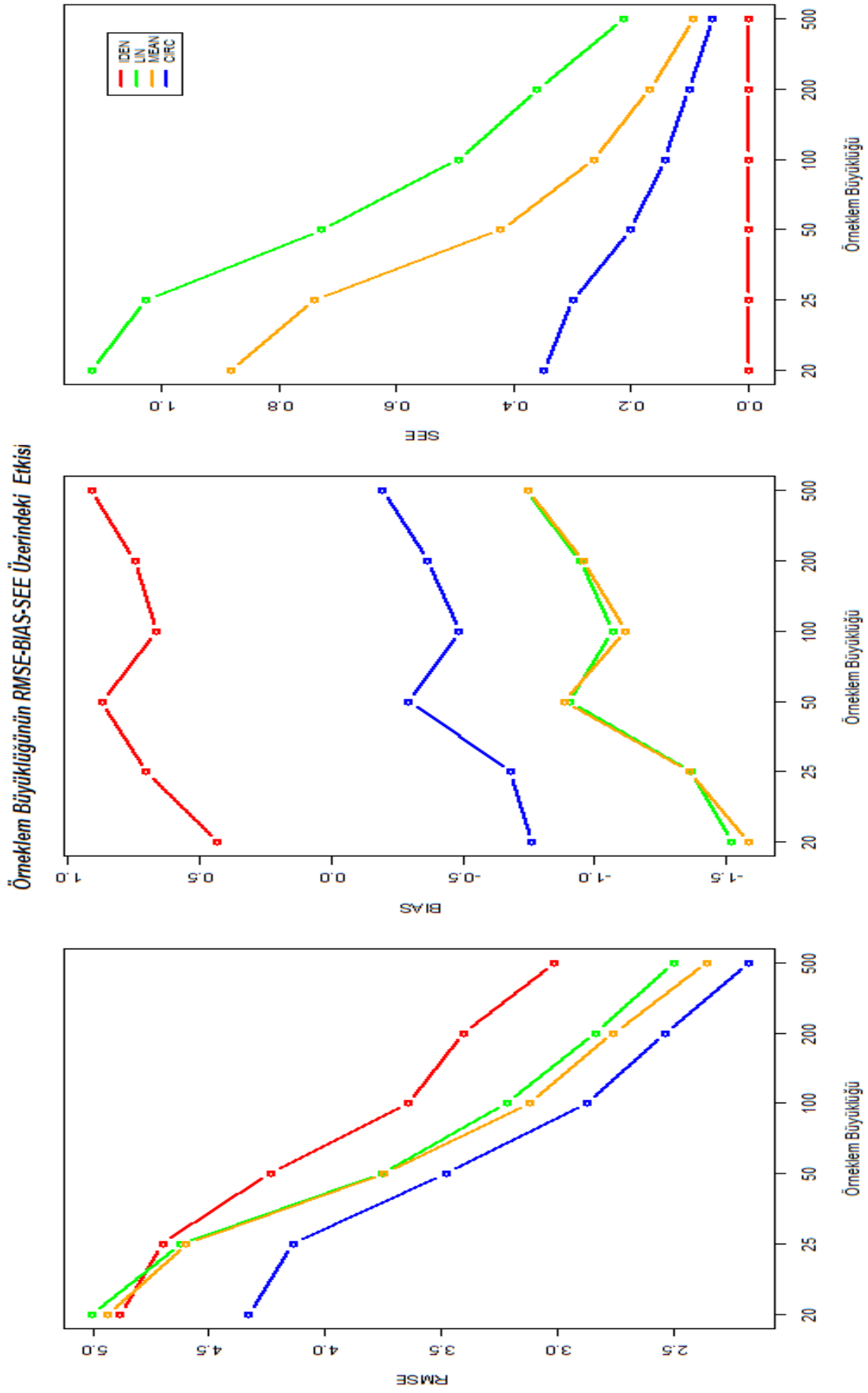
Grafik 2: Çeşitli örneklem büyüklüklerine göre, artı değer özelliğine sahip ikinci alt testlerin alt test puanları kullanılarak gerçekleştirilen eşitlemenin eşitleme yöntemlerinden elde edilen RMSE-BIAS ve SEE değerleri

1.1.1.2. Örneklem büyüklüğünün, genişletilmiş alt test puanları eşitlemelerinde, eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE) temel etkisi nasıldır?

Grafik3’de yöntemlerin RMSE değerlerindeki değişimi gösteren grafik incelendiğinde; 20 örneklem büyüklüğünde en yüksek RMSE değerini zincirleşmiş lineer eşitleme yöntemi verirken; diğer tüm örneklem büyüklüklerinde en yüksek RMSE değerine birim eşitleme yöntemi sahiptir. Tüm örneklem büyüklüklerinde en düşük RMSE değeri dairesel yay eşitleme yönteminde gözlenmektedir. Genel olarak örneklem büyüklüğü arttıkça yöntemlerin RMSE değerlerinde azalma meydana geldiği görülmektedir.

Grafik3’ün BIAS değerlerindeki değişimin yer aldığı grafik incelendiğinde; 20 örneklem büyüklüğünde en yüksek BIAS değeri Braun/Holland ve zincirleşmiş lineer eşitleme yönteminde gözlenirken diğer tüm örneklem büyüklüklerinde ise en yüksek BIAS değerini birim eşitleme yönteminin verdiği görülmektedir. En düşük BIAS değerine sahip eşitleme yöntemi dairesel yay eşitleme yöntemidir.

SEE değerlerindeki değişimlerin yer aldığı Grafik3 incelendiğinde; tüm örneklem büyüklüklerinde en yüksek SEE değerine zincirleşmiş lineer eşitleme yöntemi sahiptir. Tüm örneklem büyüklüklerinde en düşük SEE değerine ise birim eşitleme yönteminden sonra dairesel yay eşitleme yöntemi sahiptir. Genel olarak birim eşitleme yöntemi dışındaki tüm yöntemlerde örneklem büyüklüğü arttıkça SEE değerlerinde azalma meydana geldiği görülmektedir.



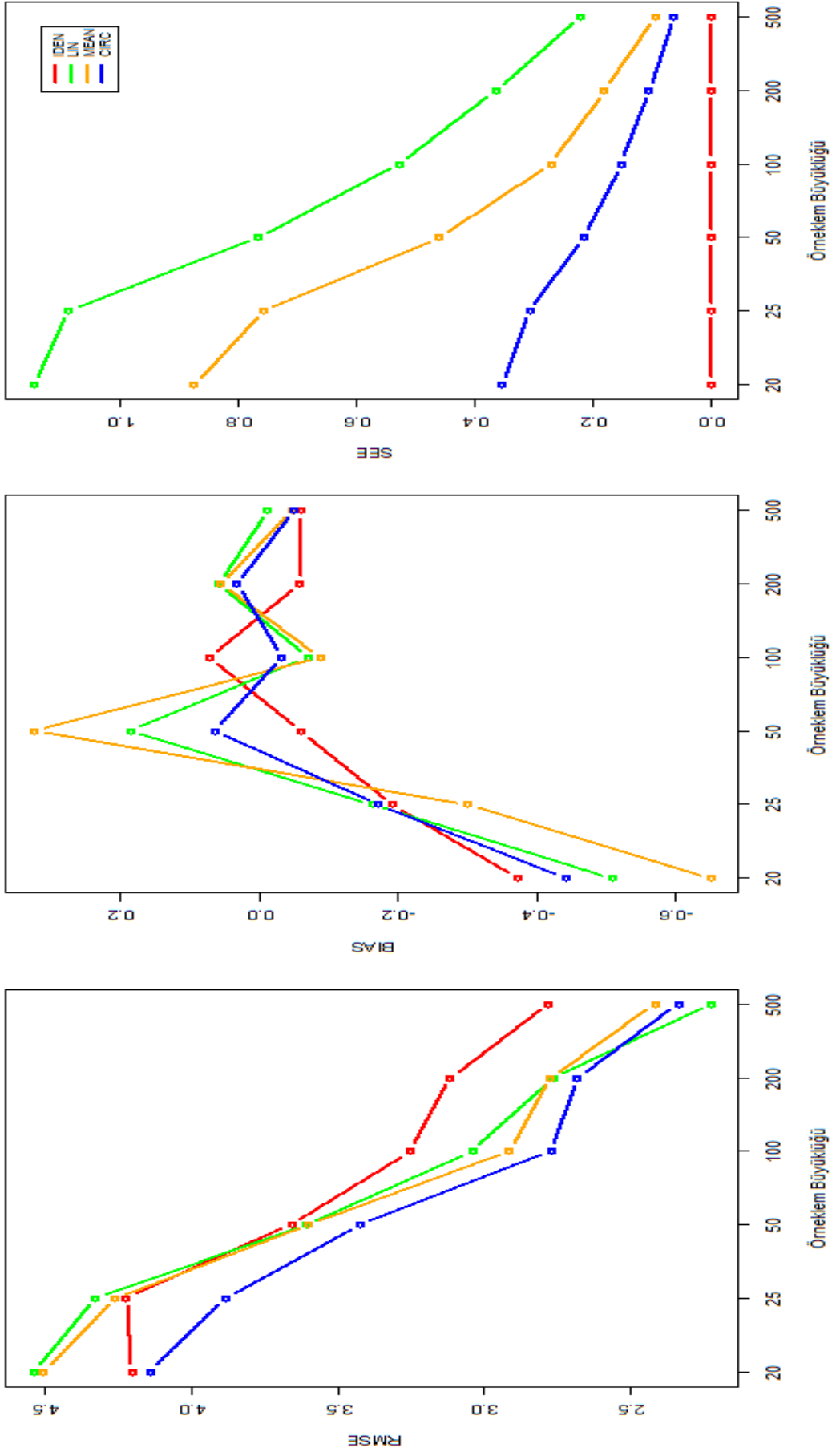
Grafik 3: Çeşitli örneklem büyüklüklerine göre, artı değer özelliklerine sahip birinci alt testin genişletilmiş alt test puanları kullanılarak gerçekleştirilen eşitlemenin eşitleme yöntemlerinden elde edilen RMSE-BIAS ve SEE değerleri

RMSE davranışlarının yer aldığı Grafik4 incelendiğinde; 20 ve 25 örneklem büyüklüklerinde en yüksek RMSE değerinin zincirleşmiş lineer eşitleme yönteminde olduğu görülmektedir. 50, 100, 200 ve 500 örneklem büyüklüklerinde en yüksek RMSE değerine sahip olan yöntem birim eşitleme yöntemidir. Genel olarak en düşük RMSE değerine sahip yöntem dairesel yay eşitleme yöntemiyken; 500 örneklem büyüklüğünde en düşük RMSE değeri zincirleşmiş lineer eşitleme yönteminde gözlenmiştir. Örneklem büyüklüğü arttıkça tüm yöntemlerin RMSE değerlerinde azalmanın meydana geldiği görülmektedir.

Grafik4 BIAS davranışlarına göre, 20 örneklem büyüklüğü dışındaki tüm değerlerde genel olarak en düşük BIAS değerine dairesel yay eşitleme yöntemi sahiptir. 20, 25, 50 ve 100 örneklem büyüklüklerinde en yüksek BIAS değeri Braun/Holland yönteminde gözlenmektedir. 200 örneklem büyüklüğünde birim eşitleme hariç diğer tüm yöntemler birbirine yakın BIAS değerleri vermektedir.

SEE'nin yer aldığı Grafik4 incelendiğinde; tüm örneklem büyüklüklerinde en düşük SEE değerine birim eşitleme yönteminin sahip olduğu görülmektedir. Birim eşitleme yönteminden sonra en düşük SEE değeri dairesel yay eşitleme yönteminde gözlenmektedir. Tüm örneklem büyüklüklerinde en yüksek SEE değerini zincirleşmiş lineer eşitleme yönteminin verdiği görülmektedir.

Örneklem Büyüklüğünün RMSE-BIAS-SEE Üzerindeki Etkisi



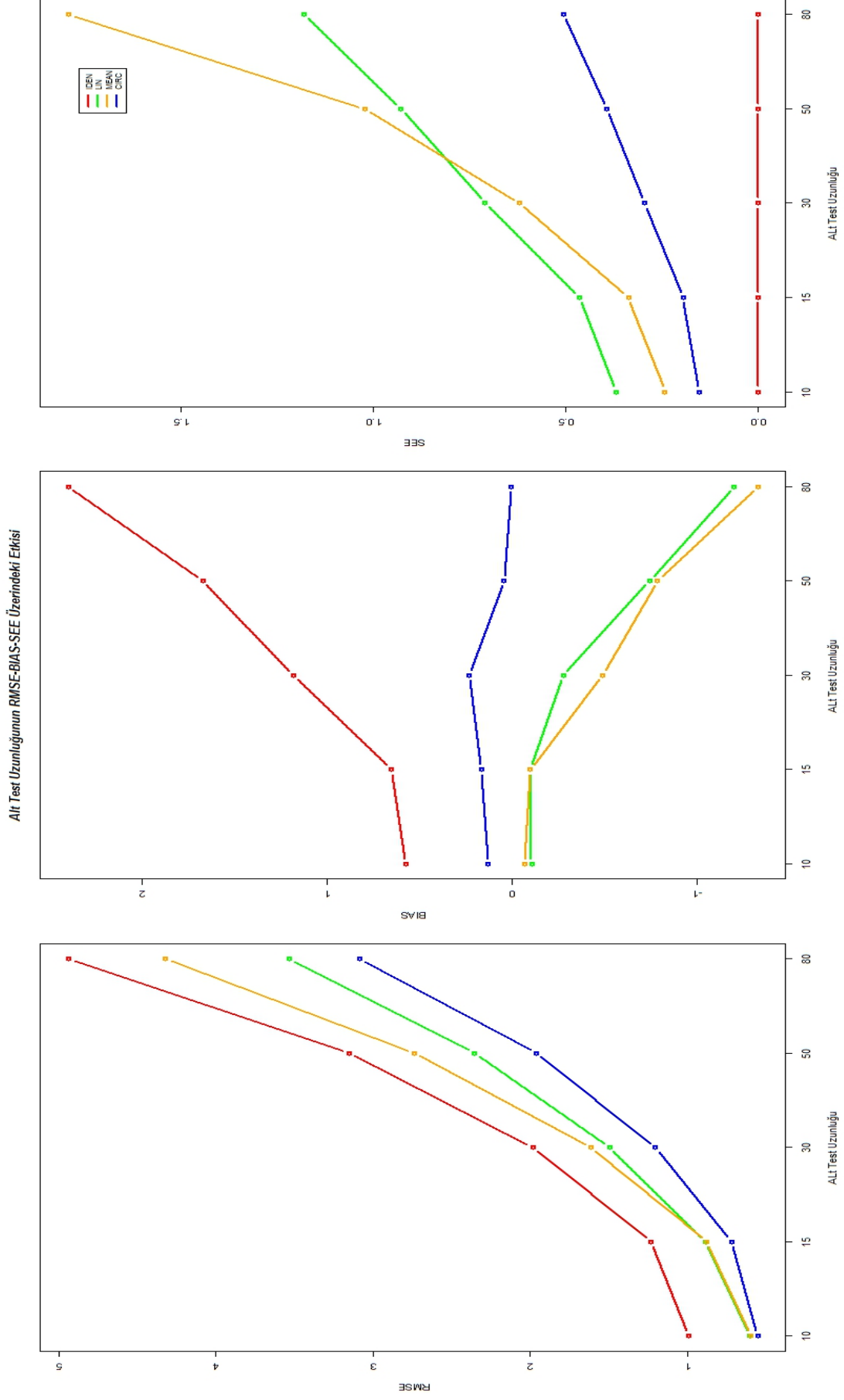
Grafik 4: Çeşitli örneklem büyüklüklerine göre, artı değer özelliğine sahip ikinci alt test genişletilmiş alt test puanla kullanılarak gerçekleştirilen eşitlemenin eşitleme yöntemlerinden elde edilen RMSE-BIAS ve SEE değerleri

1.1.2.1. Alt test uzunluğunun, alt test puanları eşitlemelerinde, eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE), eşitlemenin yanlılığına (BIAS) ve eşitlemenin standart hatasına (SEE), temel etkisi nasıldır?

Yöntemlerin RMSE değerlerinin yer aldığı Grafik5 incelendiğinde; en yüksek RMSE değeri birim eşitleme yönteminde gözlenirken, tüm alt test uzunluklarında en düşük RMSE değeri dairesel yay eşitleme yönteminde görülmektedir. Alt test uzunluğu arttıkça tüm eşitleme yöntemlerinin RMSE değerlerinde artma meydana geldiği gözlenmektedir.

Grafik5’de BIAS değerleri incelendiğinde, tüm alt test uzunluklarında en düşük BIAS değerine sahip eşitleme yöntemi dairesel yay eşitleme yöntemidir. Genel olarak 10 ve 15 test uzunluğunda yöntemlerin BIAS değerlerinde farklılaşma meydana gelmemektedir. 20, 30, 50 ve 80 test uzunluğunda en yüksek BIAS değeri birim eşitleme yönteminde görülmektedir.

Yöntemlerin SEE değerlerinin yer aldığı Grafik5 incelendiğinde, tüm alt test uzunluklarında en düşük SEE değerine dairesel yay eşitleme yönteminin sahip olduğu görülmektedir. 10, 15 ve 30 alt test uzunluklarında en yüksek SEE değerini zincirleşmiş lineer eşitleme yönteminin verdiği gözlenmektedir. 50 ve 80 test uzunluklarında ise en yüksek SEE değerine Braun/Holland eşitleme yöntemi sahiptir. Genel olarak alt test uzunluğu arttıkça birim eşitleme yöntemi dışındaki tüm yöntemlerde SEE değerlerinin arttığı gözlenmiştir.

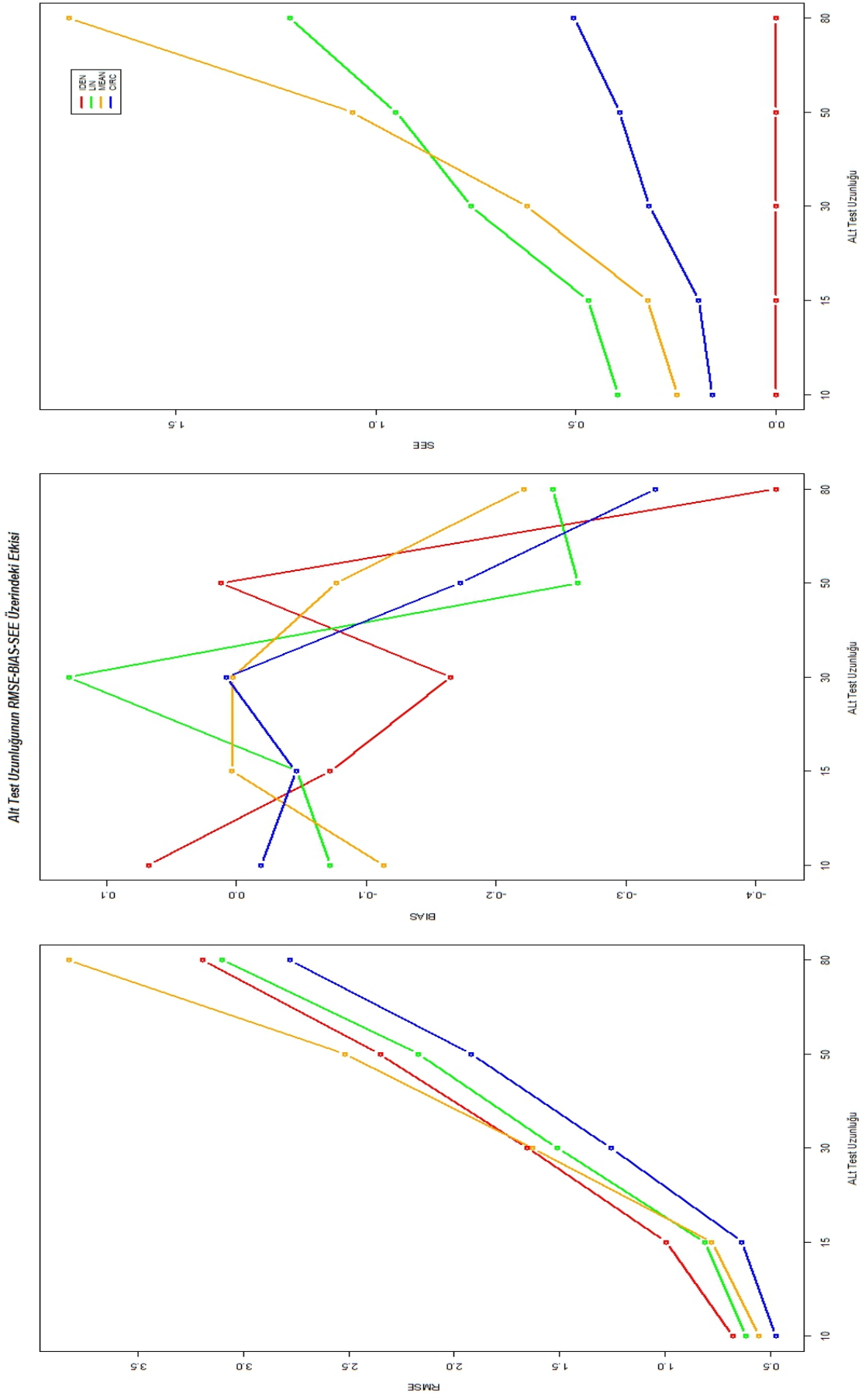


Grafik 5: Çeşitli alt test uzunluklarına göre, artı değer özelliğine sahip birinci alt testin alt test puanları kullanılarak gerçekleştirilen eşitlemenin eşitleme yöntemlerinden elde edilen RMSE-BIAS ve SEE değerleri

Grafik6'da yer alan yöntemlerin RMSE deęerleri incelendięinde; 10, 15 ve 30 alt test uzunluklarında en yüksek RMSE deęerini veren eşitleme yöntemi birim eşitlmedir. 50 ve 80 alt test uzunluklarında en yüksek RMSE deęerine Braun/Holland eşitleme yönteminin sahip olduęu görölmektedir. Tüm alt test uzunluklarında en düşük RMSE deęerine dairesel yay eşitleme yöntemi sahiptir. Genel olarak tüm yöntemlerde alt test uzunluęu arttıkça RMSE deęerlerinde artışın meydana geldięi gözlenmektedir.

Yöntemlerin BIAS deęerlerinin yer aldıę Grafik6 incelendięinde; genel olarak alt testlerde yer alan madde sayısı arttıkça yöntemlerin BIAS deęerlerinde artma görölmektedir. 50 maddeye sahip alt test uzunluęu dışında tüm alt test uzunluklarında en yüksek BIAS deęerine birim eşitleme yöntemi sahipken, 50 maddeye sahip alt test uzunluęunda en düşük RMSE deęeri birim eşitleme yönteminde görölmektedir.

Grafik6'da SEE deęerleri incelendięinde; tüm alt test uzunluklarında en düşük SEE deęerine birim eşitleme yöntemi sahiptir. Madde sayısının 10, 15 ve 30 olduęu durumlarda en yüksek SEE deęerine zincirlenmiř lineer eşitleme yöntemi sahiptir. 50 ve 80 madde sayılarında ise en yüksek SEE deęeri Braun/Holland eşitleme yönteminde gözlenmektedir. Genel olarak alt test uzunluęu arttıkça yöntemlerin SEE deęerlerinde artmanın meydana geldięi görölmektedir.



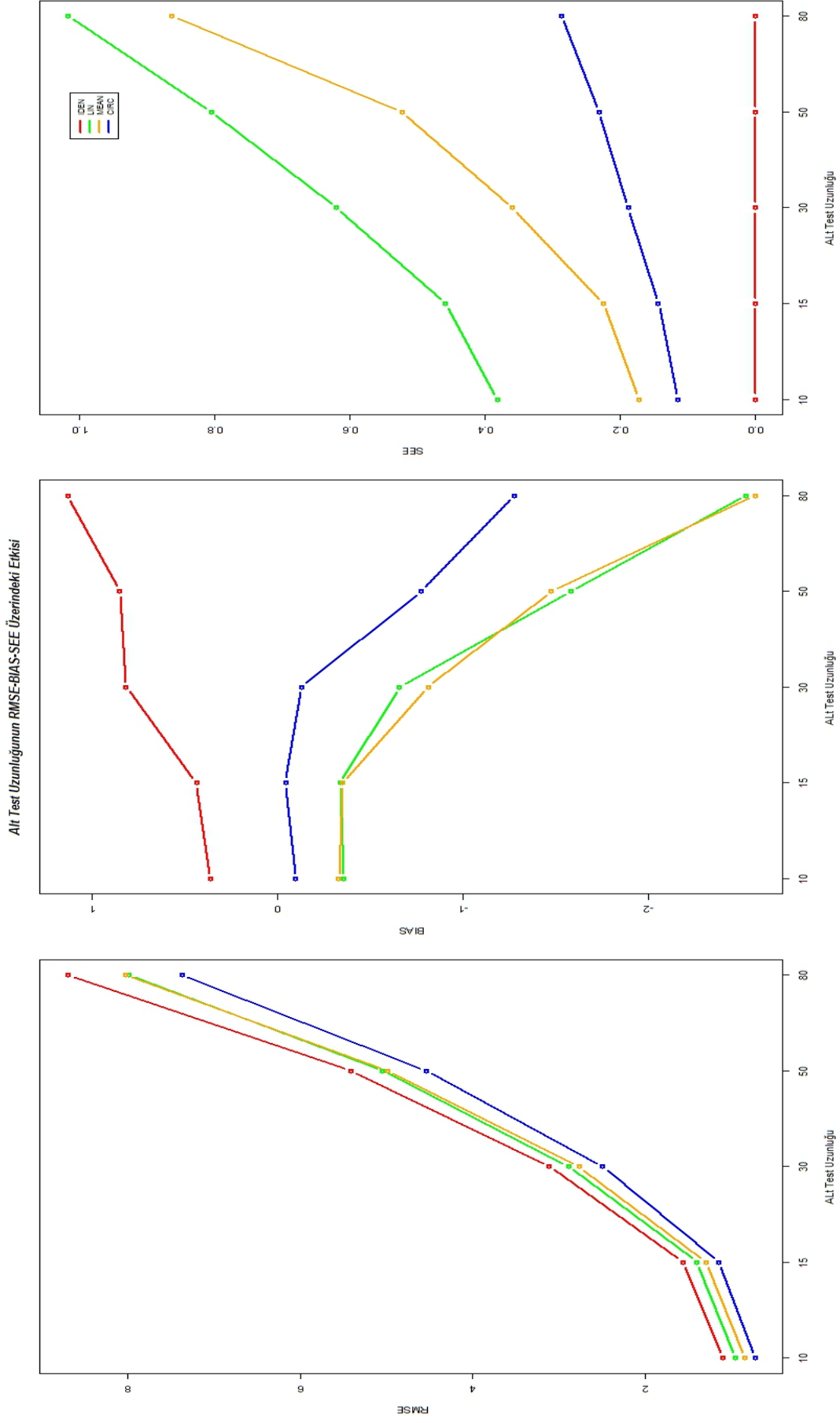
Grafik 6: Çeşitli alt test uzunluklarına göre, artı değer özelliğine sahip ikinci alt testin alt test puanları kullanılarak gerçekleştirilen eşitlemenin eşitleme yöntemlerinden elde edilen RMSE-BIAS ve SEE değerleri

1.1.2.2. Alt test uzunluğunun, alt test puanları eşitlemelerinde, eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE), eşitlemenin yanlılığına (BIAS) ve eşitlemenin standart hatasına (SEE), temel etkisi nasıldır?

Grafik7’de yöntemlerin RMSE davranışları incelendiğinde; genel olarak yöntemlerin RMSE değerleri birbirine yakındır. Tüm alt test uzunluklarında en yüksek RMSE değeri birim dönüşümde gözlenmektedir. Endüşük RMSE değeri adiresel yay eşitleme yönteminde görülmektedir. Tüm yöntemlerde alt test uzunluğu arttıkça RMSE değerlerinin arttığı görülmektedir.

Grafik7’de yöntemlerin BIAS davranışları incelendiğinde; 10 ve 15 madde sayısına sahip alt testlerde yöntemlerin BIAS değerlerinde belirgin bir farklılaşma meydana gelmediği görülmektedir. Tüm alt test uzunluklarında en düşük BIAS değerine dairesel yay eşitleme yöntemi sahiptir. 50 madde sayısına sahip alt testlerde en yüksek BIAS değerine zincirlenmiş lineer eşitleme sahipken, 80 madde sayısına sahip alt testlerde en yüksek BIAS değeri Braun/Holland eşitleme yönteminde görülmektedir.

Grafik7’de SEE değerlerindeki değişimlere bakıldığında; tüm alt test uzunluklarında en düşük SEE değerine birim eşitleme yöntemi sahiptir. Tüm alt test uzunluklarında en yüksek SEE değerine zincirlenmiş lineer eşitleme yönteminin sahip olduğu görülmektedir. Birim eşitleme haricinde, diğer yöntemlerde alt test uzunluğu arttıkça yöntemlerin SEE değerlerinde artmanın meydana geldiği gözlenmektedir.



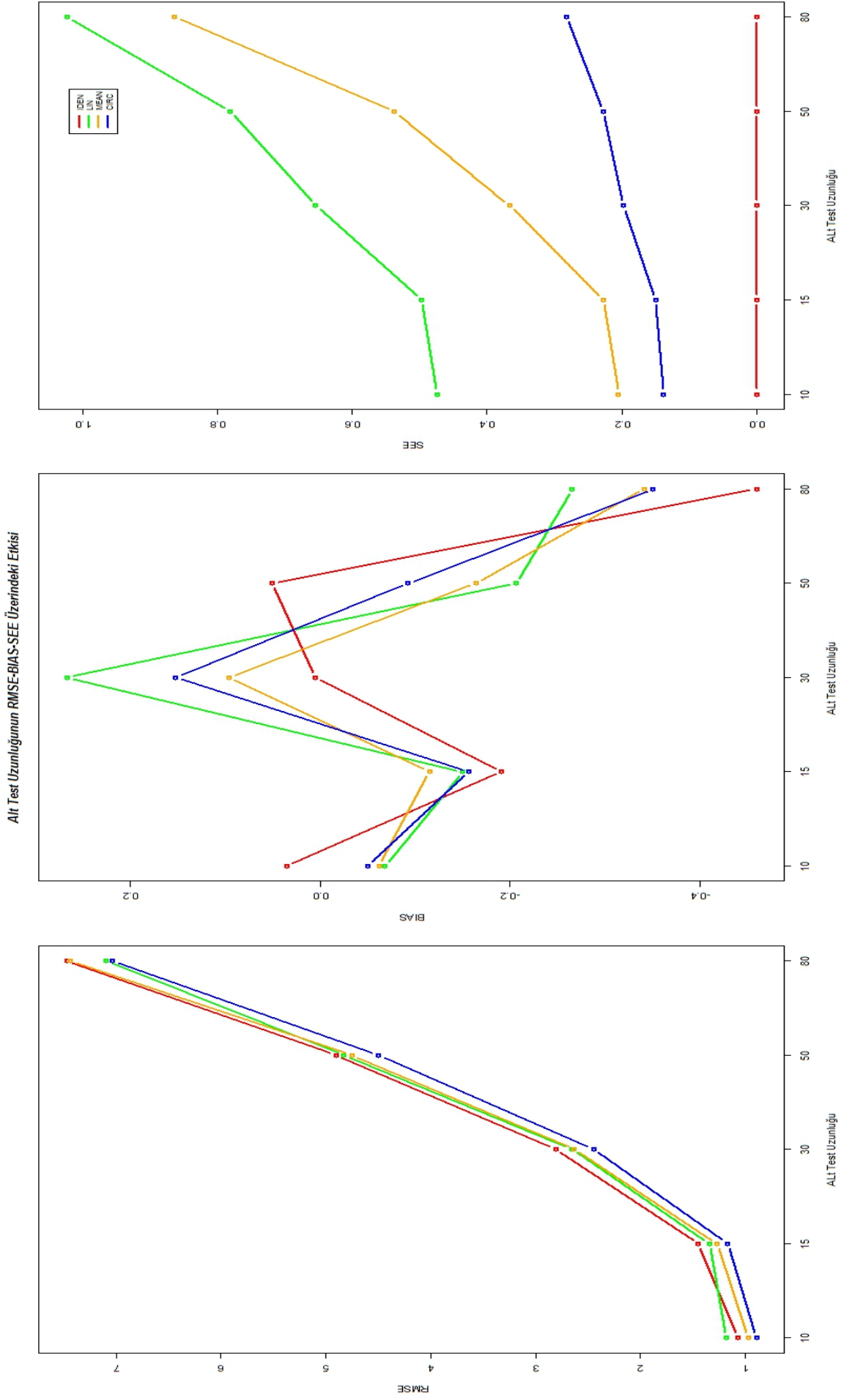
Grafik 7: Çeşitli alt test uzunluklarına göre, artı değer özelliğine sahip birinci alt test genişletilmiş alt test puanları kullanılarak gerçekleştirilen eşitlemenin eşitleme yöntemlerinden elde edilen RMSE-BIAS ve SEE değerleri

Grafik8'de yer alan yöntemlerin RMSE davranışların incelendiğinde; yöntemlerin RMSE değerleri birbirine yakın değerler aldığı görülmektedir. En yüksek RMSE değeri genel olarak birim eşitlemede gözlenmektedir. Tüm alt test uzunluklarında en düşük RMSE değerine dairesel yay eşitleme yönteminin sahip olduğu görülmektedir.

Grafik8'de yer alan yöntemlerin BIAS değişimleri incelendiğinde; alt test uzunluğu 30'a kadar arttığı durumlarda yöntemlerin BIAS değerleri belirli bir davranış göstermemektedir. Madde sayısı 50 olduğunda birim eşitleme dışındaki diğer yöntemlerin BIAS değerlerinde artma meydana geldiği görülmektedir. Madde sayısı 80 olduğunda bu artma tüm eşitleme yöntemlerinde gözlenmektedir.

Grafik8'de SEE değerlerindeki değişimlere bakıldığında; tüm alt test uzunluklarında en düşük SEE değerine birim eşitleme yöntemi sahiptir. Tüm alt test uzunluklarında en yüksek SEE değerine zincirilmiş lineer eşitleme yönteminin sahip olduğu görülmektedir. Birim eşitleme haricinde, diğer yöntemlerde alt test uzunluğu arttıkça yöntemlerin SEE değerlerinde artmanın meydana geldiği gözlenmektedir.





Grafik 8: Çeşitli alt test uzunluklarına göre, artı değer özelliğine sahip ikinci alt testin genişletilmiş alt test puanlarını kullanarak gerçekleştirilen eşitlemenin eşitleme yöntemlerinden elde edilen RMSE-BIAS ve SEE değerleri

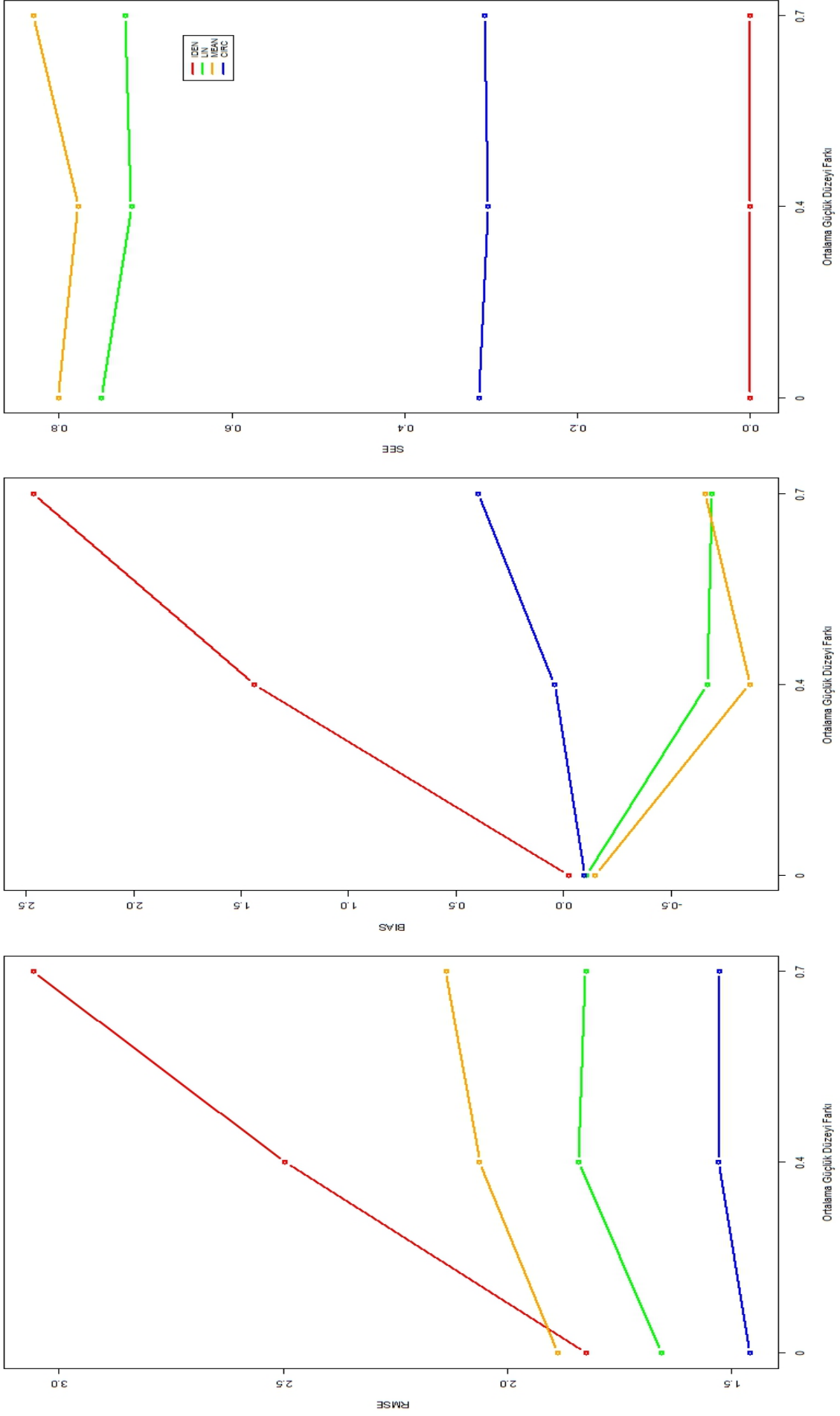
1.1.3.1. Test formları arasındaki ortalama güçlük farkının, alt test puanları eşitlemelerinde, eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE), eşitlemenin yanlılığına (BIAS) ve eşitlemenin standart hatasına (SEE), temel etkisi nasıldır?

Grafik9'da yöntemlerin RMSE davranışları incelendiğinde; test formları arasında ortalama güçlük düzeyi farkı olmadığında en yüksek RMSE değeri Braun/Holand eşitleme yönteminde gözlenmektedir. Ortalama güçlük düzeyinin 0.4 ve 0.7 olduğu durumlarda en yüksek RMSE değeri birim eşitleme yönteminde görülmektedir. Tüm ortalama güçlük düzeyi farklarında en düşük RMSE değeri dairesel yay eşitleme yönteminde gözlenmektedir. Ortalama güçlük düzeyi farkı arttıkça birim eşitleme yönteminin RMSE değerlerinde artma görülmektedir. Diğer yöntemlerde ortalama güçlük düzeyindeki farklılaşma RMSE değerlerinde belirgin bir değişimin meydana gelmediğini göstermektedir.

Grafik9' de yer alan yöntemlerin BIAS değerlerine bakıldığında; ortalama güçlük düzeyi farkı 0.0 olduğunda tüm yöntemlerin BIAS değerleri aynı ve en küçük değerindedir. Ortalama güçlük düzeyi farkının tüm değişimlerinde en düşük BIAS değerine dairesel yay eşitleme yöntemi sahiptir. Ortalama güçlük farkının 0.0 dışında tüm düzeylerinde en yüksek BIAS değerini birim eşitleme yöntemi göstermektedir.

Grafik9'de yöntemlerin SEE davranışları incelendiğinde; ortalama güçlük farkı düzeyleri değiştikçe yöntemlerin SEE değerlerinde belirgin bir farklılaşma gözlenmemektedir. En düşük SEE değerine birim eşitleme yöntemi sahiptir. En yüksek SEE değerini Braun/Holland eşitleme yöntemi göstermektedir.

Test Formları Arasındaki Ortalama Güçlük Düzeyleri Farkının RMSE-BIAS-SEE Üzerindeki E



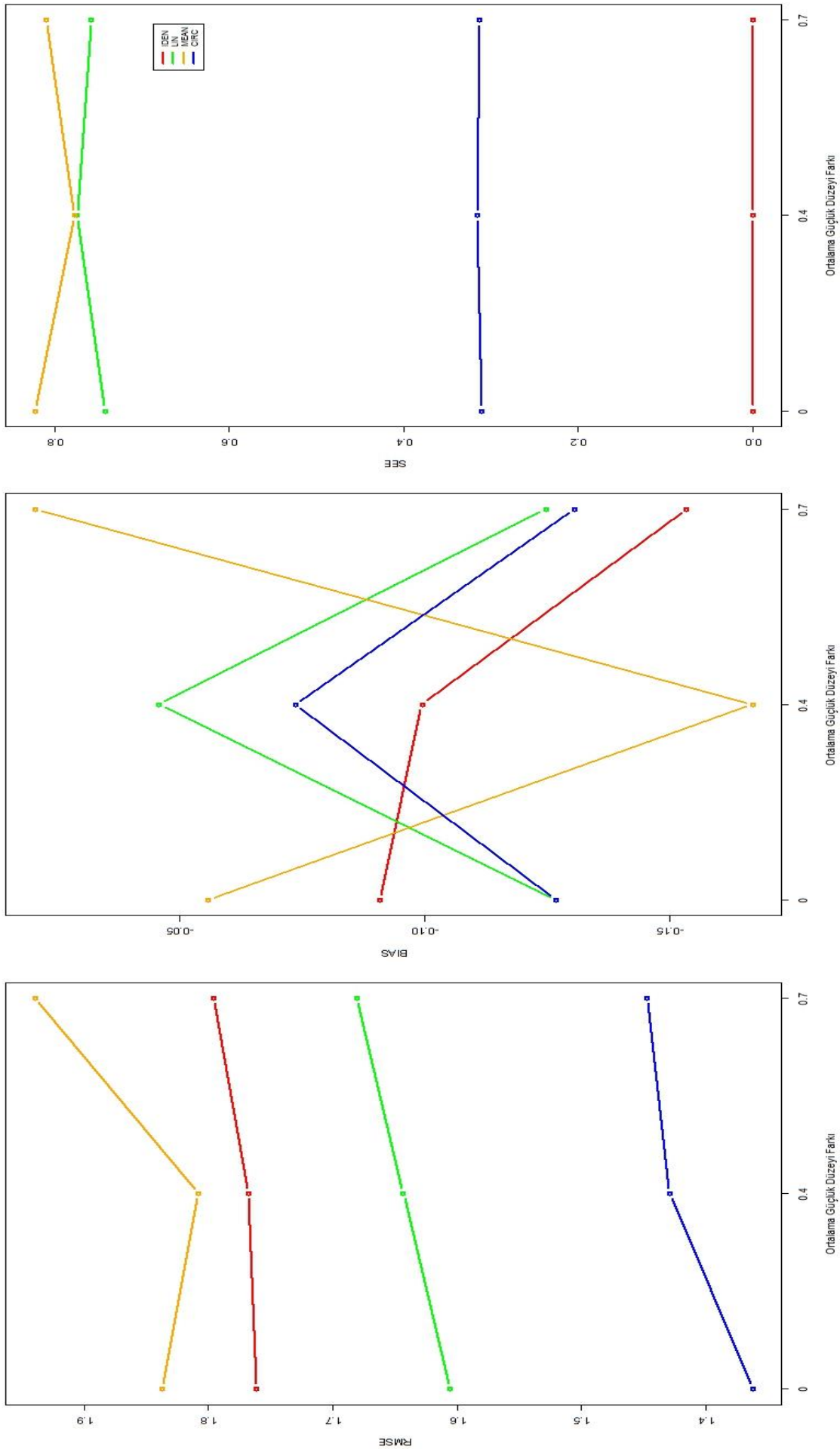
Grafik 9: Çeşitli ortalama düzeyi farklarına göre, artı değer özelliğine sahip birinci alt testin alt test puanları kullanılarak gerçekleştirilen eşitlemenin eşitleme yöntemlerinden elde edilen RMSE-BIAS ve SEE değerleri

Grafik10'da yöntemlerin RMSE davranışları incelendiğinde; ortalama güçlük düzeyi farkının tüm değişimlemelerinde yöntemlerin RMSE değerlerinde belirgin bir artış görülmemektedir. Enyüksek RMSE değerini Braun/Holland eşitleme yöntemi verirken en düşük RMSE değerini ise dairesel yay eşitleme yönteminin verdiği gözlenmektedir.

Grafik10'da yer alan BIAS değerlerinin değişimi incelendiğinde; ortalama güçlük düzeyi farkı 0.0 olduğunda en yüksek BIAS değerini dairesel yay ve zincirilmiş lineer eşitleme yöntemi verirken; en düşük BIAS değerini ise Braun/Holland eşitleme yönteminin verdiği gözlenmektedir. 0.4 ortalama güçlük düzeyi farkında en yüksek RMSE değeri Braun/Holland eşitleme yönteminde gözlenirken; en düşük BIAS değeri zincirilmiş lineer eşitleme yönteminde görülmektedir. Ortalama güçlük düzeyi farkı 0.7 olduğunda en yüksek BIAS değeri birim eşitleme yönteminde gözlenirken; en düşük BIAS değeri Braun/Holland yönteminde görülmektedir.

Grafik10'da yöntemlerin SEE değerleri incelendiğinde; tüm ortalama güçlük düzeyi farklarında en yüksek SEE değerine Braun/Holland eşitleme yöntemi sahiptir. En düşük SEE değerine birim eşitleme yöntemi sahiptir. Ortalama güçlük düzeyi farkı değiştikçe yöntemlerin SEE değerlerinde belirgin bir artmanın ya da azalmanın meydana gelmediği görülmektedir.

Test Formları Arasındaki Ortalama Güçlük Düzeyleri Farkının RMSE-BIAS-SEE Üzerindeki E



Grafik 10: Çeşitli ortalama düzeyi farklarına göre, artı değer özelliğine sahip ikinci alt testin alt test puanları kullanılarak gerçekleştirilen eşitlemenin eşitleme yöntemlerinden elde edilen RMSE-BIAS ve SEE değerleri

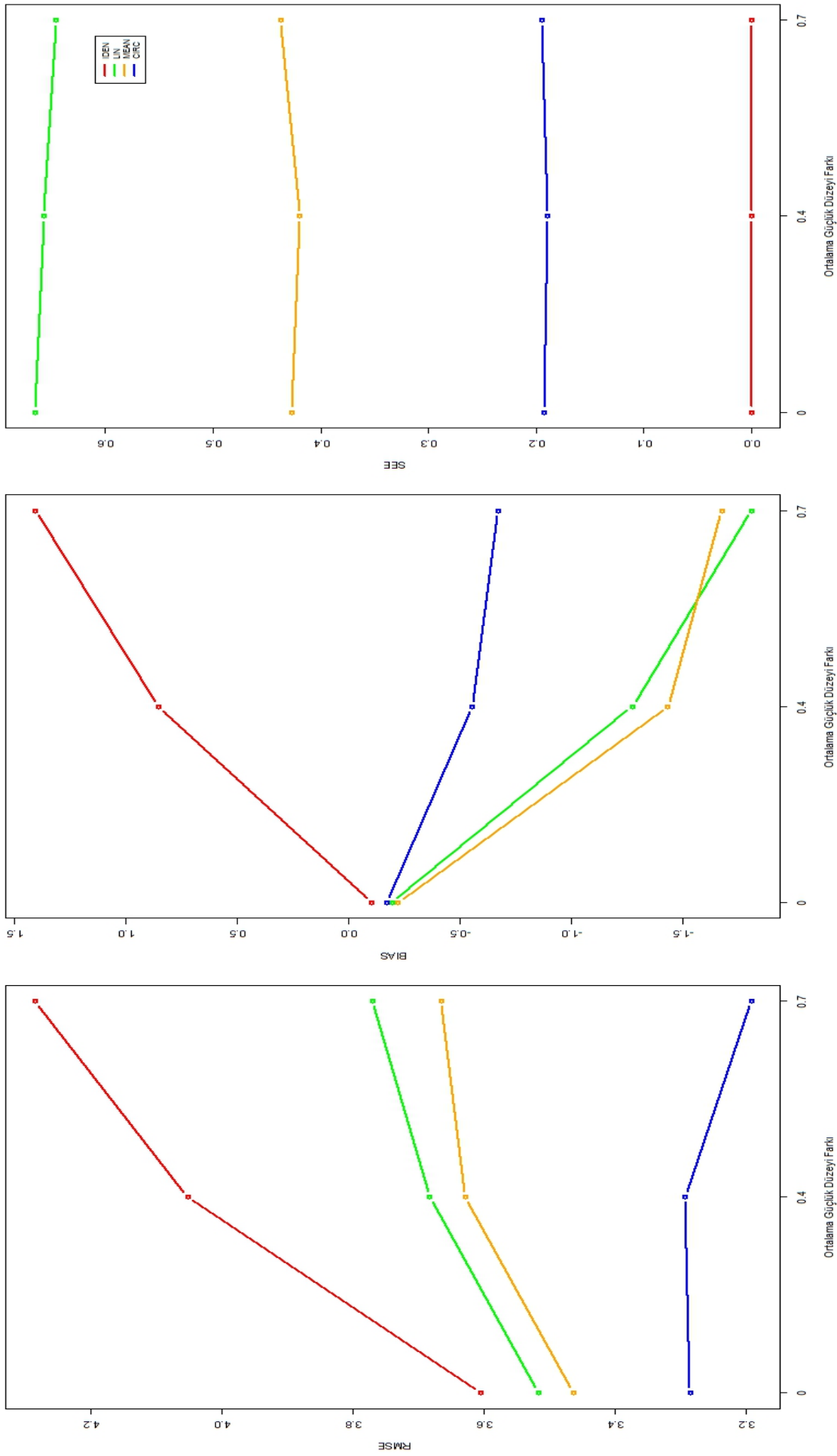
1.1.3.2. Test formları arasındaki ortalama güçlük farkının, genişletilmiş alt test puanları eşitlemelerinde, eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE), eşitlemenin yanlılığına (BIAS) ve eşitlemenin standart hatasına (SEE), temel etkisi nasıldır?

Grafik11' yöntemlerin RMSE değerleri incelendiğinde; en yüksek RMSE değeri birim dönüşümde görülmektedir. Değişimlenen ortalama güçlük düzeyi farkının tüm değerlerinde en düşük eşitleme hatasına dairesel yay eşitleme yöntemi sahiptir.

Grafik11'de yer alan yöntemlerin BIAS davranışları incelendiğinde; 0.0 ortalama güçlük düzeyi farkında tüm yöntemler yakın BIAS değerlerine sahipken; ortalama güçlük düzeyi farkı arttığında birim dönüşüm, zincirlenmiş lineer eşitleme yöntemlerinde eşitleme yanlılığı değerleri artmaktadır. 0.4 ortalama güçlük düzeyinde Braun/Holland yönteminin eşitleme yanlılığı değeri artarken 0.7 ortalama güçlük düzeyi farkında BIAS değerleri azalma göstermiştir. En düşük eşitleme yanlılığı değeri dairesel yay eşitleme yönteminde gözlenmiştir.

Grafik11'de yöntemlerin SEE değerleri incelendiğinde; ortalama güçlük düzeyi farklarında en yüksek SEE değerine zincirlenmiş lineer eşitleme yöntemi sahiptir. En düşük SEE değeri birim eşitleme yönteminde gözlenmektedir. Ortalama güçlük düzeyi farklarının değişimlendiği durumlarda sahip elde edilen eşitlemenin standart hatası değerlerinin kullanılan yöntemler için farklılaşmadığı görülmektedir.

Test Formları Arasındaki Ortalama Güçlük Düzeyleri Farkının RMSE-BIAS-SEE Üzerindeki E



Grafik 11: Çeşitli ortalama düzeyi farklarına göre, artı değer özelliğine sahip birinci alt testin genişletilmiş alt test puanları kullanılarak gerçekleştirilen eşitlemenin eşitleme yöntemlerinden elde edilen RMSE-BIAS ve SEE değerleri

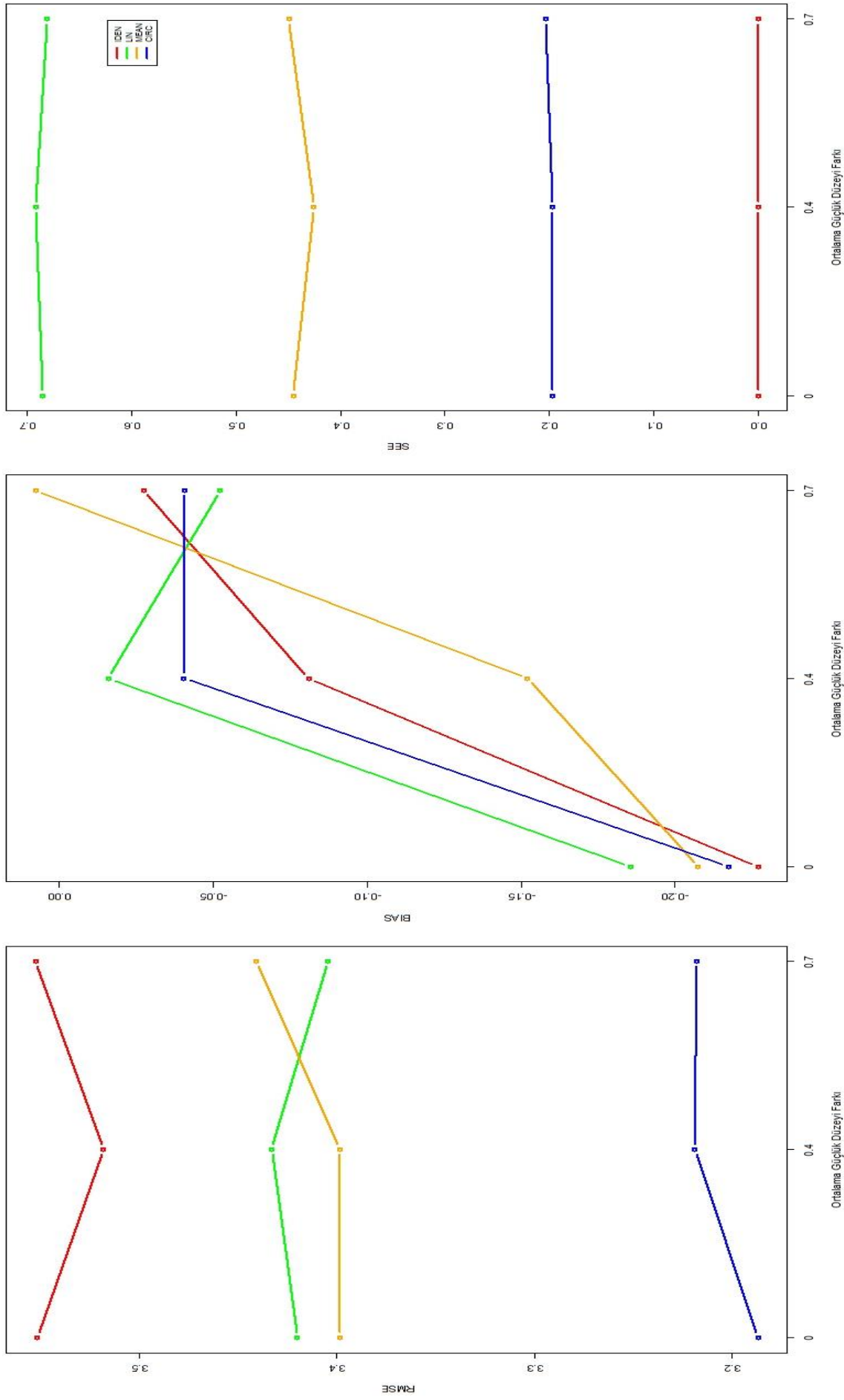
Grafik12’de yöntemlerin RMSE davranışları incelendiğinde; ortalama güçlük düzeyi farklarında eşitlemenin standart hatası değerleri tüm ortalama güçlük düzeyleri farklarında en yüksek RMSE değeri birim eşitleme yönteminde görülmektedir. Tüm ortalama güçlük düzeyleri farklarında en düşük eşitleme hatası değerleri dairesel yay eşitleme yönteminde görülmektedir.

Grafik12’de yer alan yöntemlerin BIAS değerleri incelendiğinde; ortalama güçlük düzeyi farkları arttıkça Braun/Holland eşitleme yönteminin BIAS değerlerinde azalmanın meydana geldiği görülmektedir. Braun/Holland dışındaki diğer tüm yöntemlerde, 0.7 ortalama güçlük düzeyi farkı haricindeki değerlerde artma meydana geldikçe yöntemlerin BIAS değerlerinde azalma gözlenmektedir; ayrıca en yüksek BIAS değerine birim eşitleme yöntemi sahipken, en düşük RMSE değerine ise zincirleşmiş lineer eşitleme sahiptir.

Grafik12’de yöntemlerin SEE değerleri incelendiğinde değişimlenen ortalama güçlük düzeyi farklarında genel olarak SEE değerlerinin farklılaşmadığı gözlenmektedir. En yüksek SEE değeri zincirleşmiş lineer eşitleme yönteminde görülmektedir. En düşük SEE değeri birim eşitleme yönteminde görülmektedir.



Test Formları Arasındaki Ortalama Güçlük Düzeyleri Farkının RMSE-BIAS-SEE Üzerindeki E



Grafik 12: Çeşitli ortalama düzeyi farklarına göre, artı değer özelliğine sahip ikinci alt testin genişletilmiş alt test puanları kullanılarak gerçekleştirilen eşitlemenin eşitleme yöntemlerinden elde edilen RMSE-BIAS ve SEE değerleri

2. Çeşitli faktörlerin, ortak maddelere dayalı eşitleme deseninde, artı değer özelliğine sahip olan alt testlerin, alt test /genişletilmiş alt test puanları kullanılarak yapılan eşitleme yöntemlerinden (birim eşitleme, zincirlenmiş lineer eşitleme, Braun/Holland eşitleme ve dairesel yay eşitleme yöntemleri) elde edilen eşitlemelerin eşitleme standart hatasına (SEE), eşitleme yanlılığına (BIAS) ve eşitleme hatasına (RMSE) ortak etkisi nasıldır?

Artı değer özelliğine sahip iki alt testin alt test ve genişletilmiş alt test puanları kullanılarak eşitleme yöntemlerinden elde edilen eşitleme standart hatası (SEE), eşitleme yanlılığı (BIAS) ve eşitleme hatası (RMSE) değerlerinin ortak faktörler altındaki davranışları aşağıda ortak etki grafiklerinde yer almaktadır. Ortak etki grafiklerinde yer alan yöntemlerin hata değerlerine ait şekillerin renkleri grafiklerin sağ tarafında renk cetvelinde Braun/Holland (MEAN) eşitleme, zincirlenmiş lineer eşitleme (LIN), dairesel yay eşitleme (CIRC), birim eşitleme (IDEN) şeklinde etiketlemeyle yer almaktadır.

2.1.1.1.1. Alt testleri arasında 0.70 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?

Grafik13 incelendiğinde; birinci alt testin puanlarıyla yapılan eşitleme sonucunda eşitleme yöntemlerinden elde edilen eşitlemenin standart hatası örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı olmak üzere tüm faktörlerden etkilendiği görülmektedir.

Örneklem büyüklüğü arttıkça eşitleme yöntemlerinin eşitleme standart hatasında azalma meydana geldiği gözlenmektedir. Test uzunluğu arttıkça birim dönüşüm dışındaki tüm yöntemlerde eşitlemenin standart hatası değerinde artış görülmektedir. Test uzunluğunun 80 olduğu durumda eşitleme yöntemlerinin eşitlemenin standart hataları en yüksek değerlere sahiptir. Ortalama güçlük düzeyi farkının etkisi test uzunluğunun 80 olduğu durumda Braun/Holland ve zincirlenmiş lineer eşitleme yöntemlerinde görülmektedir.



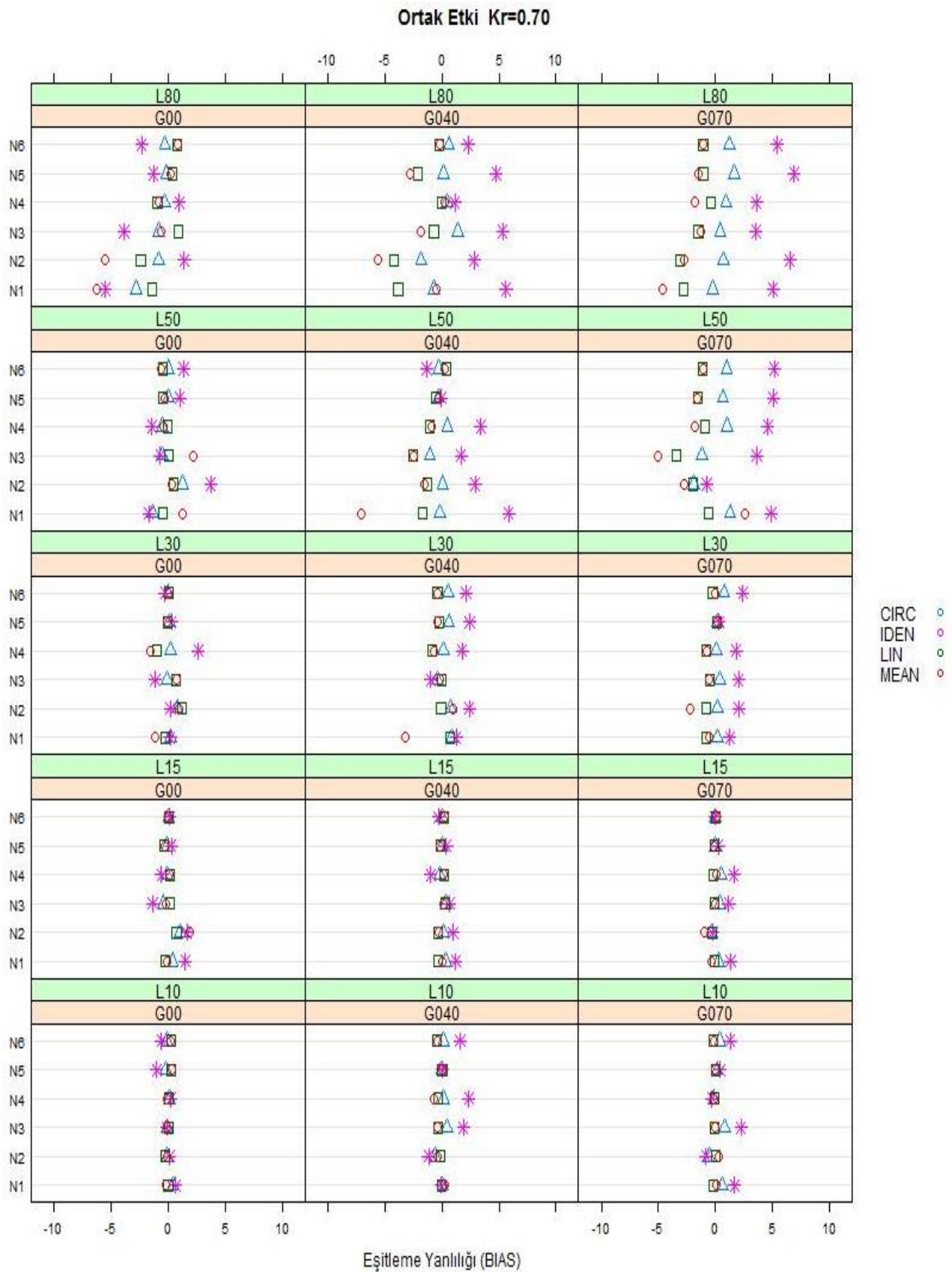
Grafik 13: 0.70 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme standart hata (SEE) değerlerine ortak etkisi

2.1.1.1.2. Alt testleri arasında 0.70 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitlemenin yanlılığına (BIAS) ortak etkisi nasıldır?

Grafik14 incelendiğinde; alt test uzunluğu ve ortalama güçlük düzeyi farkının eşitleme yanlılığında daha etkili olduğu gözlenmektedir. Eşitleme yanlılığı hatasının en çok değişkenlik gösterdiği başka bir ifadeyle ortak faktörlerden etkilenen eşitleme yöntemi birim dönüşüm yöntemidir.

Test uzunluğu arttıkça eşitleme yanlılığı değerindeki değişim artmaktadır. Bu artış en çok 50 ve 80 madde sayısına sahip alt testlerde tüm yöntemlerde görülmektedir. Ortalama güçlük düzeyi farkı arttıkça yöntemlerin BIAS değerlerinde değişim söz konusudur. Başka bir ifade ile ortalama güçlük düzeyi arttıkça yöntemlerin eşitleme yanlılığı artmaktadır.





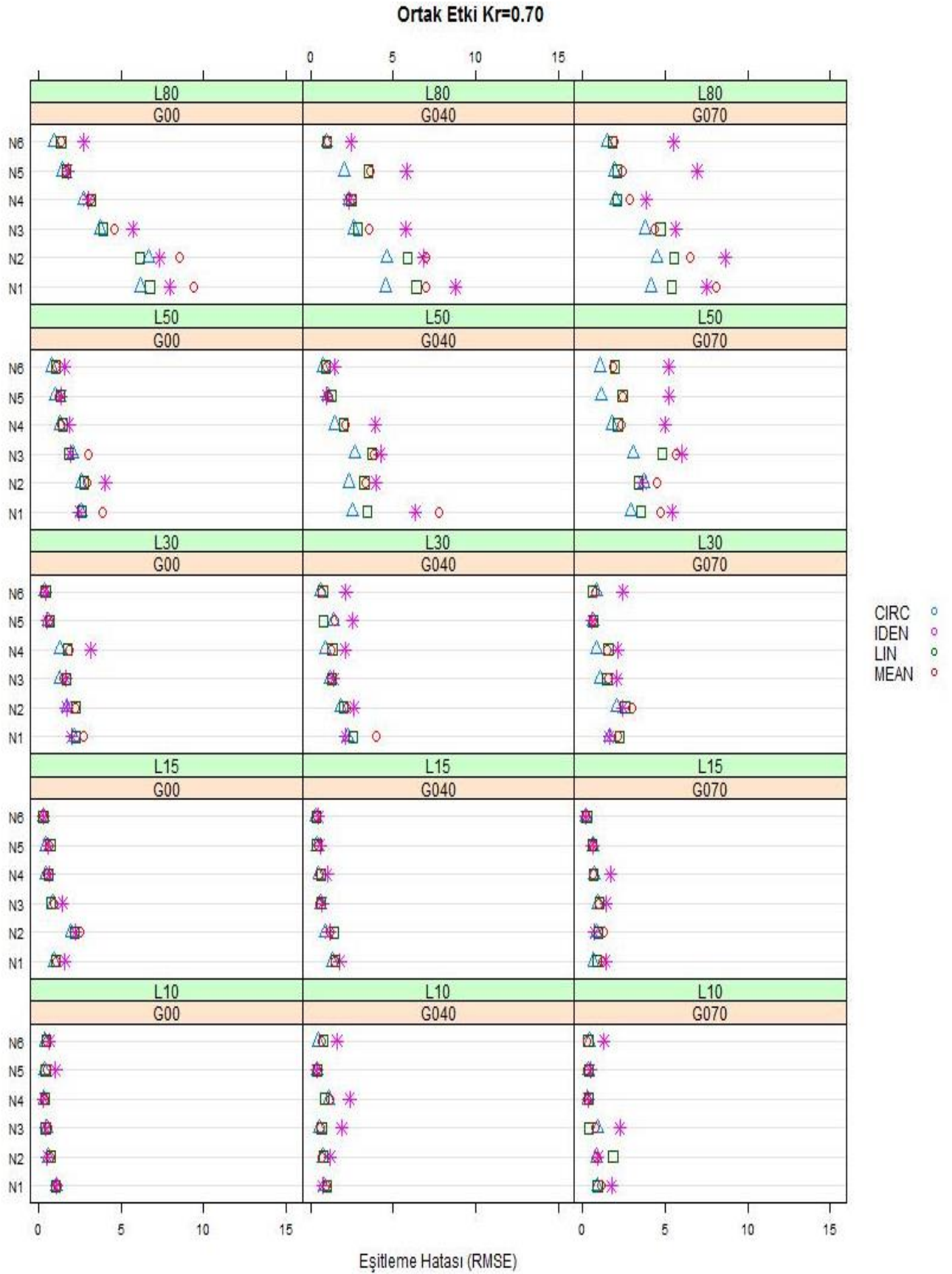
Grafik 14: 0.70 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme yanlılığı (BIAS) değerlerine ortak etkisi

2.1.1.1.3. Alt testleri arasında 0.70 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitleme hatasına (RMSE) ortak etkisi nasıldır?

Grafik15 incelendiğinde; birinci alt testlerin alt puanları kullanılarak elde edilen eşitleme hatası faktörlerin ortak etkisi altında olduğu görülmektedir. Tüm şartlar altında eşitleme hatası değerinde en çok değişkenlik gösteren birim eşitleme yöntemidir.

Örneklem büyüklüğü arttıkça birim dönüşüm dışındaki diğer yöntemlerde eşitleme hatası değeri tüm faktörler altında azalmaktadır. Test formları arasındaki ortalama güçlük düzeyi farkı değerleri artıkça eşitleme hatası değeri tüm yöntemlerde artmaktadır. Benzer şekilde test uzunluğu arttıkça RMSE değerleri de artmaktadır.





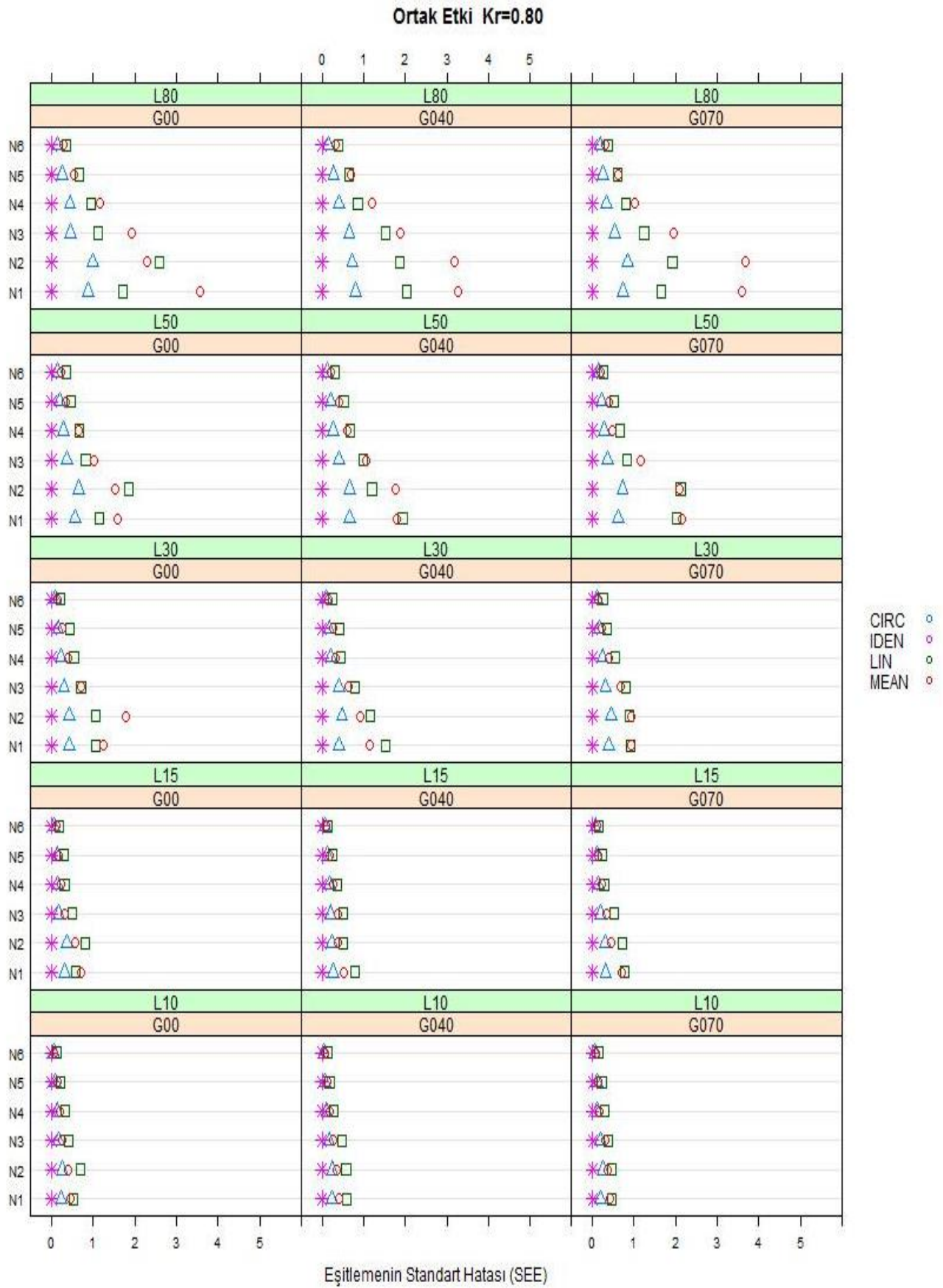
Grafik 15: 0.70 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farklı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme hatası (RMSE) değerlerine ortak etkisi

2.1.1.2.1. Alt testleri arasında 0.80 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?

Grafik16 incelendiğinde; birinci alt testin ham puanlarıyla yapılan eşitleme sonucunda eşitleme yöntemlerinden elde edilen eşitlemenin standart hatası örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı olmak üzere tüm faktörlerden etkilendiği görülmektedir.

Örneklem büyüklüğü arttıkça eşitleme yöntemlerinin eşitleme standart hatası değerlerinde azalma meydana gelirken test uzunluğu arttıkça birim dönüşüm dışındaki tüm yöntemlerde eşitlemenin standart hatası değerinde artış görülmektedir. Test uzunluğunun 80 olduğu durumda eşitleme yöntemlerinin eşitlemenin standart hataları en yüksek değerlere sahiptir.

Eşitleme yöntemlerinin eşitleme standart hatası, test formları arasında ortalama güçlük düzeyi farkı arttıkça madde sayısının fazla olduğu şartlarda artmaktadır. Genel olarak Braun/Holland, zincirleşmiş lineer eşitleme, dairesel yay eşitleme ve birim dönüşüm eşitleme yöntemi şeklinde eşitlemenin standart hatası büyükten küçüğe değerler olarak sıralandığı görülmektedir.



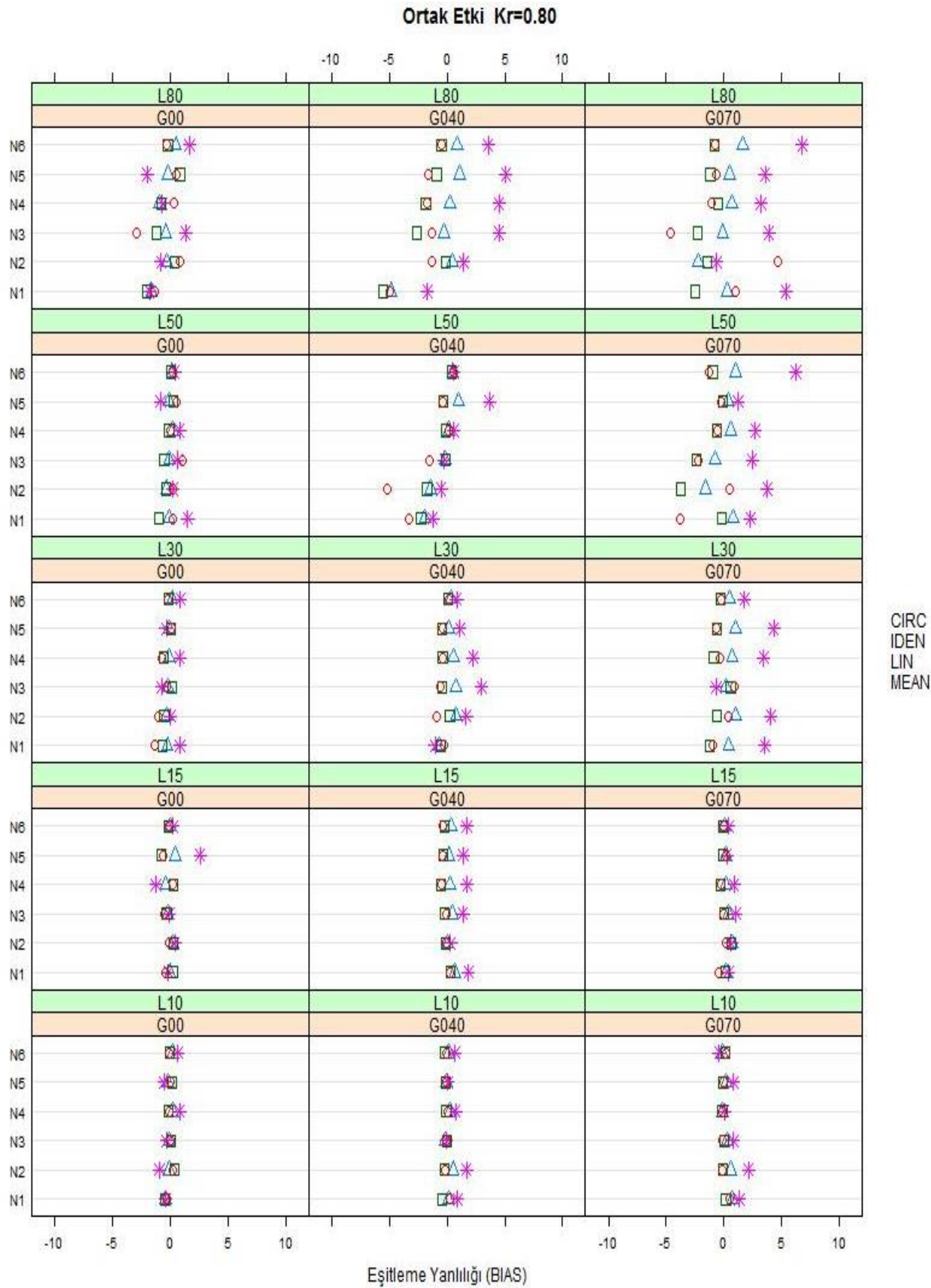
Grafik 16: 0.80 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme standart hata (SEE) değerlerine ortak etkisi

2.1.1.2.2. Alt testleri arasında 0.80 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitlemenin yanlılığına (BIAS) ortak etkisi nasıldır?

Artı değer özelliğinde sahip birinci alt testin alt puan eşitlemesiyle elde edilen Grafik17 incelendiğinde; alt test uzunluğu ve ortalama güçlük düzeyi farkının eşitleme yanlılığında daha etkili olduğu gözlenmektedir. Ayrıca Grafik17'e göre ortak faktörlerden etkilenen eşitleme yöntemi birim eşitleme yöntemidir.

Alt test uzunluğu arttıkça eşitleme yöntemlerinin eşitleme yanlılığı değerleri artmaktadır. Bu artış en çok 50 ve 80 madde sayısına sahip testlerde belirgindir. Test formları arasındaki ortalama güçlük düzeyi arttıkça özellikle birimeşitleme yönteminin eşitleme yanlılığı değerinin arttığı gözlenmektedir.





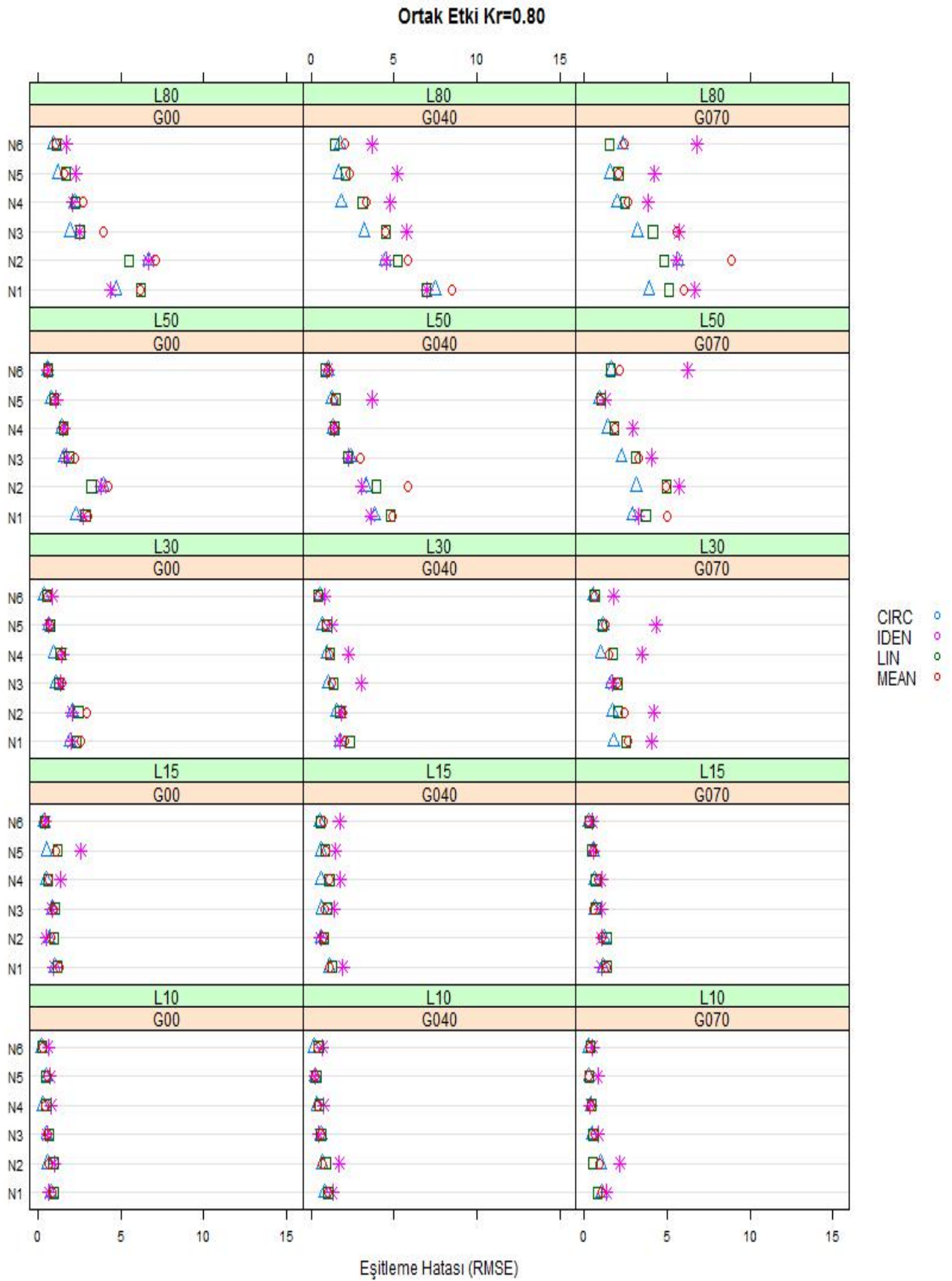
Grafik 17: 0.80 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme yanlılığı (BIAS) değerlerine ortak etkisi

2.1.1.2.3. Alt testleri arasında 0.80 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitleme hatasına (RMSE) ortak etkisi nasıldır?

Grafik18 incelendiğinde; artı değer özelliği gösteren birinci alt testlerin alt puanları kullanılarak elde edilen eşitleme örneklem büyüklüğünün, alt test uzunluğunun ve test formları arasındaki ortalama güçlük düzeyi farkı yöntemlerin eşitleme hatası değerlerini etkilemektedir.

Genel olarak örneklem büyüklüğü arttıkça yöntemlerin eşitleme hatası değerinde azalma meydana gelmektedir. Test uzunluğu arttıkça yöntemlerin eşitleme hatası değerlerinde artma görülmektedir. Test formları arasındaki ortalama güçlük düzeyi farkının artışı eşitleme hatası değerini de arttırmaktadır.



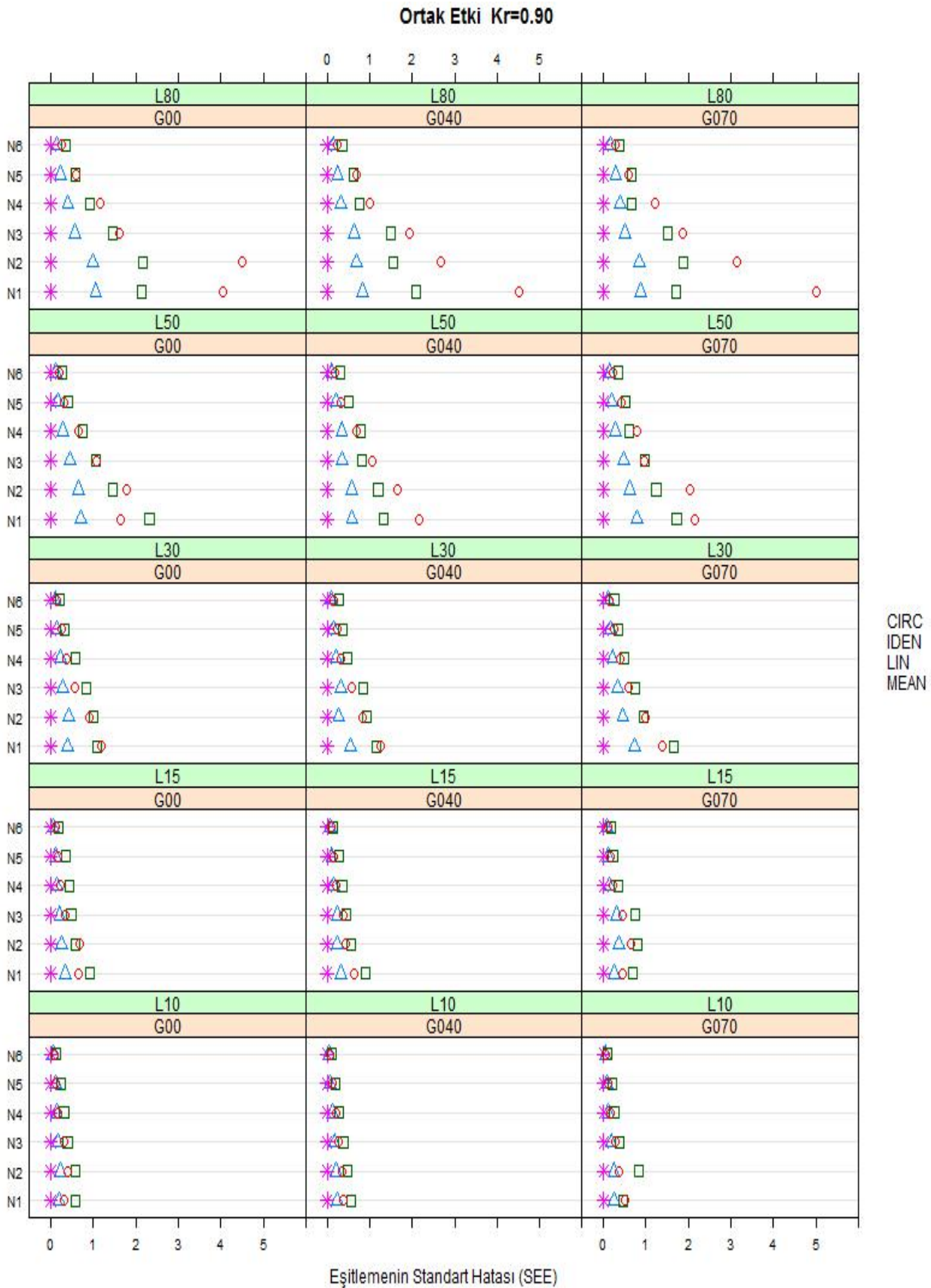


Grafik 18: 0.80 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme hatası (RMSE) değerlerine ortak etkisi

2.1.1.3.1. Alt testleri arasında 0.90 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?

Grafik19 incelendiğinde; artı değe özelliğine sahip birinci alt testin alt test puanlarıyla yapılan eşitleme sonucunda eşitleme yöntemlerinden elde edilen eşitlemenin standart hatası örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı birim eşitleme yöntemi dışındaki diğer eşitleme yöntemlerinin eşitleme standart hatası değerini değiştirmektedir.

Örneklem büyüklüğü arttıkça eşitleme yöntemlerinin eşitleme standart hatasında azalma meydana gelirken, test uzunluğu arttıkça birim eşitleme yöntemi dışındaki tüm yöntemlerde eşitlemenin standart hatası değerinde artış görülmektedir. Test uzunluğunun 80 olduğu durumda birim eşitleme yöntemi dışındaki diğer eşitleme yöntemlerinin SEE değerleri diğer şartlara göre daha çok değişme göstermektedir. Alt test formları arasında ortalama güçlük düzeyi farkı arttıkça madde sayısının fazla olduğu durumlarda eşitleme yöntemlerinin SEE değerinin arttığı görülmektedir.



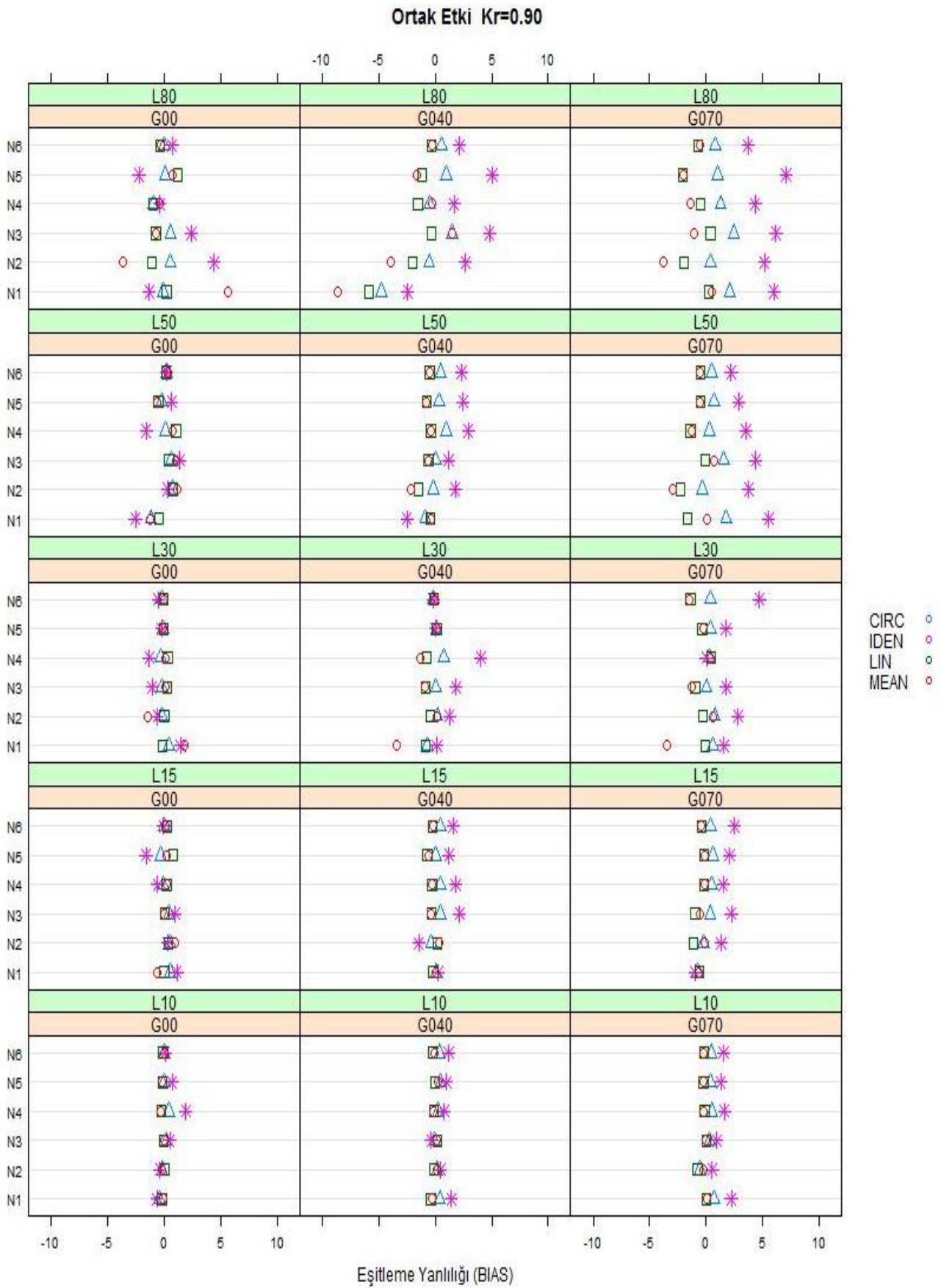
Grafik 19: 0.90 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme standart hata (SEE) değerlerine ortak etkisi

2.1.1.3.2. Alt testleri arasında 0.90 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitlemenin yanlılığına (BIAS) ortak etkisi nasıldır?

Artı değer özelliğinde sahip birinci alt testin alt puan eşitlenmesiyle elde edilen Grafik20 incelendiğinde; alt test uzunluğu ve ortalama güçlük düzeyi farkı, yöntemlerin eşitleme yanlılığı değerlerinde değişim meydana getirmektedir. Ortak faktörlerin birim eşitleme yönteminin BIAS değerlerinde daha çok değişmeye neden olduğu görülmektedir.

Alt test uzunluğu arttıkça eşitleme yöntemlerinin eşitleme yanlılığı değerleri artmaktadır. Bu artış en çok 50 ve 80 madde sayısına sahip alt testlerde daha belirgindir. Madde sayısının 80 ve test formları arasındaki ortalama güçlük düzeyi farkının 0.70 olduğu durumlarda, birim eşitleme ve dairesel yay eşitleme yöntemlerinin eşitleme yanlılığı değerlerinde fazla artış görülmektedir. Test formları arasındaki ortalama güçlük düzeyi farkı arttıkça özellikle birim eşitleme yönteminin eşitleme yanlılığı değerinin arttığı gözlenmektedir.





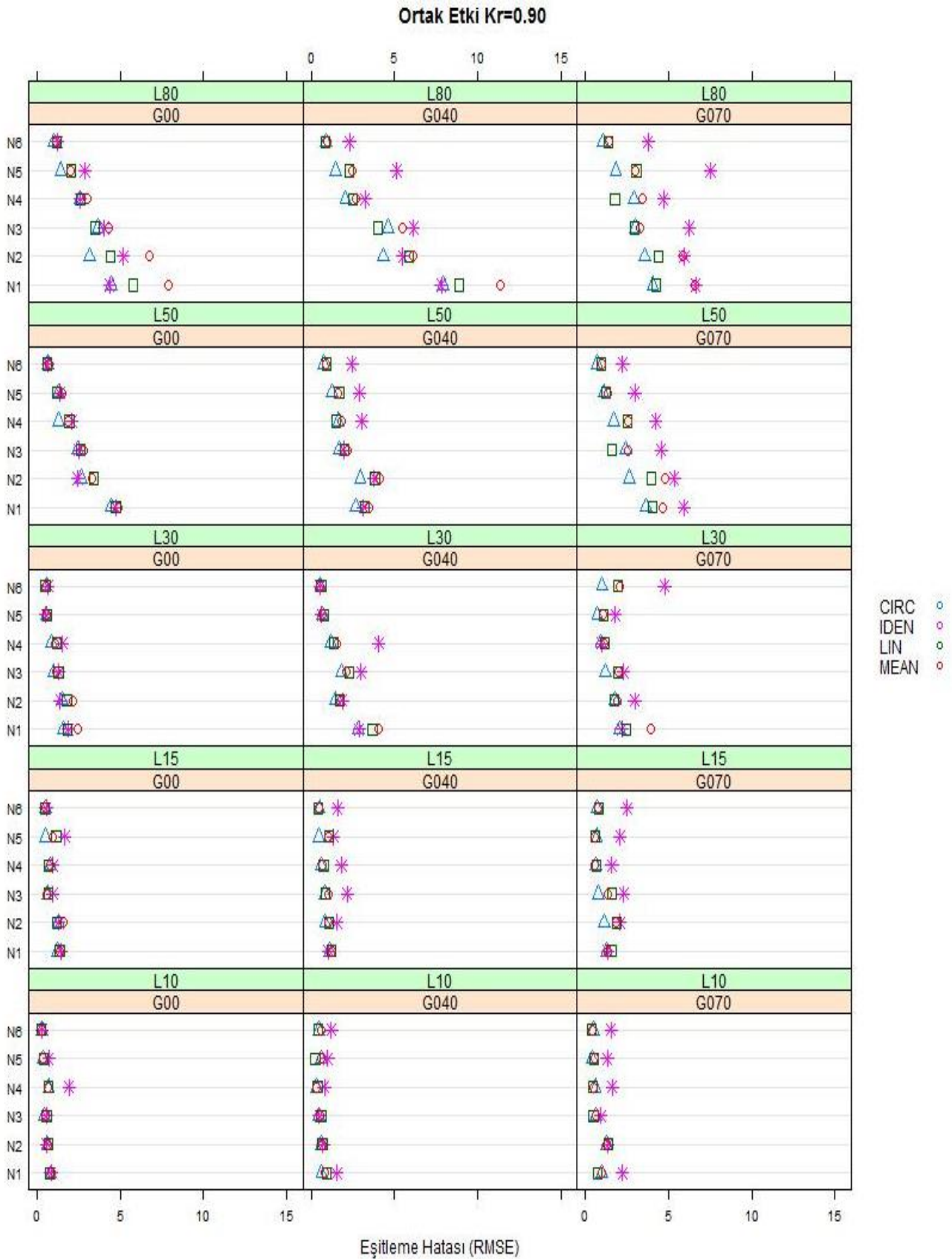
Grafik 20: 0.90 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farklı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme yanlılığı (BIAS) değerlerine ortak etkisi

2.1.1.3.3. Alt testleri arasında 0.90 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitleme hatasına (RMSE) ortak etkisi nasıldır?

Grafik21 incelendiğinde; artı değer özelliği gösteren birinci alt testlerin alt test puanları kullanılarak elde edilen eşitleme sonucunda eşitleme yöntemlerinin RMSE değerleri örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında değişim göstermektedir.

Genel olarak örneklem büyüklüğü arttıkça yöntemlerin eşitleme hatası değerinde azalma meydana gelmektedir. Alt test uzunluğu arttığında yöntemlerin eşitleme hatası değerlerinde artma görülmektedir. Test formları arasındaki ortalama güçlük düzeyi farkının artışı yöntemlerin eşitleme hatası değerini de arttırmaktadır.





Grafik 21: 0.90 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme hatası (RMSE) değerlerine ortak etkisi

2.1.2.1.1. Alt testleri arasında 0.70 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?

Grafik22 incelendiğinde; artı değer özelliğine sahip ikinci alt testin alt test puanlarıyla yapılan eşitleme sonucunda birim eşitleme yöntemi dışındaki diğer eşitleme yöntemlerinden elde edilen eşitlemenin standart hatası değeri örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı olmak üzere tüm faktörler altında değişim göstermektedir.

Örneklem büyüklüğü arttıkça eşitleme yöntemlerinin eşitleme standart hatası değerlerinde azalma meydana gelmektedir. Test uzunluğu arttıkça birim eşitleme dışındaki tüm yöntemlerde eşitlemenin standart hatası değerinde artış görülmektedir. Ayrıca alt test uzunluğu 80 ve 50 olduğunda yöntemlerin eşitleme standart hatası değerlerinin büyükten küçüğe sıralanışı; Braun/Holland, zincirlenmiş lineer eşitleme, dairesel yay eşitleme ve birim eşitleme yöntemi şeklindedir.

Alt test uzunluğu 80 olduğu durumda ortalama güçlük düzeyi farkının, Braun/Holland eşitleme ve zincirlenmiş lineer eşitleme yöntemlerinin SEE değerlerinde daha fazla değişim göstermektedir.



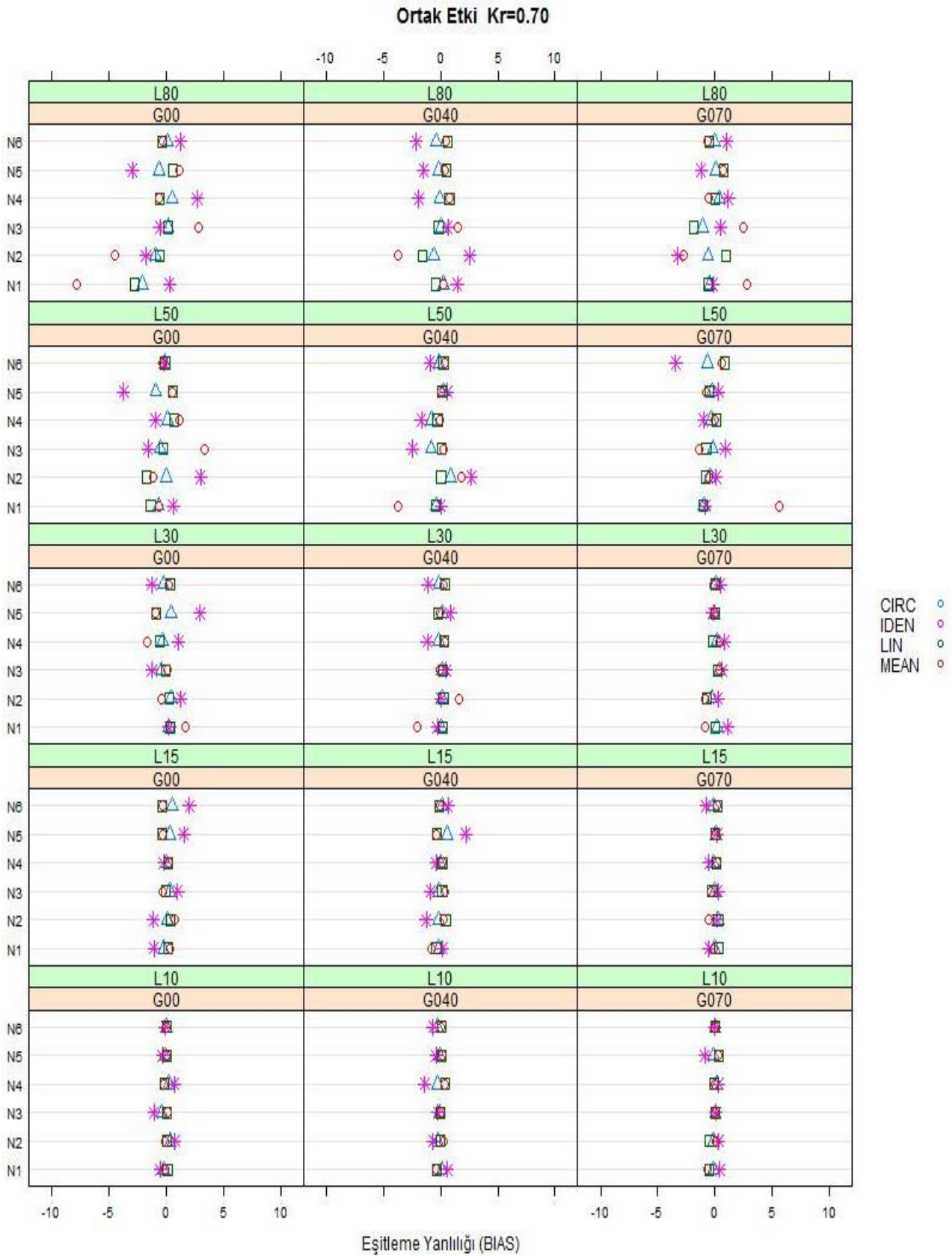
Grafik 22: 0.70 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farklı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme standart hata (SEE) değerlerine ortak etkisi

2.1.2.1.2. Alt testleri arasında 0.70 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitlemenin yanlılığına (BIAS) ortak etkisi nasıldır?

Artı değer özelliğinde sahip ikinci alt testin alt test puanları eşitlenmesiyle elde edilen Grafik23 incelendiğinde; alt test uzunluğu ve ortalama güçlük düzeyi farkının eşitleme yanlılığı değerlerinin değişiminde daha etkili olduğu gözlenmektedir. Tüm şartlar altında birim eşitleme yönteminin eşitleme yanlılığı hatası değeri en çok değişkenliği göstermektedir.

Alt test uzunluğu arttıkça eşitleme yöntemlerinin eşitleme yanlılığı değerlerindeki değişim artmaktadır. Test formları arasında ortalama güçlük düzeyinin 0.0 olduğu durumda yöntemlerin eşitleme yanlılığı değeri daha çok değişim gösterirken diğer ortalama güçlük düzeyi farklarında düzenli değişim davranışı gözlenmemektedir.





Grafik 23: 0.70 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme yanıllığı (BIAS) değerlerine ortak etkisi

2.1.2.1.3. Alt testleri arasında 0.70 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitleme hatasına (RMSE) ortak etkisi nasıldır?

Grafik24 incelendiğinde; artı değer özelliği gösteren ikinci alt testlerin alt test puanları kullanılarak eşitleme yöntemlerinden elde edilen eşitleme hatası değeri örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı faktörleri altında değişim göstermektedir.

Örneklem büyüklüğü arttıkça yöntemlerin eşitleme hatası değeri tüm faktörler altında azalmaktadır. Test formları arasındaki güçlük düzeyi farkı arttıkça eşitleme hatası değeri tüm yöntemlerde artmaktadır. Test uzunluğu 50 ve 80 olduğu durumlarda eşitleme yöntemlerinin eşitleme hatası değerinde daha fazla artış gözlenmektedir.





Grafik 24: 0.70 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme hatası (RMSE) değerlerine ortak etkisi

2.1.2.2.1. Alt testleri arasında 0.80 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?

Artı değer özelliğine sahip ikinci alt testin, alt test puanları kullanılarak gerçekleştirilen eşitleme sonucunda; eşitleme yöntemlerinden elde edilen eşitlemenin standart hatası değerleri birim eşitleme yöntemi dışındaki yöntemlerde örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı olmak üzere tüm faktörlerde değişim göstermektedir.

Örneklem büyüklüğü arttıkça eşitleme yöntemlerinin eşitleme standart hatasında azalma meydana gelirken test uzunluğu arttıkça birim eşitleme dışındaki tüm yöntemlerde eşitlemenin standart hatası değerinde artış görülmektedir. Test uzunluğu 80 olduğunda eşitleme yöntemlerinin, eşitleme standart hatası değerleri en yüksek değerlere sahiptir. 30, 15 ve 10 madde sayısı içeren testlerde zincirlenmiş lineer eşitleme, Braun/Holland, dairesel yay eşitleme yöntemlerinin eşitlemenin standart hatası değerleri büyükten küçüğe şeklinde bir sıralanış göstermektedir,.

Eşitleme yöntemlerinin eşitleme standart hatası, test formları arasında ortalama güçlük düzeyi farkı arttıkça madde sayısının fazla olduğu şartlarda artmaktadır. Genel olarak Braun/Holland, zincirlenmiş lineer eşitleme, dairesel yay eşitleme ve birim eşitleme yöntemi şeklinde eşitlemenin standart hatası büyükten küçüğe değerler olarak sıralandığı görülmektedir.



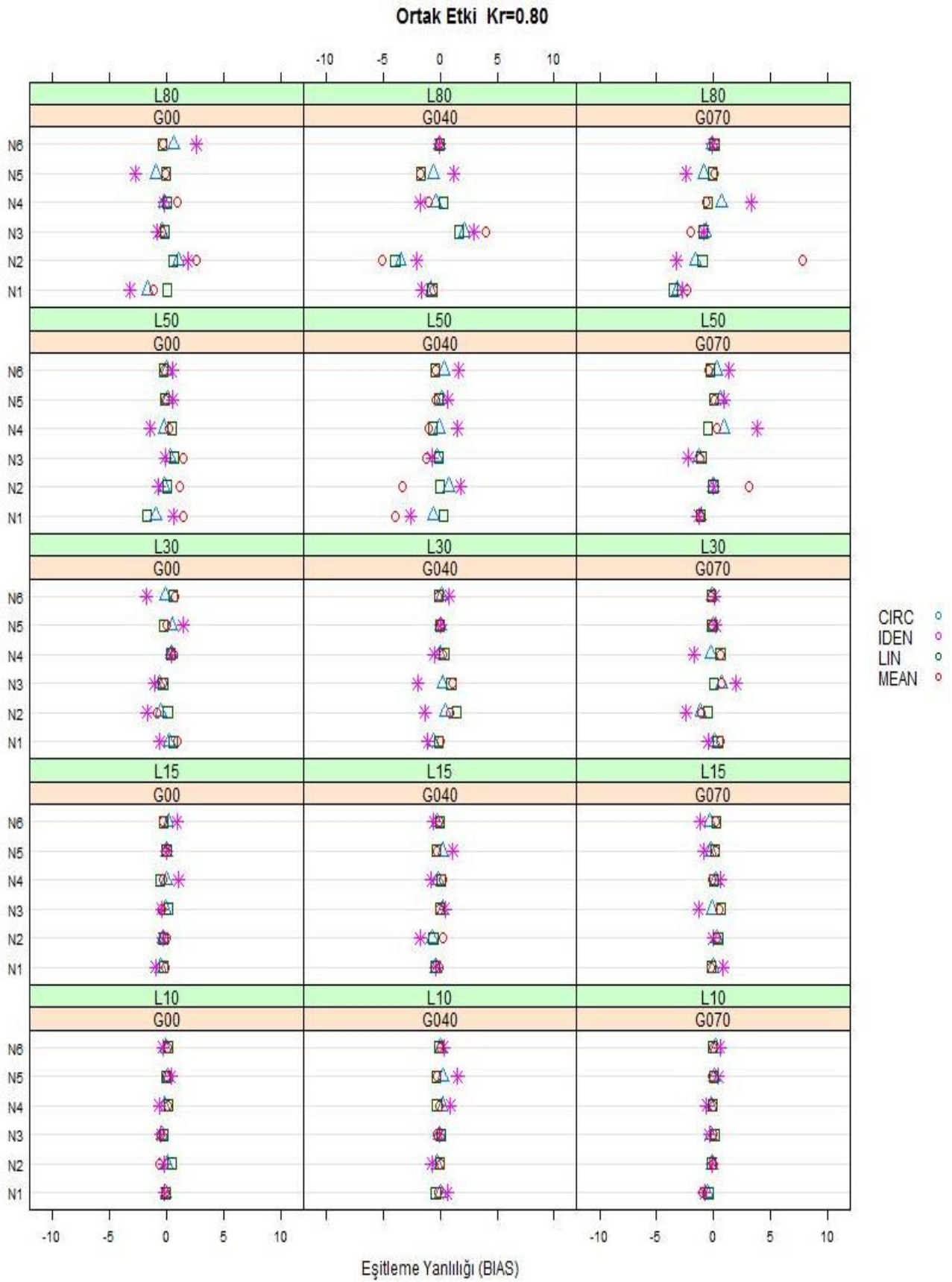
Grafik 25: 0.80 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme standart hata (SEE) değerlerine ortak etkisi

2.1.2.2.2. Alt testleri arasında 0.80 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitlemenin yanlılığına (BIAS) ortak etkisi nasıldır?

Artı değer özelliğinde sahip ikinci alt testin alt test puanları eşitlenmesiyle elde edilen Grafik26 incelendiğinde; alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörleri arttıkça yöntemlerin eşitleme yanlılığı değerlerinde düzenli bir artış görülmektedir. Grafik26'a göre ortak faktörlerden etkilenen eşitleme yöntemi birim eşitleme yöntemidir.

Test uzunluğu eşitleme yöntemlerinin eşitleme yanlılığı değerleri artmaktadır. Bu artış en çok 30, 50 ve 80 madde sayısına sahip testlerde belirgindir. Test formları arasındaki ortalama güçlük düzeyi arttıkça özellikle birim eşitleme yönteminin eşitleme yanlılığı değerlerinin arttığı gözlenmektedir.





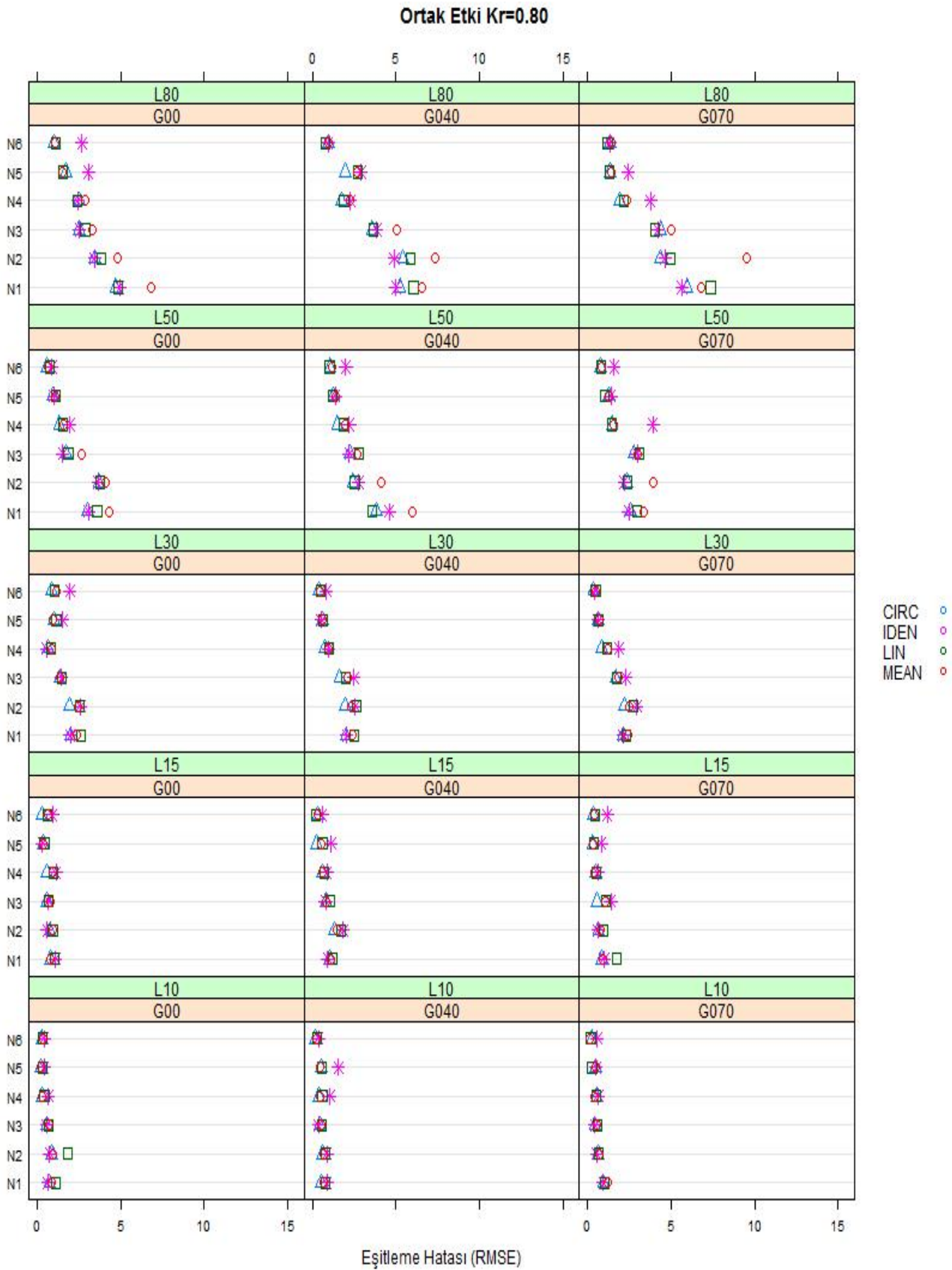
Grafik 26: 0.80 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farklı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme yanlılığı (BIAS) değerlerine ortak etkisi

2.1.2.2.3. Alt testleri arasında 0.80 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitleme hatasına (RMSE) ortak etkisi nasıldır?

Artı değer özelliği gösteren ikinci alt testlerin alt test puanları kullanılarak eşitleme yöntemlerinden elde edilen eşitleme hatası değerleri örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki ortalama güçlük düzeyi farkı faktörlerinden etkilenmektedir.

Genel olarak örneklem büyüklüğü arttıkça yöntemlerin eşitleme hatası değerinde azalma meydana gelmektedir. Test uzunluğu arttığında yöntemlerin eşitleme hatası değerlerinde artma görülmektedir. Test formları arasındaki ortalama güçlük düzeyi farkının artışı eşitleme hatası değerini de arttırmaktadır.





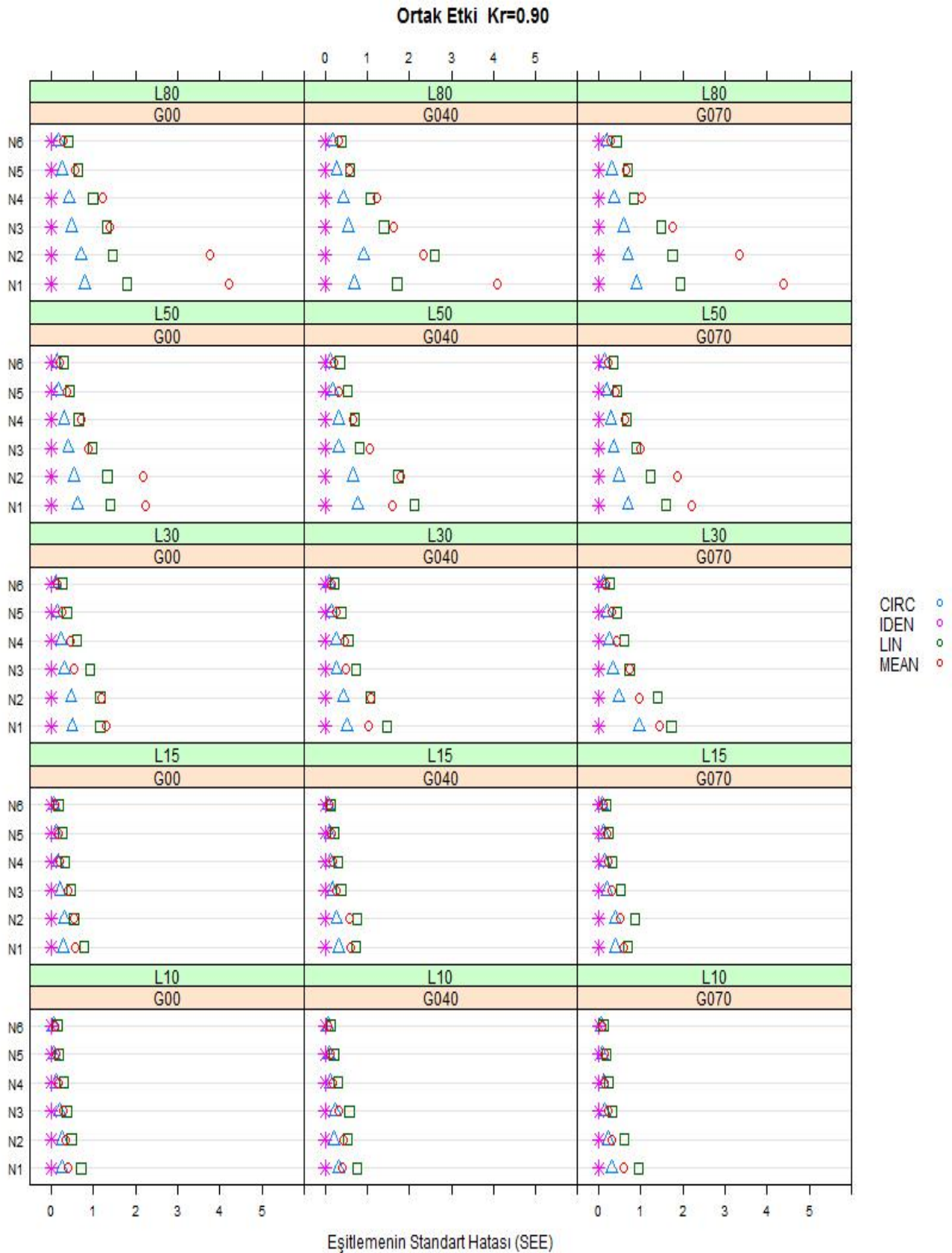
Grafik 27: 0.80 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme hatası (RMSE) değerlerine ortak etkisi

2.1.2.3.1. Alt testleri arasında 0.90 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?

Grafik28 incelendiğinde; artı deęe özellięine sahip ikinci alt testin alt test puanlarıyla yapılan eşitleme sonucunda eşitleme yöntemlerinden elde edilen eşitlemenin standart hatası örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı birim eşitleme yöntemi dışındaki dięer eşitleme yöntemlerinin eşitleme standart hatası deęerlerini deęiřtirmektedir.

Örneklem büyüklüğü arttıkça eşitleme yöntemlerinin eşitleme standart hatası deęerlerinde azalma meydana gelirken, test uzunluğu arttıkça birim eşitleme dışındaki tüm yöntemlerde eşitlemenin standart hatası deęerlerinde artış görölmektedir. Alt test uzunluğunun 80 olduęu durumlarda, eşitleme yöntemlerinin eşitlemenin standart hatası deęerlerinde dięer şartlara göre daha çok deęişme meydana geldięi görölmektedir..

Alt test formları arasında ortalama güçlük düzeyi farkı arttıkça madde sayısının fazla olduęu durumlarda eşitleme yöntemlerinin eşitleme standart hatası deęerilerinin arttıęı görölmektedir. 30 maddeye sahip alt testlerde ortalama güçlük düzeyi farkı arttıkça yöntemlerin eşitleme standart hatası deęerlerindeki büyükten küçüęe olan sıralama; Braun/Holland, dairesel yay eşitleme ve birim eşitleme yöntemi olacak şekilde düzenli bir deęişim göstermektedir.



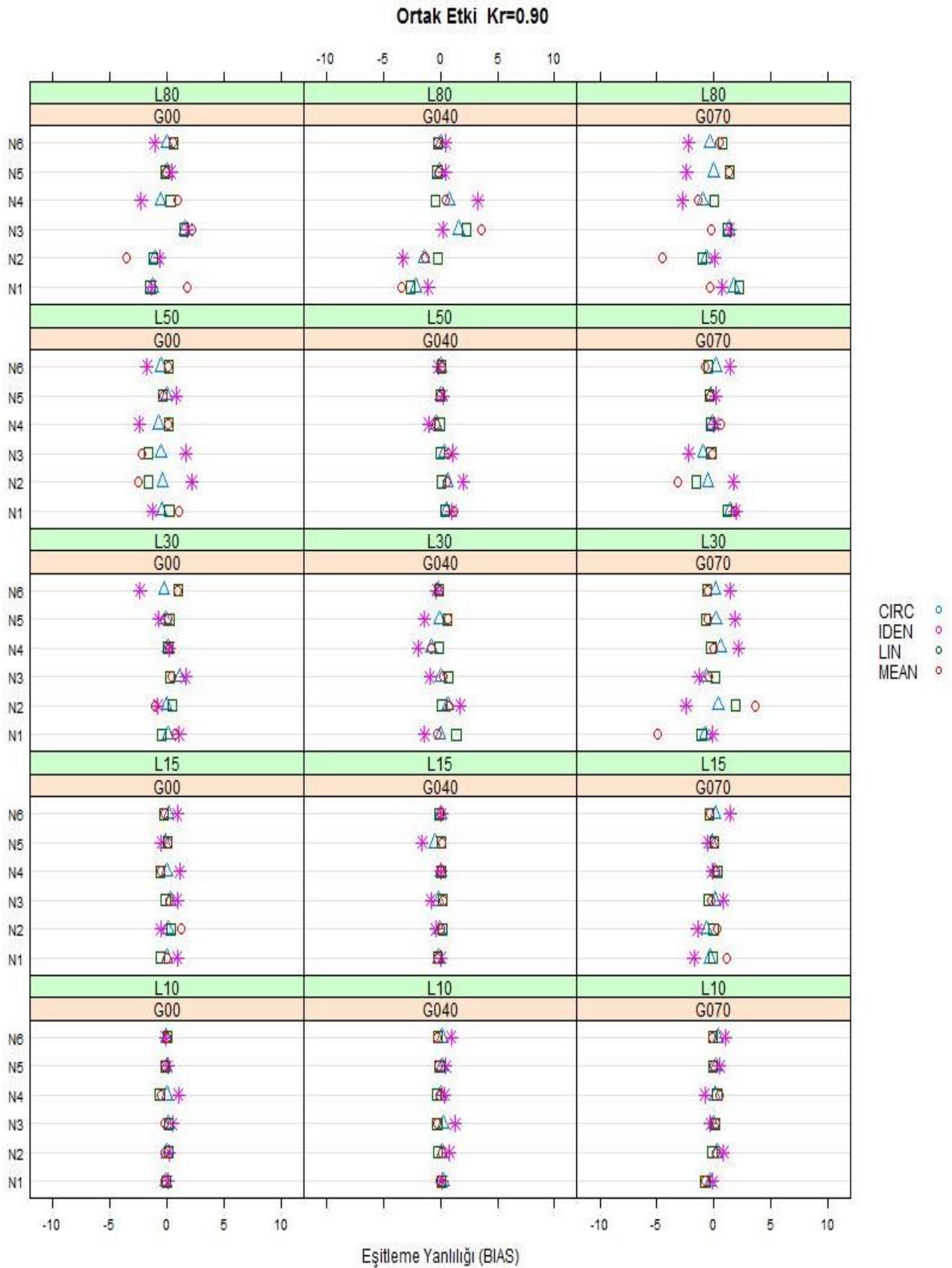
Grafik 28: 0.90 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme standart hata (SEE) değerlerine ortak etkisi

2.1.2.3.2. Alt testleri arasında 0.90 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitlemenin yanlılığına (BIAS) ortak etkisi nasıldır?

Artı değer özelliğinde sahip ikinci alt testin alt test puanları eşitlemesiyle elde edilen Grafik29 incelendiğinde; alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörleri yöntemlerin eşitleme yanlılığı değerlerinde değişim meydana getirmektedir.Grafik29'a göre ortak faktörlerin birim eşitleme yönteminin eşitleme yanlılık değerlerinde daha çok değişme meydana getirdiği görülmektedir.

Alt test uzunluğu arttıkça eşitleme yöntemlerinin eşitleme yanlılığı değerleri de artmaktadır. Bu artış en çok 30, 50 ve 80 madde sayısına sahip alt testlerde belirgindir. Test formları arasındaki ortalama güçlük düzeyi arttıkça özellikle birim eşitleme yönteminin eşitleme yanlılığı değerinde arttığı gözlenmektedir.





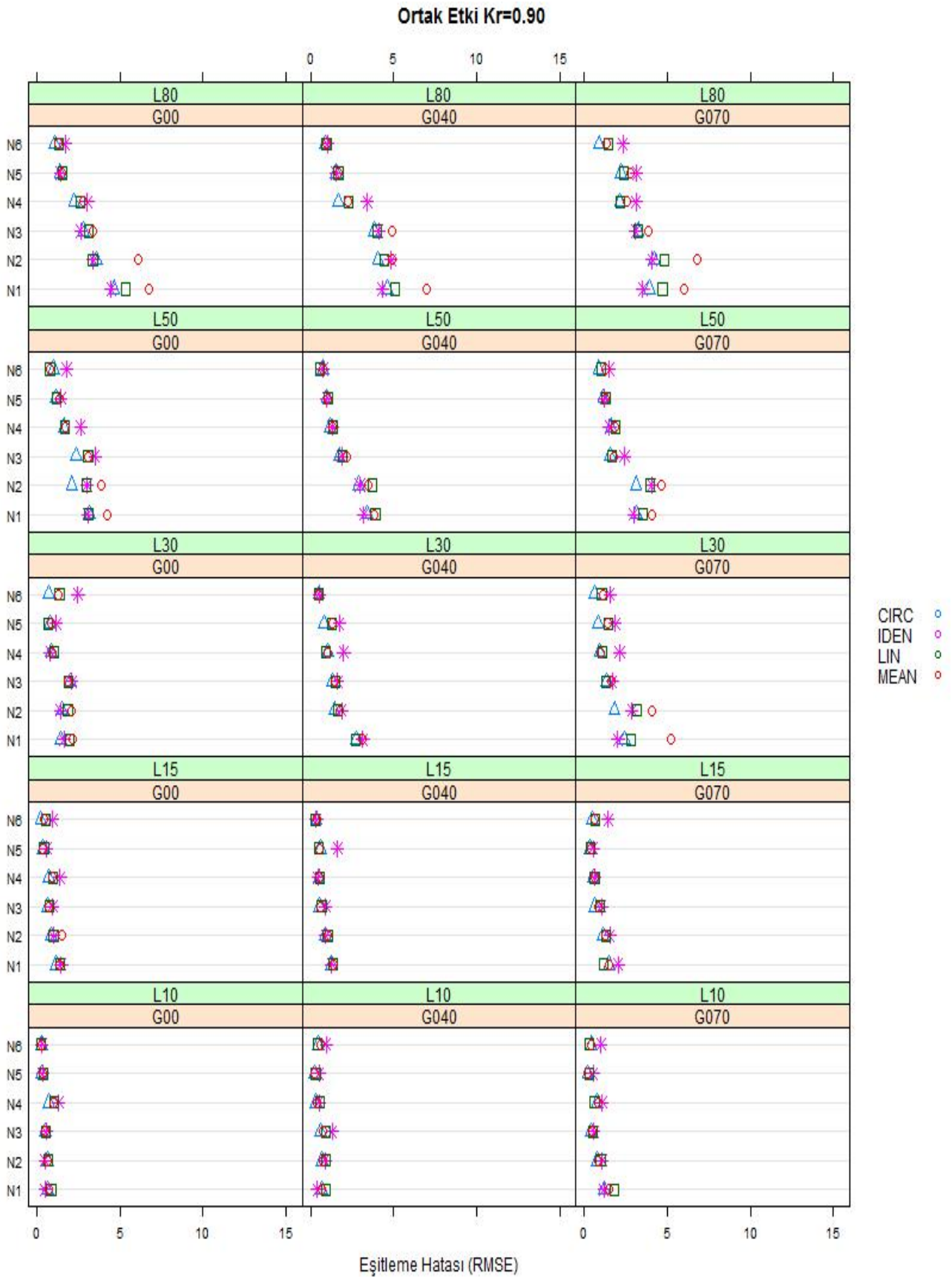
Grafik 29: 0.90 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farklı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme yanlılığı (BIAS) değerlerine ortak etkisi

2.1.2.3.3. Alt testleri arasında 0.90 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında alt test puanları eşitlenmesiyle elde edilen eşitleme hatasına (RMSE) ortak etkisi nasıldır?

Grafik30 incelendiğinde; artı değer özelliği gösteren ikinci alt testlerin alt test puanları kullanılarak eşitleme yöntemlerinden elde edilen eşitleme hatası değerleri örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında değişim göstermektedir.

Genel olarak örneklem büyüklüğü arttıkça yöntemlerin eşitleme hatası değerinde azalma meydana gelirken test uzunluğu arttığında ise yöntemlerin eşitleme hatası değerlerinde artma görülmektedir. Test formları arasındaki ortalama güçlük düzeyi farkının artışı yöntemlerin eşitleme hatası değerini de arttırmaktadır.





Grafik 30: 0.90 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme hatası (RMSE) değerlerine ortak etkisi

2.2.1.1.1. Alt testleri arasında 0.70 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde edilen eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?

Grafik31 incelendiğinde; artı değe özelliğine sahip birinci alt testin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucunda birim eşitleme yöntemi dışındaki diğer eşitleme yöntemlerinden elde edilen eşitleme standart hatası değerleri örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı olmak üzere tüm faktörler altında değişim göstermektedir.

Örneklem büyüklüğü arttıkça eşitleme yöntemlerinin eşitleme standart hatasında azalma meydana gelmektedir. Test uzunluğu arttıkça eşitlemenin standart hatası değerinde artış görülmektedir. Test uzunluğunun 80 olduğu durumda eşitleme yöntemlerinin eşitlemenin standart hataları en yüksek değerlere sahiptir. Genel olarak zincirlenmiş lineer eşitleme, Braun/Holland eşitleme, dairesel yay eşitleme ve birim eşitleme yöntemlerinin eşitleme standart hatası değeri büyükten küçüğe olacak şekilde sıralanma davranışı gösterdiği görülmektedir. Ortalama güçlük düzeyi farkı arttıkça eşitleme yöntemlerinin eşitleme standart hataları değerlerinde artış gözlenmektedir.



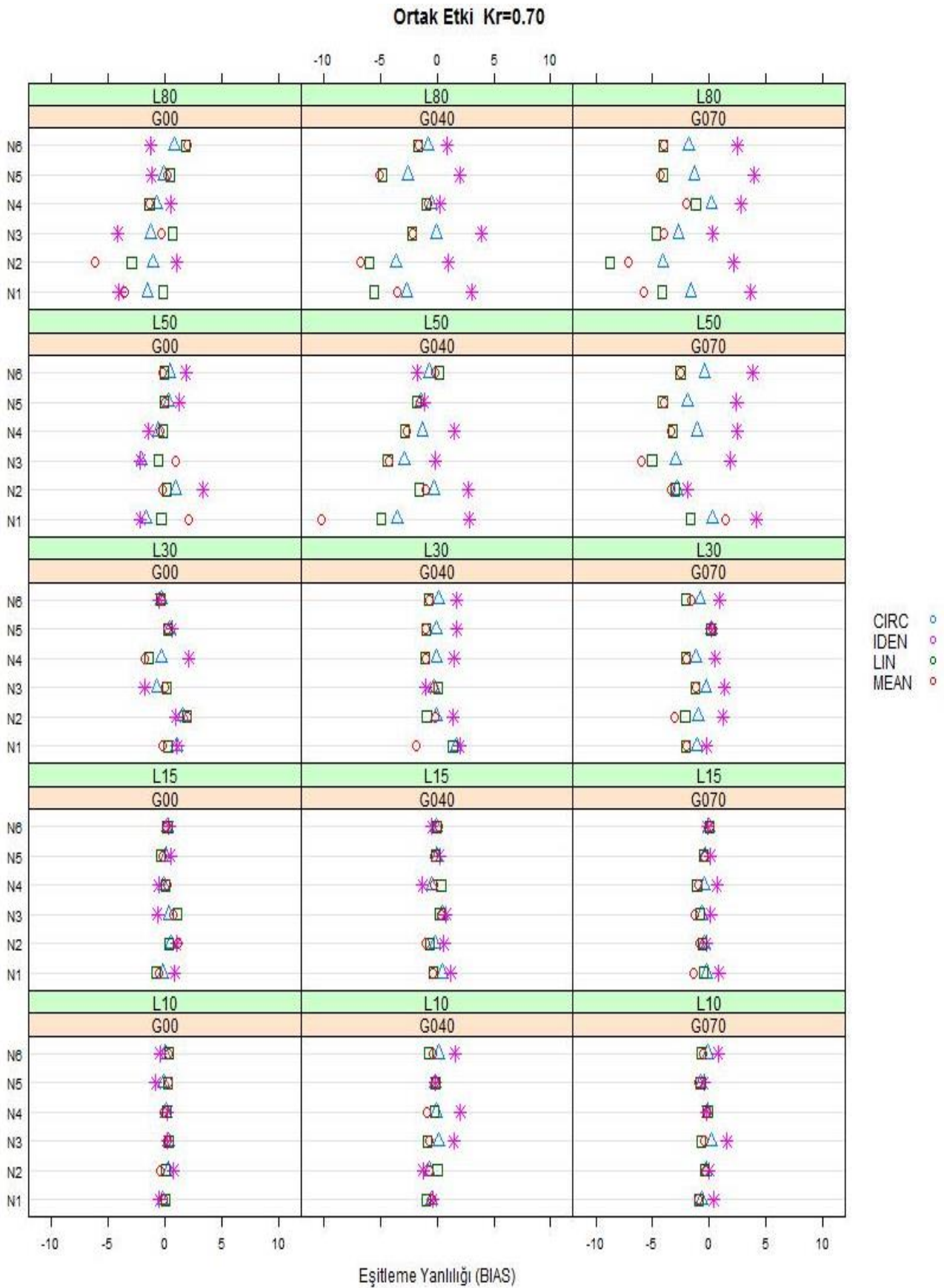
Grafik 31: 0.70 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme standart hata (SEE) değerlerine ortak etkisi

2.2.1.1.2. Alt testleri arasında 0.70 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde edilen eşitlemenin yanlılığına (BIAS) ortak etkisi nasıldır?

Artı değer özelliğinde sahip birinci alt testin genişletilmiş alt test puanı kullanılarak gerçekleştirilen eşitleme sonucunda elde edilen Grafik32 incelendiğinde alt test uzunluğu ve ortalama güçlük düzeyi farkının eşitleme yanlılığında daha etkili olduğu gözlenmektedir. Tüm şartlar altında birim eşitleme yönteminin eşitleme yanlılığı hatası en çok değişkenlik göstermektedir.

Test uzunluğu arttıkça eşitleme yanlılığı değerindeki değişim artmaktadır. Eşitleme yöntemlerinin eşitleme yanlılığında meydana gelen değişim 30, 50 ve 80 madde sayısına sahip testlerde daha fazla gözlenmektedir. Test formları arasında ortalama güçlük düzeyi farkı arttıkça yöntemlerin eşitleme yanlılık değerlerindeki değişim artmaktadır.





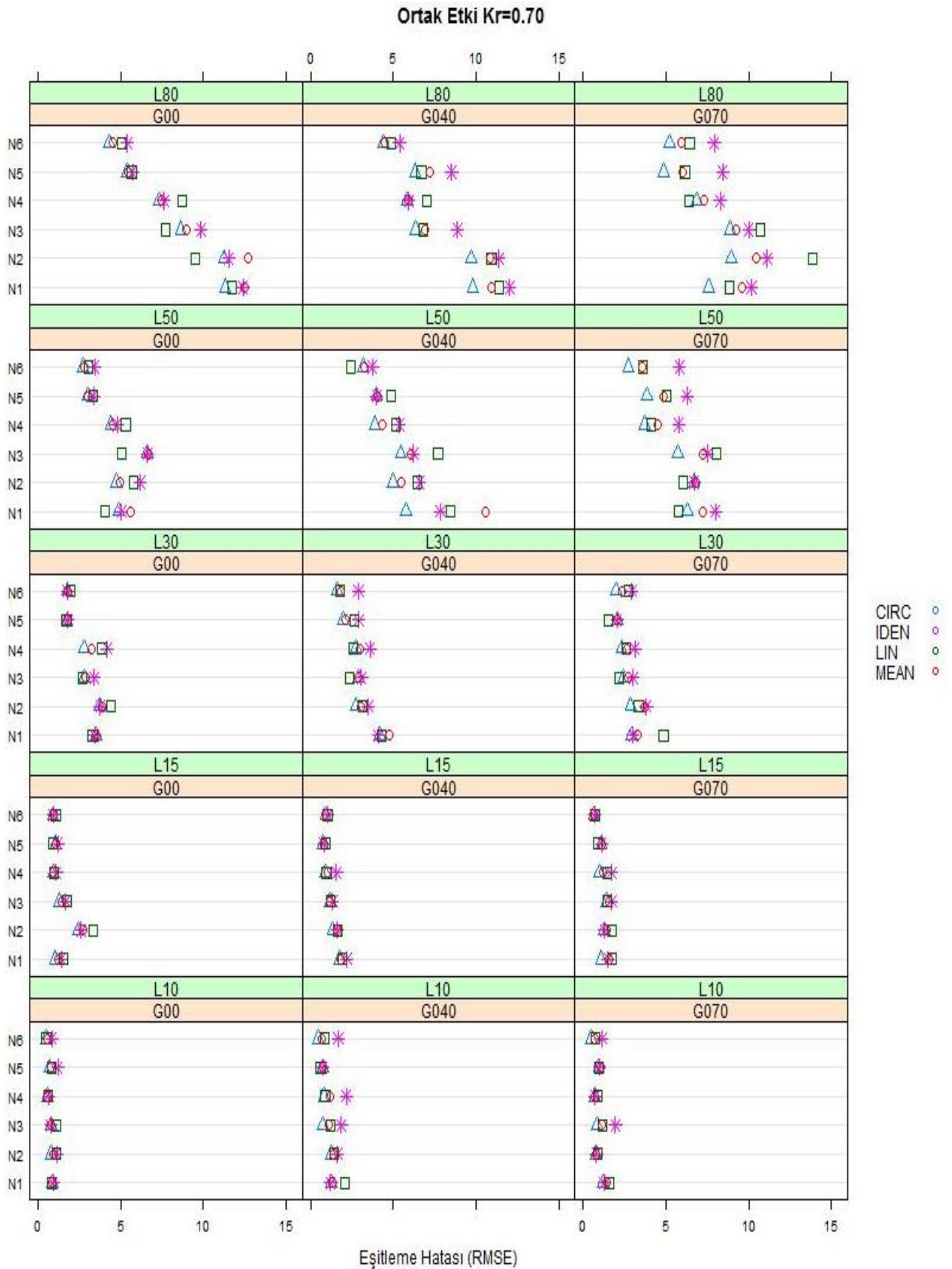
Grafik 32: 0.70 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme yanlılığı (BIAS) değerlerine ortak etkisi

2.2.1.1.3. Alt testleri arasında 0.70 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde edilen eşitleme hatasına (RMSE) ortak etkisi nasıldır?

Grafik33 incelendiğinde; artı değer özelliği gösteren birinci alt testin genişletilmiş alt puanları kullanılarak eşitleme yöntemlerinden elde edilen eşitleme hatası değeri örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı faktörleri altında değişim göstermektedir.

Genel olarak örneklem büyüklüğü arttıkça yöntemlerin eşitleme hatası değeri tüm faktörler altında azalmaktadır. Test formları arasındaki güçlük düzeyi arttıkça eşitleme hatası değeri tüm yöntemlerde artmaktadır. Test uzunluğu 50 ve 80 olduğu durumlarda eşitleme yöntemlerinin eşitleme hatası değerinde daha fazla artış gözlenmektedir.





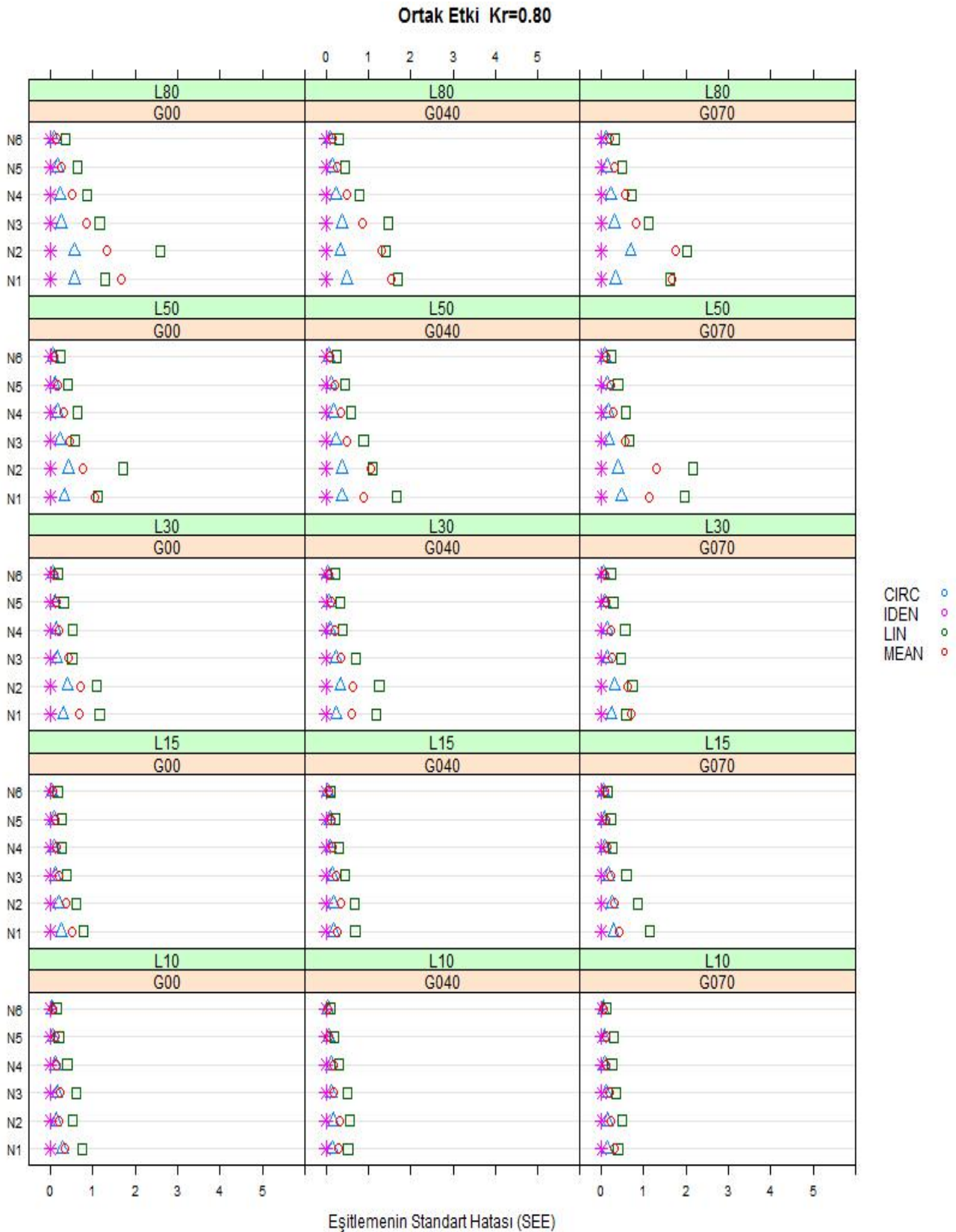
Grafik 33: 0.70 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farklı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme hatası (RMSE) değerlerine ortak etkisi

2.2.1.2.1. Alt testleri arasında 0.80 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde edilen eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?

Grafik34 incelendiğinde; artı değer özelliğine sahip birinci alt testin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucunda eşitleme yöntemlerinden elde edilen eşitlemenin standart hatası değerleri birim eşitleme yöntemi dışındaki yöntemlerde örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı olmak üzere tüm faktörlerde değişim göstermektedir.

Örneklem büyüklüğü arttıkça eşitleme yöntemlerinin eşitleme standart hatası değerlerinde azalma meydana gelirken, test uzunluğu arttığında ise SEE değerlerinde artış görülmektedir. Genel olarak zincirlenmiş lineer eşitleme, Braun/Holland eşitleme, dairesel yay eşitleme ve birim eşitleme yöntemlerinin eşitleme standart hatası değerleri büyükten küçüğe sıralanma davranışı gösterdiği görülmektedir. Eşitleme yöntemlerinin eşitleme standart hatası, test formları arasında ortalama güçlük düzeyi farkı arttıkça artmaktadır.





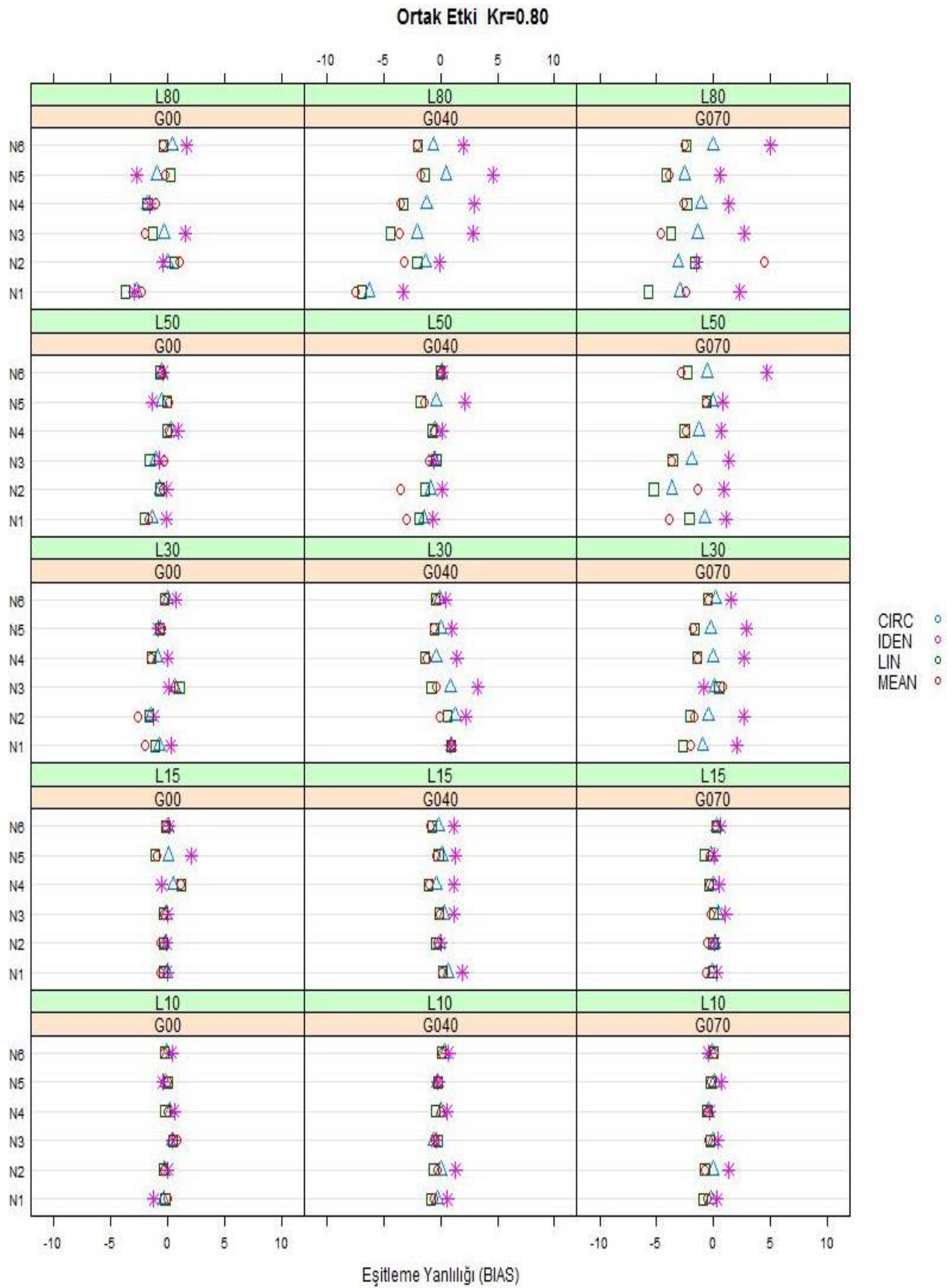
Grafik 34: 0.80 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farklı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme standart hata (SEE) değerlerine ortak etkisi

2.2.1.2.2. Alt testleri arasında 0.80 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde edilen eşitlemenin yanlılığına (BIAS) ortak etkisi nasıldır?

Artı değer özelliğinde sahip birinci alt testin genişletilmiş alt test puanları eşitlenmesiyle elde edilen Grafik35 incelendiğinde; alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin, eşitleme yanlılığı değerlerinde düzenli bir artış davranışına neden olduğu görülmektedir. Grafik35'e göre ortak faktörlerden en çok etkilenen eşitleme yöntemi birim dönüşüm yöntemidir.

Test uzunluğu arttıkça eşitleme yöntemlerinin eşitleme yanlılığı değerleri artmaktadır. Bu artış en çok 30, 50 ve 80 madde sayısına sahip testlerde belirgin görülmektedir. Test formları arasındaki ortalama güçlük düzeyi arttıkça özellikle birim eşitleme yönteminin eşitleme yanlılığı değerinin arttığı gözlenmektedir.





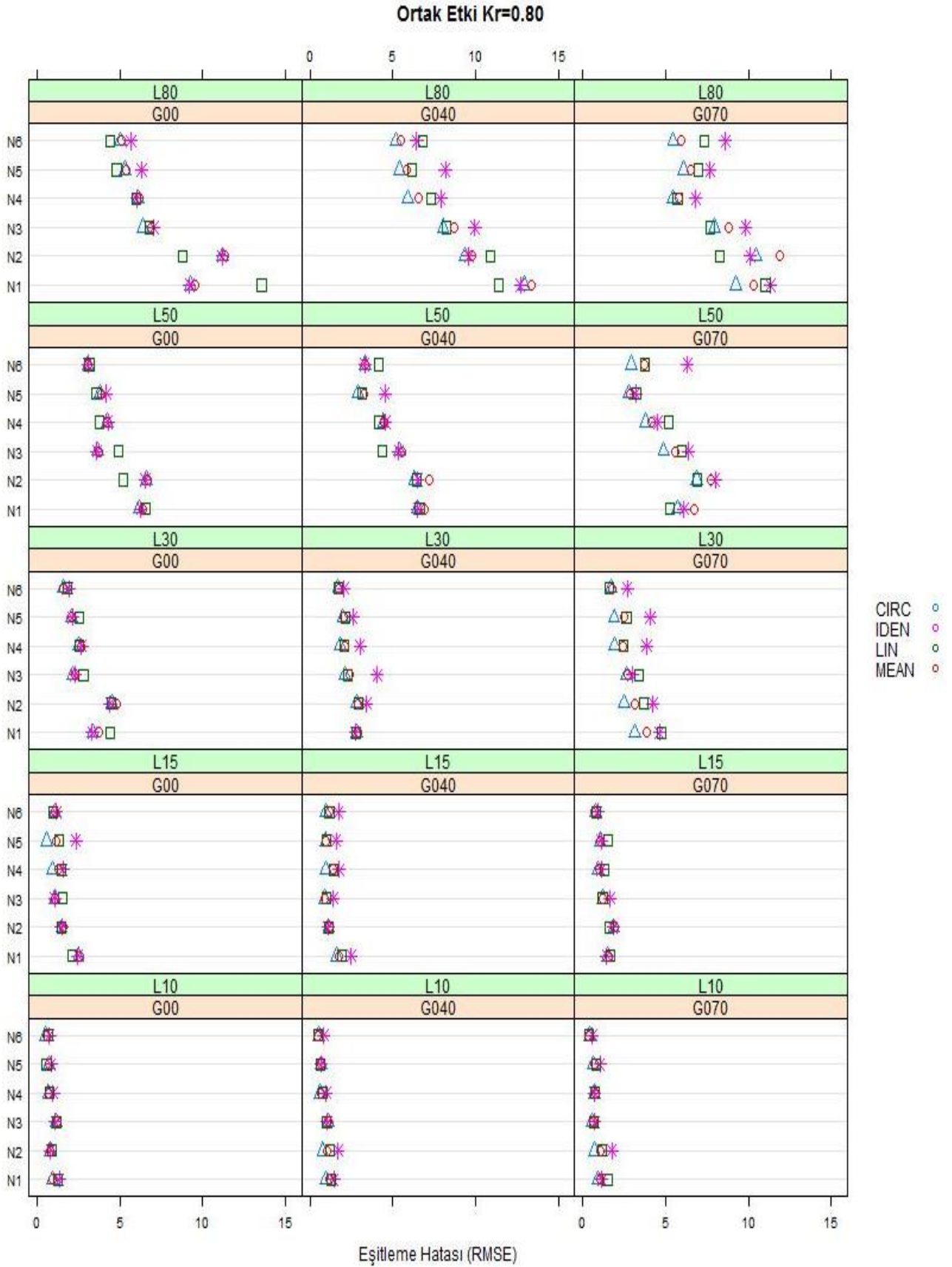
Grafik 35: 0.80 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme yanlılığı (BIAS) değerlerine ortak etkisi

2.2.1.2.3. Alt testleri arasında 0.80 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde edilen eşitleme hatasına (RMSE) ortak etkisi nasıldır?

Grafik36 incelendiğinde; artı değer özelliği gösteren birinci alt testin genişletilmiş alt test puanları kullanılarak eşitleme yöntemlerinden elde edilen eşitleme hatası değerleri örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki ortalama güçlük düzeyi farkı faktörlerinden etkilenmektedir.

Genel olarak örneklem büyüklüğü arttıkça yöntemlerin eşitleme hatası değerinde azalama meydana gelirken, test uzunluğu arttığında ise yöntemlerin eşitleme hatası değerlerinde artma görülmektedir. Test formları arasındaki ortalama güçlük düzeyi farkının artışı eşitleme hatası değerini de arttırmaktadır.





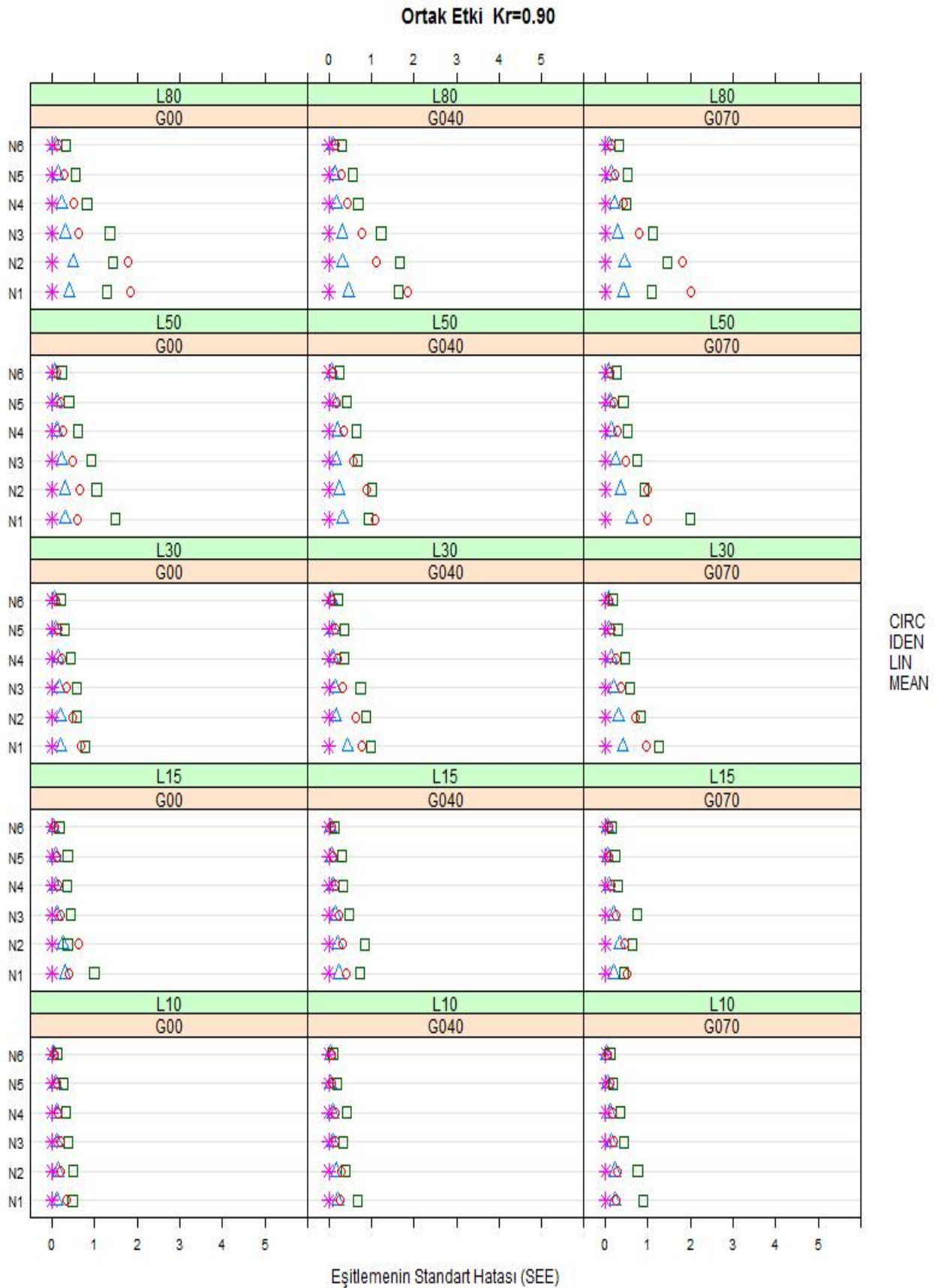
Grafik 36: 0.80 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farklı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme hatası (RMSE) değerlerine ortak etkisi

2.2.1.3.1. Alt testleri arasında 0.90 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde edilen eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?

Grafik37 incelendiğinde; artı değe özelliğine sahip birinci alt testin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucunda birim eşitleme yöntemi dışındaki diğer eşitleme yöntemlerinden elde edilen eşitlemenin standart hatası değerleri, örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı faktörlerinden etkilenmektedir.Örneklem büyüklüğü arttıkça eşitleme yöntemlerinin eşitleme standart hatasında azalma meydana gelirken, alt test uzunluğu arttıkça yöntemlerin eşitleme standart hatası değerlerinde artma görülmektedir.

Test formları arasında ortalama güçlük düzeyi farkı arttıkça eşitleme yöntemlerinin eşitleme standart hatası değerleri artmaktadır. Genel olarak ortalama güçlük düzeyi farkı ve test uzunluğu faktörlerinin düzeyleri arttıkça yöntemlerin eşitleme standart hatası değerlerinin büyükten küçüğe; zincirlenmiş lineer eşitleme, Braun/Holland eşitleme, dairesel yay eşitleme ve birim eşitleme yöntemi şeklinde düzenli bir artma gösterdiği görülmektedir.





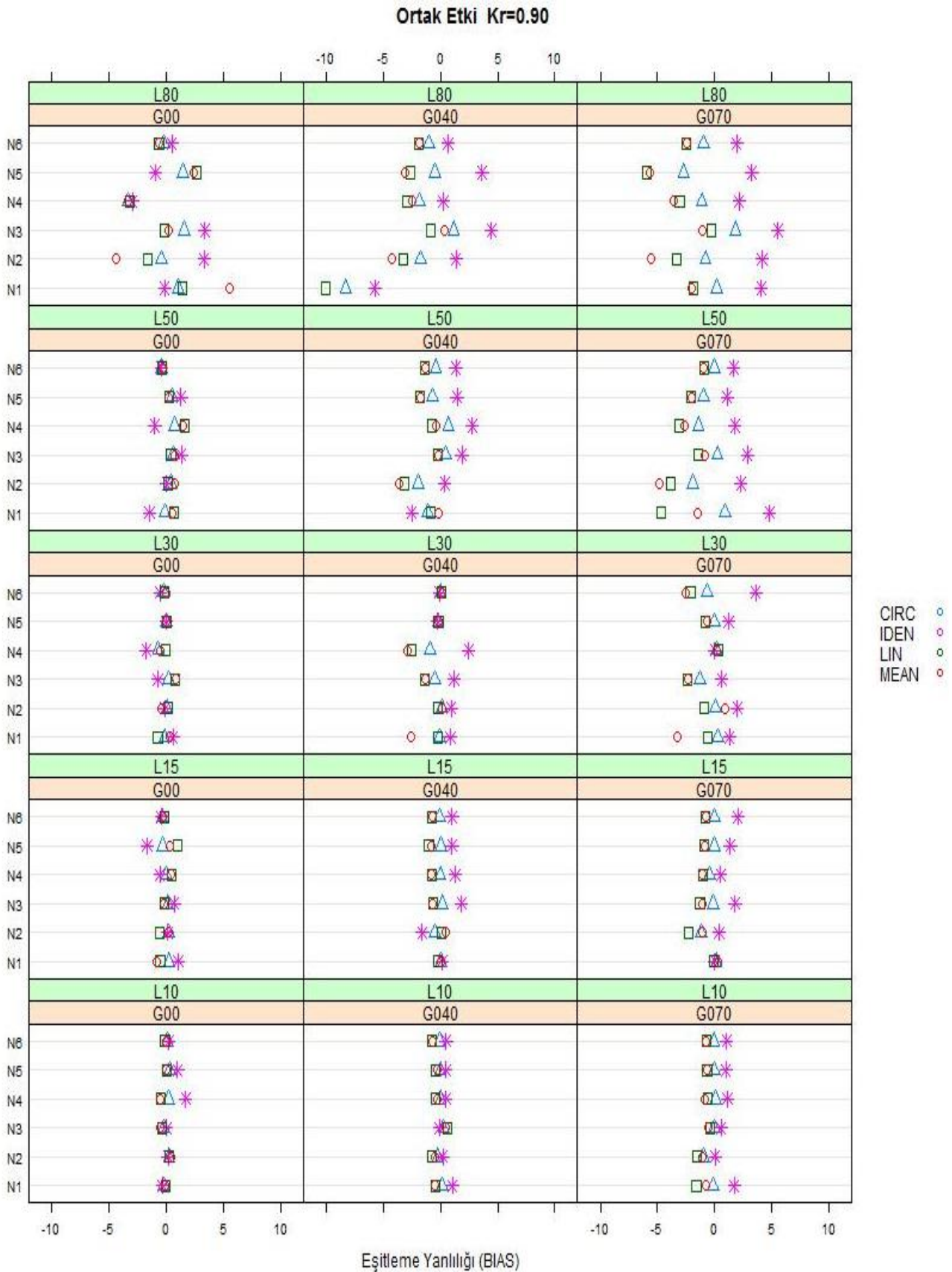
Grafik 37: 0.90 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farklı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme standart hata (SEE) değerlerine ortak etkisi

2.2.1.3.2. Alt testleri arasında 0.90 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlemesiyle elde edilen eşitlemenin yanlılığına (BIAS) ortak etkisi nasıldır?

Artı değer özelliğinde sahip birinci alt testin genişletilmiş alt test puanları eşitlemesiyle elde edilen Grafik38 incelendiğinde; alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörleri yöntemlerin eşitleme yanlılığı değerlerinde değişim meydana getirmektedir. Ortak faktörlerin birim eşitleme yönteminin eşitleme yanlılık değerlerinde daha çok değişme meydana getirdiği görülmektedir.

Test uzunluğu arttıkça eşitleme yöntemlerinin eşitleme yanlılığı değerleri artmaktadır. Test formları arasındaki ortalama güçlük düzeyi arttıkça birimeşitleme yönteminde eşitleme yanlılığı değerinin arttığı gözlenmektedir.





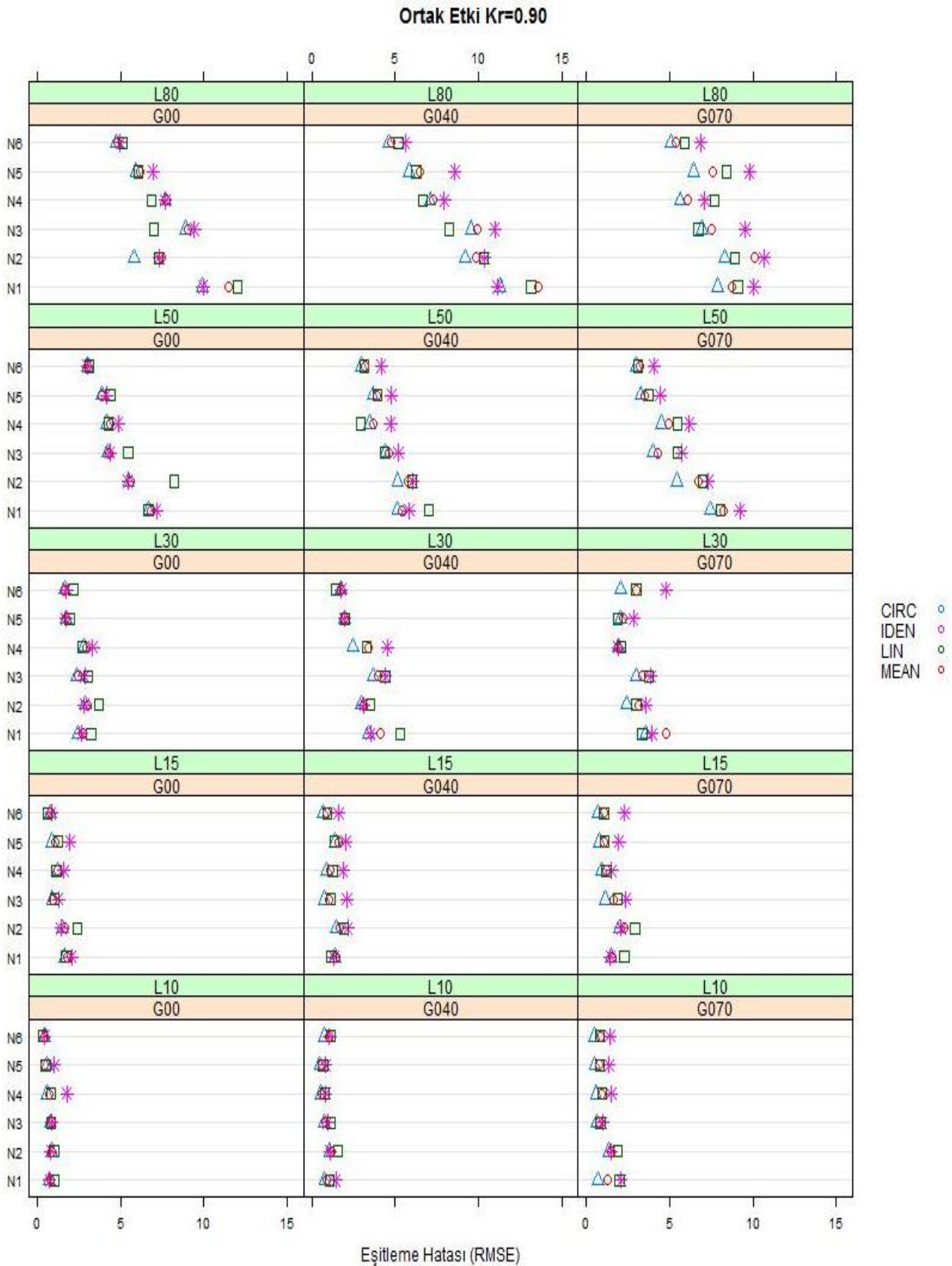
Grafik 38: 0.90 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme yanlılığı (BIAS) değerlerine ortak etkisi

2.2.1.3.3. Alt testleri arasında 0.90 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde edilen eşitleme hatasına (RMSE) ortak etkisi nasıldır?

Grafik39 incelendiğinde; artı değer özelliği gösteren birinci alt testin genişletilmiş alt test puanları kullanılarak elde edilen eşitleme sonucunda eşitleme yöntemlerinin eşitleme hatası değerleri örneklem büyüklüğünün, alt test uzunluğunun ve test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında değişim göstermektedir.

Genel olarak örneklem büyüklüğü arttıkça yöntemlerin eşitleme hatası değerlerinde azalma meydana gelirken, test uzunluğu arttığında ise yöntemlerin eşitleme hatası değerlerinde artma görülmektedir. Test formları arasındaki ortalama güçlük düzeyi farkının artışı yöntemlerin eşitleme hatası değerini de arttırmaktadır.





Grafik 39: 0.90 korelasyona sahip birinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farklı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme hatası (RMSE) değerlerine ortak etkisi

2.2.2.1.1. Alt testleri arasında 0.70 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde edilen eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?

Grafik40 incelendiğinde; artı deęe özelliđine sahip ikinci alt testin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucunda birim dönüşüm eşitleme yöntemi dışındaki diđer eşitleme yöntemlerinden elde edilen eşitlemenin standart hatası deđerleri örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı olmak üzere tüm faktörler altında deęişim göstermektedir.

Örneklem büyüklüğü arttıkça eşitleme yöntemlerinin eşitleme standart hatasında azalma meydana gelmektedir. Test uzunluğu arttıkça birim dönüşüm dışındaki tüm yöntemlerde eşitlemenin standart hatası deđerlerinde artış görülmektedir. Test uzunluğu 80 olduğunda eşitleme yöntemlerinin eşitleme standart hataları en yüksek deđerlere sahiptir. Genel olarak test formları arasındaki ortalama güçlük düzeyi arttıkça zincirlenmiş lineer eşitleme, Braun/Holland eşitleme, dairesel yay eşitleme ve birim eşitleme yöntemlerinin eşitleme standart hatası deđerleri büyükten küçüğe olacak şekilde sıralanma davranışı gösterdiği görülmektedir.



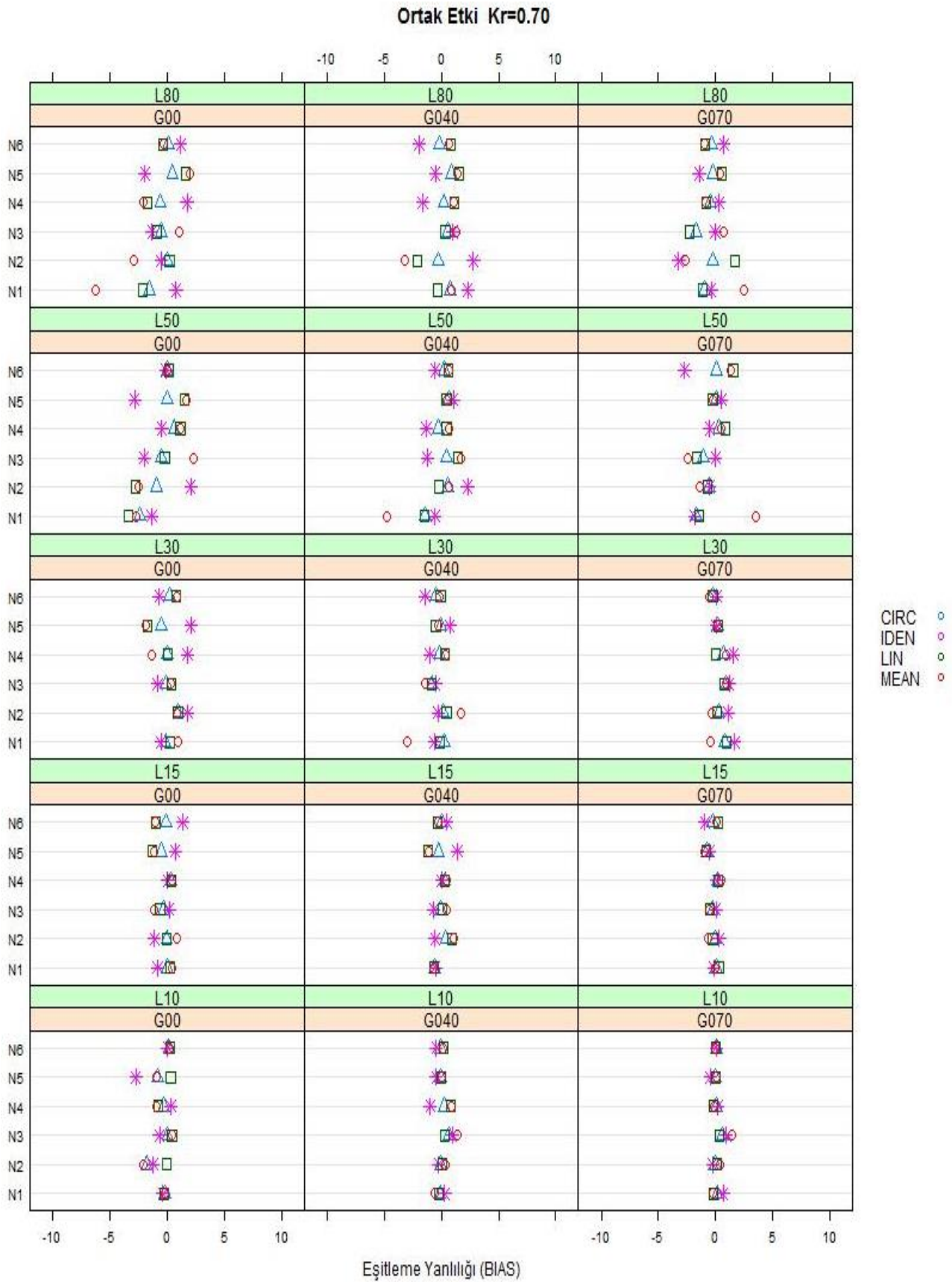
Grafik 40: 0.70 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme standart hata (SEE) değerlerine ortak etkisi

2.2.2.1.2. Alt testleri arasında 0.70 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlemesiyle elde edilen eşitlemenin yanlılığına (BIAS) ortak etkisi nasıldır?

Artı değer özelliğinde sahip ikinci alt testin genişletilmiş alt test puanları eşitlemesiyle elde edilen Grafik41 incelendiğinde; alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin eşitleme yanlılığı değeri değişiminde daha etkili olduğu gözlenmektedir. Tüm şartlar altında birim eşitleme yönteminin eşitleme yanlılığı hatası diğer yöntemlerin BIAS değerlerine göre daha çok değişkenlik göstermektedir.

Test uzunluğu arttıkça eşitleme yanlılığı değerindeki değişim artmaktadır. Test formları arasında ortalama güçlük düzeyinin 0.0 olduğu durumda yöntemlerin eşitleme yanlılığı değeri daha çok değişim gösterirken diğer ortalama güçlük düzeyi farklarında düzenli değişim davranışı gözlenmemektedir.





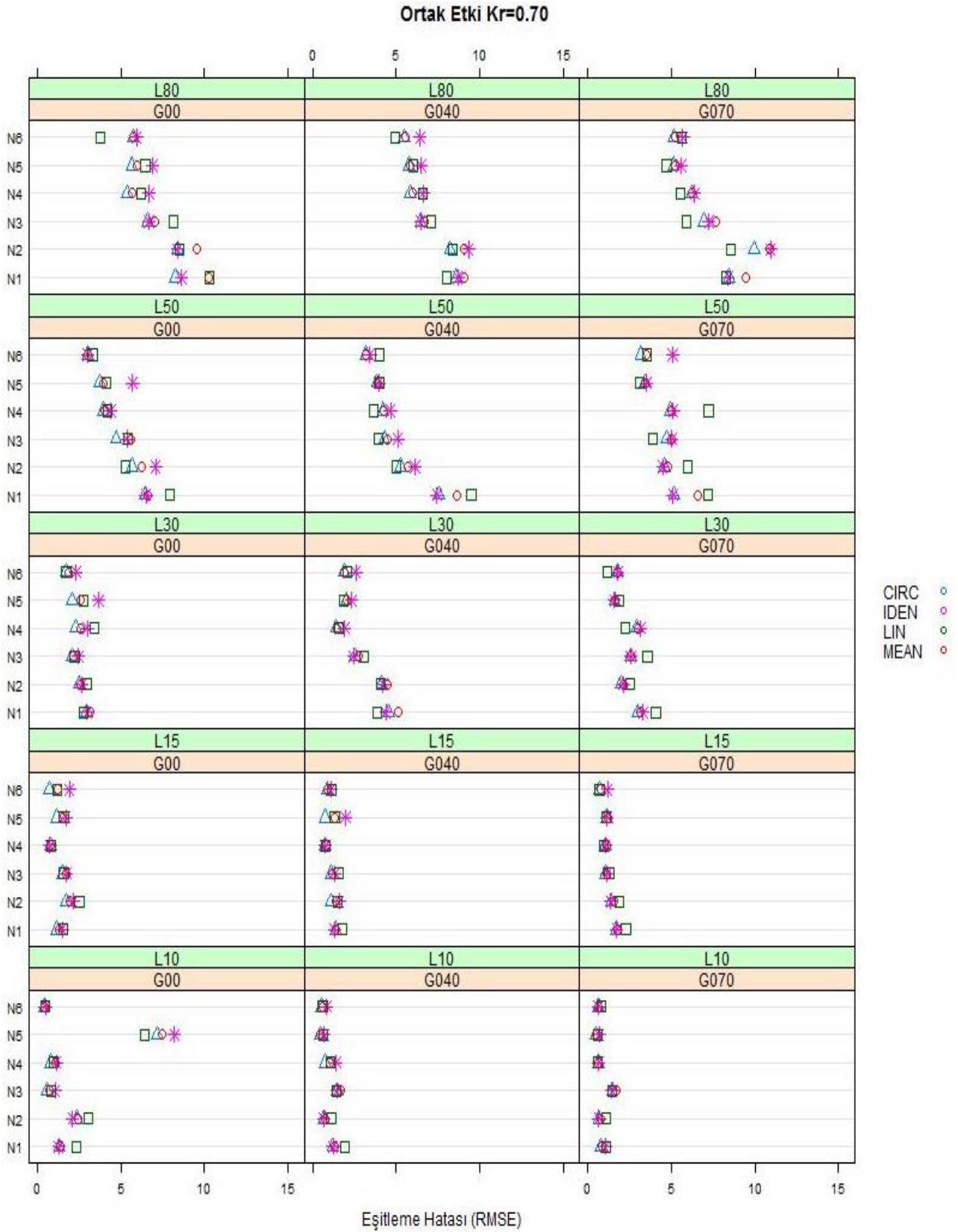
Grafik 41: 0.70 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme yanlılığı (BIAS) değerlerine ortak etkisi

2.2.2.1.3. Alt testleri arasında 0.70 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde edilen eşitleme hatasına (RMSE) ortak etkisi nasıldır?

Grafik42 incelendiğinde; artı değer özelliği gösteren ikinci alt testin genişletilmiş alt test puanları kullanılarak eşitleme yöntemlerinden elde edilen eşitleme hatası değerleri örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı faktörleri altında değişim göstermektedir.

Örneklem büyüklüğü arttıkça yöntemlerin eşitleme hatası değerleri tüm faktörler altında azalmaktadır. Test formları arasındaki güçlük düzeyi arttıkça eşitleme hatası değeri tüm yöntemlerde artmaktadır. Test uzunluğu 50 ve 80 olduğu durumlarda eşitleme yöntemlerinin eşitleme hatası değerinde daha fazla artış gözlenmektedir.





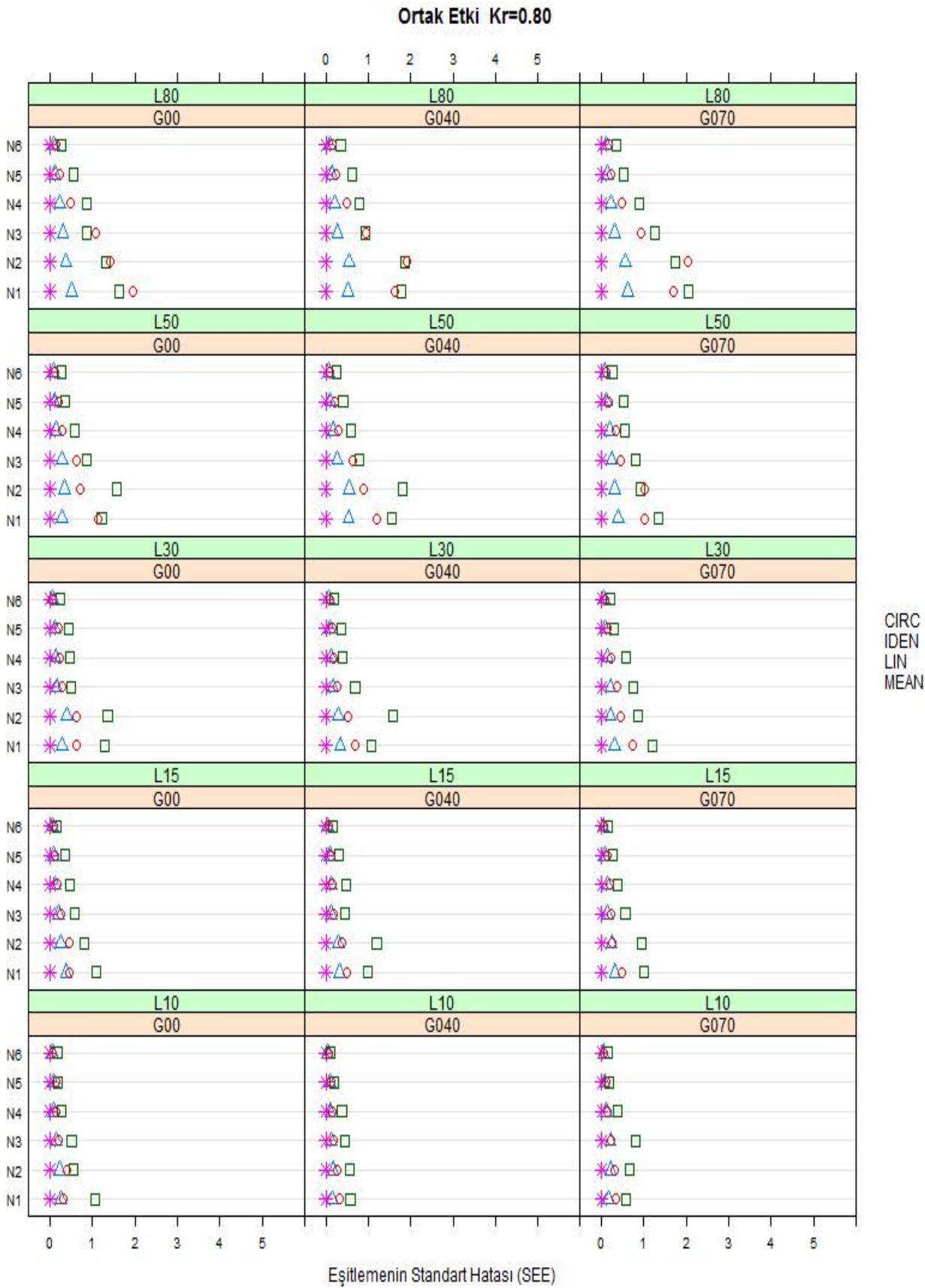
Grafik 42: 0.70 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme hatası (RMSE) değerlerine ortak etkisi

2.2.2.2.1. Alt testleri arasında 0.80 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde edilen eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?

Grafik43 incelendiğinde; Artı deęe özelliğine sahip ikinci alt testin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucunda eşitleme yöntemlerinden elde edilen eşitlemenin standart hatası deęerleri birim eşitleme yöntemi dışındaki yöntemlerde örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı olmak üzere tüm faktörlerde deęişim göstermektedir.

Örneklem büyüklüğü arttıkça eşitleme yöntemlerinin eşitleme standart hata deęerlerinde azalma meydana gelirken test uzunluğu arttıkça birim eşitleme dışındaki tüm yöntemlerde eşitlemenin standart hatası deęerinde artış görülmektedir.

Eşitleme yöntemlerinin eşitleme standart hatası deęeri, test formları arasında ortalama güçlük düzeyi farkı arttıkça artmaktadır. Genel olarak zincirlenmiş lineer eşitleme, Braun/Holland, dairesel yay eşitleme ve birim eşitleme yöntemi şeklinde eşitlemenin standart hata deęerleri büyükten küçüğe deęerler olarak sıralandığı görülmektedir.



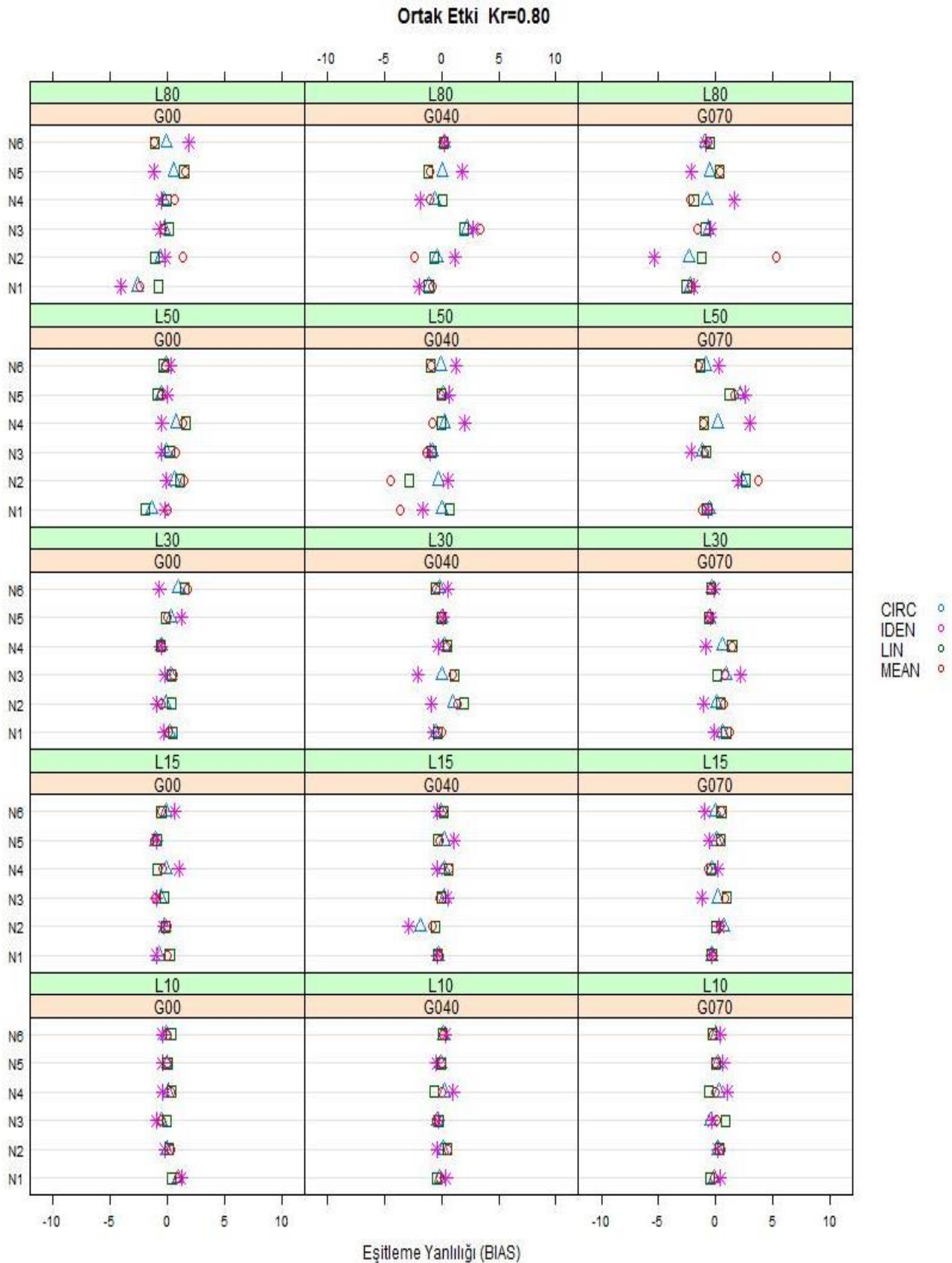
Grafik 43: 0.80 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme standart hata (SEE) değerlerine ortak etkisi

2.2.2.2.2. Alt testleri arasında 0.80 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlemesiyle elde edilen eşitlemenin yanlılığına (BIAS) ortak etkisi nasıldır?

Artı değer özelliğine sahip ikinci alt testin genişletilmiş alt test puanları eşitlemesiyle elde edilen Grafik44 incelendiğinde; alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörleri eşitleme yanlılığı değerlerinde değişime neden olmaktadır.

Alt test uzunluğu arttıkça eşitleme yöntemlerinin eşitleme yanlılığı değerleri artmaktadır. Bu artış en çok 50 ve 80 madde sayısına sahip testlerde belirgindir. Test formları arasındaki ortalama güçlük düzeyi arttıkça eşitleme yanlılığı değerinin arttığı gözlenmektedir.





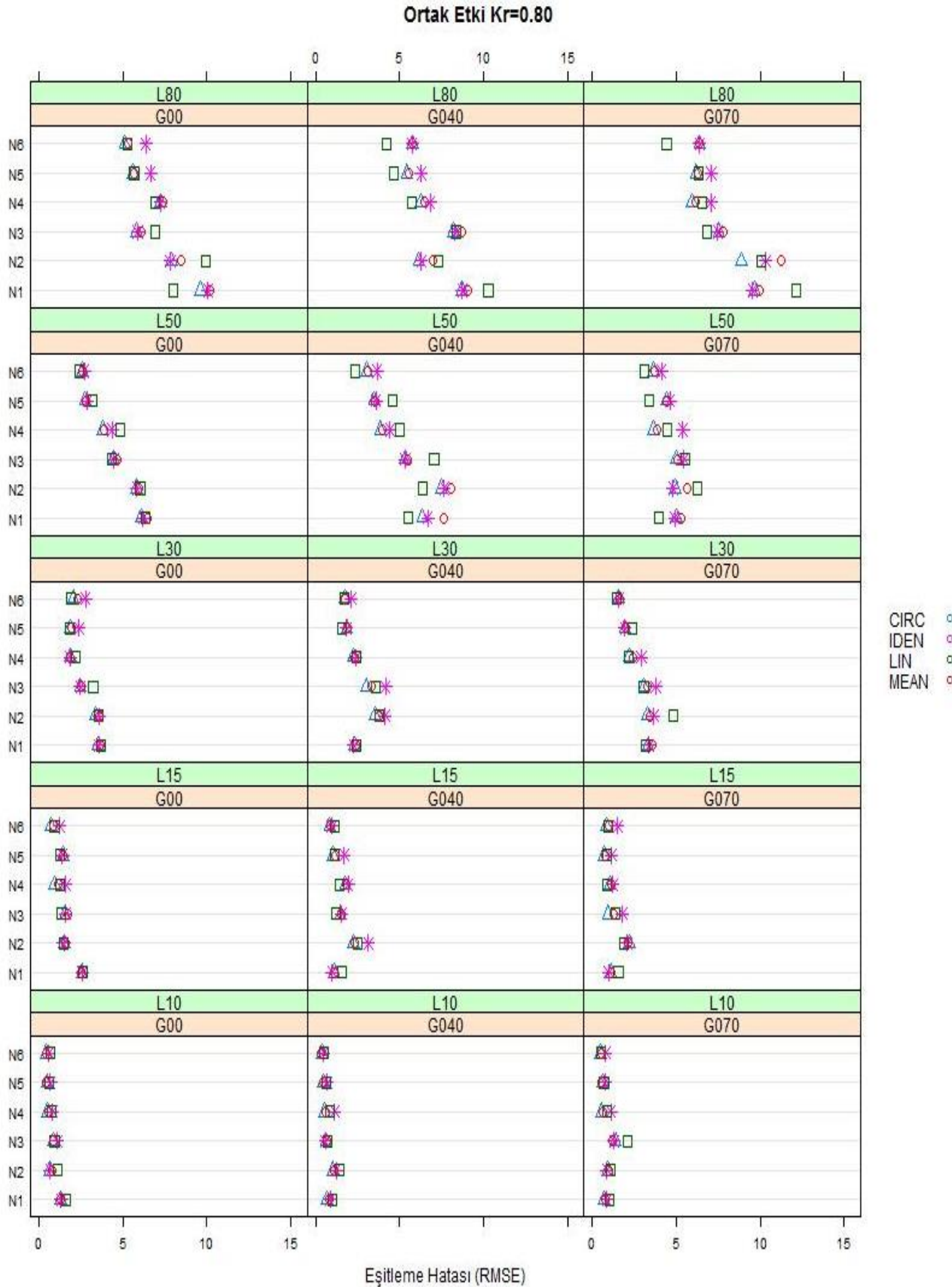
Grafik 44: 0.80 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme yanlılığı (BIAS) değerlerine ortak etkisi

2.2.2.2.3. Alt testleri arasında 0.80 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde edilen eşitleme hatasına (RMSE) ortak etkisi nasıldır?

Grafik45 incelendiğinde; artı değer özelliği gösteren ikinci alt testin genişletilmiş alt test puanları kullanılarak eşitleme yöntemlerinden elde edilen eşitleme hatası değerleri örneklem büyüklüğünün, alt test uzunluğunun ve test formları arasındaki ortalama güçlük düzeyi farkı faktörlerinden etkilenmektedir.

Genel olarak örneklem büyüklüğü arttıkça yöntemlerin eşitleme hatası değerinde azalma meydana gelmektedir. Test uzunluğu arttığında ise yöntemlerin eşitleme hatası değerlerinde artma görülmektedir. Test formları arasındaki ortalama güçlük düzeyi farkının artışı eşitleme hatası değerini de arttırmaktadır.





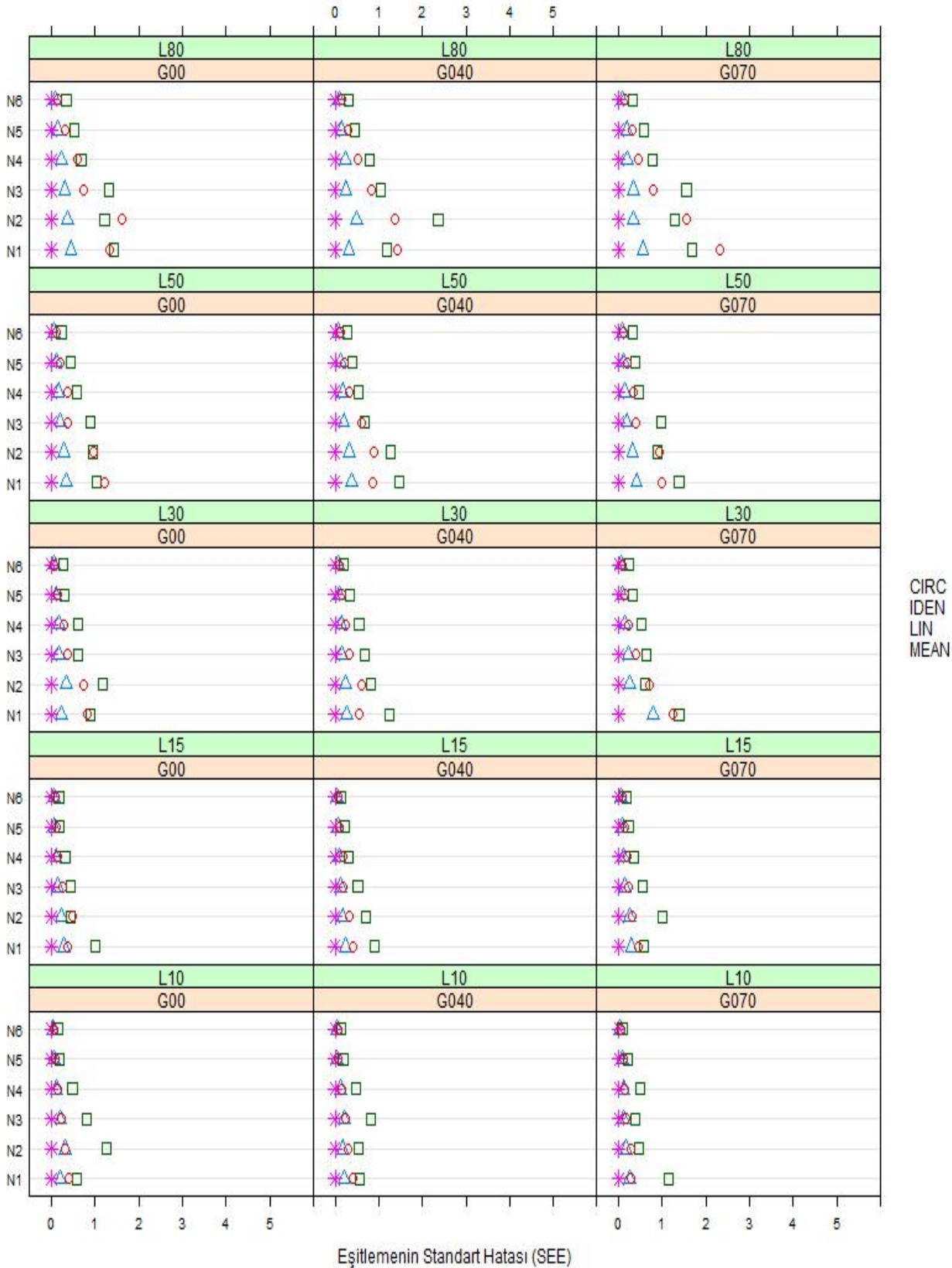
Grafik 45: 0.80 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farklı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme hatası (RMSE) değerlerine ortak etkisi

2.2.2.3.1. Alt testleri arasında 0.90 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde edilen eşitlemenin standart hatasına (SEE) ortak etkisi nasıldır?

Grafik46 incelendiğinde; artı değer özelliğine sahip ikinci alt testin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucunda eşitleme yöntemlerinden elde edilen eşitlemenin standart hatası değerleri örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki güçlük düzeyi farkı faktörleri birim eşitleme yöntemi dışındaki diğer eşitleme yöntemlerinin eşitleme standart hatası değerini değiştirmektedir.

Örneklem büyüklüğü arttıkça eşitleme yöntemlerinin eşitleme standart hatası değerlerinde azalma meydana gelirken test uzunluğu arttıkça eşitlemenin standart hatası değerlerinde artış görülmektedir. Test uzunluğunun 80 olduğu durumda eşitleme yöntemlerinin eşitlemenin standart hatası değerlerinde diğer şartlara göre daha çok değişme meydana gelmektedir.

Alt, test formları arasında ortalama güçlük düzeyi farkı arttıkça eşitleme yöntemlerinin eşitleme standart hatası değerlerinde artışın meydana geldiği görülmektedir. Genel olarak faktörlerin düzeyleri arttıkça, eşitleme yöntemlerinin SEE değerlerindeki büyükten küçüğe olan sıralama zincirleşmiş lineer eşitleme, Braun/Holland eşitleme, dairesel yay eşitleme ve birim eşitleme yöntemi şeklinde düzenli bir değişim göstermektedir.

Ortak Etki $Kr=0.90$ 

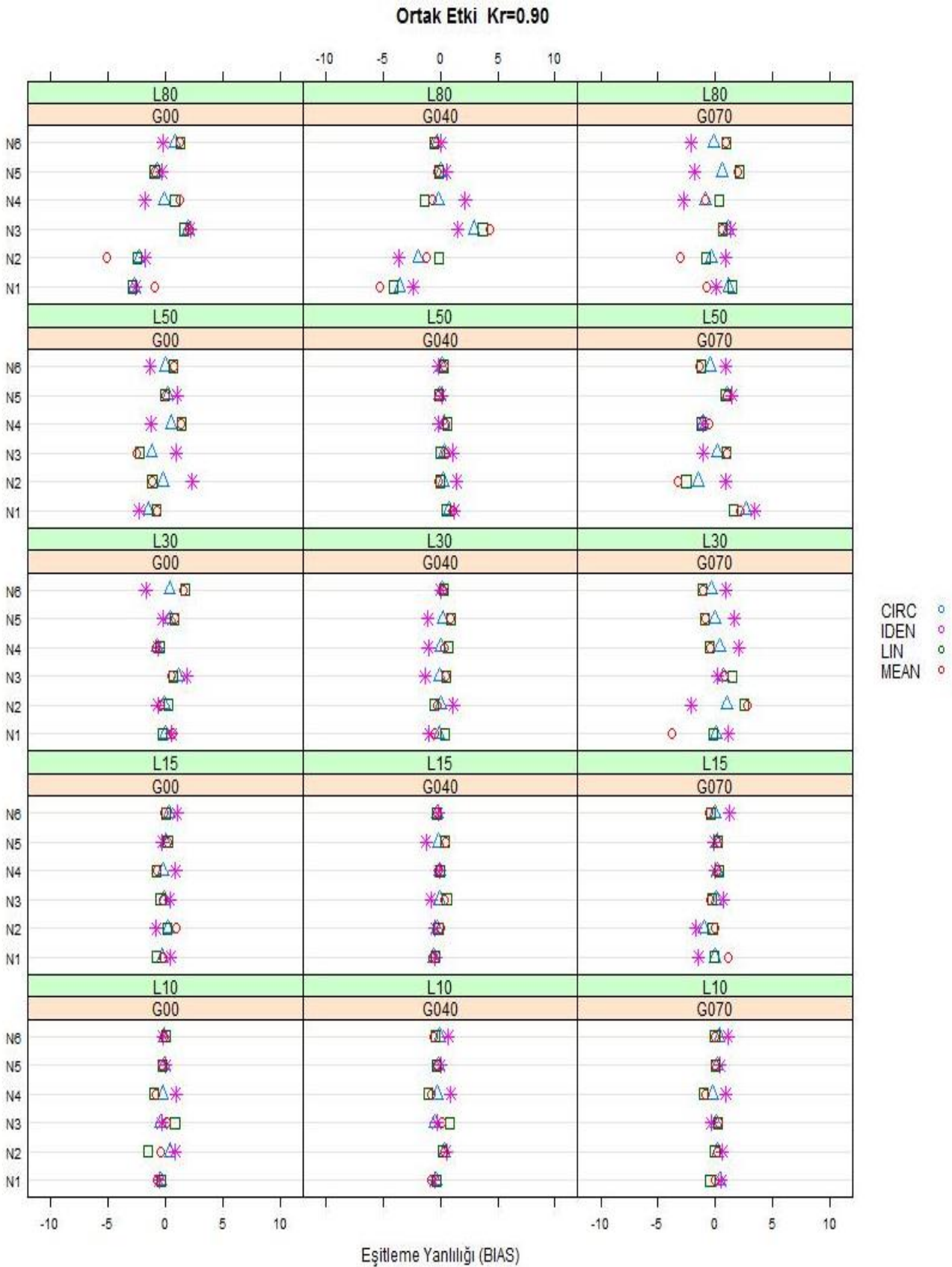
Grafik 46: 0.90 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme standart hata (SEE) değerlerine ortak etkisi

2.2.2.3.2. Alt testleri arasında 0.90 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde eşitlemenin yanlılığına (BIAS) ortak etkisi nasıldır?

Artı değer özelliğinde sahip ikinci alt testin genişletilmiş alt test puanları eşitlemesiyle elde edilen Grafik47 incelendiğinde; alt test uzunluğu ve ortalama güçlük düzeyi farkı faktörleri yöntemlerin eşitleme yanlılığı değerlerinde değişim meydana getirmektedir.

Alt test uzunluğu arttıkça eşitleme yöntemlerinin eşitleme yanlılığı değerleri artmaktadır. Bu artış en çok 30, 50 ve 80 madde sayısına sahip alt testlerde belirgindir. Test formları arasındaki ortalama güçlük düzeyi farkı arttıkça 80 madde sayısına sahip testlerde eşitleme yöntemlerinin eşitleme yanlılığı değerlerinde de artmanın meydana geldiği görülmektedir.





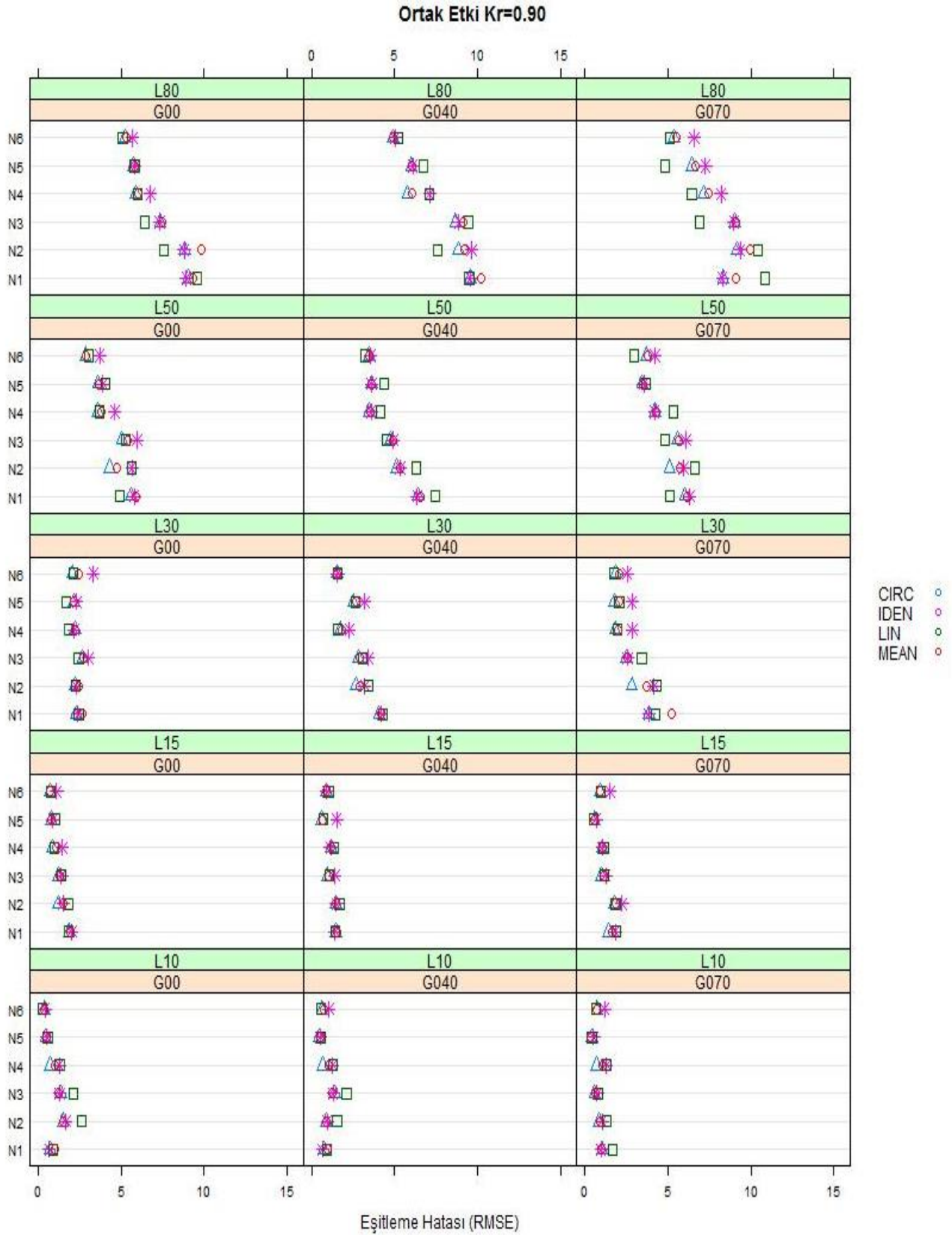
Grafik 47: 0.90 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farklı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme yanlılığı (BIAS) değerlerine ortak etkisi

2.2.2.3.3. Alt testleri arasında 0.90 korelasyon bulunan testlerin alt testlerinde; örneklem büyüklüğü, alt test uzunluğu, test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında genişletilmiş alt test puanları eşitlenmesiyle elde eşitleme hatasına (RMSE) ortak etkisi nasıldır?

Grafik48 incelendiğinde; artı değer özelliği gösteren ikinci alt testin genişletilmiş alt test puanları kullanılarak elde edilen eşitleme sonucunda eşitleme yöntemlerinin eşitleme hatası değerleri örneklem büyüklüğü, alt test uzunluğu ve test formları arasındaki ortalama güçlük düzeyi farkı faktörleri altında değişim göstermektedir.

Genel olarak örneklem büyüklüğü arttıkça yöntemlerin eşitleme hatası değerlerinde azalama meydana gelirken test uzunluğu arttığında ise yöntemlerin eşitleme hatası değerlerinde de artma görülmektedir. Test formları arasındaki ortalama güçlük düzeyi farkının artışı yöntemlerin eşitleme hatası değerini de arttırmaktadır.





Grafik 48: 0.90 korelasyona sahip ikinci alt testler için örneklem büyüklüğü, alt test uzunluğu ve ortalama güçlük düzeyi farklı faktörlerinin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucu eşitleme yöntemlerinin eşitleme hatası (RMSE) değerlerine ortak etkisi

BÖLÜM IV: TARTIŞMA VE YORUM

Bu bölümde çeşitli faktörlerin eşitleme yöntemleri hata değerlerine olan temel etkisi ve ortak etkisi üzerine elde edilen sonuçlara yer verilmiştir.

1. Örneklem büyüklüğünün alt testlerin alt test puanlarının eşitlemede kullanılmasıyla test eşitleme yöntemlerinden elde edilen RMSE, BIAS ve SEE değerlerine temel etkisinin değerlendirilmesi

Örneklem büyüklüğü arttıkça birim eşitleme yöntemleriyle birlikte diğer yöntemlerde RMSE değerlerinde azalma görülmektedir. 500 örneklem büyüklüğünde tüm eşitleme yöntemlerinin RMSE değerleri en düşük seviyededir. Ayrıca her iki alt testte de 100, 200 ve 500 örneklem büyüklüğünde en yüksek RMSE değeri birim eşitleme yönteminde görülmektedir. Bu seviyede en düşük RMSE değerini dairesel yay eşitleme yöntemi vermektedir.

Örneklem büyüklüğü arttıkça her iki alt testte eşitleme için tüm eşitleme yöntemlerinin BIAS değerlerinde azalma meydana geldiği görülmektedir. Birinci alt testte birim eşitlemenin BIAS değerleri en yüksek değere sahipken en düşük BIAS değerleri dairesel yay eşitleme yönteminde gözlenmektedir. İkinci alt test için 20, 25 ve 50 örneklem büyüklüklerinde birim eşitlemenin BIAS değerleri Braun/Holland eşitleme yöntemine göre daha düşüktür. Bu şartlar altında en düşük BIAS değerlerine zincirilmiş lineer eşitleme yöntemi sahiptir. 100, 200 ve 500 örneklem büyüklüklerinde tüm yöntemlerin BIAS değerleri birbirine çok yakınken, en düşük BIAS değeri 500 örneklem büyüklüğünde görülmektedir.

Tüm örneklem büyüklüklerinde en düşük SEE değeri birim eşitleme yönteminde görülmektedir. Her iki alt test için örneklem büyüklüğü 20, 25 ve 50 olduğunda en yüksek SEE değeri Braun/Holland eşitleme yönteminde gözlenirken, 100, 200 ve 500 örneklem büyüklüklerinde en yüksek SEE değeri zincirilmiş lineer eşitleme yönteminde görülmektedir. Birim eşitleme yönteminden sonra en düşük SEE değeri dairesel yay eşitleme yönteminde görülmektedir. Birim eşitleme yöntemi hariç, örneklem büyüklüğü arttıkça eşitleme yöntemlerinin SEE değerlerinde azalma gözlenmektedir.

Artı değer özelliği göstermeyen birinci ve ikinci alt test için birim dönüşüm yönteminin tüm örneklem büyüklüklerinde SEE değeri sıfırdır. Ayrıca birinci ve ikinci alt test için örneklem büyüklüğü arttıkça kullanılan eşitleme yöntemlerinin SEE değerlerinde azalma meydana gelmiştir.

2. Örneklem büyüklüğünün alt testlerin genişletilmiş alt puanlarının eşitlemede kullanılmasıyla test eşitleme yöntemlerinden elde edilen RMSE, BIAS ve SEE değerlerine temel etkisinin değerlendirilmesi

Birinci ve ikinci alt testin genişletilmiş alt test puanlarıyla yapılan eşitleme sonucunda; 50,100, 200 ve 500 örneklem büyüklüklerinde en yüksek RMSE değerleri birim eşitleme yönteminde görülmektedir. Tüm örneklem büyüklüklerinde en düşük RMSE değerini dairesel yay eşitleme yönteminin verdiği gözlenmektedir. Genel olarak örneklem büyüklüğü arttıkça tüm eşitleme yöntemlerinin RMSE değerlerinin azaldığı gözlenmektedir.

Her iki alt test için 20, 25 ve 50 örneklem büyüklüklerinde Braun/Holland eşitleme yönteminin BIAS değeri birim eşitleme yönteminin BIAS değerinden daha yüksektir. Dolayısıyla en yüksek eşitleme yanlılığı hatası bu örneklem büyüklüklerinde Braun/Holland eşitleme yönteminde görülmektedir. Genel olarak dairesel yay eşitleme yönteminin BIAS değeri birim eşitleme yönteminin BIAS değerinden daha düşüktür. Ayrıca eşitleme hatası değeri en az olan yöntem dairesel yay eşitleme yöntemi olduğu görülmektedir.

Her iki alt testin eşitlenmesi sonucunda en düşük SEE değerinin birim eşitleme yöntemi verirken, en yüksek SEE değeri zincirlenmiş lineer eşitleme yönteminde görülmektedir. Birim eşitleme yönteminden sona en düşük SEE değeri dairesel yay eşitleme yönteminde gözlenmektedir. Genel olarak örneklem büyüklüğü arttıkça birim eşitleme yöntemi dışındaki kimi yöntemlerin SEE değerlerinde azalma görülmektedir.

Alt testlerin ham alt puanı ve genişletilmiş alt puanlarıyla yapılan eşitleme sonucunda elde edilen sonuçlara bakıldığında örneklem büyüklüğünün artması RMSE değerlerinde azalamaya neden olmaktadır. Elde edilen sonuç Kim ve Livingston (2009) çalışmasına paraleldir. Çalışma da büyük örneklerde örneklem büyüklüğü arttıkça BIAS değerlerinde artmanın olduğu sonucuna ulaşılmıştır. Bu sonuç Skaggs (2005) sonucuyla örtüşmektedir. Birim dönüşüm eşitleme yöntemi tüm faktörlerin tüm düzeyleri altında en düşük SEE değeri verdiği görülmüştür. Örneklem büyüklüğü arttıkça birim dönüşüm dışındaki tüm yöntemlerin eşitlemenin standart hatası değerinde azalma meydana gelmiştir. Elde edilen sonuç literatürdeki çalışmalarla benzer sonuç göstermiştir (Kolen ve Brennan, 2004; Skaggs, 2005; Kim ve diğerleri, 2008).

3. Test uzunluğunun, alt testlerin alt test puanlarının eşitlemede kullanılmasıyla test eşitleme yöntemlerinden elde edilen RMSE, BIAS ve SEE değerlerine temel etkisinin değerlendirilmesi

10 , 15 ve 30 madde sayısına sahip birinci ve ikinci alt testlerin eşitlenmesi sonucunda; en yüksek RMSE değeri birim eşitleme yönteminde olduğu görülmektedir. 50 ve 80 alt test uzunluklarında birinci alt test için en yüksek RMSE değeri birim eşitleme yönteminde görülürken, ikinci alt testte ise en yüksek RMSE değeri Braun/Holland eşitleme yönteminde görülmektedir. Her iki alt testin eşitleme yöntemlerinin RMSE değerlerine bakıldığında; tüm alt test uzunluklarında en düşük eşitleme hatasını dairesel yay eşitleme yönteminin verdiği gözlenmektedir.

Birim eşitleme yöntemi 15, 30 ve 80 madde sayısına sahip her iki alt testin eşitlenmesinde en yüksek BIAS değerini vermektedir. 10, 15 ve 30 alt test uzunluğunda en düşük BIAS değeri dairesel yay eşitleme yönteminde görülmektedir. Dolayısıyla bu şartlar altında en az hata veren eşitleme yöntemi dairesel yay eşitleme yöntemidir. Genel olarak 50 ve 80 alt test uzunluklarında tüm yöntemlerin BIAS değerlerinde artmanın meydana geldiği görülmektedir.

Her iki alt testin eşitlenmesi sonucunda elde edilen bulgulara göre; tüm alt test uzunluklarında en düşük SEE değerini birim eşitleme yönteminin verdiği görülmektedir. 10 , 15 ve 30 alt test uzunluklarında en yüksek SEE değeri zincirlenmiş lineer eşitleme yönteminde görülürken, 50 ve 80 test uzunluklarında en yüksek SEE değeri Braun/Holland eşitleme yönteminde gözlenmektedir. Birim eşitleme yönteminden sonra en az eşitlemenin standart hatasına dairesel yay eşitleme yöntemi sahiptir.

4. Test uzunluğunun, alt testlerin genişletilmiş alt test puanlarının eşitlemede kullanılmasıyla test eşitleme yöntemlerinden elde edilen RMSE, BIAS ve SEE değerlerine temel etkisinin değerlendirilmesi

Her iki alt testte 10 madde sayısına sahip alt testler haricindeki tüm alt test uzunluklarında birim eşitleme yönteminin en yüksek RMSE değerine sahip olduğu görülmektedir. Tüm alt test uzunluklarında en düşük eşitleme hatası değeri dairesel yay eşitleme yönteminde gözlenmektedir. Genel olarak eşitleme yöntemlerinin RMSE değerleri birbirine yakın olmakla birlikte alt test uzunluğu arttıkça yöntemlerin RMSE değerlerinde artma görülmektedir.

Her iki alt testin eşitlenmesiyle yöntemlerden elde edilen BIAS bulgularına göre; 50 ve 80 alt test uzunluklarında tüm eşitleme yöntemlerinin BIAS değerlerinde artma gözlenmektedir. 10 alt test uzunluğunda ise tüm yöntemlerin BIAS değerleri diğer durumlara göre en düşük değerini almaktadır. Birinci alt test için 10, 15 ve 30 alt test uzunluğunda birim eşitleme en yüksek BIAS değerine sahipken, en düşük eşitleme yanlılığı değerini dairesel yay eşitleme yönteminin verdiği görülmektedir. İkinci alt test için 15 ve 80 alt test uzunluklarında birim eşitleme yöntemi en yüksek BIAS değerine sahipken, bu şartlarda en düşük eşitleme yanlılığı değerine Braun/Holland eşitleme yöntemi ve dairesel yay eşitleme yönteminin verdiği gözlenmektedir.

Her iki alt testin eşitlenmesi sonucunda yöntemlerden elde edilen SEE bulgularına göre; en düşük SEE değeri birim eşitleme yönteminde görülmektedir. Birim eşitleme yönteminden sonra en düşük eşitleme standart hatası değerini dairesel yay eşitleme yöntemi vermektedir. Zincirlenmiş lineer eşitleme yönteminin en yüksek SEE değerini verdiği görülmektedir.

Alt testlerin alt test ve genişletilmiş alt test puanlarıyla yapılan eşitleme sonucunda test uzunluğunun artmasıyla eşitleme yöntemlerinin RMSE değerlerinde artma görülmüştür. Benzer şekilde test uzunluğunun artmasıyla eşitleme yöntemlerinin BIAS değerlerinde artış gözlenmiştir. Test uzunluğundaki artış yöntemlerin SEE değerlerinin artmasına neden olmuştur. Genel olarak bakıldığında dairesel yay eşitleme yönteminin eşitleme hataları değerleri diğer yöntemlere göre daha az hata değerine sahip olduğu sonucuna ulaşılmıştır. Eşitleme hatalarında elde edilen sonuçlar Aşiret (2014) çalışmasıyla paralellik göstermektedir.

5. Test formları arasındaki ortalama güçlük farkının, alt testlerin alt test puanlarının eşitlemede kullanılmasıyla test eşitleme yöntemlerinden elde edilen RMSE, BIAS ve SEE değerlerine temel etkisinin değerlendirilmesi

Birinci alt testte 0.4 ve 0.7 ortalama güçlük düzeyi farklarında en yüksek RMSE değerini birim eşitleme yöntemi verirken, 0.0 ortalama güçlük düzeyi farkında Braun/Holland eşitleme yönteminden sonra en yüksek RMSE değerini birim eşitleme yönteminin verdiği gözlenmektedir. İkinci alt testte ise tüm ortalama güçlük düzeyi farklarında en yüksek RMSE değerini Braun/Holland eşitleme yöntemi verirken. İkinci en yüksek RMSE değerini birim eşitleme yönteminin verdiği gözlenmektedir. Dolayısıyla Braun/Holland eşitleme yönteminin eşitleme hatası değeri daha fazladır. Her iki alt test için en düşük eşitleme hatası değerini veren eşitleme yöntemi dairesel yay eşitlemedir.

Ortalama güçlük düzeyi farkı faktörüne göre; her iki alt testin eşitleme yöntemlerinden elde edilen BIAS bulguları incelendiğinde test formları arasındaki ortalama güçlük düzeyi farkı arttıkça birim eşitleme yönteminin eşitleme yanlılığı değerinin arttığı görülmektedir. Birinci alt testte en düşük eşitleme yanlılığına dairesel yay eşitleme yöntemi sahipken, ikinci alt testte 0.0 ve 0.7 düzeylerinde en düşük BIAS değerine Braun/Holland eşitleme yöntemi, 0.4 düzeyinde ise zincirlenmiş lineer eşitleme yöntemi sahiptir.

Tüm ortalama güçlük düzeyi farklarında eşitleme yöntemlerinin SEE değerlerinde belirgin bir artma ya da azalma davranışı görülmemektedir. En yüksek SEE değerine Braun/Holland eşitleme yöntemi sahiptir. En düşük eşitleme standart hatası değerine birim eşitleme yönteminden sonra dairesel yay eşitleme yöntemi sahiptir.

6. Test formları arasındaki ortalama güçlük farkının, alt testlerin genişletilmiş alt test puanlarının eşitlemede kullanılmasıyla test eşitleme yöntemlerinden elde edilen RMSE, BIAS ve SEE değerlerine temel etkisinin değerlendirilmesi

Tüm ortalama güçlük düzeylerinden en yüksek eşitleme hatası birim eşitleme yönteminden görülmektedir. En düşük eşitleme hatasını veren eşitleme yöntemi dairesel yay eşitleme yöntemidir.

Birinci alt testte ortalama güçlük düzeyi farkı arttıkça eşitleme yöntemlerinin BIAS değerlerinde artma gözlenmektedir. 0.0 ve 0.4 düzeylerinde en yüksek BIAS değerleri birim eşitleme yöntemindeyken , 0.7 ortalama güçlük düzeyi farkında ise zincirilmiş lineer eşitleme yöntemindedir. Genel olarak en düşük eşitleme yanlışlığın değerine dairesel yay eşitleme yöntemi sahiptir. İkinci alt testin BIAS bulgularına bakıldığında 0.0 ortalama güçlük düzeyi farkında en yüksek BIAS değerine birim eşitleme yöntemi sahipken , en düşük BIAS değerine zincirilmiş lineer eşitleme yöntemi sahiptir. 0.4 ortalama güçlük düzeyi farkında Braun/Holland eşitleme yönteminden sonra en yüksek BIAS değeri birim eşitleme yönteminde görülürken , en düşük BIAS değeri zincirilmiş lineer eşitleme yönteminde görülmektedir. 0.7 ortalama güçlük düzeyinde ise en yüksek eşitleme yanlışlığı değeri zincirilmiş lineer eşitleme yönteminde görülürken , en düşük BIAS değeri Braun/Holland eşitleme yönteminde görülmektedir.

Tüm ortalama güçlük düzeyleri farkında en yüksek SEE değerini veren eşitleme yöntemi zincirilmiş lineer eşitleme yöntemidir. Tüm düzeylerde en düşük SEE değerini birim eşitleme yönteminin verdiği görülmektedir. Birim eşitleme yönteminden sonra en düşük SEE değeri dairesel yay eşitleme yönteminde gözlenmektedir.

Alt test ve genişletilmiş alt test puanlarıyla yapılan eşitleme sonuçlarına bakıldığında test formları arasındaki ortalama güçlük düzeyi farkı arttıkça birim dönüşüm eşitleme yönteminin RMSE değerlerinde artış meydana gelmiştir. Elde edilen bu sonuç Kim, von Davier ve Haberman (2008) çalışmasıyla paralellik göstermektedir. Ortalama güçlük düzeyi arttıkça eşitleme yöntemlerinin SEE değerlerinde değişim gözlenmemiştir. Genel olarak bakıldığında eşitleme yöntemlerinin BIAS değerlerinde ortalama güçlük farkı düzeyi arttıkça yöntemlerin BIAS değerlerinde değişim gözlenmemiştir.

7. Çeşitli faktörlerin eşitleme yöntemlerinden elde edilen eşitlemenin standart hatasına (SEE) ortak etkisinin değerlendirilmesi

0.70 korelasyona sahip birinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede örneklem büyüklüğü arttıkça birim eşitleme yöntemi hariç eşitleme yöntemlerinin eşitlemenin standart hatası (SEE) değerinde azalma gözlenmektedir. Küçük örneklerde test formları arasındaki ortalama düzey farkı arttıkça en yüksek SEE değerini Braun/Holland eşitleme yöntemi vermektedir. Genel olarak faktör düzeylerinin artışıyla eşitleme yöntemlerinin SEE değerlerinde gözlenen artan sıralama birim eşitleme, zincirilmiş lineer eşitleme ve Braun/Holland eşitleme yöntemi şeklindedir.

0.80 korelasyona sahip birinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede örneklem büyüklüğü arttığında diğer tüm faktörler düzeylerinde eşitleme yöntemlerinin SEE değerlerinde azalmanın meydana geldiği gözlenmektedir. Alt test uzunluğu ve ortalama güçlük düzeyleri farkı arttıkça yüksek örneklem büyüklüğünde genel olarak yöntemlerin SEE değerlerinde belirgin farklılaşma görülmemektedir. Ancak örneklem büyüklüğü azaldıkça diğer faktörlerin düzeylerindeki artışla birlikte birim eşitleme dışındaki diğer eşitleme yöntemlerinin SEE değerlerinde belirgin artışın olduğu gözlenmektedir. En yüksek alt test uzunluklarında en yüksek eşitleme standart hatası değeri genel olarak Braun/Holland

eşitleme yönteminde gözlenirken, yöntemlerin SEE değerindeki azalan sıralama zincirilmiş lineer eşitleme, dairesel yay eşitleme ve birim dönüşüm yöntemleri şeklindedir.

0.90 korelasyona sahip birinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede test uzunluğu ve ortalama güçlük düzeyi farkları arttığında eşitleme yöntemlerinin SEE değerlerinde artış görülse de örneklem büyüklüğü arttıkça SEE değerlerinde azalma meydana gelmektedir. 10 ve 15 maddeye sahip alt testlerden ve 50, 25, 20 örneklem büyüklüklerinde eşitleme yöntemlerinin birbirlerine göre SEE değerlerinde farklılaşmanın meydana geldiği görülmektedir. 100, 200 ve 500 örneklem büyüklüklerinde farklılaşma belirgin değildir. Ancak 30 test uzunluğuyla birlikte yöntemlerin birbirlerine göre SEE değerlerinde farklılaşma görülmektedir. 80 ve 50 madde uzunluğuna sahip olan testlerde yöntemlerin örneklem büyüklüğüne göre SEE değerlerinde belirgin farklılaşma gözlenirken kendi aralarında da SEE değerlerinin farklılaştığı görülmektedir.

0.70 korelasyona sahip ikinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede tüm faktörler altında örneklem büyüklüğü arttıkça birim eşitleme dışındaki diğer yöntemlerin SEE değerlerinde azalma görülmemektedir. Tüm faktörler altında birim eşitlemeden sonra en az SEE değeri dairesel yay yönteminde görülmektedir. Test uzunluğu ve güçlük farkı arttığında özellikle Braun/Holland eşitleme ve zincirilmiş lineer eşitleme yöntemlerinin SEE değerlerindeki farklılaşmanın diğer yöntemlere göre daha çok olduğu görülmektedir.

0.80 korelasyona sahip ikinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede örneklem büyüklüğü arttıkça yöntemlerin SEE değerlerinde azalma görülmektedir. 10 ve 15 madde sayısına sahip alt testlerde ortalama güçlük düzeyindeki artış zincirilmiş eşitleme yönteminin SEE değerlerinde farklılaşmaya neden olsa da genel olarak bu alt test uzunluklarında yöntemlerin SEE değerleri birbirine yakın olduğu gözlenmektedir. Test uzunluğu 30, 50 ve 80 olduğu durumlarda ortalama güçlük düzeyi farkı arttıkça birim eşitleme dışındaki diğer eşitleme yöntemlerinin SEE değerlerinde belirgin artmanın olduğu görülmektedir. Ayrıca test uzunluğu arttıkça yöntemlerin SEE değerlerinde artma gözlenmektedir.

0.90 korelasyona sahip ikinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede test uzunluğu arttıkça aynı şartlar altında birim eşitleme dışındaki yöntemlerin SEE değerlerinde artma meydana gelirken, test formları arasındaki ortalama güçlük düzeyi arttıkça birim eşitleme dışındaki eşitleme yöntemlerinin SEE değerlerinde artış gözlenmektedir. Genel olarak test uzunluğu 50 ve 80 olduğunda, örneklem büyüklüğü 20, 25 ve 50 şeklindeyken en yüksek SEE değeri Braun/Holland eşitleme yönteminde görülmektedir.

0.70 korelasyona sahip birinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede değişimlenen faktörler altında örneklem büyüklüğü arttıkça yöntemlerin SEE değerlerinde artışın meydana geldiği görülmektedir. 10, 15, 30 ve 50 test uzunluklarında genel olarak tüm şartlar altında zincirilmiş lineer eşitleme yöntemi en yüksek SEE değerlerine sahiptir. SEE değerlerindeki azalan sıralama Braun/Holland eşitleme, dairesel yay eşitleme ve birim eşitleme yöntemi şeklindedir. Örneklem büyüklüğü arttıkça yöntemlerin SEE değerlerinde azalma gözlenmektedir. 100, 200 ve 500 örneklem büyüklüğünde genel olarak dairesel yay eşitleme yöntemi düşük SEE değeri göstermektedir..

0.80 korelasyona sahip birinci alt testlerin genişletilmiş alt puanları kullanılarak yapılan eşitlemede; örneklem büyüklüğü arttıkça birim dönüşüm yöntemi dışındaki yöntemlerin SEE değerlerinde azalma gözlenirken; test uzunluğu arttıkça aynı şartlar altında eşitleme yönteminin SEE değerlerinde artma meydana gelmiştir. Aynı şartlar altında test formları arasındaki güçlük düzeyi farkı arttıkça yöntemlerin SEE değerlerinde artma görülmüştür. Tüm şartlar altında eşitleme yöntemlerinden en yüksek SEE değerini zincirilmiş lineer eşitleme vermiştir.

0.90 korelasyona sahip birinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede; 20 ve 25 örneklem büyüklüğü dışında tüm şartlar altında en yüksek SEE değerini zincirlenmiş lineer eşitleme verirken yöntemlerin SEE değerlerindeki azalma sıralama Braun/Holland eşitleme, dairesel yay eşitleme ve birim eşitleme yöntemi şeklindedir. Örneklem büyüklüğü arttıkça birim eşitleme dışındaki tüm yöntemlerde azalma meydana gelirken, aynı şartlar altında test uzunluğu arttıkça SEE değerlerinde artma; aynı şartlar altında genel olarak test formları arasındaki ortalama güçlük düzeyi farkı arttıkça SEE değerlerinde artmanın meydana geldiği gözlenmektedir.

0.70 korelasyona sahip ikinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede birim dönüşüm dışındaki eşitleme yöntemlerinde örneklem büyüklüğü arttıkça SEE değerleri azalmıştır. 20 örneklem büyüklüğü hariç tüm faktörler altında en yüksek SEE değerini zincirlenmiş lineer eşitleme yöntemi vermiştir. Eşitleme yöntemlerinin SEE değerlerindeki azalan sıralama Braun/Holland, dairesel yay eşitleme ve birim eşitleme şeklindedir.

0.80 korelasyona sahip ikinci alt testlerin genişletilmiş alt puanları kullanılarak yapılan eşitlemede birim eşitleme yöntemi dışındaki tüm yöntemlerinde SEE değerlerinde örneklem büyüklüğü arttıkça azalma meydana gelirken, aynı şartlar altında test uzunluğu arttıkça eşitleme yöntemlerinin SEE değerlerinde artma görülmektedir. Aynı şartlar altında test formları arasındaki ortalama güçlük düzeyi farkı arttıkça yöntemlerin SEE değerlerinde artış gözlenmektedir. Genel olarak en yüksek SEE değeri zincirlenmiş lineer eşitleme yönteminde görülmektedir.

0.90 korelasyona sahip ikinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede genel olarak en yüksek SEE değeri zincirlenmiş lineer eşitleme yönteminde gözlenmektedir. Test uzunluğu arttıkça birim eşitleme hariç eşitleme yöntemlerinin SEE değerlerinde artma gözlenirken, test formları arasındaki ortalama güçlük düzeyi farkı arttıkça yöntemlerin SEE değerlerinde artma meydana gelmektedir.

Tüm faktörlerin eşitleme yöntemlerinin SEE değeri üzerindeki ortak etkisi incelendiğinde birim eşitleme yönteminin SEE değeri tüm şartlar altında en küçük değere sahiptir. Elde edilen bu sonuç Skaggs (2005), Babcock, Albano ve Raymond (2012) çalışmasındaki sonuçlarla paralellik göstermektedir. Ayrıca örneklem büyüklüğü arttıkça eşitleme yöntemlerinin SEE değerlerinde azalma gözlenirken, testlerin ortalama güçlük düzeyi farkı arttıkça yöntemlerin SEE değerlerinde artma gözlenmiştir. En düşük SEE değeri birim eşitleme yönteminden sonra dairesel yay eşitleme yönteminde görülmüştür. Elde edilen bu sonuçlar Aşiret (2014) çalışmasıyla paralellik göstermektedir.

8. Çeşitli faktörlerin eşitleme yöntemlerinden elde edilen eşitleme yanlılığına (BIAS) ortak etkisinin değerlendirilmesi

0.70 korelasyona sahip birinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede; genel olarak birim eşitleme yönteminin BIAS değerlerindeki farklılaşma faktör düzeylerinin artışıyla belirgin hale gelmektedir. 10, 15 ve 30 alt test uzunluğuna kadar faktörlerin düzeylerinde meydana gelen artma birim eşitleme haricindeki diğer yöntemlerin BIAS değerlerinde değişimin olmadığı görülmektedir. 50 ve 80 test uzunluklarında ortalama güçlük düzeyi farkı arttığında yöntemlerin BIAS değerlerinde belirgin farklılık görülmektedir. En çok değişkenlik Braun/Holland eşitleme ve zincirlenmiş lineer eşitleme de gerçekleşmektedir.

0.80 korelasyona sahip birinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede; genel olarak yöntemlerin BIAS değerlerinde örneklem büyüklüğünün etkisi gözlenmezken, alt test uzunluğu 30, 50 ve 80 olduğunda yöntemlerin BIAS değerlerinde farklılaşma görülmektedir. Test formları arasındaki ortalama güçlük düzeyi farkı arttıkça eşitleme yöntemlerinin BIAS değerlerindeki değişim belirginleşmektedir. Genel olarak BIAS değerlerindeki en çok değişim birim eşitleme yönteminde görülürken, BIAS değerleri en düşük dairesel yay yönteminde gözlenmektedir. Braun/Holland eşitleme ve zincirlenmiş lineer eşitleme yöntemlerinin BIAS değerleri ise birim eşitleme yönteminden sonra en çok değişkenlik gösteren eşitleme yöntemi olduğu görülmektedir.

0.90 korelasyona sahip birinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede; genel olarak ortalama güçlük düzeyi farkı arttıkça yöntemlerin BIAS değerlerinde farklılaşma gözlenmektedir. Test uzunluğu 10 ve 15 olduğunda birim eşitleme haricindeki diğer yöntemlerde belirgin değişim gözlenmezken; 30, 50 ve 80 alt test uzunluklarında yöntemlerin BIAS değerlerinde değişim meydana gelmektedir. Genel olarak BIAS değerlerindeki en çok farklılaşmanın sırasıyla birim eşitleme, Braun/Holland eşitleme, zincirlenmiş lineer eşitleme ve dairesel yay eşitleme yöntemlerinde görülmektedir.

0.70 korelasyona sahip ikinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede; 10, 15 ve 30 madde sayısına sahip alt testlerde ortalama güçlük düzeyi farkı, örneklem büyüklüğü düzeylerinin değişimi, birim eşitleme dışındaki diğer yöntemlerin BIAS değerlerinde belirgin bir fark yaratmazken, 50 ve 80 maddeden oluşan alt test uzunluklarında ortalama güçlük düzeyi farkı arttıkça yöntemlerin BIAS değerlerinde farklılaşma meydana gelmektedir. Birim eşitlemeden sonra en çok farklılaşma Braun/Holland eşitleme yönteminde gözlenirken, en düşük BIAS değeri dairesel yay yönteminde ortaya görülmektedir.

0.80 korelasyona sahip ikinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede; 10 ve 15 alt test uzunluklarında test formları arasındaki ortalama güçlük düzeyi farkı ve örneklem büyüklüğündeki değişim birim eşitleme dışındaki diğer yöntemlerin BIAS değerlerinde farklılaşmaya neden olmadığı görülmektedir. Aynı şartlar altında birim eşitleme yöntemindeki değişim diğer yöntemlere yakın değerler almaktadır. 30, 50 ve 80 madde uzunluğunda ise yöntemlerin BIAS değerlerinde değişim gözlenmekle birlikte bu testlerde formlar arasındaki ortalama güçlük düzeyi arttıkça BIAS değerlerinde belirgin farklılaşma meydana geldiği görülmektedir.

0.90 korelasyona sahip ikinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede genel olarak birim dönüşüm yönteminin BIAS değerlerinde farklılaşma görülmektedir. 30, 50 ve 80 alt test uzunluklarında diğer yöntemlerin BIAS değerlerinde belirgin farklılaşma gözlenirken, ortalama güçlük düzeyi arttıkça farklılaşma arttığı gözlenmektedir. En düşük BIAS değerine sahip eşitleme yönteminin dairesel yay eşitleme yöntemi olduğu gözlenmektedir.

0.70 korelasyona sahip birinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede; 10 ve 15 alt test uzunluklarında tüm faktörler altında eşitleme yöntemlerinin BIAS değerlerinde farklılaşmanın olmadığı görülmektedir. Belirgin olarak alt test uzunluğu 30, 50 ve 80 olduğu durumlarda ortalama güçlük düzeyi arttıkça eşitleme yöntemlerinin BIAS değerlerindeki farklılaşma artmaktadır. Bu şartlar altında en düşük BIAS değeri dairesel yay eşitleme yönteminde gözlenirken en yüksek BIAS değerleri ise genel olarak birim eşitleme, zincirlenmiş lineer eşitleme ve Braun/Holland eşitleme yöntemlerinde ortaya çıktığı gözlenmektedir.

0.80 korelasyona sahip birinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede test formları arasındaki ortalama güçlük düzeyi farkı arttıkça yöntemlerin BIAS değerlerinin arttığı görülmektedir. 0.0 ortalama güçlük düzeyi farkında test uzunluğu arttıkça belirgin BIAS değeri değişimi 30, 50 ve 80 alt test uzunluklarında gözlenirken, 0.40 ve 0.70 ortalama güçlük düzeyi farklarında bu değişimin daha da arttığı görülmektedir. En düşük BIAS değeri dairesel yay eşitleme yönteminde gözlenirken, en yüksek BIAS değeri birim eşitleme yönteminde görülmektedir. Braun/Holland eşitleme ve zincirlenmiş lineer eşitleme yöntemlerinin genel olarak benzer BIAS değerine sahip olduğu görülmektedir.

0.90 korelasyona sahip birinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede genel olarak madde sayısı 30, 50 ve 80 olan alt testlerde yöntemlerin BIAS değerlerindeki farklılaşmanın ortalama güçlük düzeyi farkının artmasıyla belirgin hale geldiği görülmektedir. En düşük BIAS değerleri dairesel yay eşitlemede gözlenirken, Braun/Holland eşitleme ve zincirlenmiş lineer eşitleme yöntemlerinin benzer değerler aldığı görülmektedir. En yüksek BIAS değerine birim eşitleme yöntemlerinin sahip olduğu gözlenmektedir.

0.70 korelasyona sahip ikinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede 0.0 ortalama güçlük düzeyi farkına sahip alt testlerde uzunluğu arttıkça eşitleme yöntemlerinin BIAS değerlerinde belirgin farklılaşma görülmektedir. 50 ve 80 madde sayısına sahip alt testlerde ortalama güçlük düzeyi arttıkça yöntemlerin BIAS değerlerinde belirgin farklılaşmanın meydana geldiği gözlenmektedir. Genel olarak en düşük BIAS değeri dairesel yay eşitleme yönteminde görülmektedir.

0.80 korelasyona sahip ikinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede; eşitleme yöntemlerinin faktörler altındaki belirgin BIAS değişimleri 50 ve 80 alt test uzunluklarında gözlenmektedir. Test formları arasındaki ortalama güçlük düzeyi farkı arttıkça yöntemlerin BIAS değerlerinde farklılaşma daha da artmaktadır.

0.90 korelasyona sahip ikinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede; değişimlenen faktörler altında eşitleme yöntemlerinin BIAS değerlerinde belirgin davranış farklılığı 30, 50 ve 80 maddeye sahip alt testlerde gözlenmektedir. Test formları arasındaki ortalama güçlük düzeyi farkı arttıkça eşitleme yöntemlerinin BIAS değerlerindeki farklılaşma daha da artmaktadır.

Elde edilen sonuçlara bakıldığında; genelde 10 ve 15 maddeden oluşan alt testlerde yapılan eşitleme sonucunda eşitleme yöntemlerinin BIAS değerlerinde farklılaşma olmazken, 80 ve 50 alt test uzunluğunda eşitleme yöntemlerinin BIAS değerlerinde değişme meydana gelmektedir. Yöntemlerin BIAS değerlerinde en çok artma birim eşitleme yönteminde olurken, birim eşitlemeyi zincirlenmiş lineer eşitleme ve Braun/Holland eşitleme yöntemi takip etmiştir. En düşük eşitleme yanlılığını veren yöntem dairesel yay eşitleme yöntemidir. Ayrıca test formlarının ortalama güçlük düzeyi farkı arttığında yöntemlerin eşitleme yanlılığı değerlerinde artma gözlenmektedir. Elde edilen sonuç, Devdass'ın (2011) elde ettiği sonuçla benzerlik göstermektedir.

9. Çeşitli faktörlerin eşitleme yöntemlerinden elde edilen eşitleme hatasına (RMSE) ortak etkisinin değerlendirilmesi

0.70 korelasyona sahip birinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede; örneklem büyüklüğü arttıkça eşitleme yöntemlerinin RMSE değerlerinde azalma gözlenirken, diğer faktörlerin düzeyleri arttıkça eşitleme yöntemlerinin RMSE değerlerinde artmanın meydana geldiği görülmektedir. Genel olarak en yüksek RMSE değeri birim dönüşüm eşitleme yönteminde görülürken, en düşük RMSE değeri dairesel yay eşitleme yönteminde gözlenmektedir.

0.80 korelasyona sahip birinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede örneklem büyüklüğü dışındaki faktörlerin düzeyleri arttıkça yöntemlerin RMSE değerlerinde artmanın meydana geldiği görülmektedir. RMSE değerlerinin büyükten küçüğe sıralanışı birim eşitleme, zincirilmiş lineer eşitleme, Braun/Holland eşitleme ve dairesel yay eşitleme yöntemi şeklindedir.

0.90 korelasyona sahip birinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede, alt test uzunluğu ve ortalama güçlük düzeyi faktörlerinin altında örneklem büyüklüğü arttıkça yöntemlerin RMSE değerlerinde azalma gözlenmektedir. Genel olarak alt test uzunluğu ve ortalama güçlük düzeyi farkı arttıkça yöntemlerin RMSE değerlerinde artma meydana gelmektedir. Ortalama güçlük düzeyi farkı arttıkça eşitleme yöntemlerinin RMSE değerlerinde artma gözlenmekle birlikte yöntemlerin RMSE değerlerinin birbirine yakın olduğu görülmektedir.

0.70 korelasyona sahip ikinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede alt test uzunluğu ve test formları arasındaki ortalama güçlük düzeyi farkı altında örneklem büyüklüğü arttıkça yöntemlerin RMSE değerlerinde azalmanın meydana geldiği görülmektedir. Alt test uzunluğu ve ortalama güçlük düzeyi farkları arttıkça yöntemlerin RMSE değerlerinde artma görülmektedir. Ayrıca 30 alt test uzunluğuna kadar genel olarak yöntemlerin RMSE değerleri birbirine yakındır.

0.80 korelasyona sahip ikinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede test uzunluğu arttıkça eşitleme yöntemlerinin RMSE değerleri artarken tüm şartlar altında örneklem büyüklüğü arttıkça eşitleme yöntemlerinin RMSE değerlerinde azalma görülmektedir. Yöntemlerin RMSE değerleri test uzunluğunda birbirinden farklılaşmaktadır. 50 alt test uzunluğuna kadar RMSE değerlerinin birbirine yakın değerler olarak arttığı gözlenmektedir.

0.90 korelasyona sahip ikinci alt testlerin alt test puanları kullanılarak yapılan eşitlemede; örneklem büyüklüğü arttıkça eşitleme yöntemlerinin RMSE değerlerinde azalma meydana gelmektedir. Ortalama güçlük düzeyi farkı ve test uzunluğu düzeyleri arttıkça eşitleme yöntemlerinin RMSE değerlerinde artma meydana gelirken, eşitleme yöntemlerinin RMSE değerleri birbirinden farklılık göstermektedir.

0.70 korelasyona sahip birinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede; örneklem büyüklüğü arttıkça eşitleme yöntemlerinin RMSE değerlerinde azalma meydana gelmektedir. Ortalama güçlük düzeyi farkı ve test uzunluğunda artma olduğunda eşitleme yöntemlerinin RMSE değerlerinde artma meydana gelirken, eşitleme yöntemlerinin RMSE değerlerinin birbirinden farklılaştığı gözlenmektedir. 10, 15 ve 30 alt test uzunluklarında diğer faktörler arttığında eşitleme yöntemlerinin RMSE değerleri birbirine yakın değerler alırken, 50 ve 80 alt test uzunluklarında diğer faktörler arttığında eşitleme yöntemlerinin RMSE değerleri birbirinden farklılık göstermektedir. Genel olarak dairesel yay eşitleme yöntemi düşük RMSE değerine sahipken, yüksek RMSE değeri birim eşitleme yönteminde görülmektedir.

0.80 korelasyona sahip birinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede örneklem büyüklüğü arttıkça eşitleme yöntemlerinin RMSE değerlerinde azalma meydana gelirken, test uzunluğu ve ortalama güçlük düzeyi farkı arttığında oldukça eşitleme yöntemlerinin RMSE değerlerinin arttığı gözlenmektedir. Ayrıca ortalama güçlük düzeyi farkı arttıkça 30, 50 ve 80 alt test uzunluklarında yöntemlerin RMSE değerleri birbirlerine göre farklılık göstermektedir. Genel olarak en yüksek RMSE değerine birim eşitleme yöntemi sahipken, en az RMSE değerine ise dairesel yay eşitleme yöntemi sahiptir.

0.90 korelasyona sahip birinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede birim eşitleme yönteminin RMSE değerleri tüm faktörler altında diğer yöntemlere göre farklılık göstermektedir. Örneklem büyüklüğü arttıkça eşitleme yöntemlerinin RMSE değerlerinde azalma gözlenmektedir. Diğer faktörlerin düzeyleri arttıkça eşitleme yöntemlerinin RMSE değerleri artarken aynı zamanda RMSE değerleri birbirinden farklılık göstermektedir. Genel olarak en yüksek RMSE değeri birim eşitleme yönteminde gözlenirken, en düşük RMSE değeri dairesel yay eşitleme yönteminde görülmektedir.

0.70 korelasyona sahip ikinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede eşitleme yöntemlerinin RMSE değerleri 30, 50 ve 80 alt test uzunluklarında birbirine göre farklılık gösterirken, 10 ve 15 madde uzunluğunda eşitleme yöntemlerinin RMSE değerleri birbirine yakın değerler almaktadır. Ortalama güçlük düzeyleri farklılaştıkça eşitleme yöntemlerinin RMSE değerleri artmakta ve birbirine yakın değerler almaktadır. Test uzunluğu arttıkça yöntemlerin RMSE değerleri artmaktadır.

0.80 korelasyona sahip ikinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede; örneklem büyüklüğü azaldıkça eşitleme yöntemlerinin RMSE değerlerinde azalama meydana gelmektedir. Alt test uzunluğu ve ortalama güçlük düzeyi farkı arttığında eşitleme yöntemlerinin RMSE değerlerinde artma gözlenmektedir.

0.90 korelasyona sahip ikinci alt testlerin genişletilmiş alt test puanları kullanılarak yapılan eşitlemede; örneklem büyüklüğü arttıkça eşitleme yöntemlerinin RMSE değerleri azalmaktadır. Diğer faktörlerin düzeyleri arttıkça eşitleme yöntemlerinin RMSE değerleri artmaktadır. Aynı zamanda birbirlerine göre farklılaşmaları da artmaktadır.

Elde edilen sonuçlara bakıldığında örneklem büyüklüğü arttıkça eşitleme yöntemlerinin RMSE değerlerinde azalma meydana gelirken diğer faktörlerde artma olduğunda yöntemlerin RMSE değerlerinde artma olduğu görülmüştür. Genel olarak en düşük RMSE değerinin dairesel yay eşitleme yönteminde olduğu gözlenmiştir. Elde edilen bu sonuç Babcock, Albano ve Raymond (2012) tarafından elde edilen bulgularla paralellik göstermiştir.

ÖNERİLER

- 1- Çalışmada eşitleme yöntemlerinin karşılaştırılmasında simülatif verilerden yararlanılmıştır. Gerçek veriler kullanılarak benzer bir çalışma yapılabilir.
- 2- Çalışmada doğrusal eşitleme yöntemleri kullanılara eşitleme gerçekleştirilmiştir. Benzer bir çalışma eşit yüzdelikli eşitleme yöntemleri kullanılarak eşitleme yöntemlerinin hata değerleri karşılaştırılabilir.
- 3- Benzer çalışma eşitleme deseni farklılaştırılarak gerçekleştirilebilir.
- 4- Bu çalışmada değişimlenen faktör düzeyleri değişimlenerek ya da faktörler değişimlenerek eşitleme yöntemlerinin eşitleme hatası değerleri karşılaştırılabilir.

KAYNAKÇA

- Albano, A.D. (2014). *equate: Observed-Score Linking and Equating. R package version 2.0-3*,
URL <http://CRAN.R-project.org/package=equate>.
- Albano, A.D. (2016). *equate: An R Package for Observed-Score Linking and Equating. R package version 2.0-3*,
URL <http://CRAN.R-project.org/package=equate>.
- Angoff, W. H. (1984). *Scales, norms, and equivalent scores*. Princeton, NJ: Educational Testing Service.
- Angoff, W. H. (1985). *An Examination Of The Assumption That The Equating of Parallel Forms is Population-Independent*. Educational Testing Service Research Report.
- Angoff, W. H. (1987). *Technical and Practical Issues in Equating: A Discussion of Four Paper*. Applied Psychological Measurement, 11.
- Aşiret, S. (2014). *Küçük Örneklemelerde Test Eşitleme Yöntemlerinin Çeşitli Faktörlere Göre İncelenmesi*. Yayımlanmamış yüksek lisans tezi, Mersin Üniversitesi, Sosyal Bilimler Enstitüsü, Mersin.
- Baykul, Y. (2010). *Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması*. Ankara: Pegem Yayınları
- Bozdağ, S. ve Kan, A. (2010a). *Şans Başarısının Test Eşitlemeye Etkisi*. Hacettepe Üniversitesi Eğitim Fakültesi Dergisi.
- Brennan, R.L and Kolen, M.J. (1987). *A Reply to Angoff*. Applied Psychological Measurement, 11.
- Brennan, R.L and Kolen, M.J. (1987). *Some Practical Issues in Equating*. Applied Psychological Measurement, 11.
- Crocker, L. and Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. New York: Holt, Rinehart and Winston.
- Cui, Z. Kolen, M.J. (2008). *Comparison of Parametric and Nonparametric Bootstrap Methods for Estimating Random Error in Equipercentile Equating*. Applied Psychological Measurement, 32.
- Çetin, E. (2009). *Dikey Ölçeklemede Klasik Test ve Madde Tepki Kuramına Dayalı Yöntemlerin Karşılaştırılması*. Yayımlanmamış yüksek lisans tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.

- Çetin, E. ve Gelbal S. (2014). *Dikey Ölçeklemede Klasik Test ve Madde Tepki Kuramına Dayalı Yöntemlerin Karşılaştırılması*. Hacettepe Üniversitesi Eğitim Fakültesi Dergisi ,29(3), s.23-34.
- Demir, S ve Güler, N. (2014). *Ortak Maddeli Denk Olmayan Gruplar Desenine İlişkin Test Eşitleme Çalışması*. International Journal Of Human Sciences. 11, s.190-206.
- Devdass, S. (2011). Conditions Affecting the Accuracy of Classical Equating Methods for Small Samples Under the NEAT Design: A Simulation Study. *Doctoral dissertation, The University of North Carolina*.
- Gafni, N. and Melamed, E. (1990). *Using the Circular Equating Paradigm for Comparison of Linear Equating Models*. Applied Psychological Measurement, 14.
- Gübeş, N.Ö. ve Kelecioğlu, H. (2015). *Karma Testlerin Eşitlenmesinde MTK Eşitleme Yöntemlerinin Eşitlik Özelliği Korunumu Ölçütüne Göre Karşılaştırılması*. Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 6(1), s. 117-125.
- Gök, B. (2012). *Denk Olmayan Gruplarda Ortak Madde Deseni Kullanılarak Madde Tepki Kuramına Dayalı Eşitleme Yöntemlerinin Karşılaştırılması*. Yayımlanmış Doktora Tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü. Ankara.
- Gök, B. ve Kelecioğlu, H. (2014). *Denk Olmayan Gruplarda Ortak Madde Deseni Kullanılarak Madde Tepki Kuramına Dayalı Eşitleme Yöntemlerinin Karşılaştırılması*. Mersin Üniversitesi Eğitim Fakültesi Dergisi, 10(1),s. 120-136.
- Hanson, B.A. and Colton D. (1994). A comparison of Presmoothing and Postsmoothing Methods in Equipercentile Equating. Iowa City, American College Testing.
- Kan, A. (2010b). *Test Eşitleme: Aynı Davranışları Ölçen, Farklı Madde Formlarına Sahip Testlerin İstatistiksel Eşitliğinin Sınanması*. Eğitim ve Psikolojide Ölçme ve Değerlendirme Dergisi:16-21.
- Kan, A. (2011c). *Test Eşitleme: OKS Testlerinin İstatistiksel Eşitliğinin Sınanması*. Eğitim ve Bilim. Vol.36:38-51.
- Gübeş, N.Ö. ve Kelecioğlu, H. (2013). *Random Grup Deseni ile Yapılan Doğrusal Eşitleme ve Eşit Yüzdellikli Eşitleme Yöntemlerinin Karşılaştırılması*. International Online Journal of Educational Sciences, 5 (1), s. 227-241.
- Kim, S., von Davier, A. A., and Haberman, S. (2008). *Small-sample equating using a synthetic linking function*. Journal Of Educational Measurement,45(4), s. 325-342.
- Kim, S. and Livingston, S. (2010). *Comparisons among Small Sample Equating Methods in a Common-Item Design*. Journal of Educational Measurement, 47(3), s.286–298
- Kim, S., von Davier, A.A. and Haberman, S. (2011). *Practical Application of a Synthetic Linking Function on Small-Sample Equating*. Applied Measurement In Education, 24, s.95–114.

- Kim, S. and Walker, M. E. (2012). *Investigating Repeater Effects on Chained Equipercentile Equating With Common Anchor Items*. *Applied Measurement In Education*, 25.
- Kolen, M. J. and Brennan R. L. (2014). *Test equating, Scaling, and Linking: Method and Practice* (Third ed.). New York, NY: Springer.
- Livingston, S.A. and Kim S. (2009). *The Circle-Arc Method for Equating in Small Samples*. *Journal of Educational Measurement*, 46(3), s.330–343
- Livingston, S. A. and Kim, S. (2010). *Random-Groups equating with samples of 50 to400 test takers*. *Journal of Educational Measurement*, 47, s. 175–185.
- Livingston, S. A. (2004). *Equating test scores (without IRT)*. Princeton, NJ: ETS.
- Mutluer, C. (2013). *Yıl İçinde Farklı Dönemlerde Yapılan Akademik Personel ve Lisansüstü Eğitime Giriş Sınavı (Ales) Puanlarına İlişkin Bir Test Eşitleme Çalışması*. Yayınlanmış Yüksek Lisans Tezi, Abant İzzet Baysal Üniversitesi, Eğitim Bilimleri Enstitüsü. Bolu.
- Puhan, G., Liang, L. (2011). *Equating Subscores Under The Nonequivalent Anchor Test (Neat) Design*. *Educational Measurement: Issues and Practice*, 30, s.23-25.
- Skaggs, G. (2005). *Accuracy of Random Groups Equating with Very Small Samples*. *Journal of Educational Measurement*, 42 (2),s.309–330.
- Sinharay,S., Haberman, S. and Puhan, G. (2007a). *Subscores Based on Classical Test Theory: To Report or Not to Report*. Educational Testing Service.
- Sinharay, S. (2010a). *When Can Subscores Be Expected To Have Added Value? Results From Operational and Simulated Data*. Educational Testing Service.
- Sinharay, S. (2010b). *How Often Do Subscores Have Added Value? Resultsfrom Operational and Simulated Data*. *Journal of Educational Measurement*, 47(2), s.150–174.
- Sinharay, S. and Rao, C.R. (2006). *The Handbook Of Statistics on Psychometrics*. North Holland.
- Sinharay, S. and Holland, P.W. (2007). *Is It Necessary to Make Anchor Tests Mini-Versions of the Tests Being Equated or Can Some Restrictions Be Relaxed?* *Journal of Educational Measurement*, 44(3), s.249–275.
- von Davier, A.A. (2010). *Statistical Models For Test Equating, Scaling and Linking*. New York: Springer.
- Wang, T. (2006). *Standard errors of equating for equipercentile equating with log-linear pre-smoothing using the delta method*. Iowa City: Center for Advanced Studies in Measurement and Assessment (CASMA).

ÖZGEÇMİŞ

Adı, Soyadı : Arzu UÇAR
Doğum Yeri, Tarihi : İzmir/ 14.10.1987
E-mail : arzukapcik@gmail.com

EĞİTİM BİLGİLERİ

2002-2006 Şehit Erkan Özcan Yabancı Dil Ağırlıklı Lisesi

2006-2010 Muğla Üniversitesi

Fen-Edebiyat Fakültesi Matematik Bölümü

2010-2013 Muğla Üniversitesi

Fen Bilimleri Enstitüsü Matematik Bilim Dalı- Cebir ve Sayılar Teorisi Ana Bilim Dalı