

ÖRNEKLEM BÜYÜKLÜĞÜNÜN TEST EŞİTLEMeye ETKİSİ

EFFECT OF SAMPLE SIZE ON TEST EQUATING

Yunus İNCİ

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

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

Yüksek Lisans Tezi

olarak hazırlanmıştır.

2014

Eđilim Bilimleri Enstitüsü M¼d¼rl¼đ¼ne,
Yunus İNCİİN hazırladığı "Örneklem Büy¼kl¼đ¼nün Test Eđitlemeye Etkisi"
başlıklı bu çalıřma, j¼rimiz tarafından **Eđitim Bilimleri Anabilim Dalı, Eđitimde
Ölçme ve Deđerlendirme Bilim Dalı'nda Yüksek Lisans Tezi** olarak kabul
edilmiştir.

Başkan

Prof. Dr. Selahattin GELBAL

Üye(Danışman)

Doç. Dr. Duygu ANIL

Üye

Doç. Dr. Şeref TAN

Üye

Doç. Dr. Burcu ATAR

Üye

Yrd. Doç. Dr. Derya ÇOBANOđI U AKTAN

ONAY

Bu tez Hacettepe Üniversitesi Lisansüstü Eđitim Öğretim ve Sınav Yönetmeliđi nin ilgili maddeleri uyarınca yukarıdaki j¼ri üyeleri tarafından **29.10.2014** tarihinde uygun gör¼lmüş ve Enstitü Yönetim Kurulunca **03.06.2014** tarihinde kabul edilmiştir.


Prof. Dr. Bertin AKMAN
Eđilim Bilimleri Enstitüsü Müdürü

ÖRNEKLEM BÜYÜKLÜĞÜNÜN TEST EŞİTLEMeye ETKİSİ

Yunus İNCİ

ÖZ

Bu arařtırmada; 2011 ve 2012 yılı SBS 8'inci sınıf matematik alt testlerine ait puanların, farklı örneklem büyüklüğü koşulu altında doğrusal ve eşit yüzdelikli eşitleme yöntemleri kullanılarak eşitlenmesi ve bu yöntemlerden en uygun olanının önerilmesi ile örneklem büyüklüğünün test eşitlemeye etkisinin gösterilmesi amaçlanmıştır.

Bu arařtırmanın evrenini, 2011 ve 2012 yıllarında SBS 8'inci sınıf alt testine tabi tutulan 517.382 kiři oluşturmaktadır. Arařtırmanın örneklemini, bilgisayar aracılığıyla tesadüfî olarak seçilerek belirlenmiş olup evrenin yaklaşık %1'ini oluşturmaktadır. Arařtırmanın örneklemini, 6000 kişiden oluşmaktadır. Bu arařtırmada, iki test formunu eşitleyebilmek için, "eşdeğer gruplar deseni" kullanılmıştır.

Verilerin analizi dört aşamada gerçekleştirilmiştir. Birinci aşamada alt testlerin betimsel istatistikleri hesaplanmış; ikinci aşamada testlerin eşitleme koşullarını sağlayıp sağlamadığı test edilmiş; üçüncü aşamada 100, 500, 1500 ve 3000 kişilik örneklem büyüklüklerinde, doğrusal ve eşit yüzdelikli eşitleme yöntemleri kullanılarak eşitlenmiş puanlar elde edilmiştir. Son aşamada 100, 500, 1500 ve 3000 kişilik örneklem büyüklüklerinde her bir eşitleme yönteminde yapılan hata miktarları Ağırlıklandırılmış Hata Kareleri Ortalaması (AHKO) ölçütü ile hesaplanmıştır.

Arařtırmada elde edilen bulgulara göre, örneklem büyüklüğünün artması ile daha az hatalı eşitlemeler yapıldığı sonucuna ulaşılmıştır. Ayrıca arařtırmada elde edilen bulgular doğrultusunda, tüm örneklem büyüklüklerinde doğrusal eşitleme yöntemi yerine eşit yüzdelikli eşitleme yönteminin tercih edilmesi sonucuna ulaşılmıştır.

Anahtar sözcükler: Test eşitleme, doğrusal eşitleme, eşit yüzdelikli eşitleme, eşdeğer gruplar deseni, örneklem büyüklüğü.

Danışman: Doç. Dr. Duygu ANIL, Hacettepe Üniversitesi, Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı

EFFECT OF SAMPLE SIZE ON TEST EQUATING

Yunus İNCİ

ABSTRACT

In this study, equating the scores belonging to SBS 8th Grade Math Subtests, which are done in 2011 and 2012, with linear and equipercentile equating methods under the condition of different sample size and suggesting the most appropriate method and also showing the effect of sample size on test equating were aimed.

Population of this study was 517.382 people who taken 2011 and 2012 SBS 8th Grade Math Subtests examinations. Sample of the study was determined by computer randomly and it was composed of 1% of the population. Sample size was 6000. In this study, in order to equate two test forms “the common-item nonequivalent groups design” was used.

Data analysis was completed in four steps. In the first step, the descriptive statistical analysis was completed for each subtests; in the second step, whether or not each of the subtests provided assumptions of equating methods was examined; in the third step, the equated scores were provided with using linear and equipercentile equating method on sample sizes of 100, 500, 1500, 3000; and in the last stage, the equating error was calculated for each methods with WMSE criterion.

According to the findings, it was concluded that with the increase in sample size less faulty equating could be made. In addition, in line with the findings obtained in the study, when using the all sample sizes used in the study the equipercentile equating method should be preferred instead of linear equating method.

Keywords: Test equating, equipercentile equating, linear equating, the random groups design, sample size.

Advisor: Assoc.Prof.Dr. Duygu ANIL, Hacettepe University, Department of Educational Sciences, Division of Educational Measurement and Evaluation.

ETİK BEYANNAMESİ

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

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


Yunus İNCİ

TEŞEKKÜR

Öncelikle, bugünlere gelmemde en büyük pay sahibi olan sevgili ailemin tüm fertlerine;

Sempatik ve sıcak tavırlarıyla yüksek lisans eğitimim ve tezimin yazım sürecinde bana vakit ayıran, yönlendiren, bilgi ve tecrübesini paylaşmaktan sakınmayan danışmanım ve değerli hocam Doç. Dr. Duygu ANIL'A,

Tezimde kullandığım verilerin elde edilmesi sürecinde, problemlerimle yakından ilgilenen ve yüksek lisans eğitimim süresince bilgi ve tecrübelerinden ziyadesiyle yararlandığım değerli hocam Prof. Dr. Selahattin GELBAL'A;

Yüksek lisans eğitimim ve tez yazım süreci boyunca yardımlarını, samimiyetlerini bana çok görmeyen birlikte ders gördüğüm değerli dostlarım, Arş. Gör. Sinan YAVUZ, Arş. Gör. Haydar KARAMAN, Arş. Gör. Bulut YILDIZTEKİN, Arş. Gör. Meltem YURTÇU ve Ankara İl Millî Eğitim Müdürü Özel Kalemi Atilla ÖZDEMİR'E;

Ankara'da bulunduğum süre zarfında benden samimiyetlerini ve sevgilerini esirgemeyen değerli büyüklerim Şuayip ve Şengül AKKURT çiftine,

Yüksek lisans eğitimi başvuru sürecimde ve meslek hayatımda değerli katkıları ve yönlendirmeleri olan Topçu Bnb. Mehmet UYSAL'A,

Şahsıma, ülkemizin en seçkin üniversitelerinden biri olan Hacettepe Üniversitesinde yüksek lisans yapma olanağı sunan değerli kurumum Türk Silahlı Kuvvetlerine;

Son olarak burada adını yazamadığım hayatım boyunca bana olumlu katkıları olan değerli öğretmenlerim, komutanlarım ve tüm dostlarıma en samimi duygularıyla teşekkürlerimi sunuyorum.

İÇİNDEKİLER

ÖZ.....	iii
ABSTRACT.....	iv
TEŞEKKÜR.....	vi
İÇİNDEKİLER.....	vii
TABLolar DİZİNİ.....	ix
ŞEKİLLER DİZİNİ.....	x
KISALTMALAR DİZİNİ.....	xi
1. GİRİŞ.....	1
1.1. Test Eşitleme.....	3
1.2. Eşitleme Koşulları.....	6
1.2.1. Eşitlik.....	6
1.2.2. Aynı Yapıyı Ölçme.....	6
1.2.3. Eşit Güvenirlik.....	7
1.2.4. Simetriklik.....	8
1.2.5. Gruptan Bağımsızlık.....	8
1.3. Eşitlemenin Aşamaları.....	9
1.4. Eşitleme Desenleri.....	9
1.4.1. Tek Grup Deseni.....	11
1.4.2. Dengelenmiş Tek Grup Deseni.....	11
1.4.3. Eşdeğer Gruplar Deseni.....	13
1.4.4. Eşdeğer Olmayan Gruplar Ortak Test Desenleri.....	14
1.5. Eşitleme Yöntemleri.....	15
1.5.1. Klasik Test Teorisine Dayalı Eşitleme Yöntemleri.....	16
1.5.1.1. Ortalama Eşitleme.....	16
1.5.1.2. Doğrusal Eşitleme.....	17
1.5.1.3. Eşit Yüzdeliikli Eşitleme.....	18
1.5.2. Örtük Özellikler Teorisine Dayalı Eşitleme Yöntemleri.....	20
1.6. Test Eşitlemede Hata Kavramı.....	21
1.7. Problem Durumu.....	23
1.8. Problem Cümlesi.....	25
1.9. Alt Problemler.....	25
1.10. Araştırmanın Amacı ve Önemi.....	25
1.11. Varsayımlar.....	26
1.12. Sınırlılıklar.....	26
2. İLGİLİ ARAŞTIRMALAR.....	27
2.1. Yurtdışında Yapılmış Araştırmalar.....	27
2.2. Yurtiçinde Yapılmış Araştırmalar.....	30
3. YÖNTEM.....	34
3.1. Araştırmanın Türü.....	34
3.2. Evren ve Örneklem.....	34
3.3. Araştırmanın Verileri.....	34
3.4. Verilerin Analizi.....	35

3.4.1. Betimsel İstatistikler.....	36
3.4.2. Eşitleme Koşullarının Test Edilmesi	37
3.4.2.1. Tek Boyutluluk	38
3.4.2.2. Alt Testlerin Aritmetik Ortalamaları ve Varyansları Arasındaki Farkın Test Edilmesi	42
3.4.2.3. Alt Testlerin Ortalama Güçlükleri Arasındaki Farkın Test Edilmesi	42
3.4.2.4. Alt Testlerin Güvenirlikleri Arasındaki Farkın Test Edilmesi.....	43
3.4.3. Eşitleme Yöntemleri	44
3.4.3.1. Doğrusal Eşitleme Yöntemi.....	44
3.4.3.2. Eşit Yüzdelikli Eşitleme Yöntemi	44
3.4.4. Değerlendirme Ölçütü	44
4. BULGULAR VE YORUMLAR	45
4.1. Alt Problem 1'e ilişkin bulgular ve yorumlar	45
4.2. Alt Problem 2'ye ilişkin bulgular ve yorumlar.....	50
4.3. Alt Problem 3'e ilişkin bulgular ve yorumlar	51
4.4. Alt Problem 4'e ilişkin bulgular ve yorumlar	55
4.5. Alt Problem 5'e ilişkin bulgular ve yorumlar	57
5. SONUÇ ve ÖNERİLER	59
5.1. Sonuçlar.....	59
5.2. Öneriler	60
KAYNAKÇA	61
EKLER DİZİNİ	65
EK-1: Hacettepe Üniversitesinin MEB' İliğine veri talebi yazısı	66
EK-2: MEB' İliğinin veri talebini uygun bulmasına ilişkin yazısı	67
EK-3: Etik Kurulu İzin Muafiyet Formu.....	68
ÖZGEÇMİŞ	69

TABLolar DİZİNİ

Tablo 1.1. Tek Grup Deseni	11
Tablo 1.2. Dengelenmiş Tek Grup Deseni	12
Tablo 1.3. İki Formun Birleştirilmesi İşlemi	13
Tablo 1.4. Birleştirilmiş Form Uygulaması ile Sıra Etkisinin Giderilmesi	13
Tablo 1.5. Eşdeğer Gruplar Deseni	14
Tablo 1.6. Eşdeğer Olmayan Gruplar Ortak Test Desenleri Tek Grup Deseni	15
Tablo 1.7. Eşit Yüzdelikli Eşitleme Yöntemi için Örnek Veri	19
Tablo 3.1. 2011 ve 2012 SBS 8'inci Sınıf Matematik Alt Testleri Betimsel İstatistikleri.....	36
Tablo 3.2. 2011 SBS 8'inci Sınıf Matematik Alt Testi Faktör Analizi Sonuçları.....	39
Tablo 3.3. 2012 SBS 8'inci Sınıf Matematik Alt Testi Faktör Analizi Sonuçları.....	40
Tablo 3.4. Alt Testlerin Aritmetik Ortalamaları ve Varyansları Arasındaki Fark Test Sonuçları	42
Tablo 3.5. Alt Testlerin Ortalama Güçlükleri Arasındaki Farkın Test Sonucu	42
Tablo 3.6. Alt Testlerin Güvenirlikleri Arasındaki Farkın Test Sonucu.....	43
Tablo 4.1. Farklı Örneklem Büyüklüklerinde Doğrusal Eşitleme Yöntemi ile Elde Edilen Eşitlenmiş Puanlar ve Puan Farkları	46
Tablo 4.2. Farklı Örneklem Büyüklüklerinde Alt Testlerin Doğrusal Eşitleme Yöntemi ile Eşitlenmesine Ait AHKO Değerleri.....	50
Tablo 4.3. Farklı Örneklem Büyüklüklerinde Eşit Yüzdelikli Eşitleme Yöntemi ile Elde Edilen Eşitlenmiş Puanlar ve Puan Farkları	52
Tablo 4.4. Farklı Örneklem Büyüklüklerinde Alt Testlerin Eşit Yüzdelikli Eşitleme Yöntemi ile Eşitlenmesine Ait AHKO Değerleri.....	56
Tablo 4.5. Farklı Örneklem Büyüklüklerinde Alt Testlerin Eşit Yüzdelikli Eşitleme ve Doğrusal Eşitleme Yöntemi ile Eşitlenmesine Ait AHKO Değerleri.....	57

ŞEKİLLER DİZİNİ

Şekil 1.1. Test Güçlüğü ve Grup Yetenek Düzeyi Bakımından Test Eşitleme Yöntemlerinin Sınıflandırılması.....	5
Şekil 1.2. Test Eşitleme Desenleri.....	10
Şekil 1.3. Test Eşitleme Yöntemleri.....	16
Şekil 3.1.2011 Yılı SBS 8'inci Sınıf Matematik Alt Testi Yamaç-Birikinti Grafiği	40
Şekil3.2. 2012 Yılı SBS 8'inci Sınıf Matematik Alt Testi Yamaç-Birikinti Grafiği	41
Şekil 4.1. 100 kişilik örneklem büyüklüğü doğrusal eşitleme puan farkı grafiği.....	46
Şekil 4.2. 500 kişilik örneklem büyüklüğü doğrusal eşitleme puan farkı grafiği.....	47
Şekil 4.3. 1500 kişilik örneklem büyüklüğü doğrusal eşitleme puan farkı grafiği	47
Şekil 4.4. 3000 kişilik örneklem büyüklüğü doğrusal eşitleme puan farkı grafiği	48
Şekil 4.5. 100 kişilik örneklem büyüklüğü eşit yüzdelikli eşitleme puan farkı grafiği	53
Şekil 4.6. 500 kişilik örneklem büyüklüğü eşit yüzdelikli eşitleme puan farkı grafiği	53
Şekil 4.7. 1500 kişilik örneklem büyüklüğü eşit yüzdelikli eşitleme puan farkı grafiği	54
Şekil 4.8. 3000 kişilik örneklem büyüklüğü eşit yüzdelikli eşitleme puan farkı grafiği	54

KISALTMALAR DİZİNİ

SBS: Seviye Belirleme Sınavı

ÖSYM: Ölçme Seçme ve Yerleştirme Merkezi

MEB: Milli Eğitim Bakanlığı

KTT: Klasik Test Teorisi

MTK: Madde Tepki Kuramı

AHKO: Ağırlıklandırılmış Hata Kareleri Ortalaması

1. GİRİŞ

Bu bölümde, Klasik Test Teorisine dayalı test eşitleme yöntemlerine ilişkin alan yazında yer alan bilgiler paylaşılmış, problem durumu, araştırmanın önemi ve amacı, problem cümlesi, alt problemler, sınırlılıklar, varsayımlar ve kısaltmalara yer verilmiştir.

İnsanların bilişsel yeteneklerinin farklı olduğu gerçeği, insanlık tarihinin başlangıcından beri kabul görmüş bir durumdur. Plato ve Aristotle, yaklaşık 2500 yıl önce, bireysel farklılıklar hakkında yazılar yazmışlardır, öyle ki onlardan da önce antik Çin tarihinde benzer çalışmalara rastlanmaktadır (Aiken, 2000, s.3). M.Ö. 2200 yıllarında Çin imparatorlarını, kamu işlerinde görev yapacak personel seçiminde bir değerlendirme aracı kullanmaya iten ihtiyacın günümüze kadar artarak geldiğini görüyoruz. İkinci dünya savaşı sonrası kapitalizmin güç kazanması ile paralel gelişen serbest piyasa ekonomisi, bireysel başarının ön plana çıkmasını sağlamıştır. Toplum mühendisliğinden yoksun küresel kapitalizm gölgesinde yönetilen ülkemizde, maddi getirisi yüksek çalışma alanları için personel yetiştiren eğitim kurumlarına olan talep yoğunluğu, insanımızı birbirleri ile yarışır hale sokmuştur. Birbirleri ile yarışır hale getirilen insanlarımız arasında, kazananı belirlemek adına uygulanan testlere olan ihtiyaç da doğal olarak artmıştır. Nitekim insanları bir yarış içerisine iten eğitim-öğretim anlayışının ürünü olan tek oturumluk test uygulamalarına, ülkemizde sıkça rastlamaktayız. Lisansüstü eğitim başvurularında önemli bir etkiye sahip olan ve kamu personelinin yurtdışı görevlerde istihdam edilmesinde de önemli bir yere sahip olan Yabancı Dil Sınavı (YDS), adayların yabancı dil yeterliliğini ölçmektedir. Buna benzer şekilde KPSS (Kamu Personeli Seçme Sınavı) olarak bilinen sınav, çeşitli kamu kurumlarının personel seçiminde bir ölçme aracı olarak kullanılmaktadır. ALES (Akademik Personel ve Lisansüstü Eğitim Sınavı), lisansüstü eğitim görmek isteyen adayların sayısal ve sözel yeteneklerinin merkezi bir şekilde ölçüldüğü diğer bir ölçme aracı olarak karşımıza çıkmaktadır. Türk Silahlı Kuvvetleri bünyesinde amacı kurmay subay yetiştirme olan Harp Akademileri Komutanlığı'nca yapılan Harp Akademileri Sınavları da genel kültür, askeri kültür ve askeri karar verme sürecinde doğru karar verme yetenekleri olmak üzere üç boyutta personelin bilgi ve becerilerini test etmekte olan bir ölçme

aracıdır. Yine Türk Silahlı Kuvvetleri bünyesinde personelin yabancı dil dinleme yeteneğini ölçmek amacıyla, Genel Dil Sınavı (GDS) yılda bir kez yapılmaktadır. Seviye Belirleme Sınavı (SBS) olarak da bilinen test uygulaması, ülkemizde bir dönem orta öğretime geçişte kullanılan bir başka merkezi test uygulamasıdır. Bu tür merkezi olarak uygulanan sınav örneklerini çoğaltmak mümkündür.

Turgut (1979), test gizliliğinin korunamaması, ölçülen değişkenin doğasından gelen sorunlar, geçerlik ve güvenilirlik araştırmaları gibi nedenlerle testlerin birden fazla formları geliştirildiğini ifade etmiştir. Yukarıda ifade edilen tüm sınavların ortak özellikleri, Turgut'un da ifade ettiği sebeplerden dolayı belirli periyotlarda bir veya iki kez yapılması ve her uygulamada farklı formların kullanılmasıdır. Harp Akademileri Sınavı (HAGS) ve Seviye Belirleme Sınavı (SBS) hariç diğer sınavların belli süre zarfında geçerlilikleri vardır. HAGS ve SBS sonuçları girildiği yıl içinde kullanılmaktadır.

İnsan hayatında sonuçları itibari ile hayati derecede öneme sahip ölçme araçlarının, güvenlik ve gizliliğin korunması uğruna farklı formlarda uygulanma zorunluluğu, beraberinde "Aynı amaç için hazırlanmış formlar zorluk açısından eşit midir?" sorusunu getirmektedir. Kamuoyunda, farklı formlar şeklinde uygulanan söz konusu sınavların birbirleri ile kıyaslandığı ve adaylarca kolay-zor olarak ifade edildiği de bilinmektedir. Doğal olarak, çeşitli sebeplerce kolay olarak ifade edilen sınav formunu kaçıran bir adayın, bilgi ve becerisi seçilmeye müsait olmasına rağmen, seçilememesi durumu ortaya çıkmaktadır.

Ölçme ve değerlendirme sürecinin ana amacı, ölçme hatasından olabildiğince arınık ölçme ve adil değerlendirme yapabilmektir. Her ne kadar birbiri ile istatistiksel özellikler ve kapsam bakımından benzerlik gösterse de, farklı iki formun tamamıyla birbirinin yerine kullanılabilir olduğunu kabul etmenin, ölçme ve değerlendirme süreci üzerinde şüpheler doğurması kaçınılmazdır. Öyle ki ülkemizde, merkezi olarak uygulanan testlerin maddelerinin, uygulama sonrasında, test geliştirmenin maliyeti düşünülerek paylaşılmama kararları alınmaya başlanmıştır. Bu durum şüphelerin daha da artmasına sebep olmaktadır. Üzerinde şüpheleri barındıran ölçme ve değerlendirme süreçleri de, beraberinde farklı sosyal problemler getirmektedir.

Farklı formlar halinde yapılması zorunlu olan uygulamalar üzerindeki şüpheleri, istatistiksel bir takım hesaplamalarla gidermek mümkündür. Söz konusu istatistiksel hesaplamaları içeren uygulama, bir testin farklı formlarından elde edilen puanların, aynı ölçek üzerinde gösterilmesine olanak sağlayan puan eşitleme veya test eşitleme uygulamasıdır. Uygulama sadece teoride kalan bir uygulama değildir. 19'uncu yüzyılın sonları ve 20'nci yüzyılın başlarında yaşamış olan Sir Francis Galton'ın istatistik alanına katkısı olan doğrusal regresyon, bir testten alınan puanların diğer testin puanlarının tahmininde kullanılan en eski metot olarak bilinmektedir (Holland, 2007). Test eşitlemeye olan istek ve ihtiyaç doğrusal regresyonun bulunmasından sonra 1938 yılında Amerika Birleşik Devletlerinde Collage Board tarafından hazırlanan SAT (Scholastic Aptitude Test, günümüzdeki açılımı; Scholastic Assesment Test) sınavının aynı yıl içinde iki form halinde yapılması ile ortaya çıkmıştır. 1941 yılında bu sınavın farklı formları arasındaki güçlük farklılıklarını giderecek anchor-test (ortak test) eşitlemenin ilk versiyonu olarak kullanılmıştır (Holland, 2007). Günümüzde ülkelerarası karşılaştırmalar yapmak için kullanılan testlerde de (TIMSS, PISA, PIRLS) çok sayıda farklı formlar kullanılmaktadır. Bu formlarda, farklı formları alan adayların puanlarının karşılaştırılmasını sağlayan ortak maddeler bulunmaktadır. Bu maddeler yardımı ile farklı formları yanıtlayan adayların puanları aynı ölçek üzerinde gösterilerek birbirleri ile kıyaslama imkânı yaratılmaktadır.

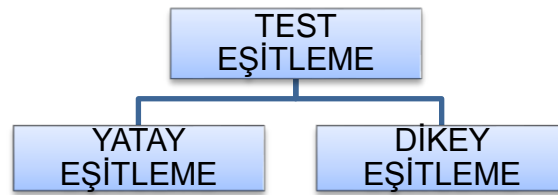
1.1.Test Eşitleme

Ölçme araçlarından yaygın olarak kullanılan ve bu araştırmanın da veri toplama aracı olarak kullanılan çoktan seçmeli testler, popülerliklerini kullanışlılıklarına, büyük grupların objektif olarak değerlendirilmesinde yarattığı kolaylığa borçludur. Ölçme araç ve yöntemlerinin nitelikleri güvenilirlik, geçerlik ve kullanışlılık kavramları altında toplanmıştır (Turgut, 1977, s.27). Klasik test teorisinde kullanılan madde istatistikleri, madde güçlük ve madde ayırıcılık indeksleri, madde güvenilirliği ve madde-madde korelasyonlarıdır. Belirtilen bu madde istatistikleri yardımıyla testten elde edilecek puanların ortalaması, standart kayması, güvenilirlik ve geçerliliği gibi test istatistikleri kestirilebilir (Anıl, 2002, s.4). Madde istatistikleri ise test geliştirmenin önemli adımlarından deneme uygulaması sonrasında en sağlıklı şekilde elde edilir. Kullanışlılık özelliğini, uygulamada sağladığı kolaylıkla

yakalayan çoktan seçmeli testlerin, güvenilirlik ve geçerlik özellikleri hakkındaki bilgiler, gizliliğin ihlal edilmemesi gerekçesi ile deneme uygulaması yapılamadığı için ancak uygulandıktan sonra görülmektedir. Test geliştirme uzmanlarının kapsam ve istatistiksel açıdan birbirine benzer test formu geliştirme istek ve çabalarına rağmen, testlerin güçlük seviyeleri açısından birbirinden farklılık göstermesi söz konusu deneme uygulamasının yapılamamasından kaynaklanır. Deneme uygulamasının yapılamamasından kaynaklanan güçlük düzeyleri arasında fark test eşitleme ile giderilebilir. Alan yazında test eşitleme ile ilgili birçok tanımlama ile karşılaşmaktayız. Angoff (1984) , test eşitlemeyi bir formun birim sistemini diğer formun birim sistemine dönüştürmek olarak tanımlarken, formların birim sistemlerinin birbirlerine dönüştürülebilir olma zorunluluğunu ön plana çıkaran Fahrenheit derece-Santigrat derece, inch-santimetre, pound-gram ilişkileri ile örneklendirmiştir.

Kolen ve Brennan (2004) test eşitlemeyi; benzer içerik ve güçlük düzeyinde geliştirilen test formları arasındaki farklılıkları düzenleyerek, bu formlardan elde edilen puanların birbiri yerine kullanılmasını sağlayan istatistiksel bir süreç olarak ifade etmiştir. Kolen ve Brennan'ın "farklı formlardan alınan puanların birbiri yerine kullanılabilmesi" ifadesi için gerekli şart Angoff'un (1984) tanımındaki formların birim cinslerinin aynı olması veya tek bir ölçek üzerinde gösterilebilir olmasıdır. Nitekim Kolen ve Brennan'ın söz ettiği söz konusu süreç içinde yapılan istatistiksel hesaplamalar da, Angoff'un (1984, s.85) tanımında belirttiği formların birim sistemlerinin, birbirlerine dönüştürme amaçlıdır. Felan (2002, s.3)'a göre test eşitleme, iki veya daha fazla test formundan elde edilmiş puanlar arasında ilişki kurmaya yarayan istatistiksel prosedürdür. Braun ve Holland (1982) ise test eşitlemeyi, zorluk düzeyleri arasında farklılık olan test formlarından alınan puanların birbirleri yerine kullanılmasını sağlayan sayısal düzenleme olarak tanımlamıştır. Sonuç olarak test eşitlemeyi, benzer içerik ve güçlük düzeyinde, aynı amaç için kullanılmak üzere geliştirilen ancak sonuçları itibari ile içerik ve güçlük düzeyleri açısından istenmedik şekilde farklılık arz eden test formlarından elde edilen puanlar arasında, istatistiksel bir ilişki kurmak suretiyle söz konusu puanların aynı ölçek üzerinde göstererek birbirleri yerine kullanılmasını sağlayan işlemler bütünü olarak tanımlayabiliriz.

Lord (1980), test eşitlemenin ancak güçlük düzeyleri eşit olan testler için kullanılabileceğini öne sürmüştür. Ancak günümüzde güçlük düzeyleri eşit olmayan testler de eşitlenebilmektedir. Testler aynı yapıyı ölçmek koşulu ile güçlük düzeyleri değişkenine bağlı olarak, yatay ve dikey olarak eşitlenebilmektedir. Aynı yapıyı ölçen benzer güçlük düzeyine sahip iki testi eşitleme olarak adlandırılan yatay eşitlemeye, farklı dönemlerde testlerin gizliliğinin korunması sebebiyle farklı formların uygulanması söz konusu olduğunda başvurulur ve paralel olması beklenen bu formlar benzer yetenek düzeyindeki adaylara uygulanır(örneğin aynı sınıf düzeyi). Farklı yetenek düzeyindeki kişilere uygulanan ve aynı yapıyı ölçen farklı güçlük düzeyindeki iki testi (örneğin farklı sınıflardaki kişilerin aldıkları testler) eşitleme ise dikey eşitleme olarak adlandırılmaktadır (Aiken, 2000,s.80; Felan, 2002, s.5). Dikey eşitleme, daha çok ilköğretimde uygulanan testlerin eşitlenmesinde kullanılmaktadır. Uygulanan bu testlerde, öğrenciler buldukları sınıf seviyelerine uygun sorulara muhatap olurlar, ancak farklı sınıf düzeylerinden elde edilen puanlar aynı ölçek üzerinde gösterilmektedir. Bu işlemler sayesinde, farklı sınıf düzeylerinden elde edilen puanlar karşılaştırılabilir ve öğrencilerin gelişimleri hakkında bilgi edinilmektedir. Aynı ölçek üzerinde gösterilmelerine rağmen, testlerin güçlük düzeylerinin ve uygulanan sınıf düzeylerinin aynı olmaması nedeniyle testlerden elde edilen puanlar birbirleri yerine kullanılabilir değildir (Kolen 1988, s.29). Grupların yetenek ve testlerin güçlük düzeylerine göre test eşitleme yöntemlerinin sınıflandırılması Şekil 1.1’de gösterilmiştir.



Şekil 1.1. Test Güçlüğü ve Grup Yetenek Düzeyi Bakımından Test Eşitleme Yöntemlerinin Sınıflandırılması

Bu araştırmada, aynı yetenek düzeyindeki gruplara uygulanan aynı zorluk düzeyindeki testlerin eşitlendiği yatay eşitleme yöntemi çalışılmıştır.

1.2. Eşitleme Koşulları

Test eşitlemenin yapılabilmesi için, eşitlenmek istenen iki test formunun bazı koşulları sağlaması gerekmektedir. Alan yazında bu koşulların ne olması gerektiği ile ilgili farklı fikirler mevcuttur. Dorans ve Holland'a (2000) göre, eşitleme koşulları; aynı yapıyı ölçme, güvenilirliklerin eşit olması, simetriklik, eşitlik ve gruptan bağımsızlıktır. Hambleton ve Swaminathan (1985, s.200) eşitleme koşullarını; eşitlik, gruptan bağımsızlık, simetriklik ve tek boyutluluk olarak belirtmişlerdir. Petersen'e (2008) göre birçok araştırmacı test eşitleme için gerekli beş koşul üzerinde mutabıktır. Bu koşullar: Aynı yapıyı ölçme, eşit güvenilirlik, simetriklik, eşitlik, gruplar arası değişmezliktir. Aşağıda, aynı yapıyı ölçme, eşit güvenilirlik, simetriklik ve gruptan bağımsızlık koşulları irdelenmiştir.

1.2.1. Eşitlik

Lord'un (1980) üzerinde eleştiriler olan eşitlik koşuluna göre; testi alan bireylerin herhangi bir gerçek puanda; gözlenen puan ortalamasının, standart sapmasının, X ve Y formlarındaki dönüştürülmüş puanların dağılımının aynı olmasıdır. Bu durumun oluşabilme ihtimali zor olmakla beraber, oluşması halinde test eşitleme işlemine olan gerekliliği de ortadan kaldıracığı aşikârdır. Test geliştirme uzmanlarının ana amacı birbirine olabildiğince paralel formlar oluşturarak, bireylerin farklı formlara tabi tutulmalarının aldıkları puanlar üzerindeki etkisini minimize etmektir. Bu durumda Lord'un test eşitleme işleminin koşulu olarak sunduğu eşitlik koşulu, nihayetinde test eşitleme uzmanının istediği ancak gerçekte ulaşması zor olan amacıdır. Test geliştirme uzmanı, amacına ulaşmış ise test eşitlemeye olan ihtiyaç yoktur. Olmayan bir ihtiyaç için koşul sunmak anlamsızdır. Bu çelişkiyi Kolen (2004, s.6) şöyle ifade etmektedir: Lord'un eşitlik koşulu eşitlemeye olan ihtiyacın olmaması durumunda sağlanmaktadır. Sonuç olarak, pratikte bu koşulu sağlamak zor ve kullanımı teoride kalmaktadır (Harris ve Crouse, 1993; Kolen ve Brennan, 1995, 2004).

1.2.2. Aynı yapıyı ölçme

Yatay test eşitlemede ana amaç, farklı iki form sonucundan alınan puanlar üzerinden bireyleri birbirleri ile kıyaslamaktır. Ana gaye, bireyleri birbirleri ile

kıyaslamak ise, bireylerin benzer özellikleri üzerinden kıyaslama yapmak en doğru kıyaslama şeklidir. İki puan arasında karşılaştırma yapılabilmesi için, puanların tek bir ölçek üzerinde aynı birim cinsinden gösterilebilir olması gerekmektedir. Sayısal yetenek ölçen bir form üzerinden alınmış puanı bulunan bir bireyin, sözel yetenek ölçen bir form üzerinden puan almış bir bireyle karşılaştırılması hatalı bir değerlendirmeyi doğuracaktır. Saatte 120 km hızla koşabilen çita ile 150 yıl gibi uzun süre yaşayabilen kaplumbağa karşılaştırmasında 120 ile 150 rakamlarının aynı sayı doğrusu üzerinde gösterilip kaplumbağanın çitaya olan üstünlüğünden söz edilemez. Woldbeck (1998) biraz daha ileri giderek, eşitlenecek testlerin aynı şeyi değil tek bir şeyi ölçmesi gerektiğini ifade ederek, tek boyutluluk şartını ifade etmiştir. Başarı testi gibi birden çok konu alanını kapsayan çok boyutlu testlerin eşitlenmesinde, eşitleme tek bir alt test düzeyinde yapılmalı ve eşitlemede Madde Tepki Kuramı (MTK) kullanıldığında, eşitlenecek testlerin tek boyutlu olması koşulu daha önemli hale gelmektedir (Woldbeck, 1998).

1.2.3. Eşit Güvenirlik

Ölçme çalışmalarında amaç, olabildiği kadar gerçek puanlara yakın gözlenen puanlar elde etmektir. Gerçek puanlara yakın ölçme sonuçları, ölçmelerdeki hata puanının azlığı oranında gerçekleşir. Ölçme sonuçları barındırdıkları tesadüfi hataların azlığı oranında güvenilirdir (Baykul, 2010, s.159). Ölçme sonuçları tesadüfi hatalardan arınık olduğu müddetçe gözlenen puanlar gerçek puanlara yakın olacaktır. Turgut (1977)'a göre güvenirlilik, ölçme sonuçlarının hatalardan arınıklık derecesidir. Eşitlemenin yapılacağı iki ölçme sonucunun, farklı miktarlarda hata barındırması, eşitlemenin sağlıklı biçimde yapılmasını engelleyecektir. Angoff (1971) , farklı güvenirlige sahip iki test formundan elde edilmiş puanların birbirleri yerine kullanılabilir olmadığını ve bu iki sonucun eşitlenemeyeceğini ifade etmiştir. Dorans ve Holland (2000), testlerin güvenirliklerinin eşit olmasına ek olarak, testlerin güvenirliklerinin büyüklüğüne de odaklanılması gerektiğini ifade etmişler ve eşitleme için yüksek güvenirliliğin daha iyi sonuçlar vereceğini belirtmişlerdir. Güvenirliliği düşük bir testin, gizlilik problemi düşünülmeden farklı zamanlarda hiç değiştirilmeden uygulanmasının bile, gerçek puanları yansıtacak sonuçlar vermeyeceği aşikâr iken, yine aynı yapıyı ölçen güvenirliliği düşük başka bir test formu ile eşitlenmesi hata üzerine hata yapmaktan başka bir anlam ifade

etmeyecektir. Güvenirliđi düşük bir test sonucunda alınan gerek puanlar, ok fazla tesadüfî hata puanı içeriyor demektir ki; söz konusu tesadüfî hataların tespiti ve kontrol altına alınması mümkün deđilken, gerek puanlarında ok fazla tesadüfî hata içeren bir başka test sonucu ile eşitleme alıřması yaparak, eşitlenmiş gerek puanlar bulmak ancak tesadüfî bir řansla mümkündür.

1.2.4. Simetriklik

Simetriklik, test eşitlemenin hangi forma X, hangi forma Y denildiđine bakılmaksızın aynı olmasıdır (Lord, 1980, s.199). Yani simetriklik, eşitlenen iki form arasından hangi formun referans form olarak alınmasının, eşitlenmiş puanlar üzerindeki etkisinin olmayışdır. Bir başka ifade ile eşitlemenin yönünün form1'den form2'ye veya form2'den form1'e olmasının eşitlenmiş puanlara olan etkisinin olmayışdır.

1.2.5. Gruptan Bađımsızlık

Eşitlemenin gruptan bađımsızlıđı, eşitleme sonucunda elde edilen puanların dönüřtürmenin yapıldıđı gruptan bađımsız olması anlamına gelir (Angoff 1971, s. 563). Holland ve Dorans (2000), önemli sayılacak alt grupların eşitlemeye bir dereceye kadar etkisi olduđunu kabul etmekle beraber, yüksek güvenirliđe sahip, format ve kapsam açısından birbirine paralel olan ve test eşitleme için uygun olarak ifade edilen testlerin eşitlenmesinde eşitlemenin yapıldıđı grubun ok fazla önemli olmayabileceđini belirtmişlerdir.

Her ne kadar eşitleme kořulları olarak yukarıda ifade edilen beř kořul teoride yer bulmuş ise de, pratikte söz konusu kořulların beřinin de aynı anda karşılandđı test formları oluřturmak güçtür. Eşitlemenin amacı, bireylerin tek bir yetenek temel alınarak karşılaştırılmalarına olanak sađlamak sureti ile ölçmenin adillliđini arttırmaktır. Şayet birey, cevapladđı her test maddesinde farklı bir yetenek kullanmak zorunda kalıyorsa, testin tek bir yetenek hakkında bilgi verdiđini söyleyemeyiz. Bu durumda test puanı, farklı yetenek düzeylerinin sayısal karşılıklarının toplanmasını ifade eder ki; bu aynı zamanda dilimizde klasik hale gelen "elma ile armudun toplanması" ile eşdeđer bir hata durumudur. Bireylerin tek bir yetenek temel alınarak karşılařtırmalarına imkân tanıyan test formları aynı

yapıyı ölçme ve tek boyutluluk koşuluna uymak zorundadır. Angoff (1971, s.563) eşitleme için sadece testte yer alan maddelerin aynı yapıyı ölçmesi ve gruptan bağımsızlık koşullarının karşılanması gerektiğini ifade etmiştir. Dorans (1990) bir testi diğerine eşitlemek için her iki testin de aynı yapıyı ölçmesi gerektiğini ve her iki testin de tek boyutlu maddeleri içermesinin zorunlu olduğunu ifade eder (Akt. Gök, 2012,s.7). Bu araştırmada öncelikle aynı yapının ölçülmesi, tek boyutluluk ve eşit güvenilirlik koşulları sınanmıştır.

1.3. Eşitlemenin Aşamaları

Kolen ve Brennan (2004, s.7) eşitlemenin yedi aşamasını şöyle sıralamışlardır.

- a. Eşitlemenin amacının belirlenmesi
- b. Alternatif formların oluşturulması
- c. Veri toplama deseninin seçilmesi (Eşitleme deseni)
- d. Veri toplama deseninin uygulanması
- e. Eşitlemenin işlevsel tanımının seçilmesi
- f. İstatistiksel kestirim yöntemlerinin seçilmesi
- g. Sonuçların değerlendirilmesi

Alternatif formların oluşturulması aşamasının, eşitleme amacının belirlenmesi aşamasından sonra verilmiş olması eşitlemeye olan ihtiyacın farklılığından ileri gelmektedir. Ülkemizde gerçek veriler ile yapılmış birçok eşitleme çalışmasında alternatif formlar eşitleme amacından önce, eşitlemeyi yapanlar dışında kişilerce oluşturulmuştur. Nitekim ülkemizde gerçek veriler ile yapılan birçok çalışmada, eşdeğer olduğu kabul edilerek uygulanan formların eşitliğinin sınanması durumuyla karşılaşılmaktadır. Aynı şekilde eşitleme deseni ve uygulanması aşamaları da eşitleme çalışmasını gerçekleştirenler tarafından yapılamamıştır. Bu araştırmada da bu aşamalar yapılmamıştır.

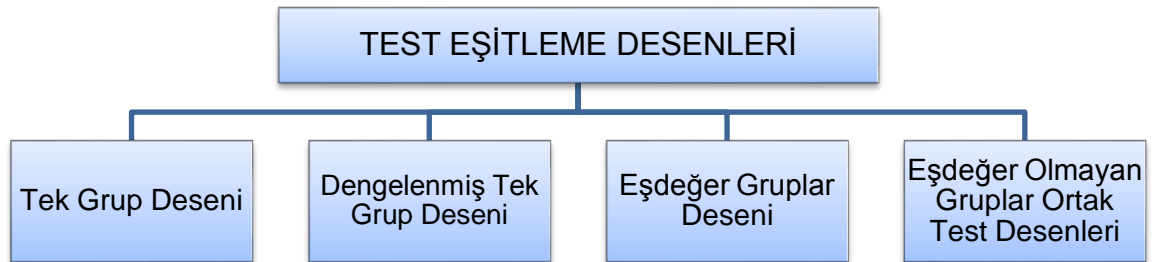
1.4. Eşitleme Desenleri

Araştırma yöntemi ve düzeneği, öncelikle araştırma probleminin yapısına ve ne bulmaya çalışacağımıza (amaca) bağlıdır (Erkuş, 2011, s.54). Eşitleme çalışmalarında, araştırma düzeneği "eşitleme deseni" olarak ifade edilir. Eşitleme

çalışması yapmaya karar vermiş bir araştırmacı, eşitleme amacına uygun eşitleme desenini belirlemelidir. Başarılı bir test eşitleme için bireylerin yetenekleri arasındaki farkı, en iyi şekilde ortaya koymaya yarayacak eşitleme deseni seçmek önemlidir. Eşitleme desenleri, bireylerin tabi tutulduğu test formu sayısı ile karşılaştırma yapılan grup sayısına göre değişkenlik göstermektedir (Holland, Dorans & Petersen, 2007, s.174). Livingston (2004, s.27), yeni ve referans formlar arasındaki bağlantıyı sağlayacak eşitleme deseni kurmanın üç yolu olduğunu ifade etmektedir. Bu yollar:

1. Eşitlenecek olan iki formun da aynı test grubuna uygulanması,
2. İki formun, ölçülen özellik bakımından eşdeğer olduğu düşünülen iki gruba uygulanması,
3. Gruplara, ölçülen nitelikle ilgili, fakat eşitlenecek formlardan farklı olan ortak bir test uygulanması.

Yine Livingston (2004), bu üç yolun kullanılmasıyla birbirleri ile kıyaslandığında avantaj ve dezavantajları olan beş farklı eşitleme deseni kurulabileceğini ifade etmektedir. İlk etapta bu beş desen, eşitlenecek formların dışında ortak test kullanılan desenler ve ortak test kullanılmayan desenler olmak üzere ikiye ayrılmaktadır. Tek grup deseni, eşdeğer grup deseni, dengelenmiş grup deseni olarak bilinen desenlerde ortak test kullanılmazken, iç ortak test ve dış ortak test eşitleme deseni olarak bilinen desenlerde ortak test kullanımı söz konusudur. Test eşitleme desenleri Şekil 1.2'de özetlenmiştir.



Şekil 1.2 Test Eşitleme Desenleri

1.4.1. Tek Grup Deseni

Tek grup deseninde, eşitlenecek iki veya daha fazla test formu tek gruba uygulanmaktadır.

Tablo 1.1 Tek Grup Deseni

Araştırma Evreni	Örnekleme (Grup)	X Formu	Y Formu
E	A Grubu	✓	✓

Nitekim ismini de tek gruba uygulanmasından almıştır.1914 yılında Amerika Birleşik Devletlerinde, iki el yazısı ölçeğinin ve aynı yıllarda Ordu Alpha ve Beta testlerinin eşitlemesinde kullanılmış en eski test eşitleme desendir (Holland ve Dorans, 2006). Eşitlenecek testlerin, tek grup üzerinden alınan veriler ile eşitlenmesi, grup farklılığının eşitleme hatasına olan etkisini azaltmaktadır. Bu durumda, Felan'ın (2002) da ifade ettiği üzere, diğer desenlere nazaran daha küçük eşitleme hatalarını bu desende görmekteyiz. Livingston (2004)'un belirttiği ve desenin kullanımının avantajı olarak görebileceğimiz özellik, testin uygulandığı grubun evreni ne derece temsil edip etmediğine bakılmaksızın, eşitleme sonuçlarının evrene genellenebilir olmasıdır. Tek grup üzerinden veri alınması sonucu testlerin farklılıkları gruptan kaynaklanmamakta ve grubun evreni temsilinin bir önemi ortadan kalkmaktadır. Başka bir avantajı ise gruptaki kişi sayısının azlığı durumunda, aynı kişi sayısı ile yapılan diğer eşitleme desenlerine oranla eşitleme kesinliğinin yüksek olmasıdır. Livingston (2004), bu desenin en önemli dezavantajını "sıra etkisi -tecrübe etkisi" veya bazı durumlarda "yorgunluk etkisi" olarak ifade etmektedir. Formların arka arkaya aynı gruba uygulanmasında, grup ilk formdan elde ettiği tecrübeyi ikinci formda kullanabilme imkânına sahiptir. Yine aynı şekilde ilk formu alan grupta, ikinci formun uygulanması esnasında yorgunluk baş gösterebilir.

1.4.2. Dengelenmiş Tek Grup Deseni

Tek grup deseninin dezavantajı olan sıra etkisinin giderilmesi, formların dengelenerek uygulanması ile giderilebilmektedir. Formların dengelenerek verildiği

grup deseni, dengelenmiş tek grup deseni olarak ifade edilmektedir. Burada grup sayısı, tek grupta olduğu gibi tektir. Sıra etkisinin giderilmesi amacı ile tek olan grup, yetenek düzeyleri aynı olacak şekilde iki gruba ayrılır. İkiye ayrılmış gruptan birine önce X formu uygulanırken, eşzamanlı olarak diğerine Y formu uygulanır. Müteakiben X formu uygulanan gruba Y formu, Y formu uygulanan gruba da X formu uygulanır. Ancak burada X formunun uygulanmasından sonra Y formunun uygulanması ile Y formunun uygulanmasından sonra X formunun uygulanmasının, uygulayıcılar üzerindeki etkisinin aynı olması gerekmektedir. Yine aynı şekilde dengelenmiş iki gruba bölme işlemi esnasında da tam olarak dengeleme işlemini yapmak kolay değildir.

Tablo 1.2 Dengelenmiş Tek Grup Deseni

Dengelenmiş Gruplar	İlk Uygulama		İkinci Uygulama	
	X Formu	Y Formu	X formu	Y Formu
Dengelenmiş Grup 1		✓	✓	
Dengelenmiş Grup 2	✓			✓

Bu durumda;

- Her iki formun da aynı zaman diliminde uygulanabilir olması halinde, X ve Y formlarında bulunan maddelerinin madde sırası takip edilerek tek sayılı maddelerde X, çift sayılı maddelerde Y formundan gelen maddeye öncelik verilmek suretiyle birleştirilmiş tek bir form oluşturulması,
- Grup ikiye bölünmeden harmanlaştırılmış tek formun uygulanması,
- Son olarak 1-0 matrisinin oluşturulması esnasında maddelerin tekrar X ve Y formlarına atanarak değerlendirme yapılması.

Basamakları sonucunda test eşitleme çalışmasının yapılması çözüm olarak değerlendirilebilir. Tablo 1.3'de örnek birleştirilmiş form uygulaması verilmiştir.

Tablo 1. 3 İki Formun Birleştirilmesi İşlemi

Birleştirilmiş Formun Oluşturulması	
1. Madde 1	X Formuna ait
2. Madde 1	Y Formuna ait
3. Madde 2	Y Formuna ait
4. Madde 2	X Formuna ait

Birleştirilmiş formun grup bölünmeden uygulanması sureti ile sıra etkisinin giderilmesi Tablo 1.4' de gösterilmektedir.

Tablo 1.4 Birleştirilmiş Form Uygulaması İle Sıra Etkisinin Giderilmesi

Araştırma Evreni	Örnekleme(Grup)	Birleştirilmiş Form
E	A Grubu	✓

1.4.3. Eşdeğer Gruplar Deseni

Test eşitleme çalışmalarında, bir gruba, eşitlenecek iki formun da uygulanması, uygulama imkânı açısından zordur. Aynı zamanda, iki formu alan tek grup üzerinde, sıra etkisinin yarattığı olumsuzluklar oluşabilmektedir. Bu durumda grup, ölçülen özellik bakımından birbirinin eşdeğeri kabul edilecek iki gruba bölünerek, eşitlenecek iki form, eşdeğer gruplara ayrı ayrı uygulanır. Alan yazında, random grup deseni olarak da bilinen eşdeğer gruplar deseni uygulamalarında, uygulamalar arası farklılığın, sadece testlerden kaynaklandığını belirtebilmek için, grup , X ve Y formlarındaki bütün puan düzeylerini temsil edecek şekilde büyük ve heterojen seçilmelidir. Livingston (2004), grubun bilgi ve yetenek düzeyi yönünden ikiye ayırmanın, zor olduğunu ancak gruptaki kişi sayısının büyük olması halinde tam olarak olmasa da eşdeğer gruplar oluşturulabileceğini ifade etmektedir. Ancak, grubun ölçülen özellik yönünden ikiye ayırmanın en iyi yolunun, paketleme yöntemi olduğunu belirtmektedir. Bu yöntemle, ölçülen özellik açısından, gruplar arası farklılıkların kontrol altına alındığını ifade etmiştir. Angoff (1984), test formlarının, önce X arkasından Y formu gelecek şekilde paketlenmesini ve formların bu sıralama bozulmadan, paketin üstünden alınarak kişilere dağıtılmasını önermiştir. Bu durumda ilk kişinin X, ikinci kişinin Y, üçüncü kişinin X formunu

aldığı sarmal bir süreç oluşacaktır. Bu dağıtma yöntemi ancak kişilerin de belirli bir sıra ile oturması ile başarısızlığa uğrayabilir(kız, erkek, kız erkek vb.). Bu desende, grupların, iki formdaki performans düzeyleri arasındaki fark, doğrudan formların güçlük düzeyleri arasındaki farkı göstermektedir (Kolen ve Brennan,2004, s.13).

Bu araştırmada, eşdeğer gruplar deseni kullanılmıştır. 2011 - 2012 yılları SBS 8'inci sınıf sınav uygulamalarına katılan her iki grubun eşdeğer olduğu varsayılmıştır. Gruptaki kişi sayılarının 258.691 gibi çok yüksek olmasından dolayı, grupların bilgi ve yetenek bakımından eşdeğer olarak varsayılabilceği değerlendirilmiştir.

Tablo 1.5 Eşdeğer Gruplar Deseni

Eşdeğer Gruplar	X Formu	Y Formu
Grup 1		✓
Grup 2	✓	

1.4.4. Eşdeğer Olmayan Gruplar Ortak Test Desenleri

Ortak test/madde desenleri, alan yazında en çok kullanılan desenlerdir (Kolen ve Brennan, 2004, 19). Bu desenlerde, grupların eşitlenmesi gerekliliği yoktur. Bu desende her iki grubun ortalama ve standart sapmalarının farklılaşmasının bir önemi yoktur (Embretson ve Reise' den aktaran Kilmen, 2010). Alan yazında eşdeğer olmayan gruplar ortak test deseni ismi ile anılması, grupların eşitlenmemesinden kaynaklanır. Eşdeğer olmayan iki grup, Tablo 1.5'de gösterildiği üzere, iki farklı test formuna tabi tutulurlar. Ancak, uygulanan iki formun yanı sıra iki grup da, gruplar arası farklılıkları kontrol edilmesi maksadı ile ortak bir test formuna tabi tutulur. Ortak test desenleri; ortak test puanlarının, testin tamamından alınan puana katkıda bulunması halinde; iç ortak test deseni, ortak test puanlarının testin tamamından alınan puana dâhil edilmediği durumlarda; dış ortak test deseni olarak ifade edilmektedir (Kolen, 1988). İç ortak test deseninde, ortak test maddeleri, eşitlenecek testlerin içerisinde yer alırken, dış ortak test deseninde ortak test maddeleri, gruplara ayrı olarak uygulanmaktadır.

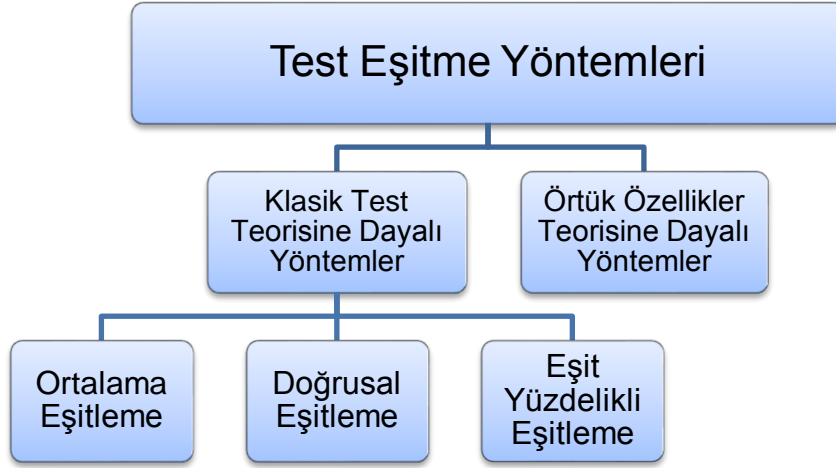
Tablo 1.6 Eşdeğer Olmayan Gruplar Ortak Test Desenleri

Eşdeğer Olmayan Gruplar	X Formu	Y Formu	Ortak Form
Grup 1		✓	✓
Grup 2	✓		✓

Ortak testlerin düzenlenmesi nasıl olursa olsun, her iki grupta “aynı işi” yapmaları gerekmektedir. Her iki grup, ortak testlerle ilgili olarak aynı içerik, sıra, hatırlama ve motivasyon etkilerine maruz kalmalıdır. Ortak testlerin uzunluğunun 20 maddeden az olmaması veya eşitlenecek testlerin uzunluğunun %20’sinden az olmaması tavsiye edilmektedir (Angoff,1984, s.107).

1.5. Eşitleme Yöntemleri

Eğitimde ve psikolojide, bireylere ait gözlenmek istenen yapı veya özelliklerin doğrudan gözlenmesi imkânsızdır. Söz konusu yapı ve özelliklerin ölçülmesi için, iki ana teori geliştirilmiştir. Gerçek puanlar teorisi olarak da bilinen klasik test teorisi (KTT) ile örtük özellikler teorisi (ÖÖT) olarak da bilinen madde tepki kuramı (MTK), eğitimde ve psikolojide kullanılan iki ölçme teorisidir. Eşitleme yöntemleri için, alan yazında, gözlenen-gerçek puanlara dayalı eşitleme yöntemleri veya parametrik-nonparametrik yöntemler gibi sınıflandırmalar mevcut ise de söz konusu sınıflandırmaların, Şekil1.3’de görüldüğü gibi ana iki ölçme teorisine dayalı olduğu görülmektedir. Bu çalışmada, KTT’ye dayalı yöntemlerin, örneklem büyüklüğü koşulu altında karşılaştırılması yapılacağından, KTT’ye dayalı yöntemler detaylı olarak ele alınırken, ÖÖT’ ye dayalı yöntemler ile ilgili özet bilgi verilmiştir.



Şekil 1.3 Test Eşitleme Yöntemleri

1.5.1. Klasik Test Teorisine Dayalı Yöntemler

Klasik test teorisinde, bireyin bir teste ilişkin puanı, gerçek puanı (True Score:T) ve hata puanının (Error Score:E) toplamını oluşturup, gözlenen puan (observed score:X) olarak ifade edilmektedir. Nitekim klasik test teorisinin temel denklemi bir numaralı eşitlikte görülmektedir.

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

Klasik test teorisine dayalı eşitleme yöntemleri de, bu temel eşitliğe dayanmakta olup Şekil1.3'de görüldüğü üzere üç kategoride ele alınmaktadır.

1.5.1.1. Ortalama Eşitleme

Sadece ortalama parametresinin kullanıldığı bu yöntemde, bir formun ortalaması diğer formun ortalamasına denk olacak şekilde dönüştürülür. Ortalama eşitlemede, eşitlenecek olan testlerin güçlük düzeylerinin, farklı olduğu ancak bu farkın puan ölçeği boyunca sabit olduğu kabul edilmektedir (Felan, 2002, s.9). Söz konusu sabitliğin ana kaynağı, klasik test teorisinin normal dağılım varsayımıdır. Standart normal dağılımda, ortalamaya göre simetriklik söz konusudur. Eşitlenecek testlerin puanlarının ortalamaya olan uzaklıklarının, eşit kabul edilmesi de normal dağılımın ortalamaya göre simetrik olmasından kaynaklanır. Ortalamaya eşit uzaklık, puanların ortalamaya eşit katkı yapmasını gerektirir. Dolayısıyla ortalama herhangi yönde yapılacak değişiklikten, tüm puanlar eşit etkilenecektir. Bu durumda X formu, Y formunu alan tüm bireylere eşit şekilde kolay veya zor gelecektir. Ortalamaları μ_x ve μ_y olan iki formun ortalama eşitleme yöntemi ile

eşitlenmesi (2) ve (3) numaralı eşitlikler yardımı ile yapılmaktadır (Kolen, 1988, s.33; Kolen ve Brenan, 2004, s.30).

$$x - \mu_x = y - \mu_y \quad (2)$$

(2) numaralı eşitlik düzenlenirse;

$$y = x - \mu_x + \mu_y \quad (3)$$

1.5.1.2. Doğrusal Eşitleme

Doğrusal eşitleme, paralel testlerin ortalama ve standart sapma haricindeki tüm özelliklerinin eşit olması varsayımına dayanmaktadır (Woldbeck, 1998, s.11; Crocker ve Algina, 1986, s.458). Standart puanlara dayalı bir yöntem olan doğrusal eşitlemede, aynı standart puanlara karşılık gelen puanların eşit olduğu kabul edilir (Angoff, 1971, s.564). Crocker ve Algina (1986, s.458) , X ve Y formlarında aynı z standart puanına denk gelen puan çiftlerinin eşit olduğunu ifade etmişlerdir. Bu durum matematiksel olarak aşağıda (4) numaralı eşitlik ile ifade edilmektedir.

$$\frac{X - \mu_x}{S_x} = \frac{Y - \mu_y}{S_y} \quad (4)$$

(4) numaralı eşitlikte Y yalnız bırakılırsa;

$$\begin{aligned} \frac{S_y(X - \mu_x)}{S_x} + \mu_y &= Y \\ Y &= \frac{S_y}{S_x}X + (\mu_y - \frac{S_y}{S_x}\mu_x) \end{aligned} \quad (5)$$

(5) numaralı eşitlik elde edilir. Burada $\frac{S_y}{S_x}$ ifadesi "a" ve $(\mu_y - \frac{S_y}{S_x}\mu_x)$ ifadesi "b" sabiti olarak ifade edilirse $Y = aX + b$ şeklinde doğru denklemi elde edilir. Burada "a" doğrunun eğimi "b" ise doğrunun Y eksenindeki kesme noktası olacaktır.

Doğrusal eşitleme yönteminin uygulanmasının kolaylığının yanında, klasik test teorisinin gruba bağımlılığından ve hesaplamasında doğru denklemi kullanımından kaynaklı bir takım problemlerinin olduğu dile getirilmektedir.

X formundan Y formuna eşitleme yapılırken, X formunda çok yüksek veya çok düşük puanların, Y formundaki karşılıkları sabit puan aralığının dışında kalabilir. Örneğin 100'er maddenin bulunduğu iki form doğrusal eşitleme ile eşitlenmek istendiğinde, zor olan X formundan 99 alan bireyin, kolay olan Y formuna

dönüştürme işleminin sonucunda 103 puan alacağı hesaplanabilir. Fakat 100 sorudan 103 puan almanın açıklaması oldukça zordur (Livingston, 2004, s.16). Tersine bir durumda da, eksi puanlarla karşılaşmak mümkündür. Bu durum eşitlenmiş puanların doğru denklemi ile elde edilmesinden kaynaklanır. Doğrunun eğiminin yüksek veya düşük olması ile kesme noktasının yeri, sabit puan aralığının dışına çıkma durumu yaratabilir, çünkü doğru denklemleri için belirlenmiş bir sabit aralık yoktur. Nitekim doğru sonsuza kadar uzanır.

Doğrusal eşitleme yönteminde karşılaşılan bir diğer problem ise, sonuçların gruba bağımlı olmasıdır. Güçlük düzeyleri farklılık gösteren iki formun, doğrusal eşitliği, başarılı bir grup ile daha az başarılı bir grupta farklılık gösterebilir. Bu sorunun çözümü için, grubun büyük ve heterojen seçilmesi önerilmektedir (Livingston, 2004, s.16). Eşitleme neticesinde oluşan doğru denkleminin eğiminin, daha başarısız gruplarda arttığı bilinmektedir. Öztürk (2010, s.11) bu sorunun kaynağının eşitleme yönteminin klasik test teorisine dayanması olabileceğini ifade etmiştir. Doğrusal eşitlemenin bir diğer sınırlılığı ise, iki teste ait ham puan dağılımının farklılaşması durumunda eşitleme hatasının artmasıdır (Bozdağ, 2007, s.15).

1.5.1.3. Eşit Yüzdellikli Eşitleme

Eşit yüzdellikli eşitleme yöntemi, bir dağılımda, aynı yüzdellik sıraya denk gelen puanların belirlenmesine dayalı bir yöntemdir (Crocker ve Algina,1986, s.461). Bu yöntem, belirli bir evrende, iki testin puan dağılımları eşit ise, söz konusu puan ölçeklerinin karşılaştırılabileceği varsayımına dayanır (Hambleton 1985, s.200). Angoff (1971)'a göre de, X ve Y testlerinde yüzdellik sırası iki grupta da eşit olan puanlar, denk olarak kabul edilebilir.

Eşit yüzdellikli eşitleme, iki aşamalı bir süreçten oluşmaktadır. İlk aşamada, eşitlenecek iki formun puan dağılımının yığılmalı frekans dağılımları tablolştırılır veya grafikleştirilir. İkinci aşamada, yığılmalı frekanslarda, aynı yüzdellik sıraya denk gelen puanlar eşitlenir. Örneğin, 50. yüzdellik dilimde form1'deki 20 puanı, form2'deki 50. yüzdellik dilime denk gelen 17 puanına eşittir (Kolen,1984, s.26).

Bu yöntem pratikte, anlatıldığı kadar kolay değildir. Çünkü; X formundaki bir ham puan ile Y formundaki bir ham puanın aynı yüzdellik sıraya gelememe ihtimali

yüksektir. Bu durumun çözümü için, aynı yüzdelik sıraya karşılık gelen ham puanları bulmak maksadı ile puan dağılımı sürekli hale getirilmektedir. Puan dağılımını sürekli hale getirmek için çeşitli düzeltme ve öteleme yöntemleri kullanılmaktadır (Cook ve Petersen, 1987, s.226).

Düzeltilme çalışmalarına eşitlemeden önce yapılıyorsa “ön-düzeltilme”, eşitleme işleminden sonra yapılıyorsa “son-düzeltilme” adı verilir. Ön-düzeltilme teknikleri, eşitleme yapılmadan önce ham puanların frekans dağılımına uygulanmaktadır. Son-düzeltilme teknikleri ise eşit yüzdelikli eşitleme sonucu elde edilen dönüşümlere uygulanmaktadır (Livingston, 2004, s.21; Heh, 2007, s.13).

Öteleme (interpolasyon) puan dağılımlarının düzeltmede kullanılan bir diğer yöntemdir. Öteleme, elde var olan değerlerden yola çıkarak bu değerler arasında, farklı bir yerde ve değeri bilinmeyen bir noktadaki olası değeri tahmin etme işlemidir. Livingston (2004, s.24), öteleme işleminin, her zaman eşit yüzdelikli eşitlemede karşılaşılan puanların aynı yüzdelik sıraya denk gelmeme probleminin çözümünde yeterli olmayacağını fakat eşit yüzdelikli eşitlemenin uygulanmasına kolaylık sağlayacağını belirtmiştir. Livingston (2004)’un verdiği öteleme işlemine ilişkin veriler Tablo1.7’de ve koyu olarak ifade edilen veriler kullanılarak çözülen örnek aşağıda verilmiştir.

Tablo1.7 Eşit Yüzdelikli Eşitleme Yöntemi için Örnek Veri

Yeni Form		Referans Form	
Ham Puan	Yüzdelik Sıra	Ham Puan	Yüzdelik Sıra
52	78.07	52	68.96
51	74.95	51	65.09
50	71.64	50	61.12
49	68.18	49	57.07
48	64.60	48	52.99
47	60.92	47	48.93
46	57.18	46	44.93
45	53.41	45	41.01
44	49.65	44	37.23
43	45.93	43	33.60
42	42.28	42	30.15

Kaynak: Livingston;2004,s.24

$$48 + \frac{53,41-52,99}{57,07-52,99} (49 - 48) = 48,10$$

Bu işlem sonucu yeni formdaki 45 ham puanının referans formdaki eşiti 48,10 olarak ifade edilmektedir. Bu işleme göre referans form (rf) , yeni form (yf) ,ötelenmiş puan (öp) olarak kodlanırsa; (15) numaralı eşitlik elde edilir.

$$\text{öp} = \text{rfaltpuan} + \frac{\text{yfyüzdeliksıra} - \text{rfaltyüzdeliksıra}}{\text{rfüstyüzdeliksıra} - \text{rfaltyüzdeliksıra}} (\text{rfüstpuan} - \text{rfaltpuan}) \quad (6)$$

Eşit yüzdelli eşitlemede formlar arasındaki güçlük farklılığının, bir eğri kullanılarak açıklanması yöntemi daha esnek ve kullanımını daha genel yapmaktadır. Örneğin, X formu Y formundan dağılımın uç puanlarında daha zor, orta puanlarında bu zorluk farkı daha düşük olabilir. Özellikle sonuçların, sadece ortalama puanlarda değil, puan ölçeğinin bütününde duyarlılığının önemli olduğu durumlarda doğrusal eşitleme yerine eşit yüzdelli eşitleme önerilmektedir. Ancak bunun içinde örneklem büyüklüğünün olabildiğince büyük tutulması gerekmektedir (Rapp, 1999, s.3). Eşit yüzdelli eşitleme yöntemi kullanılırken, örneklem yeterince büyük olmadığı zaman, puan dağılımı çarpıklaşır. Eğer dağılımlar aşırı derecede farklı ise eşit yüzdelli eşitlemenin kullanılması anlamsız hale gelir (Woldbeck, 1998, s.14). Doğrusal eşitlemede olduğu gibi, eşit yüzdelli eşitleme sonuçlarının gruba bağımlı olması bir sınırlılık olarak gösterilebilir.

1.5.2. Örtük Özellikler Teorisine Dayalı Yöntemler

Genel olarak, örtük özellikler teorisi klasik test teorisinin eksikliklerinin giderilmesi için ortaya konmuştur. Örtük özellikler teorisi ile eşitleme yöntemleri de, klasik test teorisine dayalı yöntemlerin eksikliklerinin giderilmesi amacıyla geliştirilmiştir. Klasik test teorisinde kullanılan madde istatistikleri olan, madde güçlük, madde ayırıcılık indeksi, madde güvenilirliği ve madde korelasyonları, çalışılan grubun özelliğini yansıtmaktadır. Örneğin madde güçlük indeksi, maddeye gruptan kaç kişinin doğru cevap verdiği ile ilgilidir. Maddeye doğru cevap verme durumu değiştikçe, madde ayırıcılığı da düşmekte veya artmaktadır. Ayrıca bireylerin yetenek düzeyleri de onlara uygulanan teste bağlı kalmaktadır. Yani; bireyin yetenek düzeyi, testin zorluk derecesine göre değişmektedir. Bu durumda, gruptan bağımsızlık koşulu olarak bilinen koşul sağlanamamaktadır. Örtük özellikler teorisinde ise madde parametreleri, testin uygulandığı gruptan bağımsız olarak

elde edilmektedir. Bireylerin Θ ile gösterilen yeteneği de, maddelere verdikleri cevaplardan bağımsızdır. Bireyin Θ yeteneğini ölçen ve aralarında aynı güçlük farkı bulunan iki test formundan herhangi birini alması, yetenek kestirimini değiştirmeyecektir (Hambleton ve Swaminathan, 1985, s.10). Ham puanlara dayalı test eşitleme yöntemleri, eşitlik, simetriklik ve gruptan bağımsızlık koşullarından ötürü arzu edilmemektedir. Örtük özellikler teorisine dayalı yöntemlerde kullanılan modeller, eldeki veri setine uyduğu müddetçe, bu problemleri çözmektedir (Hambleton ve Swaminathan, 1985, s.202).

Örtük özellikler teorisi ile eşitleme, karakteristik özellik teorisi veya madde karakteristik eğrileri teorisi olarak da bilinmektedir (Felan, 2002, s.10). Madde karakteristik eğrisi, belirli bir yetenek düzeyindeki cevaplayıcının beklenen test puanını belirlemede kullanılır ve maddenin doğru cevaplanma olasılığı ile yetenek düzeyi arasındaki ilişkiyi gösterir (Woldbeck, 1998, s.16). Bu ilişki, matematiksel bir fonksiyon yoluyla ortaya konmaktadır. Madde karakteristik eğrisinin, hangi matematiksel fonksiyon yoluyla gösterileceğini belirlemek için farklı modeller ortaya çıkmıştır. Bu modeller, Rasch modeli (bir parametrelili), iki parametrelili ve üç parametrelili modeller olmak üzere üçe ayrılır. Bir parametrelili modelde, madde güçlük indeksi, iki parametrelili modelde, madde güçlük indeksine ilave olarak madde ayırıcılık gücü indeksi, üç parametrelili modelde, madde güçlük ve ayırıcılık indeksine ilave olarak şans parametresi kullanılmaktadır.

1.6. Test Eşitlemede Hata Kavramı

Eşitleme hatası olarak nitelendirilen hata kavramı, bireyin bizzat katıldığı test sonucunda ortaya çıkan yetenek düzeyi ile eşitleme sonucunda katılmadığı testteki yetenek düzeyi arasındaki fark olarak tanımlanabilir. Eşitleme yöntemlerinin seçiminde, önemli bir parametredir. Eşitleme hatasının düşüklüğü, eşitlemede seçilen yöntemin doğru bir seçim olduğunu göstermektedir. Seçkisiz ve sistematik olmak üzere iki çeşit eşitleme hatası vardır (Kolen, 1988; Felan, 2002).

- ✓ **Seçkisiz eşitleme hatası;** evrenden örneklem alınarak çalışıldığında söz konusu olan hata türüdür. Seçkisiz eşitleme hatası, eşitlemenin standart hatası olarak da tanımlanır. Eşitlemenin hatası da eşitlenmiş puanların standart sapmasıdır. Eşitleme hatası, örneklem gruplarının küçüklüğünden

kaynaklanmaktadır. Örneklemedeki birey sayısının arttırılmasıyla, bu hatalar azaltılabilir (Kolen, 1988, s.35).

- ✓ **Sistemik eşitleme hatası**; test eşitleme yönteminin ve eşitleme deseninin uygun olmaması veya varsayımların ihlal edildiği durumlarda ortaya çıkmaktadır. Örneğin, tek grup kullanılarak yapılan test eşitleme çalışmalarında yorgunluk ve sıra etkisi kontrol edilmiyorsa, sistemik hatanın varlığından söz edilebilir. Eşdeğer gruplar deseninde grupların oluşturulmasında kullanılan sarmal sürecin yanlış kullanımı da, sistemik hatanın kaynağı olabilir. Öte yandan karşılaştırılan formların içerik, güçlük ve güvenilirliklerinin farklı olması tüm eşitleme desenlerinde sistemik hataya sebep olmaktadır. Sistemik hata, özellikle eşdeğer olmayan gruplar deseninde, önemli problemler yaratmaktadır. Eşitlemenin yürütüldüğü gruptaki bireylerin eşitlenen formu alan bireylerden, yetenek düzeyi olarak önemli ölçüde farklılık gösterdiği durumlarda oluşabilir (Kolen, 1988, s.35).

Test eşitleme yöntemlerinden kestirilen eşitleme hataları karşılaştırılırken, bir ölçüte ihtiyaç duyulur. WMSE (Weighted Mean Square Error- Ağırlıklandırılmış Hata Kareleri Ortalaması-AHKO) en yaygın kullanılan ölçütlerden biridir. Bu araştırmada AHKO değerleri ölçüt olarak kullanılmıştır. Bu ölçütün matematiksel ifadesi aşağıdaki gibidir;

- **AHKO**; her bir ham puan ile bu ham puana karşılık gelen eşitlenmiş puanları karşılaştırır ve matematiksel olarak eşitlik (16) da belirtildiği şekilde bulunur.

$$AHKO = \frac{\sum_{i=1}^{k-1} f_i X_E - X_{crit}^2}{\sum_{i=1}^k f_i S_Y^2} \quad (7)$$

k : Y testindeki madde sayısı

S_Y^2 : Y testindeki ham puanların varyansı

X_{crit} : Y testindeki i. ham puan.

X_E : Eşitleme yöntemleriyle elde edilen ve X testindeki i. ham puana eşit olan puan.

f_i : Y testindeki i. ham puan frekansı

Seçkisiz eşitleme hatası, eşitlenmiş puanların standart sapması ile kolaylıkla bulunabilirken, sistemik eşitleme hatasını bulmak zordur. Eşitlemede amaç her

iki hatanın da minimum olmasını sağlamak olmalıdır. Yukarıda ifade edilen ölçütün değerinin düşük olması, söz konusu seçkisiz hataların az olduğunu gösterecektir. Bu durumda AHKO değerlerinin küçük olduğu eşitleme yöntemi, daha uygun eşitleme yöntemi olacak veya ilgili örneklem büyüklüğünde eşitlemenin daha doğru yapıldığı sonucu çıkarılmaktadır.

1.7. Problem Durumu

Ülkemizde farklı amaçlar doğrultusunda, merkezi olarak uygulanan test uygulamaları mevcuttur. Her merkezi uygulamada, güvenlik ve gizliliğin korunması gerekçeleri ile farklı formların uygulanması söz konusudur. İnsan hayatında sonuçları itibari ile, hayati derecede öneme sahip ölçme araçlarının, güvenlik ve gizliliğin korunması uğruna farklı formlarda uygulanma zorunluluğu, beraberinde “Aynı amaç için hazırlanmış formlar zorluk açısından eşit midir?” sorusunu getirmektedir. Kamuoyunda, farklı formlar şeklinde uygulanan söz konusu sınavların birbirleri ile kıyaslandığı ve adaylarca “kolay-zor” olarak ifade edildiği de bilinmektedir. Doğal olarak, çeşitli sebeplerce kolay olarak ifade edilen sınav formunu kaçıran bir adayın, bilgi ve becerisi seçilmeye müsait olmasına rağmen, seçilememesi durumu ortaya çıkmaktadır. Ölçme ve değerlendirme sürecinin ana amacı, ölçme hatasından olabildiğince arınık ölçme ve adil değerlendirme yapabilmektir. Her ne kadar birbiri ile istatistiksel özellikler ve kapsam bakımından benzerlik gösterse de, farklı iki formun tamamıyla birbirinin yerine kullanılabilir olduğunu kabul etmenin, ölçme ve değerlendirme süreci üzerinde şüpheler doğurması kaçınılmazdır.

Öyle ki; ülkemizde merkezi olarak uygulanan testlerin maddelerinin, uygulama sonrasında, test geliştirmenin maliyeti düşünülerek, paylaşılmama kararları alınmaya başlanmıştır. Bu durum, şüphelerin daha da artmasına sebep olmaktadır. Üzerinde şüpheleri barındıran ölçme ve değerlendirme süreçleri, beraberinde farklı sosyal problemler getirmektedir. Ayrıca, merkezi olarak uygulanan test uygulamalarında, testlerin güçlük durumu, testlerin uygulanmasından sonra ortaya çıkmaktadır. Örneğin; YDS (Yabancı Dil Sınavı) ile ilgili maddelerin paylaşılmamasına yönelik yapılan eleştirilere cevaben, ÖSYM tarafından 17.09.2013 tarihli bir duyuru (<http://www.osym.gov.tr/belge/1-19350/baskanligimiz-tarafindan-yapilan-sinavlarda-sorulan-sor-.html>) Erişim

[tarihi:13](#) Ekim 2013) yayınlanmıştır. Bu duyuruda, maddelerin gerçek kolaylık/zorluk/ölçme kalitesi derecelerinin, bilimsel anlamda, sadece adayların sınav anında verdikleri cevaplar ile belirlenebildiği belirtilmiştir. Anıl ve Öztürk (2012) de, test geliştirme uzmanları tarafından hazırlanan test formlarının, içerik ve istatistiksel özellikler bakımından tam olarak birbirine eşdeğer olamayacağını ve formların en azından güçlük düzeyi olarak birbirinden farklılık gösterebileceğini ifade etmişlerdir. Bu sorunun çözümü için de, bir testin farklı formlarından elde edilen puanların karşılaştırılabilmesi için puanların eşitlenmesi gerektiğini belirtmişlerdir.

Sosyal bir problem haline gelmiş, adaylarca “kolay veya zor” olarak ifade edilen test uygulamalarında, ölçme ve değerlendirme sürecinin adilliğinin artırılması için, çözümler üretilmelidir. Bu çözümlerden biri eşitleme çalışmalarıdır. Ayrıca, ülkemizde yapılan eşitleme çalışmaları incelendiğinde, çalışmalarda çeşitli örneklem büyüklüklerinin kullanıldığı ve daha çok sanal veriler üzerinde çalışıldığı da görülmektedir. Bu durumda, ölçme ve değerlendirme süreçlerinin adilliğinin artırılmasına çözüm olarak gösterilen puan eşitleme yöntemlerinin, ülkemizde uygulanmış bir test uygulaması verileri kullanılmak sureti ile farklı örneklem büyüklükleri koşulu altında incelenmesi ihtiyacı vardır. 2011 ve 2012 yıllarında uygulanmış SBS (Seviye Belirleme Sınavı) 8’inci sınıf matematik alt testleri ülkemizde uygulanmış test uygulamaları olarak seçilmiştir.

Araştırma, yukarıda ifade edilen ihtiyacın karşılanmasına yönelik olarak, örneklem büyüklüğü koşulu altında eşitleme yöntemlerinin karşılaştırılmasını amaçlamaktadır. Örneklem büyüklüklerinin belirlenmesinde, çalışmada kullanılacak veri analiz yöntemleri, eşitleme çalışmalarında kullanılan örneklem büyüklükleri ve ülkemizde pratikte kullanılması olası örneklem büyüklükleri dikkate alınmıştır. Sonuç olarak araştırmada, 100, 500, 1500, 3000 kişilik örneklem büyüklükleri seçilmiştir.

Bu çalışmanın ana amacı, yukarıda ifade edilen ihtiyacı, çalışmanın sınırlılıkları dâhilinde karşılamaya çalışmaktır. Bu amaç doğrultusunda, aşağıdaki problem cümlesine ve alt problemlere cevap aranacaktır.

1.8. Problem Cümlesi

Farklı örneklem büyüklüğü koşulu altında 2011 ve 2012 yılları SBS 8'inci sınıf matematik alt test puanlarının, doğrusal ve eşit yüzdellikli eşitleme yöntemleri kullanılarak eşitlenmesinde uygun olan eşitleme yöntemi hangisidir?

1.9. Alt Problemler

1. Farklı örneklem büyüklüğü koşulu altında 2011 ve 2012 yılı SBS 8'inci sınıf matematik alt test puanlarının doğrusal eşitleme yöntemi kullanılarak elde edilmiş eşitlenmiş puanları nasıldır?
2. Örneklem büyüklüğünün doğrusal eşitleme yöntemi üzerindeki etkisi nasıldır?
3. Farklı örneklem büyüklüğü koşulu altında 2011 ve 2012 yılı SBS 8'inci sınıf matematik alt test puanlarının eşit yüzdellikli eşitleme yöntemi kullanılarak elde edilmiş eşitlenmiş puanları nasıldır?
4. Örneklem büyüklüğünün eşit yüzdellikli eşitleme yöntemi üzerindeki etkisi nasıldır?
5. Farklı örneklem büyüklüğü koşulu altında 2011 yılı SBS 8'inci sınıf matematik alt testi puanlarının, 2012 yılı SBS 8'inci sınıf matematik alt testi puanlarına eşitlenmesi için, doğrusal eşitleme ve eşit yüzdellikli eşitleme yöntemlerinden hangisi daha uygun eşitleme yöntemidir?

1.10. Araştırmanın Amacı ve Önemi

Bu araştırmanın amacı, 2011 ve 2012 yıllarında uygulanan SBS 8'inci sınıf Matematik alt testlerinin verilerini kullanarak, doğrusal ve eşit yüzdellikli eşitleme yöntemlerini, örneklem büyüklüğü koşulu altında karşılaştırmak ve örneklem büyüklüğünün eşitleme yöntemleri üzerinde etkisini göstermektir.

Ülkemizde yapılmış yatay test eşitleme çalışmalarında, örneklem büyüklüğünün test eşitleme yöntemleri üzerindeki etkisi yalnızca Madde Tepki Kuramına dayalı eşitleme yöntemleri için ve sanal veriler kullanılarak yapılmıştır. Bu açıdan bakıldığında, örneklem büyüklüğünün test eşitlemeye etkisinin gerçek veriler kullanılmak suretiyle bulunması konusu ülkemizde ilk defa çalışılacak bir araştırma konusudur. Bu manada alan yazına katkı sağlayacağı açıktır.

Ayrıca ülkemizde merkezi olarak uygulanan test uygulamaları üzerinde yoğun eleştiriler ve şüpheler bulunmaktadır. Ölçme ve değerlendirme süreçleri üzerinde, şüphelerin bulunması toplumsal bir problemdir. Bilindiği üzere Hacettepe Üniversitesi; bilim, teknoloji, görsel ve işitsel sanat alanlarında yürüttüğü araştırmalar ve verdiği eğitim-öğretimle, değişime ve gelişime açık bireyler yetiştirmeyi, her alandaki birikimini toplum yararına sunmayı görev edinmiştir. Bu manada, yapılan bu araştırma, toplumsallaşmış bir probleme bilimsel bir çözüm sunma gayretidir.

Sonuç olarak, alan yazına katkı sağlaması ve üniversitenin misyonuna uygun olarak, mevcut bilgi birikiminin toplum yararına kullanılması araştırmayı önemli kılmaktadır.

1.11. Varsayımlar

- ✓ 2011yılı SBS 8'inci sınıf sınav oturumuna katılan bireylerin, 2012 yılı SBS 8'inci sınıf sınav oturumuna katılan bireyler ile eşdeğer olduğu varsayılmıştır.
- ✓ Adaylara her iki uygulamada da aynı koşullar sunulmuştur.
- ✓ Adaylar sınav uygulamalarında, maddeleri içtenlikle cevaplamışlardır.

1.12. Sınırlılıklar

- ✓ Araştırma, 2011 ve 2012 yıllarında uygulanmış SBS 8'inci sınıf Matematik alt testine tabi tutulan adaylarla sınırlıdır.
- ✓ Araştırma, Klasik Test Teorisine Dayalı Eşitleme Yöntemleri ile sınırlıdır.
- ✓ Araştırma, örneklem büyüklüğünün etkisini görmek maksatlı kullanılan 100, 500, 1500 ve 3000 kişilik örneklem büyüklükleri ile sınırlıdır.

2. İLGİLİ ARAŞTIRMALAR

Alan yazında, özellikle yurtdışında yapılmış birçok çalışma mevcuttur. Puan eşitleme çalışmalarına, ülkemizde 90'lı yıllardan sonra rastlanmakta ve son yıllarda çalışmaların arttığı gözükmetedir. Aşağıda öncelikle yurtdışında yapılan çalışmalar özetlenmiş ardından ülkemizde yapılan çalışmaların özetlerine yer verilmiştir.

2.1. Yurtdışında Yapılmış Araştırmalar

Petersen, Cook ve Stocking (1983) tarafından yapılan araştırmada, KTT' ne ve MTK' na dayalı eşitleme yöntemleri karşılaştırılmıştır. Araştırma sonucunda, Madde Tepki Kuramına dayalı eşitleme yöntemlerinin, diğer yöntemlere nazaran daha iyi sonuçlar verdiği sonucuna ulaşılmıştır.

Yine 1983 yılında Stocking ve Lord tarafından yapılan araştırmada, Stocking-Lord ve Ortalama-Sigma yöntemleri karşılaştırılmış ve sonuç olarak Stocking-Lord yönteminin daha uygun sonuçlar verdiği görülmüştür.

1984 yılında Caldwell tarafından yapılan araştırmada, ortak maddeli formlar kullanılmak sureti ile KTT' ye ve Rasch modeline dayalı eşitleme yöntemleri karşılaştırılmıştır. Sonuç olarak, Rasch modelinin KTT' ye dayalı eşitleme yöntemlerine nazaran daha iyi eşitleme sonuçları verdiği bulunmuştur.

Skaggs ve Lisstz (1986), doğrusal, eşit yüzdelikli, Rasch ve üç parametrelili modelin karşılaştırılmasını içeren çalışmalarında, dış ankor test düzeneği kullanmışlardır. Ayrıca çalışmada, yatay eşitlemenin yanı sıra dikey eşitlemede yapılmıştır. Eşitleme yöntemlerinin sonuçlarının karşılaştırılmasında, ağırlıklandırılmış ve ağırlıklandırılmamış hata kareleri ortalamaları indeksleri kullanılmıştır. Sonuç olarak, eşit yüzdelikli eşitleme ve 3 parametrelili model yöntemlerinin diğer iki yöntemle göre, daha uygun sonuçlar verdiği görülmüştür.

1986 yılında Angoff ve Cowell tarafından yapılan çalışmada, gruptan bağımsızlık koşulu test edilmiştir. Desen olarak eşdeğer gruplar deseni kullanılan bu çalışmada veriler, homojen GRE-Sayısal Testi ve heterojen GRE Sözel+Sayısal testlerinden toplanmıştır. Çalışma sonunda, homojen paralel testlerin gruptan bağımsızlık koşuluna uygun olduğunu, paralel fakat heterojen testlerle eşitlemenin

aynı olduğu ve eşitleme sonuçlarının evrene genellenebileceği sonucuna varılmıştır.

1987 yılında yapmış olduğu çalışmada doğrusal ve eşit yüzdelikli eşitleme yöntemlerinin karşılaştırılması için yeni bir yöntem öneren Budescu'nun çalışması, deneysel bir çalışma niteliği taşımaktadır. Yapılan bu deneysel çalışma neticesinde doğrusal eşitlemenin yetersiz ve uygun olmayan durumları ayırt etmede daha başarılı sonuçlar verdiği ifade edilmiştir.

Zeng (1991), normallik koşulunun karşılanması ve karşılanmaması durumlarında, büyük örneklem için tek grup deseninde doğrusal eşitlemeye yöntemi için farklı eşitleme hatalarını karşılaştırmıştır. Çalışmada simülasyon ve gerçek veriler kullanılmıştır. Araştırmada yer alan eşitleme hataları; bootstrapt yöntemi, iki delta yöntemi ile bulunan eşitleme hatalarıdır. Bulgular, dağılımın normal olduğu durumlarda, iki delta yöntemine ait eşitleme hatalarının yakın sonuçlar verdiğini göstermiştir. Sonuçlara bakıldığında, delta yönteminin bootstrapt yöntemine tercih edilebileceğini, neden olarak ise; bootstrapt yönteminin daha zaman alıcı olduğu ve delta yönteminin az hesaplama gerektirdiği gösterilmiştir. Ayrıca eşitlemenin standart hatasının, normallik koşulu altında daha küçük olduğu bulunmuştur.

1993 yılında Kromrey ve Du Bose yaptıkları Monte Carlo çalışmasında, tek ve iki bağlantılı eşitleme durumunda, eşitleme kararlılığını ve istatistiksel yanlılığı küçük gruplar seçerek incelemiştir. Örneklem; 25, 50 ve 100 kişiden oluşmuştur. Çalışmada Angoff'un IV (nonrandom groups) eşdeğer olmayan doğrusal modeli kullanılmıştır. Ankor madde kullanımının da yapıldığı çalışmada, ankor maddeleri testlerin %30'u ve üzerinde olacak şekilde kullanılmıştır. Eşitleme yönteminin kararlılığına, "Bootstrapt" eşitlemenin standart hatası ile bakılmıştır. Eşitlemenin standart hatasının, puan dağılımının ortasında daha küçük iken, ortalamadan uzaklaştıkça hatanın büyüdüğü görülmüştür. Ayrıca örneklem büyüdükçe, eşitlemenin standart hatası azalmaktadır sonucuna ulaşılmıştır.

Motika ve Chason (1995), İspanyolca ve Fransızca öğretmenlik sertifika sınavlarını ve bu sınavların alt testlerini, Angoff'un IV (nonrandom groups) modelini kullanarak, doğrusal eşitleme yöntemi ile eşitlemiştir. Çalışmada, 200 İspanyol ve 75 Fransız öğretmenin sınav sonuçları, veri olarak kullanılmıştır. Form A, B, C olmak üzere her bir testin üç farklı formu kullanılmıştır. Çalışmada, Form A, Form

B'ye devamında eşitlenmiş B puanları, Form C'ye eşitlenmiştir. Daha sonra eşitlenmiş C puanları Form A'ya eşitlenmiştir. Eşitlenmiş A puanları ile Form A'nın ham puanları karşılaştırılmıştır. Çalışmada, hata kareleri ortalamasının karekökü (RMSE) ölçüt olarak kullanılırken, eşitlemenin yanlılığını da bakılmıştır. Çalışmada, alt testlere ait ham puanlar, kendi aralarında eşitlenirken aynı zamanda tüm test puanları da bütün olarak eşitlenmiştir. Neticede, tüm test puanlarının eşitleme hatasının ve yanlılığının, alt testlere ait eşitleme hatasından ve yanlılığından daha küçük olduğu görülmüştür.

Yang (1997) tarafından, ortak maddeli eşitlenmemiş gruplar eşitleme deseninin kullanıldığı araştırmada, KTT'ye ve MTK'ya dayalı eşitleme yöntemleri karşılaştırılmıştır. Araştırmada ortak maddelerin güçlüğü'nün, testin güçlüğü'nü yansıttığı durumlarda eşitlemenin daha doğru yapıldığı sonucuna ulaşılmıştır. Bu bulgulara ek olarak, Madde Tepki Kuramına dayalı eşitleme yöntemlerinin, diğer eşitleme yöntemlerine göre daha hatasız sonuçlar verdiği saptanmıştır.

Tsai (1997), seçkisiz grup deseninde minimum örneklem büyüklüğünü bulmayı amaçladığı çalışmasında eşitleme hatasını ölçüt olarak kullanmıştır. Araştırmada, KTT'ye dayalı eşitleme yöntemleri karşılaştırılmıştır. Veriler, American College Testing (ACT)'nin bilim alt testine ait üç formdan elde edilmiştir. Araştırmada 25, 50, 100, 200, 500 ve 1000 kişilik örneklem kullanılmıştır. Araştırmanın sonucunda örneklem büyüklüğü arttıkça, eşitleme hatasının azaldığı, doğrusal eşitleme ve eşit yüzdelikli eşitlemenin puan dağılımının ortasında daha az hata içerdiği bulunmuştur. Ayrıca, ortalama eşitleme için daha küçük örneklem gerekirken, eşit yüzdelikli eşitleme için daha büyük örneklem gerekmekte olduğu bulgusuna ulaşılmıştır.

Kim ve Cohen (2002) tarafından yapılan araştırmada, 300 ve 1000 kişilik örneklem büyüklüklerinde eşitleme hataları incelenmiştir. Araştırma sonucunda, 1000 kişilik örneklem büyüklüklerinden elde edilen eşitleme hatalarının, 300 kişilik örneklem büyüklüklerine nazaran daha düşük olduğu saptanmıştır.

Lee ve Ban (2010) tarafından yapılan test eşitleme araştırmasında, örneklem büyüklüğü 500 ve 3000 olmak üzere iki kategoride ele alınmıştır. Araştırma sonucunda, 3000 kişilik örneklemelerden elde edilen eşitleme hatalarının, 500 kişilik örneklemelerden elde edilenlere nazaran daha düşük olduğu saptanmıştır.

2.2. Yurtiçinde Yapılmış Araştırmalar

Türkiye’de test eşitlemeye ilişkin olarak yapılan çalışmaların ilki olan ve Kelecioğlu (1994) tarafından yapılan araştırmanın amacı; 1990, 1991 ve 1992 yılları arasında yapılan Öğrenci Seçme Sınavı (ÖSS) puanlarını, Klasik Test Teorisi ve Madde Tepki Kuramına dayalı eşitleme yöntemlerine göre eşitlemek olmuştur. Çalışmada, Klasik Test Teorisine dayalı yöntemlerden doğrusal eşitleme yöntemi ve eşit yüzdelikli eşitleme yöntemi, Madde Tepki Kuramına dayalı yöntemlerden Rasch modeli ve İki parametrelili lojistik yöntem kullanılmıştır. Kelecioğlu(1994), testleri sözde ortak test desenini kullanarak eşitlemiş ve eşitleme yöntemlerindeki hata miktarlarını, eşitlenmiş puanlarla test puanlarının hata kareleri ortalamalarını hesaplayarak bulmuştur. Araştırma sonucunda Klasik Test Teorisine dayalı eşitleme yöntemlerinin Türkçe ve Fen Bilimleri alt testleri için uygun yöntemler olduğu; Sosyal Bilimler ve Matematik alt testleri için ise Madde Tepki Kuramına dayalı eşitleme yöntemlerinin uygun yöntemler olduğu saptanmıştır.

Kelecioğlu (1995), yaptığı çalışmada, faktör analizi, madde gücü ve uzman kanılarına dayanılarak sözde ortak test oluşturma yöntemlerinden elde edilen sözde ortak testlerin eşitleme hatası üzerindeki etkisini incelemiştir. Araştırmanın sonucunda, ÖSS matematik alt testlerinin eşitlenmesinde Rasch Modeli ve madde güçlüklerine dayanılarak oluşturulmuş ortak testle yapılan eşitleme hatasının diğer eşitleme hatalarından daha küçük olduğu görülmüştür.

Şahhüseyinoğlu (2005) tarafından yapılan başka bir test eşitleme araştırmasında, 2000, 2001 ve 2002 yıllarında Hacettepe Üniversitesince düzenlenen İngilizce Yeterlik Sınavı puanları üzerinde, Madde Tepki Kuramı ve Klasik Test Kuramına dayalı eşitleme yöntemleri karşılaştırılmıştır. Araştırmada ortak test desenine dayalı eşit yüzdelikli, doğrusal ve Rasch modeli yöntemleri kullanılmıştır. Yöntemlerin karşılaştırılmasında ölçüt olarak eşitleme hatası kullanılmıştır. Şahhüseyinoğlu (2005) ,üç yöntemle ait hata miktarını, eşitlenmiş puanlarla test puanları arasındaki hata kareleri ortalamalarını hesaplayarak bulmuş ve hata miktarı en küçük olanı en uygun yöntem olarak belirlemiştir. Araştırma sonucunda eşit yüzdelikli eşitleme yönteminin, Rasch yöntemine dayalı eşitleme kadar uygun bir yöntem olduğu sonucuna ulaşılmıştır.

Şans başarısının test eşitlemeye etkisini inceleyen Bozdağ (2007), çalışmasında şans başarısından arındırılmış ve arındırılmamış 2003 ve 2005 OKS Türkçe alt testlerine ait puanları, tek grup desenine dayalı doğrusal ve eşit yüzdelikli eşitleme yöntemlerini kullanarak eşitlemiştir. Araştırma verileri, 2005–2006 eğitim- öğretim yılında Mersin ili sınırları içinde yer alan farklı ilköğretim okullarında öğrenim gören 1031 8. sınıf öğrencisinden elde edilmiştir. Bozdağ (2007), yöntemlere ait eşitlenmiş puanlar ile test puanları arasındaki hata kareleri ortalamalarını hesaplayarak, hata miktarı en küçük olan yöntemi, en uygun yöntem olarak önermiştir. Araştırma sonucunda, şanstın arındırılmamış test puanlarının eşitlenmesinde, en uygun yöntemin doğrusal eşitleme, şanstın arındırılmış test puanlarının eşitlenmesinde ise eşit yüzdelikli eşitlemenin en uygun yöntem olduğu görülmüştür. Tüm hatalar göz önünde bulundurulduğunda, şanstın arındırılmış eşit yüzdelikli eşitleme yöntemine ait hatanın daha düşük olduğu görülmüştür.

Çetin (2009), dikey ölçeklemede, KTT ve MTK yöntemlerini karşılaştırmıştır. Araştırmanın verilerini, 2005 yılı ilköğretim Öğrencilerinin Başarılarının Belirlenmesi Sınavına ait 6, 7 ve 8. Sınıf öğrencilerinin test puanları oluşturmuştur. Bu araştırma sonucunda, 250, 500 ve 1000 kişilik örneklem büyüklüklerinden, en az hata veren örneklem büyüklüğünün 1000 kişilik örneklem olduğu saptanmıştır. Bu çalışma, ülkemizde yapılan dikey eşitleme çalışmalarında örneklem büyüklüğünün test eşitlemeye etkisini araştıran ilk araştırmadır.

Öztürk (2010), 2008 ilkbahar ve sonbahar dönemlerinde ÖSYM tarafından yapılan Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı (ALES) puanlarını, Klasik Test Teorisi'ne dayalı doğrusal ve eşit yüzdelikli eşitleme yöntemlerini kullanarak eşitlemiştir. Araştırmada tek grup deseni kullanılmıştır. Araştırmanın sonucunda; sayısal-1 ve sayısal-2 alt testlerinin eşitlenmesi için en uygun yöntemin, ağırlıklandırılmış hata kareleri ortalamaları katsayısı daha küçük olan eşit yüzdelikli eşitleme yöntemi olduğu görülmüştür.

Kilmen (2010), çalışmasında Madde Tepki Kuramına dayalı "ortalama-ortalama", "ortalama-standart sapma", "Haebara" ve "Stocking-Lord" eşitleme yöntemlerinden kestirilen eşitleme hatalarının, yetenek dağılımı (benzer ve farklı yetenek dağılımı) ve örneklem büyüklüğü (500-1000 kişilik) değişkenlerine dayalı olarak karşılaştırılmasını amaçlamıştır. Araştırmada kullanılan veriler, simülasyon verileridir. Araştırma sonucunda Stocking-Lord yöntemi ile yapılan test eşitleme

uygulamasının, daha az hatalı eşitlemeler yaptığı saptanmıştır. Örneklem büyüklüğüne dayalı olarak test eşitleme hataları incelendiğinde ise, örneklem büyüklüğü arttıkça eşitleme hatalarının azaldığı saptanmıştır. Yetenek dağılımı değişkenine göre test eşitleme yöntemleri karşılaştırıldığında ise, benzer yetenek dağılımına sahip grupların eşitlenmesinden elde edilen eşitleme hatalarının, farklı yetenek dağılımına sahip grupların eşitlenmesinden elde edilen eşitleme hatalarına nazaran daha düşük olduğu bulunmuştur.

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, iki test formunu eşitleyebilmek için “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. Araştırmada ele alınan koşullar doğrultusunda, büyük örneklem ile daha uzun testler kullanıldığında ve benzer yetenek dağılımına sahip gruplarda, yöntemlerin daha az hatalı olduğu sonucuna ulaşılmıştır.

Mutluer (2013), 2011 ilkbahar ve sonbahar dönemlerinde ÖSYM tarafından merkezi olarak yapılan Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı (ALES) Puanlarını, doğrusal ve eşit yüzdelikli 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 yüzdelikli eşitleme yöntemi olduğu görülmüştür.

Genel olarak yapılan çalışmalar incelendiğinde, ülkemizde yapılan çalışmalarda eşitleme aşamalarının bire bir uygulanmadığı, diğer bir ifade ile eşitleme çalışmalarında kullanılan gerçek verilerin eşitlenmek üzere oluşturulan testler olmadığı görülmektedir. Ayrıca ülkemizde yapılan eşitleme çalışmalarında daha çok sanal verilerin tercih edildiği görülmektedir. Eşitlenmek istenen testlerin eşitleme koşullarını sağlayamaması, ülkemizde yapılan eşitleme çalışmalarında sanal verilerin kullanılmasının sebebi olarak söylenebilir. Gerek ülkemizde gerekse yurtdışında yapılan çalışmalarda örneklem büyüklüğünün test eşitleme yöntemleri

zerindeki etkisi incelendiđinde, rneklem byklđnn artmasnn eđitlemenin random hatasn dđrdđ gzkmektedir. Eđitleme yntemlerinin karđılađtırlmasnda ise, MTK'ya dayalı yntemlerin genel olarak KTK'ya dayalı yntemlerden daha az hatalı sonular verdiđi grlmektedir.

3. YÖNTEM

Bu bölümde; araştırmanın türü, araştırma evreni ve araştırmanın örnekleme, verilerin toplanması ile verilerin analizi açıklanmaktadır.

3.1. Araştırmanın Türü

Bu araştırmada, 2011 ve 2012 yıllarında uygulanmış SBS 8'inci sınıf matematik alt testlerinden elde edilmiş ham puanlar, farklı örneklem büyüklükleri koşulu altında, doğrusal ve eşit yüzdelli eşitleme yöntemleriyle eşitlenmiş ve elde edilen eşitleme hataları birbiriyle karşılaştırılmıştır. Araştırmada var olan yöntem ve tekniklerin gerçek veri üzerinden sınanması yapıldığından araştırma temel araştırma niteliğindedir (Karasar, 2010).

3.2. Evren ve Örneklem

Bu araştırmanın evrenini, 2011 ve 2012 yıllarında SBS 8'inci sınıf alt testine tabi tutulan 517.382 kişi oluşturmaktadır. Araştırma, gerçek veri üzerinden yürütülmüş ve evreni temsil edecek örnekleme oluşturan kişiler, bilgisayar aracılığıyla, evrenin yaklaşık %1'i tesadüfi olarak seçilerek belirlenmiştir. Araştırmanın örnekleme, her iki yıla ait 3000'er kişiden toplamda 6000 oluşmaktadır. Araştırmada kullanılan diğer örneklem büyüklükleri (100, 500, 1500), 3000'er kişilik örneklemelerden tesadüfi olarak seçilmiştir.

3.3. Araştırma Verileri

Araştırmanın verilerini, 2011 ve 2012 yıllarında yapılan SBS 8'inci sınıf matematik alt testine tabi tutulan bireylere ait test ve madde puanları oluşturmaktadır. SBS, ülkemizde 2008 ve 2013 yılları arasında, Ortaöğretime geçiş sisteminin bir parçasını oluşturmuştur. Ortaöğretime geçiş sistemi ile ilköğretim düzeyindeki eğitim sürecinin değerlendirilmesi sonucu öğrencilerin başarı puanlarına uygun olarak, ortaöğretim kurumlarına yerleştirme amaçlanmıştır. Bu sistemin bir parçası olan SBS uygulamasında, matematik, türkçe, ingilizce, sosyal bilgiler, fen ve teknoloji alt test uygulamaları mevcuttur. Söz konusu alt test uygulamalarından,

eşitleme koşullarını (tek boyutluluk, aynı yapıyı ölçme, eşit güvenilirlik) en iyi sağlayan matematik alt testi, bu araştırmada kullanılmak üzere seçilmiştir. Matematik alt test uygulaması, 20 maddeden ibarettir. Matematik alt testi ile sınava girenlerin sayısal ve mantıksal akıl yürütme becerilerini ölçmek amaçlanmaktadır. Maddeler, çoktan-seçmeli ve dört seçeneklidir. Araştırmada kullanılan veriler düzenlenirken, farklı kitapçık türüne göre elde edilen kişilerin cevapları, tek kitapçık türü temel alınarak 1–0 puan matrisine çevrilip birleştirilmiştir. Verilerin elde edilmesi ile ilgili yazışmalar, ekler dizininde verilmiştir.

3.4.Verilerin Analizi

Verilerin analizi, SPSS.20 ve Excel (2007) paket programları yardımı ile dört aşamada gerçekleştirilmiştir. Birinci aşamada, 2011 ve 2012 SBS 8'inci sınıf matematik alt testlerinin betimsel istatistikleri hesaplanmış; ikinci aşamada, testlerin, eşitleme koşullarını sağlayıp sağlamadığı test edilmiş; üçüncü aşamada doğrusal ve eşit yüzdelikli eşitleme yöntemleri kullanılarak eşitlenmiş puanlar elde edilmiş ve son aşamada ise her bir eşitleme yönteminde ortaya çıkan hata miktarları hesaplanmıştır. Analizlerin üçüncü ve dördüncü aşamaları, 100, 500, 1500 ve 3000 kişilik örneklem büyüklükleri için tekrar edilmiş ve örneklem büyüklüğünün, eşitleme hatası üzerinde etkisinin tespiti yapılmıştır. Her aşamada kullanılan yöntem ve istatistikler aşağıda sıralanmıştır.

- ✓ Birinci aşamada; eşitlenecek testlerin ortalama, ortanca, tepe değer, standart sapma, varyans, çarpıklık ve basıklık katsayıları hesaplanmıştır.
- ✓ İkinci aşamada; eşitleme koşullarından tek boyutluluk koşulu için *açımlayıcı faktör analizi*, testlerin aritmetik ortalamalarının karşılaştırılmasında *t testi*, ortalama güçlüklerinin karşılaştırılmasında *z testi*, testlerin varyanslarının karşılaştırılmasında *Levene's testi*, güvenilirliklerinin karşılaştırılmasında Fischer'ın z istatistiği kullanılmıştır.
- ✓ Eşitleme için gerekli varsayımların test edilmesinden sonra, üçüncü aşamada eşitlemeye geçilmiştir. Bu aşamada, farklı örneklem büyüklüklerinde, doğrusal ve eşit yüzdelikli eşitleme yöntemleri kullanılmıştır. Araştırmada kullanılan eşitleme deseninin, eşdeğer gruplar deseni olmasından dolayı, bu eşitleme desenine uygun olan eşitleme

modelleri kullanılmıştır. Doğrusal eşitleme yöntemi için, Angoff (1971)'un doğrusal *eşitleme modeli*, eşit yüzdelikli eşitlemede aynı yüzdeler dilime giren puanlar nadiren denk geldiğinden dolayı Livingstone (2004) öteleme formülünden yararlanılmıştır. Her iki modele ait işlem süreçleri, birinci bölümde detaylı olarak anlatılmıştır.

- ✓ Son aşamada, farklı örneklem büyüklüklerinde yapılan her iki eşitleme yöntemine ait eşitleme sonuçları için hata miktarları tespit edilmiştir. Hata miktarlarının tespitinde, alan yazında sıkça kullanılan Ağırlıklandırılmış Hata Kareleri Ortalamaları(AHKO) indeksi kullanılmıştır. AHKO değerlerinin bulunmasında kullanılan formül, birinci bölümde açıklanmıştır. Bulunan AHKO değerleri, örneklem büyüklüğünün, kullanılan her iki eşitleme yöntemi üzerindeki etkisini görmek ve her iki eşitleme yönteminden uygun olanının tespitinde kullanılmıştır.

3.4.1. Betimsel İstatistikler

Verilerin analizinin ilk aşamasında hesaplanmış olan, 2011 ve 2012 yılları SBS 8'inci sınıf Matematik alt testlerinin betimsel istatistikleri Tablo 3.1'de verilmiştir.

Tablo3.1 2011 ve 2012 SBS 8'inci sınıf matematik alt testleri betimsel istatistikleri

İstatistikler	2011 SBS 8'inci Sınıf Matematik Alt Testi	2012 SBS 8'inci Sınıf Matematik Alt Testi
K	20	20
N	3000	3000
Ortalama	6,39	7,37
Ortanca	6	5
Tepe Değer	4	4
Std. Sapma	4,92	4,85
Varyans	24,23	23,57
Çarpıklık K.	1,201	0,89
Basıklık K.	0,590	-0,096
KR-20	0,92	0,91

K: Testteki Madde Sayısı, N:Örneklem Büyüklüğü

Tablo 3.1 incelendiğinde, her iki yıla ait testlerin ortalamasının, ortancadan büyük olduğu görülmektedir. Bu durumda, her iki yıla ait alt test puan dağılımlarının, sağa çarpık olduğu söylenebilir. Alt testlerin çarpıklık katsayılarının pozitif olması da, bu durumu desteklemektedir. Tablo 3.1'de verilen basıklık katsayıları incelendiğinde,

2011 yılına ait alt testin basıklık katsayısının pozitif olması, bu teste ait puan dağılımlarının, normalden daha sivri olduğu, diğer taraftan 2012 yılına ait alt testin basıklık katsayısının 0 yakın fakat negatif olması bu teste ait puan dağılımının normalden basık olduğu söylenebilir. Bu istatistikler sonucunda, her ne kadar 2011 yılına ait alt testin çarpıklık katsayısı +1 bandını aşmış olsa da, örneklem sayısının 30'dan çok büyük olması ve basıklık katsayılarının, her iki test içinde normallik için istenilen değerler arasında tespit edilmesinden dolayı, her iki alt testin puan dağılımı, normal dağılım olarak değerlendirilebilir. Dolayısıyla, bu aşamadan sonra yapılacak istatistiksel analizler, normallik varsayımının arandığı parametrik testler üzerinden yürütülmüştür. Öte yandan, her iki alt test puanlarına ait hesaplanan KR-20 güvenirlik katsayılarının 0,92 ve 0,91 gibi oldukça yüksek değerlere sahip olduğu gözükmektedir.

3.4.2. Eşitleme Koşullarının Test Edilmesi

Eşitleme işlemlerine geçilmeden önce, testlerin eşitleme yöntemlerinin koşullarına uygunluğu gözden geçirilmiştir. Bu araştırmada doğrusal ve eşit yüzdelikli eşitleme yöntemlerinin karşılaştırılması yapılmıştır. Doğrusal eşitleme yönteminin kullanılabilmesi için eşitlenecek iki testin ortalamaları ve standart sapmaları dışında tüm özelliklerinin aynı olması gerekmektedir (Woldbeck,1998). Öte yandan doğrusal eşitlemeye nazaran biraz daha esnek olan eşit yüzdelikli eşitleme yöntemi için, eşit güvenirlik ve tek yapının ölçülmesi özellikleri yeter şartlardır (Woldbeck,1998,s.13).

Alan yazında eşitleme koşulları için her ne kadar 5 koşul ifade edilmiş olsa da, Klasik Test Teorisine dayalı yapılan eşitleme çalışmalarında test edilen koşullarının, aynı ve tek yapının ölçülmesi, eşit güvenirlik, güçlük düzeyleri arasında paralellik olduğu görülmektedir. Bu çalışmada da, benzer şekilde eşitlenecek testler için tek yapıyı ölçme, güvenirliklerin eşitliği, varyansların ve ortalamalarının eşitliği, güçlük düzeylerinin paralelliği şartları test edilmiştir. 2011 ve 2012 yılına ait verilerin birbirinden bağımsız graplardan elde edilmiş olması nedeniyle, analizlerde ilişkisiz veya bağımsız örneklem için geçerli olan testler kullanılmıştır. Söz konusu testlere ilişkin analiz sonuç ve yorumları aşağıda sıralanmıştır.

3.4.2.1. Tek Boyutluluk

Eşitleme çalışmalarında, bireylere ait puanların karşılaştırılması ve bu karşılaştırmalar sonucunda bir takım modeller kurulmak sureti ile eşitlenecek iki test arasındaki ilişkinin modellenmesi yapılmaktadır. Söz konusu modelin, anlamlı olabilmesi için modelin kurulumunda kullanılan verilerin kaynağının aynı özellik veya yapıyı ölçüyor olması gerekmektedir. Aksi takdirde kurulan model her ne kadar istatistiksel açıdan bazı değerlere ulaşılmasını mümkün kılsa da, ulaşılan değerler anlamsız olacaktır. Woldbeck (1998), bu durumu eşitlenecek iki testin aynı ve tek bir özelliği ölçmesi gerektiğini vurgulayarak özetlemiştir.

Faktör analizi, aynı yapıyı ya da niteliği ölçen değişkenleri bir araya toplayarak ölçmeyi az sayıda faktör ile açıklamayı amaçlayan istatistiksel tekniktir (Büyüköztürk, 2012, s.123). Tek boyutluluğun ölçülmesinde, faktör analizi, sıklıkla kullanılan bir yöntemdir. Faktör analizi sonuçlarına bakılarak, değişkenlerin kaç faktörü ölçtüğüne karar verebilmek için, Büyüköztürk (2012), üç değerlendirme ölçütü önermiştir. Özdeğer, açıklanan varyans yüzdesi ve yamaç-birikinti grafiğinin incelenmesi neticesinde değişkenlerin tek yapıyı ölçüp ölçmediklerine karar verilebilir.

Yamaç-birikinti grafiği, Cattell (1966) tarafından faktör sayısına karar vermek amacıyla geliştirilen, baskın faktörleri ortaya koyarak, faktör analizinin asıl amacı olan faktör azaltmaya yardımcı olan bir grafikdir. Grafikte Y eksenini öz değer miktarını, X eksenini de bileşenlerin sayısını gösterir. Y eksenindeki bileşenler, X eksenine doğru bir iniş yaparlar. Bu iniş eğilimi, varyansa yaptıkları katkı çerçevesinde noktalarla gösterilir ve belli bir noktada genellikle keskin bir şekilde plato yapar. Bu nokta kesme noktası olarak düşünülebilir ve bu noktadan sonraki bileşenlerin varyansa yaptıkları katkı hem küçük hem de yaklaşık olarak aynıdır. Platoya kadar olan noktalar, faktör sayısı hakkında araştırmacıya bilgi verir (Gorsuch, 1983).

Özdeğerlerin incelenmesi esnasında, ilk iki özdeğer arasındaki oranın 3 veya 3,5 kat olması, ilk faktöre ait açıklanan varyans yüzdesinin tek başına %30 veya üzeri olması, değişkenlerin tek yapıyı ölçtüğü sonucunu verdiği bilinmektedir. Faktör analizine geçilmeden önce, verilerin faktör analizine uygunluğunu gösteren *Kaiser-Meyer-Olkin (KMO)* ve *Barlett test* sonuçlarına da bakılması gerekmektedir. Söz

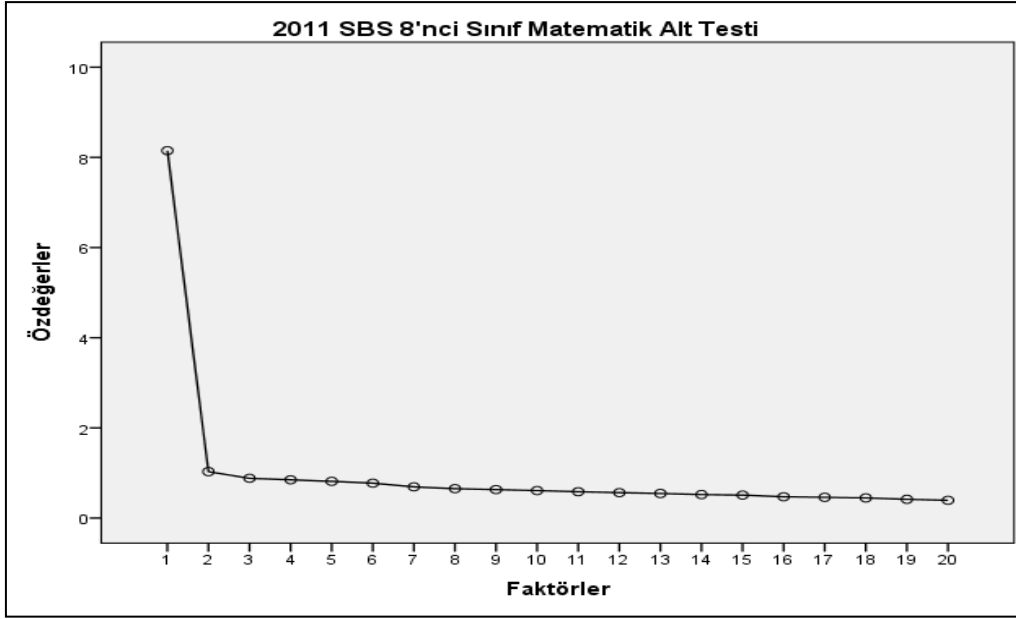
konusu deęerin, 0,60'dan byk olması ve *Barlett testi* sonucunda *ki-kare* istatistięinin anlamlı ıkması verilerin faktr analizine uygunluęunun kanıtı olarak deęerlendirilir.

Bu alıřmada da, deęiřken olan test maddeleri, aynı yapıyı lp lmediklerini grmek maksadıyla ncelikle KMO ve Barlett testine ve akabinde faktr analizine tabi tutulmuřtur. Yapılan analizler neticesinde, 2011 yılına ait SBS 8'inci sınıf matematik alt testinde KMO deęeri; 0,97, 2012 yılına ait SBS 8'inci sınıf matematik alt testinde KMO deęeri ise 0,96 olarak bulunmuřtur. Elde edilen KMO deęerlerinin 0,60'dan aık ve net bir řekilde byk olması verilerin faktr analiz iin uygun olduęunu gstermektedir. Barlett testi, deęiřkenler arasında iliřki olup olmadıęını, kısmi korelasyonlar temelinde inceler. Hesaplanan ki-kare istatistięinin anlamlı ıkması, veri matrisinin uygunluęunun gstergesidir. Test sonucunun anlamlı ıkması, puanların normallięinin bir kanıtı olarak da grlebilir (Bykztrk, 2012, s.126). Barlett testi sonucu elde edilen ki-kare istatistiklerinin, her iki yıla ait alt testlerde anlamlı olduęu ($p < 0.05$) grlmřtr. Bu sonu, KMO deęerlerini destekler nitelikte olup, verilerin faktr analizine uygunluęunun bir dięer gstergesidir.

Her iki yıla ait alt test verilerinin faktr analizine uygunluęu sılandıktan sonra faktr sayısı 5 ile sınırlandırılarak, faktr analizi yapılmıřtır. Faktr analizi sonucunda, her iki teste ait ortaya ıkan zdeęerler, aıklanan varyans oranları Tablo 3.2 ve Tablo 3.3'de, yama-birikinti grafikleri ise řekil 3.1 ve řekil 3.2'de verilmiřtir.

Tablo 3.2 2011 SBS 8'inci Sınıf Matematik Alt Testi Faktr Analizi Sonuları
2011 SBS 8'inci Sınıf Matematik Alt Testi

Faktr	zdeęer	Varyans Yzdesi	Kmltif Yzde
1	8,14	40,74	40,74
2	1,03	5,15	45,89
3	0,88	4,40	50,30
4	0,84	4,24	54,55
5	0,81	4,07	58,62

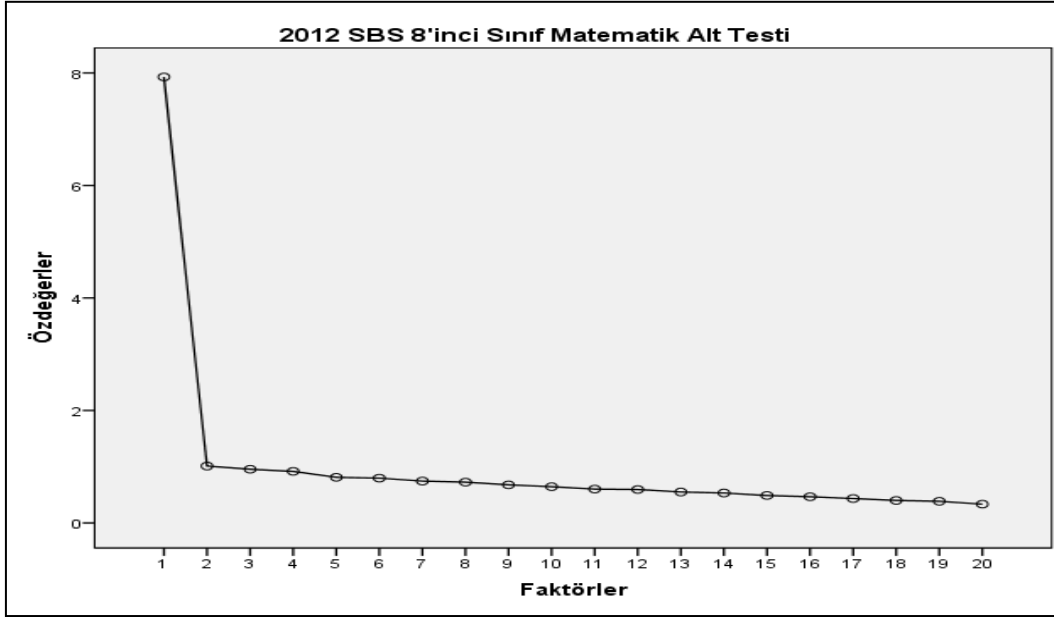


Şekil 3.1 2011 Yılı SBS 8'inci Sınıf Matematik Alt Testi Yamaç-Birikinti Grafiği

Tablo 3.2 incelendiğinde, 2011 yılına ait alt testin varyansının %40,74'ünün, birinci faktör tarafından açıklandığı görülmüştür. Buna ilave olarak, birinci faktör ile ikinci faktör özdeğerleri arasında keskin düşüş görülmekte, birinci ve ikinci özdeğer arasındaki oran yaklaşık 8 kat olarak gözükmektedir. Bu bulgular, yukarıda Şekil3.1'de verilen yamaç-birikinti grafiği ile de desteklenmektedir. Bu bulgulara dayalı olarak 2011 yılına ait SBS Matematik alt testinin, tek bir yapıyı ölçtüğü söylenebilir.

Tablo 3.3 2012 SBS 8'inci Matematik Alt Testi Faktör Analizi Sonuçları

2012 SBS 8'inci Sınıf Matematik Alt Testi			
Faktör	Özdeğer	Varyans Yüzdesi	Kümülatif Yüzde
1	7,92	39,64	39,64
2	1,01	5,05	44,70
3	0,95	4,77	49,47
4	0,91	4,58	54,06
5	0,81	4,05	58,12



Şekil 3.2 2012 Yılı 8'inci Sınıf Matematik Alt Testi Yamaç-Birikinti Grafiği

Tablo 3.3 incelendiğinde, 2011 yılına ait alt testin varyansının %39,64'ünün, birinci faktör tarafından açıklandığı görülmüştür. Buna ilave olarak, birinci faktör ile ikinci faktör özdeğerleri arasında keskin düşüş görülmekte, birinci ve ikinci özdeğer arasındaki oran, yaklaşık 8 kat olarak gözükmektedir. Bu bulgular, yukarıda Şekil3.2'de verilen yamaç-birikinti grafiği ile de desteklenmektedir. Bu bulgulara dayalı olarak, 2012 yılına ait SBS matematik alt testinin, tek bir yapıyı ölçtüğü söylenebilir.

Yukarıdaki sonuçlara dayalı olarak, her iki yıla ait alt testlerin, tek boyutlu olduğu veya tek bir yapıyı ölçtüğünü söylemek mümkündür. Ayrıca, her iki yıla ait testlerin aynı yapıyı ölçüp ölçmediklerine ilişkin uzman kanısına başvurulmuştur. Tecrübe olarak birbirinden farklı 5 ortaokul matematik öğretmeni ile 3 ölçme ve değerlendirme uzmanının görüşlerine başvurulmuştur. Görüşüne başvuru alan uzmanların tamamı, her iki testin de aynı yapıyı ölçtüğü konusunda olumlu fikir beyan etmişlerdir. Tüm bu sonuçlar ışığında, her iki yıla ait alt testlerin, aynı ve tek yapıyı ölçtüğünü, dolayısıyla eşitleme için gerekli olan tek boyutluluk koşulunu sağladığı söylenebilir.

3.4.2.2. Alt Testlerin Aritmetik Ortalamaları ve Varyansları Arasındaki Farkın Test Edilmesi

Testlerin ortalamaları arasındaki farkın, anlamlı olup olmadığı, *ilişkisiz gruplar t-testi* ile varyanslarının eşit olup olmadığı ise, *Levene's testi* sonucuna göre değerlendirilmiştir. Tablo 3.4'de ilgili test sonuçları verilmiştir.

Tablo3.4 Alt Testlerin Aritmetik Ortalamaları ve Varyansları Arasındaki Fark Test Sonuçları

Alt Testler	Levene's Testi			t Testi	
	S_x^2	F	p	X	t
2011 SBS 8'inci Sınıf Matematik Alt Testi	24,22	0,79	0,37	6,39	-7,76*
2012 SBS 8'inci Sınıf Matematik Alt Testi	23,58			7,37	

*p < 0,05

Tablo 3.4 incelendiğinde, her iki yıla ait alt testlerin ortalamaları arasında, 0.05 anlamlılık düzeyinde, anlamlı bir fark olduğu dolayısıyla her iki alt testin ortalamalarının eşit olmadığı görülmektedir. Levene's testi sonuçlarına göre ise, alt testlerin varyansları arasında 0.05 anlamlılık düzeyinde, anlamlı bir fark olmadığı başka bir ifade ile varyansların homojen ve eşit olduğu görülmektedir.

3.4.2.3. Alt Testlerin Ortalama Güçlükleri Arasındaki Farkın Test Edilmesi

Ortalama güçlükler arasında anlamlı bir farkın olup olmadığını, test edebilmek amacıyla *z testinden* yararlanılmıştır. Tablo3.5'de ilgili test sonuçları verilmiştir.

Tablo3.5 Alt Testlerin Ortalama Güçlükleri Arasındaki Farkın Test Sonucu

Alt Testler	K	p	z
2011 SBS 8'inci Sınıf Matematik Alt Testi	20	0,32	-3,27*
2012 SBS 8'inci Sınıf Matematik Alt Testi	20	0,36	

*p < 0,05

Tablo 3.5 incelendiğinde, her iki yıla ait alt testlerin ortalama güçlükleri arasında 0.05 anlamlılık düzeyinde, anlamlı bir fark olduğu dolayısıyla iki alt testin eşit zorluk düzeyine sahip olmadığı söylenebilir.

3.4.2.4. Alt Testlerin Güvenirlikleri Arasındaki Farkın Test Edilmesi

Öncelikle eşitlenecek 2011 ve 2012 SBS 8'inci sınıf matematik alt testlerinin, her biri için KR-20 güvenirlilik katsayıları bulunmuştur. Testin güvenirlilik katsayısı olarak hesaplanan korelasyon (r), test puanlarına ilişkin bireysel farklılıkların ne derece gerçek ne derece hata faktörüne bağlı olduğunu yorumlamak amacıyla kullanılır (Büyüköztürk, 2012). KR-20 güvenirlilik katsayısı, madde puanlarının test puanı ile korelasyonunu ifade eden iç tutarlılık katsayısıdır. Bu bilgiler ışığında her iki yıla ait güvenirlilik katsayıları korelasyon katsayısı olarak değerlendirilmiştir. İlişkisiz iki korelasyon katsayısı (r_1 ve r_2) arasındaki farkın manidarlığı, Fischer'in z_r dönüşümünün kullanılması ile kolayca test edilebilir (Akhun, 1984; Hovardaoğlu, 1995). Bu bilgiler ışığında, iki güvenirlilik katsayısı arasında fark olup olmadığı Fischer'in z istatistiği ile test edilerek iki testin güvenirlilikleri arasında anlamlı bir fark olup olmadığı sonucu aranmıştır.

Tablo 3.6 Alt Testlerin Güvenirlikleri Arasındaki Farkın Test Sonucu

Alt Testler	KR-20	z_r	z
2011 SBS 8'inci Sınıf Matematik Alt Testi	0,920	1,589	1,12
2012 SBS 8'inci Sınıf Matematik Alt Testi	0,915	1,528	

Tablo 3.6 incelendiğinde, her iki yıla ait alt test güvenirlilikleri arasında 0.05 anlamlılık düzeyinde, manidar bir fark olmadığı görülmektedir. Bu sonuca dayalı olarak, her iki yıla ait alt testlerin eşit güvenirliliğe sahip olma koşulunu sağladığı söylenebilir.

Yapılan analizler neticesinde, 2011 SBS 8'inci sınıf matematik alt testi ile 2012 SBS 8'inci sınıf matematik alt testinin, aynı yapıyı tek boyutta ölçtüğü, eşit güvenirlilik, eşit varyansa sahip olduğu bulunurken, puan ortalamaları ve güçlükleri arasında farkın olduğu görülmektedir. Tüm bu analizler neticesinde, eşitleme koşullarının sağlandığı sonucuna varılarak testlerin eşitlemesine geçilmiştir.

3.4.3. Eşitleme Yöntemleri

Bilindiği gibi eşitleme koşullarından biri de simetrikliktir. Simetriklik koşuluna göre, eşitleme yönünün eşitleme üzerinde etkisinin olmaması gerekmektedir. Bu çalışmada simetriklik koşulu sınanmadığından, eşitleme yönü araştırmacı tarafından belirlenmiştir. Buna göre, 2011 yılı SBS 8'inci sınıf matematik alt testi, 2012 yılı SBS 8'inci sınıf matematik alt testine eşitlenmiştir. Çalışmada, Klasik Test Teorisine dayalı doğrusal eşitleme yöntemi ve eşit yüzdellikli eşitleme yöntemleri kullanılmıştır.

3.4.3.1. Doğrusal Eşitleme Yöntemi

Doğrusal eşitleme yönteminde, Angoff (1971)'un doğrusal eşitleme modeli kullanılmıştır. Modele ait eşitlikler çalışmanın birinci bölümünde açıklanmıştır.

3.4.3.2. Eşit Yüzdellikli Eşitleme Yöntemi

Eşit yüzdellikli eşitleme yönteminde, Angoff (1971)'un aynı eşit yüzdellik dilime denk gelen puanların eşit kabul edileceği tanımından yararlanılmıştır. Öncelikle her iki yıla ait puanların yığılmalı frekans tabloları çıkarılmış ve aynı yüzdellik dilime denk gelen puanların eşit kabul edilmiştir. Ancak yığılmalı frekans tablosunda puanların aynı yüzdellik dilime denk gelmesi nadiren olduğundan Livingstone (2004) öteleme formülünden yararlanılmıştır. Burada kullanılan işlem aşamaları, çalışmanın birinci bölümde detaylı olarak açıklanmıştır.

3.4.4. Değerlendirme Ölçütü

Test eşitleme yöntemlerinin karşılaştırılması veya aynı eşitleme yönteminin farklı örneklem büyüklükleri koşulu altında karşılaştırılması için, eşitleme yöntemlerinden kestirilen eşitleme hataları kullanılmaktadır. Test eşitleme yöntemlerinden kestirilen eşitleme hataları karşılaştırılırken, bir ölçüte ihtiyaç duyulur. Bu çalışmada, ölçüt olarak, AHKO (Ağırlıklandırılmış Hata Kareleri Ortalaması) indeksi kullanılmıştır. İndekse ait eşitlik ve açıklamalar birinci bölümde açıklanmıştır.

4. BULGULAR VE YORUMLAR

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

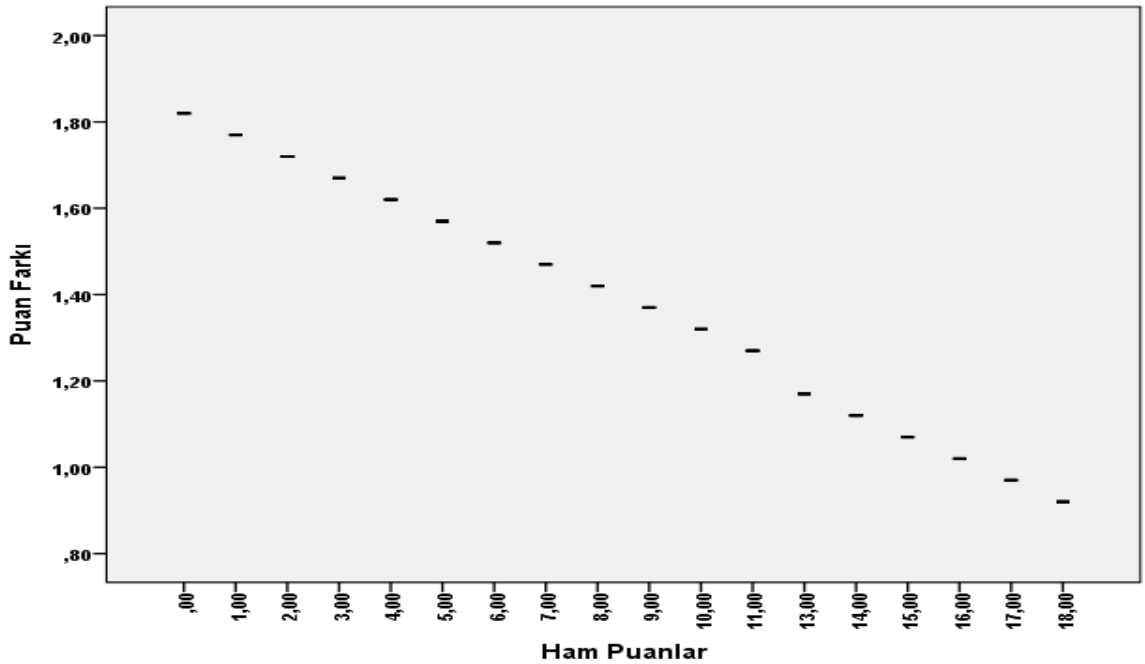
4.1. Alt Problem 1'e İlişkin Bulgular Ve Yorumlar

Alt Problem 1: Farklı örneklem büyüklüğü koşulu altında, 2011 ve 2012 yılı SBS 8'inci sınıf matematik alt test puanlarının doğrusal eşitleme yöntemi kullanılarak elde edilmiş eşitlenmiş puanları nasıldır?

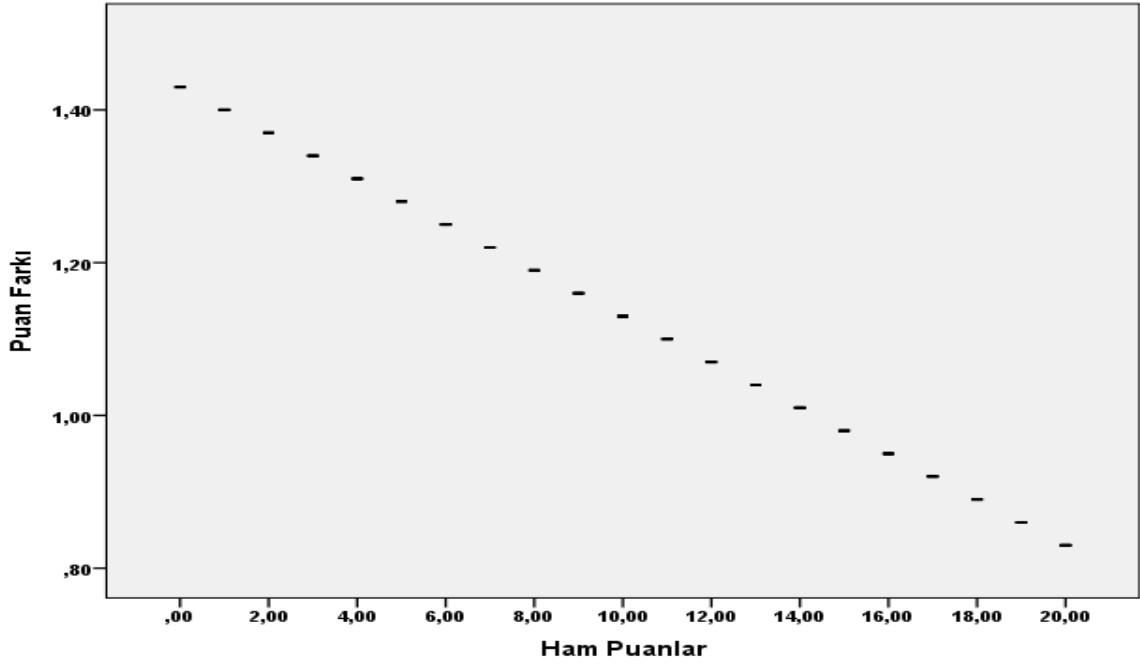
Bu alt problemin çözümü için 2011 ve 2012 yılı SBS 8'inci sınıf matematik alt test verilerinden, öncelikle 3000'er kişilik veri setleri random olarak seçilmiştir. Eşitleme koşulları, seçilen bu 3000'er kişilik veri seti üzerinden sınıanmıştır. Bundan dolayı, 1500,500 ve 100'er kişilik veri setleri 3000'er kişilik veri setlerinden random olarak seçilmiştir. Böylece eşitleme koşullarının tüm örneklem büyüklüklerinde geçerli olması sağlanmıştır. Alt problemin çözümünde, doğrusal eşitleme yöntemi eşitlikleri kullanılarak, 2011 SBS 8'inci sınıf matematik alt testi ham puanları, 2012 yılı SBS 8'inci sınıf matematik alt testi ham puanlarına eşitlenmiştir. Eşitleme işlemi sonucunda, 3000 kişilik örneklem büyüklüğünde $Y=0,98X+1,10$, 1500 kişilik örneklem büyüklüğünde $Y=0,99X+1,15$, 500 kişilik örneklem büyüklüğünde $Y=0,97X+1,43$, 100 kişilik örneklem büyüklüğünde $Y= 0,95X+1,82$ olan doğru denklemleri bulunmuştur. Elde edilen doğru denklemleri yardımı ile tüm örneklem büyüklükleri için, ham puanlara karşılık gelen eşitlenmiş puanlar bulunmuştur. Her örneklem büyüklüğünde, ham puanlara karşılık gelen eşitlenmiş puanlar ve ham puanlar ile eşitlenmiş puanlar arasındaki farklar, Tablo4.1'de gösterilmiştir. Ayrıca, her örneklem büyüklüğü için puan farkları ile ham puanların matematiksel ilişkisini gösteren grafikler aşağıda Şekil4.1, Şekil4.2, Şekil4.3, Şekil4.4'de verilmiştir.

Tablo 4.1 Farklı örneklem büyüklüklerinde doğrusal eşitleme yöntemiyle elde edilen eşitlenmiş puanlar ve puan farkları

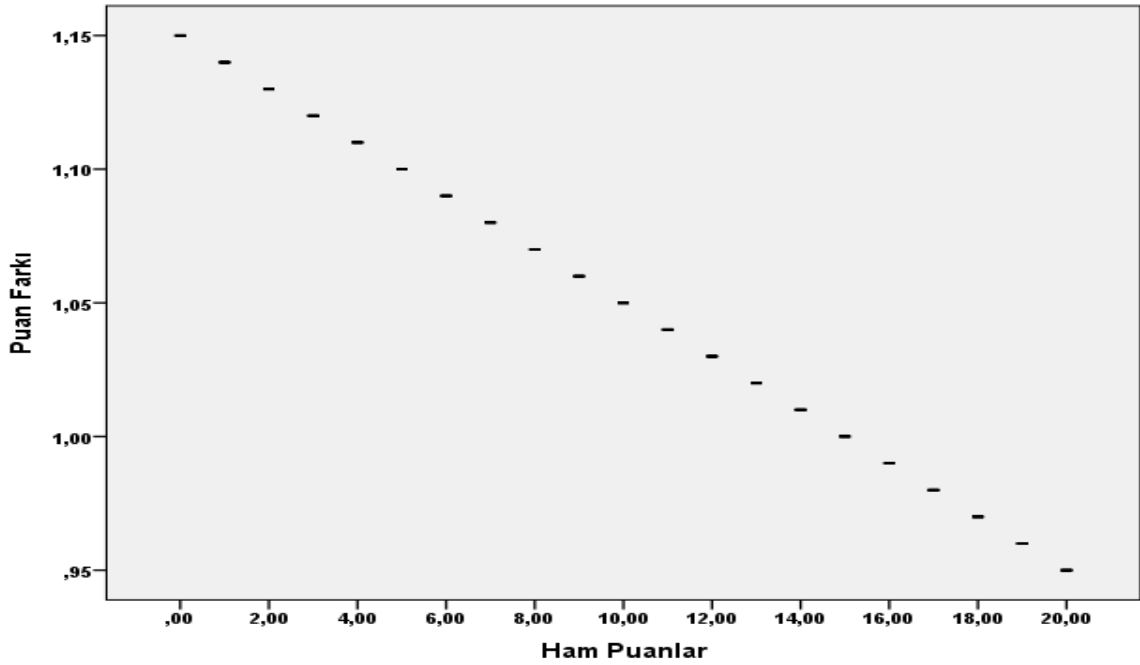
Örneklem Büyüklüğü	100		500		1500		3000	
Ham Puan	Eşitlenmiş Puan	Puan Farkı	Eşitlenmiş Puan	Puan Farkı	Eşitlenmiş Puan	Puan Farkı	Eşitlenmiş Puan	Puan Farkı
0	1,82	1,82	1,43	1,43	1,15	1,15	1,10	1,10
1	2,77	1,77	2,40	1,40	2,14	1,14	2,08	1,08
2	3,72	1,72	3,37	1,37	3,13	1,13	3,06	1,06
3	4,67	1,67	4,34	1,34	4,12	1,12	4,04	1,04
4	5,62	1,62	5,31	1,31	5,11	1,11	5,02	1,02
5	6,57	1,57	6,28	1,28	6,10	1,10	6,00	1,00
6	7,52	1,52	7,25	1,25	7,09	1,09	6,98	0,98
7	8,47	1,47	8,22	1,22	8,08	1,08	7,96	0,96
8	9,42	1,42	9,19	1,19	9,07	1,07	8,94	0,94
9	10,37	1,37	10,16	1,16	10,06	1,06	9,92	0,92
10	11,32	1,32	11,13	1,13	11,05	1,05	10,90	0,90
11	12,27	1,27	12,10	1,10	12,04	1,04	11,88	0,88
12	-	-	13,07	1,07	13,03	1,03	12,86	0,86
13	14,17	1,17	14,04	1,04	14,02	1,02	13,84	0,84
14	15,12	1,12	15,01	1,01	15,01	1,01	14,82	0,82
15	16,07	1,07	15,98	0,98	16,00	1,00	15,80	0,80
16	17,02	1,02	16,95	0,95	16,99	0,99	16,78	0,78
17	17,97	0,97	17,92	0,92	17,98	0,98	17,76	0,76
18	18,92	0,92	18,89	0,89	18,97	0,97	18,74	0,74
19	-	-	19,86	0,86	19,96	0,96	19,72	0,72
20	-	-	20,83	0,83	20,95	0,95	20,70	0,70
Ortalama Puan Farkı		1,38		1,13		1,05		0,90



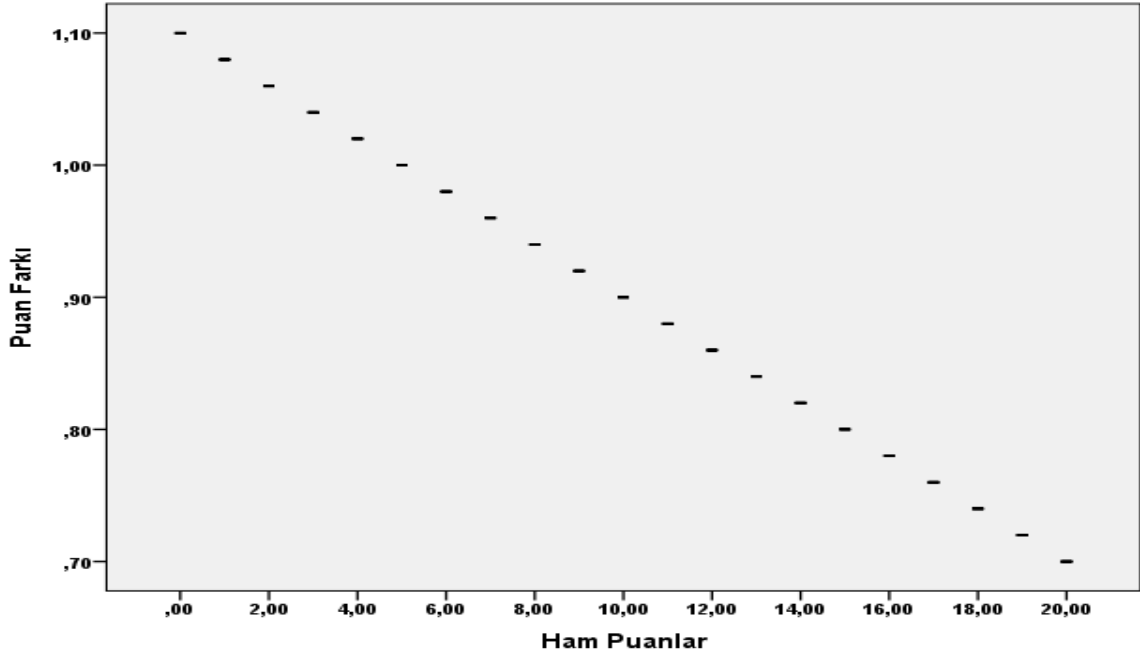
Şekil 4.1 100 kişilik örneklem büyüklüğü doğrusal eşitleme puan farkı grafiği



Şekil 4.2 500 kişilik örneklem büyüklüğü doğrusal eşitleme puan farkı grafiği



Şekil4.3 1500 kişilik örneklem büyüklüğü doğrusal eşitleme puan farkı grafiği



Şekil 4.4 3000 kişilik örneklem büyüklüğü doğrusal eşitleme puan farkı grafiği

Tablo 4.1 incelendiğinde, 100 örneklem büyüklüğünde 12, 19, 20 puan alan adayın olmaması 12, 19, 20 ham puanlarına karşılık gelen eşitlenmiş puanların ve puan farkının olmasına sebep olmuştur. Tüm örneklem büyüklüklerinde eşitlenmiş puanların ham puan üst sınırını, az bir farkla geçtiği görülmektedir. Bu durum eşitlenmiş puanların, doğrusal bir denklem ile elde edilmesinin doğal sonucudur. Her ne kadar ham puan alt sınırını negatif yönde bir aşma gözükme de, doğru denkleminin kesme noktasının negatif olması durumunda, ham puan sınırının, negatif yönde aşılması da olası sonuçlar arasındadır. Nitekim Livingston (2004), Jaeger (1981) söz konusu ham puan aralığının aşılmasını, doğrusal eşitleme yöntemine has bir özellik olduğunu ve doğrusal eşitleme yöntemi için beklendik bir sonuç olduğunu belirtmişlerdir. Ancak, ham puan aralığının dışındaki puanların değerlendirilmesi, doğal olarak sorun teşkil eder. Kolen ve Brennan (2004), bu sorunun çözümü için, iki alternatif çözüm sunmuştur. İlki ham puan sınırını aşan puanların, olduğu gibi kabulü, diğeri ise eşitlenmiş puanların ham puan sınırlarında kesilerek, ham puan sınırını pozitif yönde aşan eşitlenmiş puanların, en yüksek ham puana denkliğini, ham puan sınırını negatif yönde aşan eşitlenmiş puanların ise en düşük ham puana denkliğini kabul etmektir.

Değerlendirme bir karar verme işidir ve ölçme sonuçlarının bir ölçütle kıyaslanarak bir karara varılması işi olarak tanımlanmaktadır (Baykul, 2010, s.98). Burada değerlendirme yapan ölçme ve değerlendirme uzmanı, 100 ve 0 gibi iki ölçüt koymakta ve buna göre karar vermektedir. Ölçme ve değerlendirme uzmanı, 100 tam puan alan tüm bireylerin eşit yetenek düzeyinde ve aynı şekilde 0 alan tüm bireylerin aynı yetenek düzeyinde olduğu kararına ulaşılır. Gerçekte ise, tüm bu bireylerin eşit yetenek düzeyinde olduklarını kabul etmek zordur. Yukarıda örneği verilen durum, 100 tam puan alan bireylerin aynı yetenek seviyesinde olduklarını söylenemeyeceğinin kanıtıdır. Dolayısıyla burada problem olarak ifade edilmesi gereken durum, 100'ü aşan her hangi bir puanın açıklanması değil, 100 tam puan alan bireylerin aynı yetenek seviyesinde olduklarının kabulüdür. Birbirleri ile karşılaştırılmak istenen her iki testte de aynı kabul söz konusu olduğuna göre, eşitleme sonucunda 100 ve üzeri puanların 100, 0 ve altı puanların 0 olarak değerlendirilmesinde bir sakınca yoktur. Dolayısıyla Kolen ve Brennan'ın ikinci çözüm yolunun, daha uygun gözüktüğü söylenebilir.

Kolen ve Brennan'a (2004) göre, doğrusal eşitleme formlar arasındaki güçlük farklılığını puan ölçeği boyunca sabitlemekten ziyade, söz konusu güçlük farklılığını puan ölçeği boyunca düzenlemektedir. Yani başarısız öğrenciler için form X, form Y'den daha zor iken, başarılı öğrenciler için form X, form Y'den daha kolay gözükabilir. Puan dağılımı boyunca ham puanlar yükseldikçe, eşitlenmiş puanlar ile ham puanlar arasındaki farkın azaldığı gözükmektedir. Puan farkları grafikleri de bu değişkenliği göstermektedir. Bu durumda, eşitlenen iki test formunun güç farklılığının, puan dağılımı boyunca sabit olmadığı aksine değişkenlik gösterdiği söylenebilir.

Tüm örneklem büyüklüklerinde, eşitlenmiş puanların ham puanlardan, puan dağılımı boyunca büyük olduğu ve söz konusu puan farkının, puanlar yükseldikçe azaldığı görülmektedir. Puan dağılımı boyunca oluşan farkın, iki testin aritmetik ortalamaları arasındaki farktan kaynaklandığı söylenebilir. Her ne kadar iki testin varyansları arasında, manidar bir fark olmasa da, aritmetik ortalamalar arasındaki manidar farkın, ham puanlar ile eşitlenmiş puanların aynı olmamasına sebep olduğu söylenebilir. Bu alt probleme ait tüm örneklem büyüklüklerindeki bulguların ortak özellikleri, eşitlenmiş puanların, ham puan sınırını pozitif yönde aşmış olması ve puan dağılımı boyunca güç farklılığının sabit olmamasıdır. Bu bulgular,

Keleciođlu (1994), Bozdađ (2007), Öztürk (2010), Kan (2010), Mutluer (2013) tarafından yapılan alıřmaların bulguları ile tutarlılık göstermektedir.

4.2. Alt Problem 2'ye İliřkin Bulgular Ve Yorumlar

Alt Problem 2: Örnekleme büyüklüğünün doğrusal eşitleme yöntemi üzerindeki etkisi nasıldır?

Bu alt problemin çözümü için, ilk alt probleme ait doğrusal eşitleme sonuçlarının, ağırlıklandırılmış hata kareleri ortalamaları ölçütüne göre karşılaştırılması yapılmıştır. Her örneklem büyüklüğü için hesaplanan AHKO değerleri Tablo4.2'de verilmiştir.

Tablo4.2 Farklı örneklem büyüklüklerinde alt testlerin doğrusal eşitleme yöntemi ile eşitlenmesine ait AHKO değerleri

Eşitleme Yöntemi	Örnekleme Büyüklüğü	AHKO
Doğrusal Eşitleme	100	0,098
	500	0,069
	1500	0,066
	3000	0,047

AHKO değerleri için standart hata bulunmadığı için, AHKO değerleri arasında istatistiksel açıdan bir farkın manidarlığının ne olduğuna ilişkin bir değerlendirme yapmak zordur. Ancak pratikte 0,05'lik veya daha fazla olan bir fark, anlamlı olarak değerlendirilmektedir(Skaggs ve Lissitz, 1986, s.311). Tablo4.2 incelendiğinde, AHKO değerleri arasında sadece 3000 ve 100 örneklem büyüklükleri arasında 0,05'lik bir farkın olduğu gözükmektedir. Bununla birlikte örneklem büyüklüğü arttıkça, AHKO değerlerinin de manidar bir şekilde azaldığı gözükmektedir. AHKO formülü incelendiğinde, eşitlenmiş puanlar ile ham puanlar arasındaki farkın hata değeri olarak nitelendirildiği gözükmektedir. Bu hata değerlerinin, ağırlıklandırılmış ortalamalarını ifade eden AHKO değerlerinin küçük olması, eşitleme random hatasının az olduğu, bir başka ifade ile eşitlenmiş puanlar ile ham puanlar arasındaki farkın az olduğu anlamına gelmektedir. Dolayısı ile AHKO değerleri küçüldükçe, eşitleme çalışmasının random hatalardan arınlığı artmaktadır.

Alan yazında doğrusal eşitleme için önerilen örneklem büyüklüğü, minimum 400'dür(Zhu,1998, s.17; Kolen ve Brennan, 2004, s.288). Kolen ve Brennan, bu örneklem büyüklüğünün eşdeğer olmayan grup desenleri için de önerilebileceğini

ifade etmiştir(Kolen ve Brennan, 2004). Tablo4.2 incelendiğinde, 1500 ve 500 kişilik örneklem büyüklüklerine ait AHKO değerlerinin birbirlerine kısmen yakın olduğu, 100 kişilik örneklem büyüklüğüne inildiğinde, AHKO değerinin ciddi bir biçimde arttığı görülmektedir. Bu durum, doğrusal eşitleme için örneklem büyüklüğünün, 400 ve üstü olduğu durumlarda daha kararlı sonuçlar verdiği sonucunu desteklemektedir.

Bu bulgular, Gök (2012), Kilmen (2010), Çetin (2009), Lee ve Ban (2010), Kim ve Cohen (2002), Tsai (1997) tarafından yapılan çalışmaların bulguları ile tutarlılık göstermektedir. Kolen ve Brennan (2004) da örneklem büyüklüğünün eşitleme hatası üzerinde, direkt etkisi olduğunu belirtmişlerdir. Harris ve Crouse (1993),örneklem büyüklüğünün eşitlemeye etkisini gösteren bir çalışmalarında, geniş örneklem büyüklüklerinin daha iyi eşitleme sonuçlarını doğurduğunu ifade etmişlerdir. Bu bilgiler ışığında, örneklem büyüklüğü arttıkça doğrusal eşitleme yöntemi için eşitlemenin doğrulunun arttığı söylenebilir.

4.3. Alt Problem 3'ye İlişkin Bulgular Ve Yorumlar

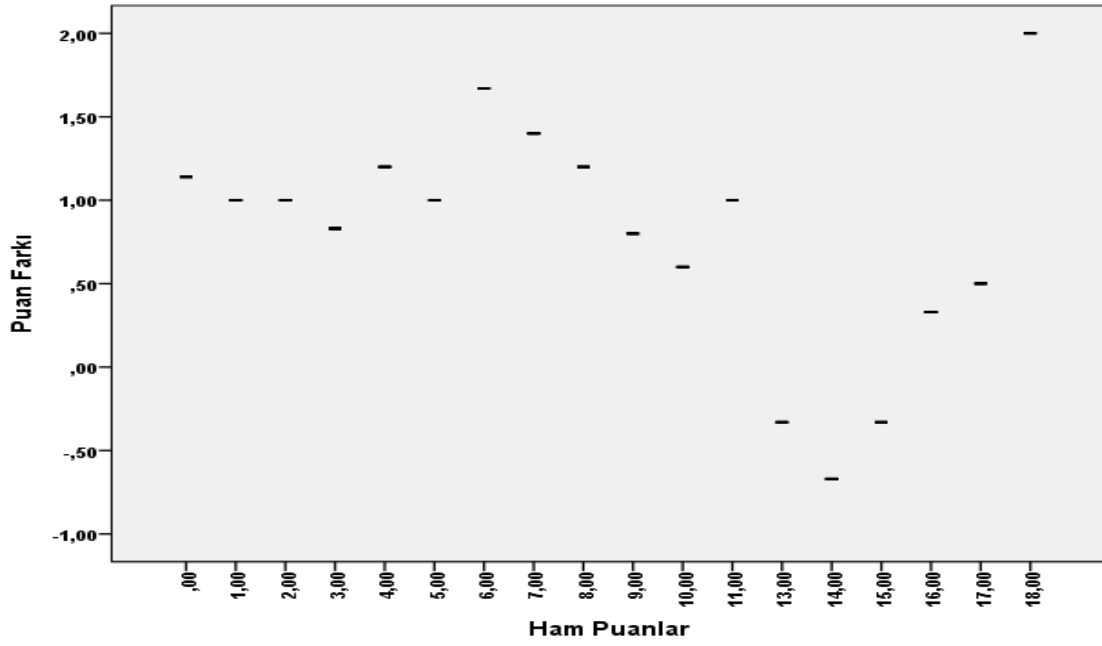
Alt Problem 3: Farklı örneklem büyüklüğü koşulu altında, 2011 ve 2012 yılı SBS 8'inci sınıf matematik alt test puanlarının eşit yüzdellikli eşitleme yöntemi kullanılarak elde edilmiş eşitlenmiş puanları nasıldır?

Bu alt problemin çözümü için 2011 ve 2012 yılı SBS 8'inci sınıf matematik alt test verilerinden, öncelikle 3000'er kişilik veri setleri random olarak seçilmiştir. Eşitleme koşulları, seçilen bu 3000'er kişilik veri seti üzerinden sınıanmıştır. Bundan dolayı, 1500,500 ve 100'er kişilik veri setleri, 3000'er kişilik veri setlerinden random olarak seçilmiştir. Böylece; eşitleme koşullarının tüm örneklem büyüklüklerinde geçerli olması sağlanmaya çalışılmıştır. Alt problemin çözümünde, eşit yüzdellikli eşitleme yöntemi aşamaları takip edilerek, 2011 SBS 8'inci sınıf matematik alt testi ham puanları, 2012 yılı SBS 8'inci sınıf matematik alt testi ham puanlarına eşitlenmiştir. Her iki alt testin yığılmalı frekans tablolarında puanların aynı yüzdellik dilime nadiren denk gelmesinden dolayı Livingstone (2004)'un öteleme formülünden yararlanılmıştır. Her örneklem büyüklüğünde, ham puanlara karşılık gelen eşitlenmiş puanlar ve ham puanlar ile eşitlenmiş puanlar arasındaki farklar Tablo4.3'de gösterilmiştir. Ayrıca her örneklem büyüklüğü için, puan farkları ile

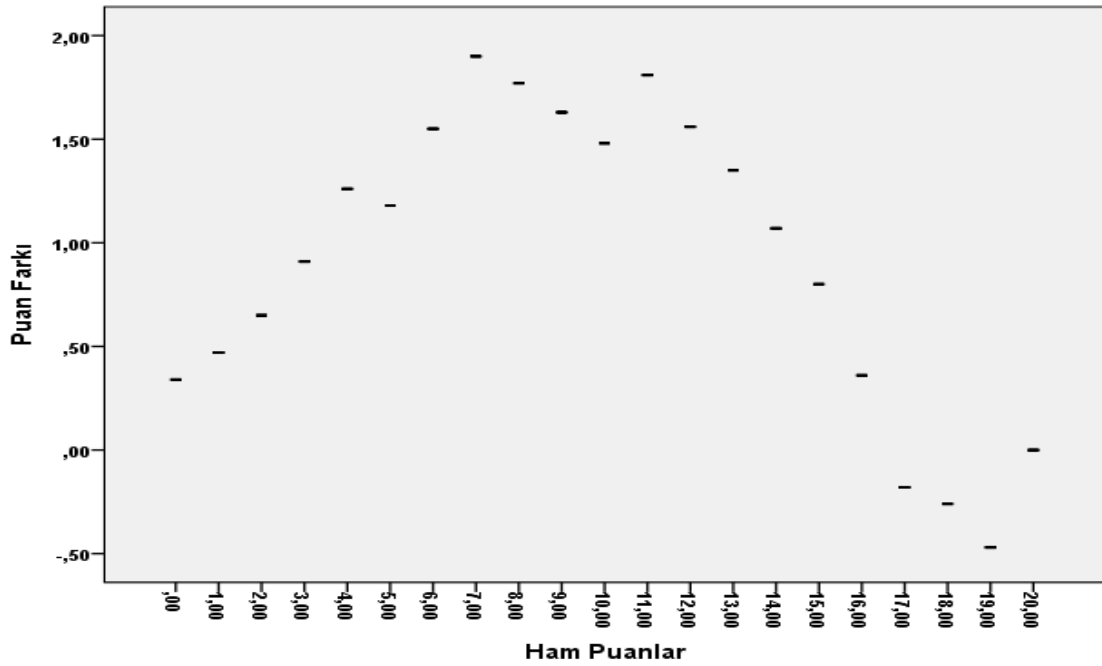
ham puanların matematiksel ilişkisini gösteren grafikler Şekil4.5, Şekil4.6, Şekil4.7, Şekil4.8'de verilmiştir.

Tablo 4.3 Farklı örneklem büyüklüklerinde eşit yüzdeliği eşitleme yöntemiyle elde edilen eşitlenmiş puanlar ve puan farkları

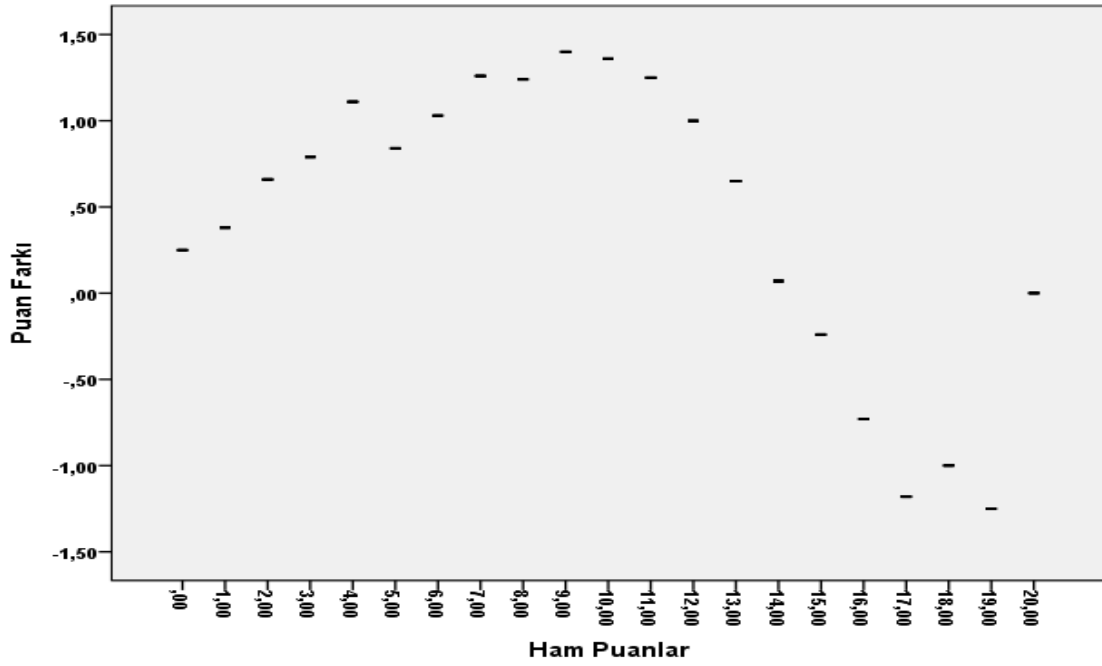
Örneklem Büyüküğü	100		500		1500		3000	
Ham Puan	Eşitlenmiş Puan	Puan Farkı	Eşitlenmiş Puan	Puan Farkı	Eşitlenmiş Puan	Puan Farkı	Eşitlenmiş Puan	Puan Farkı
0	1,14	1,14	0,34	0,34	0,25	0,25	0,27	0,27
1	2,00	1,00	1,47	0,47	1,38	0,38	1,27	0,27
2	3,00	1,00	2,65	0,65	2,66	0,66	2,32	0,32
3	3,83	0,83	3,91	0,91	3,79	0,79	3,18	0,18
4	5,20	1,20	5,26	1,26	5,11	1,11	5,02	1,02
5	6,00	1,00	6,18	1,18	5,84	0,84	6,12	1,12
6	7,67	1,67	7,55	1,55	7,03	1,03	7,50	1,50
7	8,40	1,40	8,90	1,90	8,26	1,26	8,10	1,10
8	9,20	1,20	9,77	1,77	9,24	1,24	9,21	1,21
9	9,80	0,80	10,63	1,63	10,40	1,40	10,11	1,11
10	10,60	0,60	11,48	1,48	11,36	1,36	11,27	1,27
11	12,00	1,00	12,81	1,81	12,25	1,25	12,13	1,13
12	-	-	13,56	1,56	13,00	1,00	13,37	1,37
13	12,67	-0,33	14,35	1,35	13,65	0,65	14,16	1,16
14	13,33	-0,67	15,07	1,07	14,07	0,07	14,54	0,54
15	14,67	-0,33	15,80	0,80	14,76	-0,24	15,56	0,56
16	16,33	0,33	16,36	0,36	15,27	-0,73	16,26	0,26
17	17,50	0,50	16,82	-0,18	15,82	-1,18	16,99	-0,01
18	20,00	2,00	17,74	-0,26	17,00	-1,00	17,96	-0,04
19	-	-	18,53	-0,47	17,75	-1,25	18,90	-0,10
20	-	-	20,00	0,00	20,00	0,00	20,00	0,00
Ortalama Puan Farkı		0,94		1,00		0,85		0,69



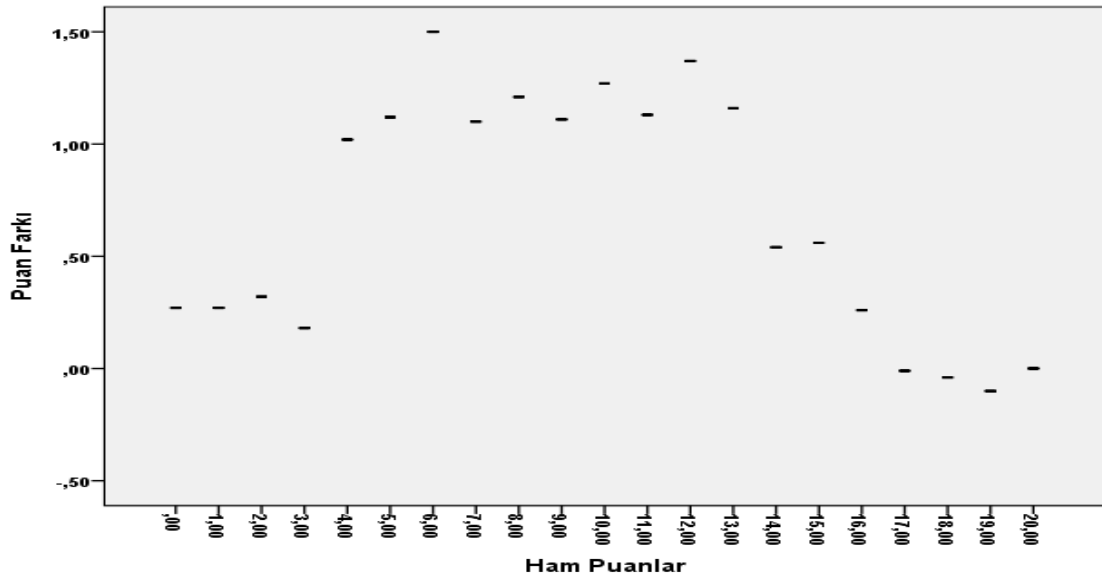
Şekil 4.5 100 kişilik örneklem büyüklüğü eşit yüzelikli eşitleme puan farkı grafiği



Şekil 4.6 500 kişilik örneklem büyüklüğü eşit yüzelikli eşitleme puan farkı grafiği



Şekil 4.7 1500 kişilik örneklem büyüklüğü eşit yüzdeliği eşitleme puan farkı grafiği



Şekil 4.8 3000 kişilik örneklem büyüklüğü eşit yüzdeliği eşitleme puan farkı grafiği

Eşit yüzdeliği eşitleme, aynı yüzdeliği sırada bulunan puanların eşit olduğunu kabul etmektedir. Örneklem büyüklüğünün küçültülmesi, yüzdeliği dilim aralıklarını daha geniş kılmakta ve dağılımı çarpıklaştırmaktadır. Tablo4.3 incelendiğinde, 100 kişilik örneklem büyüklüğünde 12, 19, 20 ham puanlarının 0 frekansa sahip olduğu

görülmektedir. 0 frekansa sahip puanların doğal olarak eşitlenmiş puanları da bulunamamaktadır.

Tablo4.3 incelendiğinde eşitlenmiş puanların, 0,25 ile 20 arasında değerler aldığı ve ham puan aralığının dışına çıkmadığı görülmektedir. Kolen ve Brennan (2004), doğrusal eşitlemede görülen eşitlenmiş puanların ham puan aralığının dışında değerler alma sorununun eşit yüzdelikli eşitlemede oluşmadığını belirtmiştir. Elde edilen bulgular da bu sonucu destekler niteliktedir.

Kolen ve Brennan (2004), eşit yüzdelikli eşitlemede, formlar arası güçlük farklılıklarının bir eğri kullanarak açıklandığını ifade etmişlerdir. Üst ve alt seviye puanlarda X formu, Y formundan daha zor olabilirken, orta seviye puanlarda Y formu, X formundan daha zor olabilmektedir. 100 kişilik örneklem büyüklüğünde 13-15 puan aralığında eşitlenmiş puanların ham puanlardan daha düşük olduğu, diğer puanlarda ham puanların eşitlenmiş puanlardan daha yüksek olduğu görülmektedir. 500 ve 3000 kişilik örneklem büyüklüklerinde, 0-16 puan aralığında eşitlenmiş puanların ham puanlardan yüksek olduğu ancak 16-19 puan aralığında ham puanların eşitlenmiş puanlardan daha yüksek olduğu görülmektedir. 1500 kişilik örneklem büyüklüğünde ise 0-14 puan aralığında eşitlenmiş puanların ham puanlardan yüksek olduğu, 15-19 puan aralığında ham puanların eşitlenmiş puanlardan daha yüksek olduğu görülmektedir. Puan farkları grafikleri de değişkenliği göstermektedir. Bu bulgulara dayalı olarak, Kolen ve Brennan (2004) ifade ettiği gibi, güçlük düzeylerinin puan dağılımı boyunca farklılık gösterdiği söylenebilir.

4.4. Alt Problem 4'e İlişkin Bulgular Ve Yorumlar

Alt Problem 4: Örneklem büyüklüğünün eşit yüzdelikli eşitleme yöntemi üzerindeki etkisi nasıldır?

Bu alt problemin çözümü için, üçüncü alt probleme ait eşit yüzdelikli eşitleme sonuçlarının, ağırlıklandırılmış hata kareleri ortalamaları ölçütüne göre karşılaştırılması yapılmıştır. Her örneklem büyüklüğü için hesaplanan AHKO değerleri, Tablo4.4'de verilmiştir.

Tablo4.4 Farklı örneklem büyüklüklerinde alt testlerin eşit yüzdellikli eşitleme yöntemi ile eşitlenmesine ait AHKO değerleri

Eşitleme Yöntemi	Örneklem Büyüklüğü	AHKO
Eşit Yüzdellikli Eşitleme	100	0,052
	500	0,059
	1500	0,044
	3000	0,019

AHKO değerleri için standart hata bulunamadığı için, AHKO değerleri arasında istatistiksel açıdan bir farkın manidarlığının ne olduğuna ilişkin bir değerlendirme yapmak zordur. Ancak, pratikte 0,05'lik veya daha fazla olan bir fark anlamlı olarak değerlendirilmektedir(Skaggs ve Lissitz, 1986, s.311). Tablo4.4 incelendiğinde değerler arasında 0,05'lik bir fark gözükmemektedir. AHKO formülü incelendiğinde eşitlenmiş puanlar ile ham puanlar arasındaki farkın, hata değeri olarak nitelendirildiği gözükmemektedir. Bu hata değerlerinin, ağırlıklandırılmış ortalamalarını ifade eden AHKO değerlerinin küçük olması, eşitlemedeki random hataların az olduğu; bir başka ifade ile eşitlenmiş puanlar ile ham puanlar arasındaki farkın az olduğu anlamına gelmektedir. 100 kişilik örneklem büyüklüğüne ait AHKO değerinin 500 kişilik örneklem büyüklüğüne ait AHKO değerinden düşük çıkması, ham puanlar ile eşitlenmiş puanlar arasındaki farkın, gerçekte azlığından ziyade, düşük örneklem büyüklüklerinde eşit yüzdellikli eşitlemenin sınırlılıklarından dağılımın çarpıklaşması probleminin yaşanmasından kaynaklandığı söylenebilir. Nitekim 100 kişilik örneklem büyüklüğünde bazı ham puanların 0 frekansa sahip olduğu gözlenmiştir. 3000, 1500, 500 kişilik örneklem büyüklüklerinde ise, örneklem büyüklüğünün artmasının, AHKO değerlerini düşürdüğü görülmektedir. Özellikle 500 kişilik örneklem büyüklüğünden 1500 kişilik örneklem büyüklüğüne geçişte AHKO değeri ciddi oranda azalmaktadır. 3000 kişilik örneklem büyüklüğünde ise AHKO değerinin oldukça düştüğü görülmektedir.

Kolen ve Brennan (2004) da, örneklem büyüklüğünün eşitleme hatası üzerinde direkt etkisi olduğunu belirtmişlerdir. Harris (1993), örneklem büyüklüğü eşitlemeye etkisini gösteren bir çalışmada, geniş örneklem büyüklüklerinin daha iyi eşitleme sonuçlarını doğurduğunu ifade etmiştir. Rapp (1993), eşit yüzdellikli eşitleme yöntemi için örneklem büyüklüğünün olabildiğince büyük tutulması

gerektiğini ifade etmiştir. Alan yazında, eşit yüzdellikli eşitleme için önerilen örneklem sayısı 1500 ve üstüdür (Felan, 2002, s.12; Kolen ve Brennan, 2004, s.288; Zhu,1998, s.17). Yukarıdaki bulgular, alan yazında, eşit yüzdellikli eşitleme için önerilen örneklem sayısını doğrulamaktadır.

Bu bulgular, Gök (2012), Kilmen (2010), Çetin (2009), Lee ve Ban (2010), Kim ve Cohen (2002),Tsai (1997) tarafından yapılan çalışmaların bulguları ile tutarlılık göstermektedir.

4.5. Alt problem 5'e ilişkin bulgular ve yorumlar

Alt Problem 5: Farklı örneklem büyüklüğü koşulu altında 2011 yılı SBS 8'inci sınıf matematik alt testi puanlarının, 2012 yılı SBS 8'inci sınıf matematik alt testi puanlarına eşitlenmesi için, doğrusal eşitleme ve eşit yüzdellikli eşitleme yöntemlerinden hangisi daha uygun eşitleme yöntemidir?

Bu alt problemin çözümü için, ikinci ve dördüncü alt problemlerde her iki eşitleme yöntemi için bulunan AHKO değerlerinde faydalanılmıştır. Tüm örneklem büyüklüklerinde, hangi yöntemin daha uygun olduğuna karar verebilmek için, her iki yöntemle ait ilgili örneklem büyüklüklerinde bulunan ve aşağıda Tablo4.5'de verilen AHKO değerleri karşılaştırılmıştır.

Tablo4.5 Farklı örneklem büyüklüklerinde alt testlerin eşit yüzdellikli ve doğrusal eşitleme yöntemleri ile eşitlenmesine ait AHKO değerleri

Örneklem Büyüklüğü	Eşitleme Yöntemi	AHKO
100	Doğrusal Eşitleme	0,098
	Eşit Yüzdellikli Eşitleme	0,052
500	Doğrusal Eşitleme	0,069
	Eşit Yüzdellikli Eşitleme	0,059
1500	Doğrusal Eşitleme	0,066
	Eşit Yüzdellikli Eşitleme	0,044
3000	Doğrusal Eşitleme	0,047
	Eşit Yüzdellikli Eşitleme	0,019

Tablo4.5 incelendiğinde, tüm örneklem büyüklüklerinde, eşit yüzdellikli eşitleme yönteminin, doğrusal eşitleme yönteminden daha düşük AHKO değerleri aldığı görülmektedir. Bu bulgulara dayalı olarak, tüm örneklem büyüklüklerinde, 2011 yılı

SBS 8'inci sınıf matematik alt testi puanlarının, 2012 yılı SBS 8'inci sınıf matematik alt testi puanlarına eşitlenmesi için eşit yüzdelikli eşitleme yönteminin daha uygun olduğu söylenebilir. Bu bulgu, Tsai(1997), Bozdağ(2007), Öztürk(2010) ve Mutluer(2013) tarafından yapılan çalışmaların bulguları ile tutarlılık göstermektedir.

5. SONUÇ VE ÖNERİLER

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

5.1. Sonuçlar

Araştırmadan elde edilen bulgulara dayalı olarak, ulaşılan sonuçlar aşağıda verilmiştir.

1. Farklı örneklem büyüklüklerinde (3000, 1500, 500, 100) 2011 yılı SBS 8'inci sınıf matematik alt testinin 2012 yılı SBS 8'inci sınıf matematik alt testine, Klasik Test Teorisine dayalı eşitleme yöntemlerinden doğrusal eşitleme yöntemi ile eşitlenmesi sonucu, eşitlenmiş puanların, ham puan sınırlarını pozitif yönde aştığı görülmüştür. Tüm örneklem büyüklüklerinde, eşitlenmiş puanlar ile ham puanlar arasındaki farkın, ham puanlar yükseldikçe, azaldığı tespit edilmiştir. Tüm örneklem büyüklüklerinde, testler arası güçlük düzeylerinin, puan ölçeği boyunca değişkenlik gösterdiği görülmüştür.
2. Farklı örneklem büyüklüklerinde, 2011 yılı SBS 8'inci sınıf matematik alt testinin, 2012 yılı SBS 8'inci sınıf matematik alt testine, Klasik Test Teorisine dayalı eşitleme yöntemlerinden doğrusal eşitleme yöntemi ile eşitlenmesi sonucu hesaplanan AHKO değerlerinin, örneklem büyüklüğü arttıkça, küçüldüğü tespit edilmiştir. Örneklem büyüklüğünün artmasının, doğrusal eşitleme yönteminin doğruluğunu arttırdığı sonucuna ulaşılmıştır. Ayrıca doğrusal eşitleme yöntemi için alan yazında yer bulan 400 ve daha büyük örneklem büyüklüğünün kullanılması önerisi, bu çalışmanın bulguları ile tutarlılık göstermiştir.
3. Farklı örneklem büyüklüklerinde, 2011 yılı SBS 8'inci sınıf matematik alt testinin, 2012 yılı SBS 8'inci sınıf matematik alt testine, Klasik Test Teorisine dayalı eşitleme yöntemlerinden eşit yüzdelli eşitleme yöntemi ile eşitlenmesi sonucu, 500, 1500 ve 3000 kişilik örneklem büyüklüklerinde örneklem büyüklüğü arttıkça, AHKO değerlerinin azaldığı görülmüştür. 100 kişilik örneklem büyüklüğündeki AHKO değerinin 500 kişilik örneklem büyüklüğündeki AHKO değerinden küçük olmasının, 100 kişilik örneklem

büyükliğünde puan dağılımının çarpıklaşmasından kaynaklandığı sonucuna ulaşılmıştır. Tüm örneklem büyüklüklerinde testler arası güçlük düzeylerinin puan ölçeği boyunca değişkenlik gösterdiği görülmüştür. Alan yazında eşit yüzdelikli eşitleme yöntemi için 1500 ve üzeri örneklem büyüklüğünün kullanılması önerisi, bu çalışmanın bulguları ile tutarlılık göstermiştir.

4. 2011 yılı SBS 8'inci Sınıf Matematik Alt Testinin 2012 yılı SBS 8'inci sınıf matematik alt testine eşitlenmesinde, tüm örneklem büyüklüklerinde eşit yüzdelikli eşitleme yönteminin, doğrusal eşitleme yöntemine nazaran daha uygun olduğu sonucuna ulaşılmıştır.
5. Bu araştırma, eşitleme çalışmalarında kullanılacak yöntemin belirlenmesinde, örneklem büyüklüğünün mutlaka göz önünde bulundurulması gerekliliğini göstermektedir.

5.2 Öneriler

1. Araştırmada, SBS verileri kullanılmıştır. Benzer çalışmalar, ülkemizde farklı dönemlerde, birbirlerinin yerine geçen formlar halinde uygulanan YDS, ALES, vb. sınav uygulamaları için yapılarak, sınavı hazırlayan kurumların ölçme ve değerlendirme süreçlerini gözden geçirmeleri sağlanabilir.
2. Bu araştırmada, 100, 500, 1500 ve 3000 kişilik örneklem büyüklükleri kullanılmıştır. Benzer çalışmalar farklı örneklem büyüklükleri için yapılabilir.
3. Bu araştırmada, doğrusal eşitleme ve eşit yüzdelikli eşitleme yöntemleri kullanılmıştır. Farklı eşitleme yöntemleri kullanılarak benzer çalışmalar yapılabilir.
4. Araştırmada, eşdeğer gruplar deseni kullanılmıştır. Farklı eşitleme desenleri kullanılarak, benzer çalışmalar yapılabilir.
5. Bu araştırmada, örneklem büyüklüğünün test eşitlemeye etkisi incelenmiştir. Eşitlenecek testlerdeki madde sayılarının, eşitleme üzerindeki etkisi incelenebilir.

KAYNAKÇA

- Aiken, Lewis.R.(2000). Psychological Testing and Assessment. Allyn and Bacon.
- Akhun, İ. (1984). İki korelasyon katsayısı arasındaki farkın manidarlığının test edilmesi. *Ankara Üniversitesi Eğitim Fakültesi Dergisi*. 17, 1–7. PegemA Yayıncılık.
- Angoff, W. H. (1971). Scales, norms and equivalent scores. In R. L. Thorndike (Ed.) Educational measurement (pp.508–600). Washington, DC: American Council on Education.
- Angoff, W.H.(1984). Scales, norms and equivalent scores. New Jersey: Educational Testing Service.
- Anıl, D.(2002). Deneme Uygulamasının Yapılamadığı Durumlarda Madde ve Test Parametrelerinin Klasik ve Örtük Özellikler Test Teorisine Göre Kestirilmesi. Yayınlanmış doktora tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Anıl, D. ve Öztürk, N.(2012). Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı Puanlarının Eşitlenmesi Üzerine Bir Çalışma. *Eğitim ve Bilim*, 37, 181-193.
- Baykul, Y.(2010). Eğitimde ve Psikolojide Ölçme. Ankara: Pegem Yayınları.
- Braun, H. I & Holland, P. W. (1982) Observed- score test equating: A mathematical analysis of some ETS equating procedures. In P. W. Holland & D. B. Rubin (Eds.), Test equating (pp. 9-49). New York: Academic Press.
- Bozdağ, S.(2007). Şans başarısının test eşitlemeye etkisi. Yayınlanmamış yüksek lisans tezi, Mersin Üniversitesi, Sosyal Bilimler Enstitüsü. Mersin.
- Büyüköztürk, Ş. (2012). Sosyal bilimler için veri analizi el kitabı. Ankara: Pegem Yayınları.
- Cook, L.L., & Petersen, N. S. (1987). Problems related to the use of conventional and item response theory equating methods in less than optimal circumstances. *Applied Psychological Measurement*, 11, 225-244.
- Crocker, L. & Algina, J. (1986). Introduction to classical & modern test theory. New York: Harcourt Brace Jovanovich College Publishers.
- Çetin, E. (2009). Dikey ölçekte klasik test ve madde tepki kuramına dayalı yöntemlerin karşılaştırılması. Yayınlanmamış doktora tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Dorans, Neil J. (1990). "Equating Methods and Sampling Designs." *Applied Measurement in Education*, 3(1), 3-17.

- Dorans, J. N. & Holland, P. W. (2000). Population invariance and the equitability of tests: Basic theory and the linear case. *Journal of Educational Measurement*, 37, 281-306.
- Erkuş, A.(2011). Bilimsel Araştırma Süreci. Ankara: Seçkin Yayıncılık.
- Felan, G.(2002). Test equating: Mean, linear, equipercentile, and item response theory. Paper presented at the annual meeting of the Southwest Educational Research Association, Austin.
- Gorsuch, R. L. (1983). *Factor Analysis*. Philadelphia: Saunders.
- 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.
- Hambleton, R. K. & Swaminathan, H. (1985). Item response theory: Principles and applications. Baston: Kuluwer-Nijhoff Publisihing.
- Harris, D. J. & Crouse, J. D. (1993). A study of criteria used in equating. *Applied Measurement in Education*, 6 (3),195-240.
- Heh, V. K.(2007). *Equating accuracy using small samples in the random groups design. Doctoral Dissertation*, Ohio University.
- <http://www.osym.gov.tr/belge/1-19350/baskanligimiz-tarafindan-yapilan-sinavlarda-sorulan-sor-.html> Erişim Tarihi:13Ekim 2013
- Holland, P. W. & Dorans, N. J. (2006). Linking and equating. In R. L. Brennan (Ed.), *Educational measurement* (pp. 187–220). Westport, CT: Praeger Publishers.
- Holland, P.W., Dorans, N. J. & Petersen, N. S.(2007). Equating test scores. In Rao, C.R., Sinharay, S. (Eds.), *Handbook of statistics: Psychometrics* (Vol. 26, pp. 169-197). Amsterdam: Elsevier B. V.
- Holland, P. W. (2007). A framework and history for score linking. N.J. Dorans, M.Pommerich & P.W. Holland. (Ed.) *Linking and aligning scores and scales* (pp. 5–29). USA: Springer Scince&Business Media.
- Hovardaoğlu, Selim (1995). *Doğrusal Regresyon ve Korelasyona Giriş*, Ankara: Habitoğlu Basım ve Yayın.
- Jaeger, R. M. (1981). Some explatory indices for selection of a test equating method. *Journal of Educational Measurement*, 18(1), 23-38.
- Karasar, N. (2010). Bilimsel araştırma yöntemi. Ankara: Nobel Yayınları.

- Kilmen S. (2010). Madde Tepki Kuramı'na dayalı test eşitleme yöntemlerinden kestirilen eşitleme hatalarının örneklem büyüklüğü ve yetenek dağılımına göre karşılaştırılması. Yayımlanmış Doktora Tezi, Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Kolen, M. J. (1984). Effectiveness of analytic smoothing in equipercntile equating. *Journal of Educational Statistics*, 9(1), 24-44.
- Kolen, M. J. (1988). An NCME instructional module on traditional equating methodology. *Educational Measurement: Issues and Practice*, 7, 29-36.
- Kolen, M. J. ve Brennan, R. L. (1995). *Test equating methods and practices*. New York: Springer.
- Kolen, M. J. & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices* (2nd. ed.). New York: Springer.
- Kolen, M. J. (2004). Population invariance in equating and linking: Concept and history. *Journal of Educational Measurement*, 41(1), 3-14.
- Livingston, S. A. (2004). Equating test scores (Without IRT). Educational Testing Service.
- Lord, F. M. (1980). *Applications of Item Response Theory To Practical Testing Problems*. Hillsdale, NJ: Erlbaum.
- Öztürk, N. (2010). Akademik Personel Ve Lisansüstü Eğitimi Giriş Sınavı puanlarının eşitlenmesi üzerine bir çalışma. Yayımlanmış Yüksek Lisans Tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü. Ankara.
- Petersen, N. S. (2008). A discussion of population invariance of equating. *Applied Psychological Measurement*, 32, 98-101.
- Rapp, J. (1999). *Linear and equipercntile methods for equating PET*. National Institute for Testing and Evaluation.
- Skagg, Garry & Lissitz Robert W. (1986). "An Exploration of the Robustness of Four Test Equating Models" *Applied Psychological Measurement*. 10, 303-317.
- Tsai, T. H. (1997). Estimating minumum sample sizes in random groups equating. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago.
- Turgut, F. (1977). *Eğitimde Ölçme ve Değerlendirme Metodları*, Ankara.
- Turgut, F. (1979). "Ölçekleme, Norm Hazırlama, Normları Eşitleme ve Ayarlama". *Test Geliştirme Teknikleri Basılmamış Ders Notları*, Ankara.

Woldbeck, T. (1998). Basic concepts in modern methods of test equating. Paper presented at the annual meeting of the Southwest Psychological Association, New Orleans.

Zeng, L. (1991). Standard errors of linear Equating fort the single- group design (Report No: 91-4). ACT Research Report Series.

Zhu, W. (1998). Test equating: What, why and how?. *Research Quarterly for Exercises and Sport*, 69(1), 11-23.

EKLER DİZİNİ

EK 1: Hacettepe Üniversitesinin MEB' İliğına veri talebi yazısı



T.C.
HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

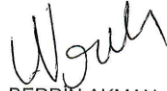
Sayı: 51944218-010.99/271
Konu: Yunus İNCİ'nin veri isteđi

12/02/2014

MİLLİ EĞİTİM BAKANLIđI
Yenilik ve Eğitim Teknolojileri Genel Müdürlüğü
Ölçme, Deđerlendirme ve Yerleřtirme Gurup Başkanlıđına

Enstitümüz Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Deđerlendirme Bilim Dalı yüksek lisans programı öğrencilerinden Yunus İNCİ'nin "Örneklem Büyüklüğünü Test Eşitlemeye Etkisi Üzerine Bir Çalışma" konulu tez çalışmasında kullanılmak üzere veri isteđine ilişkin dilekçesi ve ekleri ilişikte gönderilmiştir.

Bilgilerinizi ve geređini saygılarımla rica ederim.


Prof.Dr. BERRİN AKMAN
Enstitü Müdürü 4.

EKLER :
1 Tez Öneri Formu
2 Öğrenci Belgesi
3 Taahhütname
4 Örnek Veri Deseni

Enstitü Sekreter V. : O. ERUYSAI (Paraf)

Hacettepe
Üniversitesi Eğitim Bilimleri Enstitüsü



2924150857

EK2: MEB' İğının veri talebini uygun bulmasına ilişkin yazısı



T.C.
MİLLÎ EĞİTİM BAKANLIĞI
Yenilik ve Eğitim Teknolojileri Genel Müdürlüğü

Sayı : 39339088/42/1032074
Konu : İstatistiki Bilgiler

10/03/2014

HACETTEPE ÜNİVERSİTESİ REKTÖRLÜĞÜNE
Eğitim Bilimleri Enstitüsü

İlgi : 12.02.2014 tarih ve 51944218-010.99/271 sayılı yazınız.

Üniversiteniz Eğitim Bilimleri Enstitüsü, Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı Yüksek Lisans öğrencisi Yunus İNCİ'nin "Örneklem Büyüklüğünü Test Eşitlemeye Etkisi Üzerine Bir Çalışma" konulu tez çalışmalarında kullanmak üzere, ilgi yazınızla talep edilen istatistiki bilgilerden verilmesi uygun görülen veriler hazırlanmıştır.

2011 ve 2012 yıllarında Seviye Belirleme Sınavı sınav soruları ve cevap anahtarı Milli Eğitim Bakanlığı İnternet sitesinde yayımlanmaktadır.

İstenilen istatistiki bilgiler; adı geçen öğrencinin tez çalışması haricinde kullanılmaması ve başka amaçlarla paylaşılmaması şartıyla CD ortamında yazımız ekinde gönderilmektedir.

Bilgilerinizi rica ederim

Mustafa KOÇ
Bakan a.
Genel Müdür

EKLER

1- İstatistiki Bilgileri CD'si (2 adet)



Güvenli Elektronik İmza
Aslı ile Aynıdır
11.03.2014

Bu belge, 5070 sayılı Elektronik İmza Kanununun 5 inci maddesi gereğince güvenli elektronik imza ile imzalanmıştır. Evrak teyidi <http://evraksorgu.meb.gov.tr> adresinden ac13-cc02-39cd-90fa-227c kodu ile yapılabilir.

Atatürk Biv. 06648 Kızılay/ANKARA
Elektronik Ağ : www.meb.gov.tr
e-posta : etg_ds_itiraz@meb.gov.tr

Ayrıntılı bilgi için : M. Erdal YALÇIN Birim Sorumlusu
Tel : (0 312) 497 40 84
Faks : (0 312) 223 87 36

EK3: Etik Kurulu İzin Muafiyet Formu

	HACETTEPE ÜNİVERSİTESİ EĞİTİM BİLİMLERİ ENSTİTÜSÜ TEZ ÇALIŞMASI ETİK KURUL İZİN MUAFİYETİ FORMU	Form: 40
HACETTEPE ÜNİVERSİTESİ EĞİTİM BİLİMLERİ ENSTİTÜSÜ EĞİTİM BİLİMLERİ ANABİLİM DALI BAŞKANLIĞINA		Tarih: .../.../.....
Tez Başlığı / Konusu: Örneklem Büyüklüğünün Test Eşitlemeye Etkisi		
Yukarıda başlığı/konusu gösterilen tez çalışmam:		
<ol style="list-style-type: none">1. İnsan ve hayvan üzerinde deney niteliği taşımamaktadır,2. Biyolojik materyal (kan, idrar vb. biyolojik sıvılar ve numuneler) kullanılmasını gerektirmemektedir.3. Beden bütünlüğüne müdahale içermemektedir.4. Gözlemsel ve betimsel araştırma (anket, ölçek/skala çalışmaları, dosya taramaları, veri kaynakları taraması, sistem-model geliştirme çalışmaları) niteliğinde değildir.		
Hacettepe Üniversitesi Etik Kurulları ve Komisyonlarının Yönergelerini inceledim ve bunlara göre tez çalışmamın yürütülebilmesi için herhangi bir Etik Kuruldan izin alınmasına gerek olmadığını; aksi durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.		
Gereğini saygılarımla arz ederim.		
		Tarih ve İmza
Adı Soyadı: Yunus İNCİ	_____	
Öğrenci No: N11222482	_____	
Anabilim Dalı: Eğitim Bilimler Anabilim Dalı	_____	
Programı: Eğitimde Ölçme ve Değerlendirme Bilim Dalı	_____	
Statüsü: <input type="checkbox"/> Y.Lisans <input type="checkbox"/> Doktora <input type="checkbox"/> Bütünleşik Dr.	_____	
<u>DANIŞMAN GÖRÜŞÜ VE ONAYI</u>		
Tez çalışmasında Milli Eğitim Bakanlığının onayı ile MEB veri tabanından yararlanılmıştır. Tez çalışması hiçbir kişisel bilgi niteliğinde veri içermemektedir. Hacettepe Üniversitesi Etik Kurulları ve Komisyonlarının Yönergelerine göre herhangi bir Etik Kuruldan izin alınmasına gerek yoktur.		
_____ Doç. Dr. Duygu ANIL		

ÖZGEÇMİŞ

Adı Soyadı	Yunus İNCİ
Doğum Yeri	İZMİR
Doğum Yılı	01.03,1981

Eğitim ve Akademik Durumu

Lise	Maltepe Askeri Lisesi	1995-1999
Lisans	Kara Harp Okulu	1999-2003
Yabancı Dil	İngilizce (Çok İyi),Boşnakça(Temel),Arapça(Temel)	
İş Deneyimi	Türk Silahlı Kuvvetleri	2003-