YAZMA BECERİLERİNİN DEĞERLENDİRMESİNDE ÖZ, AKRAN VE ÖĞRETMEN PUANLARININ ÇOK YÜZEYLİ RASCH ÖLÇME MODELİYLE İNCELENMESİ

RANA SALMANER

YÜKSEK LİSANS TEZİ

EĞİTİMDE ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI

GAZİ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

EYLÜL, 2015

TELİF HAKKI ve TEZ FOTOKOPİ İZİN FORMU

Bu tezin tüm hakları saklıdır. Kaynak göstermek koşuluyla tezin teslim tarihinden itibaren .....(….) ay sonra tezden fotokopi çekilebilir.

YAZARIN

Adı : Rana

Soyadı : SALMANER

Bölümü : Eğitimde Ölçme ve Değerlendirme

İmza :

Teslim tarihi :

TEZİN

Türkçe Adı: Yazma Becerilerinin Değerlendirmesinde Öz, Akran ve Öğretmen Puanlarının Çok Yüzeyli Rasch Ölçme Modeliyle İncelenmesi.

İngilizce Adı: An Investigation of Self, Peer and Teacher Scores within the Context of Assessment of Writing Skills with Many Facets Rasch Measurement Model.

ETİK İLKELERE UYGUNLUK BEYANI

Tez yazma sürecinde bilimsel ve etik ilkelere uyduğumu, yararlandığım tüm kaynakları kaynak gösterme ilkelerine uygun olarak kaynakçada belirttiğimi ve bu bölümler dışındaki tüm ifadelerin şahsıma ait olduğunu beyan ederim.

Yazar Adı Soyadı: Rana SALMANER

İmza: ……..

Jüri onay sayfası

Rana SALMANER tarafından hazırlanan "Yazma Becerilerinin Değerlendirmesinde Öz, Akran ve Öğretmen Puanlarının Çok Yüzeyli Rasch Ölçme Modeliyle İncelenmesi" adlı tez çalışması aşağıdaki jüri tarafından oy birliği / oy çokluğu ile Gazi Üniversitesi Eğitimde Ölçme ve Değerlendirme Anabilim Dalı'nda Yüksek Lisans tezi Yüksek Lisans tezi olarak kabul edilmiştir.

Danışman: Doç. Dr. İsmail KARAKAYA

(Eğitimde Ölçme ve Değerlendirme, Gazi Üniversitesi)

Başkan: Yrd. Doç. Ömer KUTLU

(Ölçme ve Değerlendirme, Ankara Üniversitesi)

Üye: Prof. Dr. Mehtap ÇAKAN

(Eğitimde Ölçme ve Değerlendirme, Gazi Üniversitesi)

Tez Savunma Tarihi: 17/09/2015

Bu tezin Eğitimde Ölçme ve Değerlendirme Anabilim Dalı'nda Yüksek Lisans Tezi olması için şartları yerine getirdiğini onaylıyorum.

Prof. Dr. Tahir ATICI

Eğitim Bilimleri Enstitüsü Müdürü

TEŞEKKÜR

Çalışmam boyunca değerli yardımlarıyla ve katkılarıyla beni yönlendiren, faydalı önerileriyle ve yorumlarıyla bu çalışmanın tamamlanmasında katkısı olan danışmanım Doç. Dr. İsmail KARAKAYA'ya teşekkürlerimi sunarım.

Görüşleriyle ve eleştirileriyle çalışmama değerli katkılarda bulunan jüri üyelerim Prof. Dr. Mehtap ÇAKAN ve Yrd. Doç. Ömer KUTLU'ya şükranlarımı sunarım.

Tez yazım sürecinde çeşitli şekillerde desteğini gördüğüm Gazi Üniversitesi Eğitimde Ölçme ve Değerlendirme Anabilim Dalı'ndaki saygı değer hocalarıma, araştırma görevlisi arkadaşlarıma ve özellikle bana sabırla katlanan ve hep yanımda olan canım arkadaşımlarım Arş. Gör. Halime YILDIRIM, Arş. Gör. Menekşe UYSAL ve Arş. Gör. Sibel ADA'ya teşekkür ederim.

Sevgisiyle ve desteğiyle, çalışmamın ortaya çıkmasındaki sabırlarıyla ve yardımlarıyla hiçbir zaman beni yalnız bırakmayan canım aileme çok teşekkür ederim.

YAZMA BECERİLERİNİN DEĞERLENDİRMESİNDE ÖZ, AKRAN VE ÖĞRETMEN PUANLARININ ÇOK YÜZEYLİ RASCH ÖLÇME MODELİYLE İNCELENMESİ

(Yüksek Lisans Tezi)

Rana Salmaner

GAZİ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eylül, 2015

ÖZ

Bu çalışmanın amacı, 2014-2015 eğitim öğretim yılı ikinci döneminde beşinci sınıflarla yapılan yazma becerileri ile ilgili çalışmalarda öz, akran ve öğretmen puanlarını incelemektir. Çalışmada çok yüzeyli Rasch ölçme modeli kullanılmıştır. Birinci analiz sonuçlarına göre yazma konusunda en başarılı öğrencinin 48 numaralı öğrenci iken, en başarısız öğrenci 59 numaralı öğrencidir. Öğrencilerinin çoğunun orta başarı düzeyinde toplandığı ve uyumsuz öğrencilerin genelde yüksek başarı düzeyine sahip olduğu söylenebilir. En cömert puanlanan kategori "yazım", en katı puanlanan kategori ise "bütünsellik"tir. Öz değerlendirmelerin en cömert ve öğretmen değerlendirmelerin ise en katı değerlendirmeler olduğu görülmektedir. İkinci analiz sonuçlarına göre en başarılı öğrencinin 48 numaralı öğrenci olduğu ve bu konuda en başarısızın ise 19 numaralı öğrenci olduğu görülmektedir. Uyumsuz öğrenci olarak 11 numaralı öğrenci belirlenmiştir. En cömert puanlanan kategori "yazım" ve "bütünsellik" iken, en katı puanlanan "çıkarımda" bulunma kategorisidir Öz değerlendirmeler en cömertken, öğretmen ve akran değerlendirmelerin yaklaşık eşit düzeyde ve daha katı olduğu görülmektedir. Üçüncü analize göre en başarılı öğrencinin 4 numaralı öğrenci olduğu ve yine bu konuda en başarısızın ise 19 numaralı öğrenci olduğu görülmektedir. Uyumsuz öğrenci olarak 54 ve 22 numaralı öğrenciler belirlenmiştir. En cömert puanlanan kategorinin "çıkarımda bulunma" ve en katı puanlananın "görsellerden faydalanma"dır. Son olarak yapılan bütün verilerin alındığı analizde ise en başarılı öğrencinin 48 numaralı öğrenci olduğu ve bu konuda en başarısızın ise 19 numaralı öğrenci olduğu görülmektedir. Uyumsuz öğrenci belirlenememiştir. En cömert puanlanan kategoriler "çıkarımda bulunma", "düzen" ve "yazım" iken en katı puanlanan kategoriler "görsellerden faydalanma" ve "anlatım"dır.

Verilerin ayrı ayrı girildiği bütün analiz ve ayrı ayrı analizlerde birinci uygulamadaki öz puanlama ile ikinci uygulamadaki öz puanlamanın aynı katılık düzeyinde olduğu, fakat üçüncü uygulamadaki öz puanlamanın daha katı olduğu görülmektedir. Akranlarda birinci puanlama en cömertken, ikinci puanlama en katı ve üçüncü ise orta katılıktadır. Öğretmenlerde en katının birinci uygulama olduğu daha sonra ikinci uygulamanın geldiği ve üçüncü uygulamanın orta katılıkta olduğu görülmüştür. Ayrı ayrı yapılan analizlerle ayrı ayrı girilen verilerle yapılan bütün analiz sonuçları arasında farklılaşma olmadığı fakat ayrı ayrı verilerle yapılan bütün analizde durum hakkında daha kesin ve ayrıntılı bilgiler elde edilebileceği ve bu doğrultuda daha kesin ve doğru yorumlar yapılabileceği sonucuna ulaşılmıştır. Ayrıca puanlayıcı türlerinin sürecin sonuna doğru orta katılığa daha fazla yaklaştığı ve doğal olarak birbirlerine de yaklaştıkları gözlemlenmiştir.

Bilim kodu: 6.015

Anahtar Kelimeler: Öz değerlendirme, Akran değerlendirme, Yazma becerilerinin değerlendirilmesi, Çok yüzeyli Rasch ölçme modeli, Puanlayıcı güvenirliği.

Sayfa Adedi: xiv+108

Danışman: Doç. Dr. İsmail KARAKAYA

INVESTIGATION OF SELF, PEER AND TEACHER SCORES WITHIN THE CONTEXT OF ASSESSMENT OF WRITING SKILLS WITH MANY FACETS RASCH MODEL

(M.S Thesis)

Rana Salmaner

GAZI UNİVERSITY

GRADUATE SCHOOL OF EDUCATIONAL SCIENCES

September, 2015

ABSTRACT

The purpose of this study is to investigate self, peer and teacher scores during the writing skill assessment process of two fifth classes in the second semester of the year 2015. Many Facets Rasch Measurement Model was used in the study. According to the first analysis' results, the most successful student was 48th one, and the least successful one was 59th. The 11st student was determined as misfit element. While the most lenient scored category was "spelling", the severest scored one was "word choice". The results indicated that many self-raters assessed their own writing higher than predicted by the Rasch model. Self-raters were the most lenient raters and teachers were the severest. For the results of second analysis, the most successful student was 48th one, and the least one was, 19th. It could be said that most of the students were generally in the middle of success level. In addition, misfit students were at the top of success level. While the most lenient scored categories are "spelling" and "word choice", the severest scored one is "inference". The results indicated that many self-raters assessed their own writing higher than predicted by the Rasch model. Self-raters were the most lenient raters. Peers and teachers were in equal severity level and severer than self-raters. For the third one's results, the most successful student was 4th one, and the least successful one was 19th. The 54st and 22nd students were determined as misfit elements. While the most lenient scored category was "inference", the severest scored one was "using visuals". In the last and general analysis, the most successful student was 48th one, and the least one was 19th. No student was determined as misfit element. While the most lenient scored categories were "inference", "order" and "spelling", the severest scored ones were "using visuals" and "narration". In the general

analysis which the data was separately coded and previous separate analysis, it could be seen that the first and second self-rating was at the same level severity, but the third one was severer than them. While the first rating was the lenient for peers, the second one was the severest and the third one was in the middle of severity level. Looking teacher, it could be seen that the severest one was the first application, and the most lenient one was the third. Between the result of previous and separate analysis and the analysis, separately coded, there was no visual difference, but for the analysis, separately coded, it could provide both more detailed and more certain information, and more certain and truer comments about the situations. Furthermore, rater types became in the middle of severity level during the process and naturally they got closer to each other.

Science Code : 6.015

Key Words: Self-assessment, Peer-assessment, Assessment of Writing Skills, Many Facets Rasch Measurement Model, Inter and Intra -Raters Reliability.

Page Number: xiv+108

Supervisor: Assoc. Prof. İsmail KARAKAYA

İÇİNDEKİLER

ÖZ ..... **v**

ABSTRACT ..... **vii**

İÇİNDEKİLER ..... **ix**

TABLOLAR LİSTESİ ..... **xii**

ŞEKİLLER LİSTESİ ..... **xiii**

SİMGELER VE KISALTMALAR LİSTESİ ..... **xiv**

BÖLÜM I **1**

GİRİŞ **1**

1.1. Problem Durumu **1**

1.2. Kuramsal Çerçeve **10**

1.2.1. Puanlayıcı Güvenirliği **10**

1.2.2. Puanlama Güvenirliğinde Kullanılan Yöntemler **10**

1.2.2.1. KTK'na Dayalı olan Yöntemler **10**

1.2.2.2. Genellenebilirlik Kuramı **12**

1.2.2.3. Çok Yüzeyli Rasch Ölçme Modeli **13**

1.3. Araştırmanın Amacı **17**

1.4. Problem Cümlesi **18**

1.4.1. Araştırma Soruları **18**

1.5. Araştırmanın Önemi **18**

1.6. Araştırmanın Sınırlılıkları **20**

1.7. İlgili Araştırmalar **20**

BÖLÜM II **27**

YÖNTEM **27**

2.1 Araştırmanın Modeli **27**

2.2. Çalışma Grubu **27**

2.3 Veri Toplama Araçları **29**

2.3.1. Yazma Görevlerinin Geliştirilme Süreci **29**

2.3.2. Dereceli Puanlama Anahtarlarının Geliştirilme Süreci **30**

2.3.3. Geliştirilen Ölçme Araçlarının Geçerlik Ve Güvenirliğine İlişkin Toplanan Kanıtlar **31**

2.3.4. Veri toplama süreci 33
2.4. Verilerin Analizi 35
2.4.1. Model Veri Uyumu 36
BÖLÜM III 37
BULGULAR 37
3.1. Birinci Araştırma Sorusuna İlişkin Bulgular 37
3.1.1. Alt Problem 1.a'ya İlişkin Bulgular 37
3.1.2. Alt Problem 1.b 'ye İlişkin Bulgular 42
3.1.3. Alt Problem 1.c 'ye İlişkin Bulgular 47
3.1.4. Alt Problem 1.d'ye İlişkin Bulgular 49
3.2. İkinci Araştırma Sorusuna İlişkin Bulgular 51
3.2.1 Alt Problem 2.a'ya İlişkin Bulgular 51
3.2.2. Alt Problem 2.b'ye İlişkin Bulgular 53
3.2.3. Alt Problem 2.c'ye İlişkin Bulgular 55
3.2.4. Alt Problem 2.d 'ye İlişkin Bulgular 57
3.3. Üçüncü Probleme İlişkin Bulgular 58
BÖLÜM IV 61
SONUÇ VE TARTIŞMA 61
4.1. Tartışma 61
4.2. Sonuçlar 64
4.2.1 Birinci Araştırma Sorusuna İlişkin Analiz Sonuçları 64
4.2.2 İkinci Araştırma Sorusuna İlişkin Analiz Sonuçları 65
4.2.3. Üçüncü Araştırma Sorusuna İlişkin Analiz Sonuçları 66
4.2.4. Genel Sonuçlar 67
4.3. Öneriler 68
4.3.1. Araştırma Sonuçlarından Çıkan Öneriler 68
4.3.2. Araştırmacılara Yönelik Öneriler 69
KAYNAKÇA 70
EKLER 78
EK 1: Görevleri ve Dereceli Puanlama Anahtarlarını Değerlendirme Formu 79
EK 2: 1. Hafta Kullanılan Performans Görevi ve Dereceli Puanlama Anahtarı 83
EK 3: 2. Hafta Kullanılan Performans Görevi ve Dereceli Puanlama Anahtarı 86
EK 4: 3. Hafta Kullanılan Performans Görevi ve Dereceli Puanlama Anahtarı 89
EK 5: 4. Hafta Kullanılan Performans Görevi ve Dereceli Puanlama Anahtarı 92
EK 6: 5. Hafta Kullanılan Performans Görevi ve Dereceli Puanlama Anahtarı 95
EK 7: Uygulamalar için Milli Eğitim Bakanlığı'ndan Alınan İzin Belgesi 98
EK 8: Bir Bütün Olarak Alınan Analizin Öğrenci Yazma Beceri Düzeylerini Gösteren Yüzeyin Uyum İçi, Uyum Dışı Değerleri ve Uyum İstatistikleri Tablosu 99

EK 9: Bir Bütün Olarak Alınan Analizin DPA'daki Kategorilerin Katılık ve Cömertlik Düzeylerini Gösteren Yüzeyin Uyum İçi, Uyum Dışı Değerleri ve Uyum İstatistikleri Tablosu.....100
EK 10: Bir Bütün Olarak Alınan Analizin Puanlayıcı Türünün Katılık Ve Cömertlik Düzeylerini Gösteren Yüzeyin Uyum İçi, Uyum Dışı Değerleri ve Uyum İstatistikleri Tablosu.....101
EK 11: Bir Bütün Olarak Alınan Analizin Uygulanan Görevlerin Zorluk Düzeylerini Gösteren Yüzeyin Uyum İçi, Uyum Dışı Değerleri ve Uyum İstatistikleri Tablosu.....102
EK 12: Öğrencilerin Kompozisyonlarından Bir Örnek 103
EK 13: Öz, Akran ve Öğretmen Puanlama Örnekleri 104
EK 14: Öğrencilerin Uygulama ile İlgili Görüş Örnekleri 107

TABLOLAR LİSTESİ

Tablo 1.	Puanlama Güvenirliğinde Kullanılan Yöntemlerin Verinin Ölçek Türüne ve Puanlayıcı Sayısına Göre Sınıflandırılması	11
Tablo 2	Çalışma Grubunun Cinsiyet ve Sınıf Değişkenlerine Göre Dağılımı	28
Tablo 3.	Puanlayıcıların Bölüm ve Cinsiyete Göre Dağılımları	31
Tablo 4.	Uzmanların Alanlarına Göre Dağılımları	33
Tablo 5.	Performans Görevlerinin ve Dereceli Puanlama Anahtarlarının Kapsam Geçerlik İndeksleri	35
Tablo 6.	Birinci Analizdeki Öğrencilerin Beceri Ölçüm Raporu	42
Tablo 7.	Birinci Analizdeki DPA'daki Kategorilerin Puanlanma Katılıkları Ölçüm Raporu	44
Tablo 8.	Birinci Analizdeki Puanlayıcı Türü Katılık Ölçüm Raporu	45
Tablo 9.	İkinci Analizdeki Öğrencilerin Beceri Ölçüm Raporu	47
Tablo 10.	İkinci Analizdeki DPA'daki Kategorilerin Puanlanma Katılıkları Ölçüm Raporu	48
Tablo 11.	İkinci Analizdeki Puanlayıcı Türü Katılık Ölçüm Raporu	49
Tablo 12.	Üçüncü Analizdeki Öğrencilerin Beceri Ölçüm Raporu	49
Tablo 13.	Üçüncü Analizdeki DPA'daki Kategorilerin Puanlanma Katılıkları Ölçüm Raporu	50
Tablo 14.	Üçüncü Analizdeki Puanlayıcı Türü Katılık Ölçüm Raporu	50
Tablo 15.	Bütün Analizin Puanlayıcı Türleri Katılıklarının Ölçüm Raporu	60

ŞEKİLLER LİSTESİ

Şekil 1.	Birinci analizdeki öğrenci becerisi, kategori puanlama katılığı, puanlayıcı türü katılığı ve Likert ölçek işlevinin kalibrasyon haritası	38
Şekil 2.	İkinci analizindeki öğrenci becerisi, kategori puanlama katılığı, puanlayıcı türü katılığı ve Likert ölçek işlevinin kalibrasyon haritası	39
Şekil 3.	Üçüncü analizdeki öğrenci becerisi, kategori puanlama katılığı, puanlayıcı türü katılığı ve Likert ölçek işlevinin kalibrasyon haritası	41
Şekil 4.	Bir bütün olarak alınan analizlerdeki öğrenci becerisi, kategori puanlama katılığı, puanlayıcı türü katılığı ve Likert ölçek işlevinin kalibrasyon haritası	52
Şekil 5.	Uygulamanın sırasıyla bir bütün olarak alınan analiz, birinci, ikinci ve üçüncü analizlerindeki öğrenci yüzeylerinin kalibrasyon haritası	54
Şekil 6.	Uygulamanın sırasıyla bir bütün olarak alınan analiz, birinci, ikinci ve üçüncü analizlerinde dereceli puanlama anahtarındaki kategorilerin puanlanma durumlarının kalibrasyon haritası	56
Şekil 7.	Uygulamanın sırasıyla bir bütün olarak alınan analiz, birinci, ikinci ve üçüncü analizlerindeki puanlayıcı türünün katılığının kalibrasyon haritası	57
Şekil 8.	Uygulamalar boyunca yapılan analizlerdeki puanlama türleri ile ilgili verilerin ayrı kodlanarak yapılmış olduğu analizinin kalibrasyon haritası	59

SİMGELER VE KISALTMALAR LİSTESİ

A	Adlandırma Ölçeği
ANOVA	Varyans Analizi
ÇYRÖM	Çok Yüzeyli Rasch Ölçme Modeli
EA	Eşit Aralıklı Ölçek
ESLPE	İkinci Dil Olarak İngilizce Yerleştirme Sınavı
EO	Eşit Oranlı Ölçek
G	Ayırma Oranı
GK	Genellenebilirlik Kuramı
ICC	Inter-Class Correlation Coefficient
KGİ	Kapsam Geçerlik İndeksi
KGÖ	Kapsam Geçerlik Ölçütü
KTK	Klasik Test Kuramı
MTK	Madde Tepki Kuramı
PAG	Puanlayıcılar Arası Güvenirlik
S	Sıralama Ölçeği
TOEFL	Test of English as a Foreign Language

BÖLÜM I
GİRİŞ

Bu bölümde araştırmanın problem durumuna, amacına, sorularına, önemine ve sınırlılıklarına ilişkin bilgiler verilmiştir. Son olarak da çalışmayla ilgili araştırmalara değinilmiştir.

1.1. Problem Durumu

Öğrenme ve öğretme süreci etkinliklerinin hedeflenen kazanımların elde edilmesini sağlayacak şekilde ilerleyebilmesini ve öğrenen özelliklerini dikkate alarak sürecin biçimlendirilmesini sağlayacak temel öge ölçme ve değerlendirme uygulamalarıdır (Göçer, 2010). Diğer bir anlatımla, sistemdeki eksiklik ve yanlışlıkların ortaya konmasını ve bu yöntemle sistemin varlığını etkili bir şekilde devam ettirebilmesini sağlayan sistemin kontrol mekanizmasıdır. Eğitim sisteminde kontrol mekanizmasını ölçme ve değerlendirme çalışmalarının oluşturduğu düşünüldüğünde, eğitim sisteminin nitelikli bir şekilde kendini sürdürebilmesi, öncelikle nitelikli ölçme ve değerlendirme çalışmalarının yapılmasıyla ilgilidir (Alıcı, 2014). Eğitim sürecinde ölçme ve değerlendirmenin bu önemli rolü dikkate alındığında, bu amaçla kullanılan ölçme araçlarının da eğitimdeki yerinin tartışılmaz olduğu söylenebilir. Okullarda genel olarak hatırlama ve anlama düzeyindeki kazanımlarla ölçme ve değerlendirmeler yapıldığından, geleneksel ölçme araçlarının yaygın olarak kullanıldığı belirtilebilir. Geleneksel ölçme araçlarındaki soru dağılımının bilişsel sınıflamalara uygun olarak yapılması, ölçme ve değerlendirme açısından önemlidir. Hatırlama düzeyinde bilgilerin yoklanması yanında kavrama, uygulama, analiz, sentez ve değerlendirme gibi üst düzey zihinsel becerileri belirlemeye yönelik soruların da dengeli bir biçimde seçilmesi gerekmektedir (Haladyna, 1997).

Üst düzey zihinsel becerler bireyin birden fazla beceriyi, kendi bireysel özellikleriyle ilişkilendirerek kullanması olarak tanımlanır (Marzona, 2001). Üst düzey zihinsel

becerilerin kazandırılması ve ölçülmesi, kişinin hayat boyu daha başarılı olması açısından önemlidir. Üst düzey zihinsel beceriler ele alındığında yazma becerisinin önemli bir yere sahip olduğu belirtilebilir. Yazma, bilginin kazanılması ve bilginin ifade edilmesi ile ilgili süreçlerden oluşan bir beceridir (Carter, Bishop ve Kravits, 2002, s.20). Bu beceri kişilerin kariyerlerinde başarılı olmaları için gerekli olan becerilerin başında gelmektedir. Hayatın her alanında (okulda, işyerinde ve toplumda) yazma becerilerine ihtiyaç duyulduğu söylenebilir. Yazma süreci pek çok alt beceriyi gerektiren çok yönlü, karmaşık ve zor bir süreçtir (Evans, 2001, s.1). Bunun nedeni bir metni oluştururken birçok bilişsel, dilbilimsel ve fiziksel işlemle birlikte metnin nasıl yazılacağı (metin yapısı), kime yazılacağı (hedef kitle) ya da niçin yazılacağı ilişkilendirilerek bir kompozisyon oluşturulmaya çalışılmasının zorluğu olabilir. Bu zor görevi bireylerin isteyerek ve severek yapabilecekleri alışkanlığa dönüştürmenin yolunun, ilköğretimde verilen yazma becerisine yönelik eğitime bağlı olduğu söylenebilir.

Deniz (2003, s. 242)'e göre okullarda kompozisyon ve yazılı anlatım kavramlarıyla belirtilen yazı ve yazma çalışmasının başarıyla gerçekleşebilmesi için doğru ve zengin bir kelime hazinesine sahip olmak, bol ve dikkatli okumak, iyi bir gözlemci olmak, başarıya niyetli, azimli ve bir amaca sahip olmak, dilin özelliklerini ve inceliklerini bilmek, geniş düşünmek ve bol hayal kurmak, duygu ve düşünceleri mantıklı ve ahenkli bir düzen içine sokabilmek, orijinal (özgün) olmak, yazma işini zevkle yapabilmek gibi hususların bir arada olması gerekmektedir. Ancak, geleneksel anlayışın bir sonucu olarak kompozisyon atasözü ya da özdeyişin açıklanması olarak düşünülmekte ve uygulanmaktadır. Bu anlayışın dışına çıkarak farklı tür ve içerikte yazma çalışmalarının yaptırılması gerektiği söylenebilir. Öğrencilerden yalnızca kompozisyon yazmalarını istemek, bir süre sonra sıkılmalarına yol açabilir. Bunun için öğrencilere çeşitli yazma fırsatları yaratmak ve bu türler hakkında bilgiler vermek gerekir (Yangın, 2002, s. 118). Bu şekilde öğrenciler kendi düşünce dünyalarını kurabilecek ve özgün fikirler üretebileceklerdir. Ayrıca bu becerinin gelişmesi ile birlikte bireylerin bilgiyi transfer etme, kendi düşünceleri ile bilgiler arasında bağlantı kurma ve metnin yapısında tutarlılık sağlama imkânları da artmaktadır (Ungan, 2007). Bunların yanı sıra, okul ortamında yazılı anlatım becerisinin gelişmiş olması, öğrencilerin akademik başarılarını olumlu yönde etkileyebilmektedir. Yazmanın, okul ortamında iki açıdan önemli olduğu söylenebilir: birincisi, bir fikri kanıtları ile birlikte ifade etme, rapor yazma, metni gözden geçirme, plân yapma ve değerlendirmedir; ikincisi ise, öğrencinin bilgisini derinleştirmesinin ve ilerletmesinin bir yoludur. Yazma, konu alan

bilgisini öğrenmek için bir araçtır (Sperling ve Freedman, 2001). Öğrencilerin, üretken, yaratıcı ve etkili bir yazma becerisi kazanabilmeleri için konu seçiminden kontrol ve düzeltmeye kadar yazma sürecini içselleştirmiş olmaları gerektiği belirtilebilir. Bu gerekliliklere ek olarak etkili bir yazma çalışmasının en önemli parçalarından birisi de değerlendirme süreçleridir. Bu doğrultuda ölçme ve değerlendirmenin eğitimdeki en büyük rollerinden birinin öğrencilerin sınıf içi veya dışında gerçekleştirdiği performansı ölçmek olduğu düşünüldüğünde, bu tarz üst düzey zihinsel becerilerin belirlenmesinde performansa dayalı durum belirleme araçlarının yaygın olarak kullanılması gerektiği düşünülebilir.

Eğitimdeki öğrenme çıktılarını kazanma başarısı olarak da tanımlanan performans; bilişsel, duyuşsal ve psikomotor davranışlarla üst düzey zihinsel beceriler olmak üzere bütün öğrenme ürünlerini içerir (Turgut ve Baykul, 2010, s. 259). Performansa dayalı durum belirleme, öğrencinin gerçekçi koşullarda, karmaşık ödevleri yaparken öğrendiği temel bilgileri ne kadar iyi kullandığını ölçmeye çalışır (Erman Aslanoğlu ve Kutlu, 2003). Performans ölçülmesinde öğrencilerin bir etkinlik yapması (örn, bir model kurması) ya da orijinal bir yanıt oluşturmasına, öğrencilerin yüksek düzeyde düşünme ve problem çözme becerileri geliştirmesine ve problem çözme yeteneklerini gerçek dünyaya uyarlamasına ihtiyaç vardır. Bunlara ek olarak, birden fazla çözüm ve stratejiyi göz önünde bulundurmasına, bilgiye erişebilmesine ve bu becerilerin oluşabileceği birkaç dakikadan birkaç güne genişleyebilen zaman periyoduna gereksinimi vardır (Lane ve Stone, 2006). Öğrencilerin aktif öğrenme yoluyla süreç sonunda ortaya koydukları ürünlerin değerlendirilmesi performans değerlendirme olarak ifade edilirken, öğrencilerin aktif öğrenme yoluyla süreç içerisinde elde ettikleri çalışmaların, etkinliklerin değerlendirilmesi ise performansa dayalı durum belirleme olarak adlandırılabilir. Başka bir deyişle, öğrenci başarısının gelişimini sağlamak amacıyla daha çok üst düzey düşünme gerektiren karmaşık yapıdaki becerileri ölçmek için yapılan çalışmalar performansa dayalı durum belirleme çalışmaları olarak tanımlanır (Kutlu, Doğan ve Karakaya, 2014). Bu durumun sonucunda performans değerlendirme sürecin değerlendirilmesinde oluşabileceği gibi, sadece ürünün ya da hem süreç hem de ürünün değerlendirilmesini de kapsayabileceği görülmektedir (Alıcı, 2014).

Performansa dayalı durum belirlemenin en önemli yararlarından birisi, diğer yollarla ölçülmesi zor olan karmaşık öğrenme çıktılarını ölçebiliyor olmasıdır. Aynı zamanda bilişsel süreçlerin gelişimine de katkı sağladığı söylenebilir. Diğer bir yararı ise, öğrencinin

öğrendiği ya da yapılandırdığı bilgileri gerçek yaşam durumlarında aktarmasına olanak sağlamasıdır. Bunlara ek olarak öğrenciyi merkeze alıp, öğretim ve değerlendirme sürecinde söz sahibi olmasını sağlayan modern öğrenme kuramlarına uyumlu olması da avantajları arasında belirtilebilir. Ayrıca bu performansa dayalı durum belirleme çalışmaları öğrencileri derse motive ederek, öğrencilerin ders çalışma alışkanlıklarını olumlu etkileyebilir. Performansa dayalı durum belirleme çalışmalarının gerçekleşmesinde temel bazı sorunlarla karşılaşılmaktadır. Bu sorunlar, öğretmenlerin bu konuda öğrencilere yeterince rehberlik yapmaması; öğrencilerin internet üzerinden ulaşabilecekleri hazır çalışmaları performans görevi olarak sunması; velilerin öğrenciye verilen performans görevini kendilerinin yapması; performans çalışmalarının zaman alıcı olması, öğrenci düzeyine, ilgi ve ihtiyaçlarına ve dersle ilgili beceri ve kazanımlarla uygun olmaması şeklinde belirtilebilir (Alıcı, 2014). Bunların yanında en önemli sınırlılığı ise güvenirlik sorunudur (Kutlu vd., 2014).

Performans değerlendirmelerde öğrencilerin nasıl puanlanacağını belirlemek için dereceleme ölçeklerinden, diğer ifadeyle dereceli puanlama anahtarlarından (DPA) faydalanılabilir. Dereceli puanlama anahtarları öğrencilerin performanslarını ve cevaplarını belirlenen ölçülere göre puanlamada kullanılan kılavuzlardır (Turgut ve Baykul, 2010, s. 266). Diğer bir ifadeyle, öğrencilerin sergiledikleri performansların sürecini ve sonucunu analiz edebilmek için öğretmenler veya ölçme uzmanları tarafından hazırlanan açıklayıcı puanlama tablolarıdır (Brookhart, 1999). Bu tablolar, performans görevleri gibi öznel puanlanan araçların nesnel puanlanabilmesi için hazırlanmaktadır. Dereceli puanlama anahtarındaki her bir ölçüt hakkındaki cevap gerekliliklerinin tanımlanmış olması, birbirinden bağımsız puanlayıcıların bir soruya verilen cevaba aynı puanı verme olasılıklarını yükseltmektedir (Moskal, 2000). Görevin ve değerlendirmenin amacına bağlı olarak bütünsel (holistic) veya analitik puanlama anahtarı şeklinde hazırlanabilir. Bütüncül puanlama anahtarı performans sürecine veya ürüne yönelik tek bir puanlama yapılacağı durumlarda kullanılır. Analitik puanlama anahtarı ise performans parçalarının nasıl puanlanacağına ilişkin yönergeler içerir. Bu yönüyle analitik puanlama anahtarı ile öğrencinin hangi noktalarda istenen performansı sergilediği, hangi noktalarda performansında zayıf noktaların ve eksikliklerin bulunduğu ortaya çıkarılabilir. Öğrencilerin gelişmesi için geri bildirim vermek ve öğrencilerin belirli ölçütlere göre ne düzeye ulaştığına dair açık bir tablo sergilemek amaçlarıyla da dereceli puanlama anahtarları kullanılabilmektedir.

Öğrencinin öğrenmelerini, yeterliklerini ve bilgilerini değerlendirmek için genel olarak eğiticiler tarafından kabul gören çok sayıda puanlama ve değerlendirme yöntemi bulunmaktadır (Matsuno, 2006). Ama geleneksel sınıf ortamlarında yalnızca öğretmen değerlendirici rolü oynamaktadır. Bu geleneksel değerlendirme biçimi öğretmen için oldukça yorucu olmasının yanı sıra, öğrencinin kendi hatalarını fark etmesine engel olabilir. Yazma becerisinin de dâhil olduğu üst düzey zihinsel becerilerin değerlendirilmesi için uzun ve kısa yanıtlı, boşluk doldurma, çoktan seçmeli gibi maddelerden oluşan geleneksel ölçme araçlarının yetersiz kaldığı bazı araştırma sonuçlarından görülebilmektedir (Baştürk, 2008; Berberoğlu, 1988; White, 1999; Wiliam, 2003). Performans değerlendirmede ise bir değerlendirici belli bir öğrenci ya da değerlendirme ölçütünde gereğinden fazla katı olabilir; bunun sonucunda, bazı öğrenciler diğerleriyle karşılaştırıldığında daha düşük notlar alabilirler. Dolayısıyla performans belirleme çalışmalarının amacına tam olarak hizmet edebilmesi ve öğrencilerin özellikle çeşitli beceriler kazanabilmeleri için başta öz değerlendirme olmak üzere, akran ve grup değerlendirmelerinin oldukça önem taşıdığı söylenebilir. Bunun sonucu olarak da öz ve akran değerlendirme gibi öğretmenin iş yükünü azaltıp, öğrencilere daha iyi öğrenebilme yolunu açabilecek ve öğrenen odaklı etkinlikler olan (Bayat, 2010) alternatif değerlendirme konularına odaklanılmıştır (Hargreaves, Earl ve Schmidt, 2001). Andrade, Du, ve Mycek (2010) öz değerlendirmeyi, öğrencilerin açıkça belirlenmiş ölçütler ve amaçlar doğrultusunda kendi çalışmalarını değerlendirdikleri ve bunlara göre düzenlemeler yaptıkları çalışmaların kalitelerini yansıtan biçimlendirici değerlendirme süreci olarak açıklamışlardır. Bunun dışında, Noonnan ve Randy (2005) bu değerlendirme çeşidini öğrencilerin, öğrenme sürecinde gerçekleştirdikleri çalışmaları, öğretmen ya da öğrenciler tarafından belirlenmiş ölçütler doğrultusunda değerlendirmeleri ve öğrenmeleri hakkında kendi kararlarını vermeleri olarak belirtmişlerdir. Öz değerlendirmenin öncelikli olarak öğrencileri kendi öğrenmeleri hakkında karar vermede ve amaçlarını belirlerken daha fazla sorumluluk almalarında cesaretlendirmek ve performanslarını yansıtma gerekliliğiyle öğrencileri daha aktif bir şekilde eğitim sürecine dâhil etmektir (Ballantyne, Hughes ve Mylonas, 2002; Dochy, Segers ve Sluusmans, 1999; Matsuno, 2006). Boud ve Lublin (1983)'e göre öz değerlendirme, öğrencinin kendisini doğru şekilde değerlendirme ve öğrenmelerini kontrol edebilme yeteneğini kazandırdığı için eğitimde olabilecek en önemli süreçtir. Brown ve Hudson (1998) da ayrıca öz değerlendirme sürecindeki öğrencinin

etkinliğinin fazla olmasının öğrenme konusunda motivasyon arttırabileceğini belirtmişlerdir.

Bir diğer değerlendirme yöntemi olan akran değerlendirme ise Boud, Cohen ve Sampson (1999) tarafından, öğrencilerin sınıf arkadaşlarının çalışmalarını belirli ölçütler doğrultusunda değerlendirilmesi olarak tanımlanmıştır. Topping (1998) daha ayrıntılı bir şekilde, akran değerlendirmeyi "bireylerin, benzer statüdeki akranlarının öğrenme işi ile ilgili ürünlerine başarı, nitelik, değer, düzey açısından eleştirel gözle bakmalarını sağlayan bir planlama işi" olarak belirtmiştir. Bu değerlendirme planı doğrultusunda, akranlar birbirlerinin öğretimsel ürünlerini puanlayarak değerlendirebilecekleri gibi yorumlayarak, eleştirerek, görüş belirterek geri dönüt verebilirler ya da her ikisini de kullanarak sözlü ve yazılı biçimde değerlendirebilirler. Bu özellikler göz önüne alındığında, akran değerlendirme "hem geri bildirim verme amacıyla yapılan izlemeye dayalı (formative) değerlendirmenin hem de başarı belirleme yönüyle düzey belirlemeye dayalı (summative) değerlendirmenin kapsamına girmektedir (Temizkan, 2009). Ayrıca puanlayıcıların daha etkili puanlama yapabilmeleri için bir anlamda onlara dönüt vermesi açısından önem taşımaktadır. Öz ve akran değerlendirme yöntemlerini temel olan öğretmen değerlendirme yöntemleri yerine kullanmaktansa, öğretmen değerlendirme yöntemlerin eksikliklerini kapatmak için onlarla ahenkli bir şekilde ek olarak kullanmak uygulayıcılara daha fazla yarar sağlayabilir. Ayrıca, modern öğretim yöntemleri başta işbirliğine dayalı öğrenme, proje tabanlı öğrenme olmak üzere öğrencilerin daha aktif olduğu öğrenme yöntemlerinin tam anlamıyla kullanılabilmesi öz, arkan ve grup değerlendirme yöntemlerinin kullanımını bir anlamda zorunlu kıldığı söylenebilir.

Ek olarak, öğrencilerin kullandığı dereceli puanlama anahtarının, öğretmenin kullandığı puanlama anahtarı ile aynı olması öğrencinin kendini değerlendirebilme becerisinin kazandırılması açısından yararlı olabilmektedir. Böylece öğrenci kendine yönelik yaptığı puanlamalar ile öğretmenin yaptığı puanlamaları karşılaştırma olanağına sahip olabilmektedir (Karakaya, 2015). Dereceli puanlama anahtarları öz ve akran değerlendirme kapsamında kullanılırken performans düzeylerinin tanımlamalarında nicel tanımlama yerine nitel tanımlamalara ağırlık verilebilir. Öğrencinin kullanacağı araçta nicel değerler yerine nitel değerler tercih edilmesi öğrencilerin kendilerini daha doğru değerlendirmelerine katkı sağladığı söylenebilir. Nicel değerler genellikle not kavramını çağrıştırdığı için ölçme araçlarında kullanımına dikkat edilmesi önem taşımaktadır (Karakaya, 2015a).

Bu kapsamda performansa dayalı durum belirleme sürecinde öğrenciden beklenen davranışlar dereceli puanlama anahtarları veya kontrol listeleri ile belirlenmesine rağmen, puanlayıcılar, puanlayıcıların görüş ayrılıkları, objektiflik eksikliği, anlaşılır olmayan puanlama anahtarı ve çevredeki değişiklikler gibi muhtemel hatalar ile karşılaşılır (Kim, 2000). Bu hata kaynakları puanlayıcılar arası tutarlığı diğer bir ifadeyle puanlayıcı güvenirliğini düşürür. Her ölçmenin güvenilir olması istenmekle birlikte, ölçmeye karışan sabit, sistematik veya tesadüfî hatalar güvenirliğin veya geçerliğin düşmesine sebep olabilir. Sabit ve sistematik hata kaynakları ölçmeden uzaklaştırılabilirken, tesadüfî hata kaynağı ölçmeden tam olarak arındırılamamaktadır. Özellikle bir yanıtı olmayan kompozisyon tipi gibi sınavlarda ölçmenin güvenirliğinin ve geçerliğinin arttırılması için puanlayıcı güvenirliğinin de arttırılması gerektiğine işaret eder. Ölçmeler, bütün puanlayıcıların aynı cevaba aynı puanı vermesi sağlandığı müddetçe, objektiftir (Turgut ve Baykul, 2010, s. 132). Bu sebeple yapılan ölçme işleminde güvenirliği arttırmak amacıyla puanlayıcı sayısı arttırılabilir ve böylece puanlayıcı güvenirliği sağlanmış olur. Ölçmenin güvenirliğini arttırmak için puanlayıcı güvenirliğini arttırmaya çalışmak, ölçmenin konusunu duyuşsal, bilişsel ve psikomotor davranış alanlarının birkaçını içinde barındırabilen performans ölçmelerde daha büyük önem kazanmaktadır (Turgut ve Baykul, 2010, s. 259). Bu sebeplerden dolayı, yapılan performans ölçümlerinde ölçme sonuçları değerlendirilmeden önce, puanlayıcılar arası tutarlığın incelenmesi gereklidir (Çakıcı Eser ve Gelbal, 2012).

Puanlayıcılar arası güvenirliğin (PAG) analizi, bir takım konuların özellikleri hakkında bağımsız puanlamalar yapan iki ya da daha fazla puanlayıcı arasındaki uyumun derecesini belirlemenin bir yolunu sağlar (Hallgren, 2012). Ana dil eğitiminde, akran puanlamalarındaki varyansı bulmak için "gerçek puan" yaklaşımını benimsenmiş ve puanlayıcılar arası güvenirliği bulmak için faktör analitik yaklaşım, varyans analizi (ANOVA) ve t-testleri kullanılmıştır (Matsuno, 2006). Bu yöntemlere ek olarak Klasik Test Kuramı (KTK) bağlamında Pearson Momentler Çarpımı, Spearman-Brown Düzeltme Formülü, Uyum Oranı, Uzlaşma Katsayısı, Kappa, Sınıf İçi Korelasyon Katsayısı istatistiği gibi farklı yöntemlerden de yararlanılabilir. Bu yaklaşımların her birinde, tüm puanlayıcılar belli bir bireyin performansına tam olarak aynı puanı verirse, istenilen ideal sonuçlara ulaşılabilir. Puanlayıcılar bireysel olarak farklı değer sistemlerine sahiptir. Bu sonucunda en dikkatli, detaylı puanlayıcı eğitiminin bile puanlayıcılar arasındaki birliği sağlamada başarısız olabileceği bir gerçektir ve elde edilen sonuçlar tam olarak ideal gerçeği

yansıtmamaktadır (Borman, 1977). Ayrıca klasik yaklaşımda bireylerin ham puanları hangi puanlayıcının onları puanladığına bağlı olarak değiştiğinden, puanlayıcılar arasındaki ayrılıklar adaletsizliğe sebep olabilir. Matsuno (2006)'a göre öğrenciler başkalarını puanlama konusunda genelde tecrübesizdir ve bu durum özellikle öz ve akran değerlendirmeleri için geçerlidir. Klasik yaklaşımın bu zayıf yönlerini yenmek için son yıllarda Genellenebilirlik Kuramı (GK) ve Çok Yüzeyli Rasch Ölçme Modeli (ÇYRÖM) (Linacre 1993) kullanılmaktadır.

Genellenebilirlik Kuramı gözlenen puanlarda var olan veya var olabilecek olan tutarsızlıkların güçlü istatistiksel yöntemler ile belirlenmesini ve ölçülmesini sağlayan, ANOVA üzerine kurulu bir kuram (Brennan, 2001, s. 4) olarak tanımlanırken, gözlenen puanların logaritmik olarak lojit ölçeğine dönüştürülmesine dayalı doğrusal bir modeli olan ÇYRÖM ikiden fazla değişkenlik kaynağından faydalanan, Rasch ölçme halinin daha genişletilmiş hali olarak tanımlanmıştır (Nakamura, 2002). ÇYRÖM, ele aldığı yüzeyler ile bireylerin başarı seviyelerini, görev ya da maddelerin güçlük düzeylerini ve puanlayıcıların katılıklarını kestirmek için kullanılan istatistiksel bir tekniktir (Güler, 2008).

ÇYRÖM, geleneksel ölçme yöntemlerine göre olduğu gibi, GK'ya göre de birtakım üstün özelliklere sahiptir (Linacre, 1993; Rasch, 1980). Seçilen örnekleme dayanan KTK ve GK'dan farklı olarak, ÇYRÖM ham puanlar yerine kişilere ait (ölçme hatalarından arındırılmış) ölçme değerlerini kullanmaktadır. ÇYRÖM, aynı zamanda her bir yüzey için yüzeyler arası olması gereken doğrusal bir bağlantının kurallarını belirlemekte, bu bağlantıları oluşturmakta (Hambleton, Swaminathan ve Rogers, 1991, s.128) ve yüzeylerin her birine ilişkin olarak örneklemlerden bağımsız parametre kestiriminde bulunmaktadır. Değişkenlik kaynakları eş zamanlı ve istatistiksel olarak bağımsız bir şekilde analiz edilir. GK değişkenlik kaynakları arasındaki etkileşimi dikkate alan bir yaklaşımken, Madde Tepki Kuramı (MTK) etkileşimin olmadığı varsayımını göz önünde bulundurur (Güler, 2008). Kısaca bu model, yansız ve etkili bir ölçme elde etmek için yüzeyleri ortak bir düzlemde birleştirerek standart hale getirmektedir. Buna ek olarak ÇYRÖM'de her tür sistematik hata kaynağı bir değişkenlik kaynağı olarak görülmektedir. Bunun sonucu olarak GK'da değişkenlik kaynağı olarak ele alınmayan bireyler ya da öğrenciler ÇYRÖM'de değişkenlik kaynağı olarak modele dâhil edilmektedir. Bir diğer farklılık ise GK'da ölçmenin kalitesi, bir grup ya da bir bütün olarak araştırılırken ÇYRÖM, her bir değişkenlik kaynağının her bir elemanına ilişkin araştırma yapmaya imkân verilir ve bu durum bize hangi öğrencinin kendisini veya akranını ne derece etkili olarak

puanlayabildiğini gösterebilmektedir. Bu ise öğrencilere yönelik bire bir nesnel puanlama yapabilme, kendilerini ve akranlarını eleştirel gözle değerlendirebilmelerine önemli bir derecede katkı sağlamaktadır. GK'da bir ölçme için bir güvenirlik katsayısı kestirilirken, ÇYRÖM'de her bir değişkenlik kaynağı için aynı anda ayrı güvenirlik katsayıları kestirilebilir. Böylece her bir değişkenlik kaynağındaki elemanların beklenmedik ya da aykırı durumlardaki performanslarının nasıl olduğunu belirlemeye imkân verir (Alharby, 2006). Ayrıca bireylerin becerilerini, soruların güçlüğünü ve puanlayıcı katılık ya da cömertlik davranışlarını aynı anda karşılaştırma olanağı sunmaktadır (Baştürk, 2010).

Alanyazına bakıldığında, ÇYRÖM'nin bazı araştırmacılar (Baştürk, 2008; Baştürk ve Akın, 2012; Gyagenda ve Engelhard, 1998; McNamara ve Adams, 1991; Weigle, 1998; Weigle, 1999) tarafından belli bir durumdaki bireylerin o konuda başarılarını ve puanlayıcıların katılık durumlarını belirlemek için kullanılırken, bazı çalışmalarda (Farrokhi ve Esfandiari, 2011; Schaefer, 2008) ise puanlama yanlılıkları ve bunları etkileyen faktörleri araştırmak için kullanılmıştır. Bunlara ek olarak bazı araştırmacılar (Güler, 2008; Macmillan, 2000; Sudweeks, Reeve ve Bradshaw, 2005) diğer kuramlarla olan karşılaştırma çalışmalarını içerirken, bazıları (Farrokhi, Esfandiari ve Dalili, 2011; Matsuno, 2006; Nakamura, 2002; Semerci, 2011) ise puanlayıcı çeşitlerini (öz, akran ve öğretmen) araştırmayı amaçlamıştır. Bu çalışma da benzer olarak öz, akran ve öğretmen puanlarını ÇYRÖM ile karşılaştırmayı amaçlamaktadır. Bunun yanı sıra, Nakamura (2002) çalışmasında sadece akran puanlarını kullanırken, örneklem sayısı çok düşük tutarak her puanlayıcının herkesi puanlamasını sağlamıştır. Matsuno (2006)'ın çalışmasında öz, akran ve öğretmen puanlamalarını birlikte kullanmış, fakat her birey sınıftaki her akranı tarafından puanlanamayacağından her akran belli birkaç arkadaşını puanlamış ve öğretmen puanlayıcılar da aynı şekilde yine belli bir kısım öğrenciyi puanlamışlardır. Bu konuda Linacre (2002) Rasch ile ilgili çalışmalarda her birey her kategoride her puanlayıcı tarafından puanlandığında daha kesin sonuçlar çıkacağını belirtmiştir. Bu sebeple bu çalışmada benzer olarak akranlar konusunda aynı sorunlarla karşılaşılmakta ve her akran bir arkadaşını puanlamaktadır. Fakat öğretmen puanlayıcıların farklı olarak her bireyi her kategoride puanlayarak bu sorunun giderilmesi amaçlanmıştır. Ayrıca çalışmalarda genel olarak öz akran ya da öğretmen puanlarından biri ya da birkaçı ayrı yapı ya da yüzeylerde incelenerek farklı yöntemlerle yorumlama yollarına gidilmiştir (Farrokhi, Esfandiari ve Dalili, 2011; Matsuno, 2006; Nakamura, 2002; Semerci, 2011). Bu doğrultuda bu çalışmada veriler hem ayrı ayrı hem de ortak bir yüzeyde girilerek sonuçlar arasındaki

farklılaşmalar ve bir bütün olarak alınan analizin yorumlanma kolaylıkları belirlenmeye çalışılmıştır.

Bu çalışmada öğrencilerin öz ve akran puanlamalarının öğretmen puanlarıyla karşılaştırılmasını zaman içerisindeki değişimlerini de dikkate alarak incelenmesi amaçlanmıştır. Bunun için de Türkçe derslerindeki yazma performansına yönelik görevlere verilen öz, akran ve öğretmen puanları çok yüzeyli Rasch ölçme modeliyle incelenmiştir.

1.2. Kuramsal Çerçeve

1.2.1. Puanlayıcı Güvenirliği

Puanlayıcı güvenirliği genel olarak puanların tutarlılığı olarak adlandırılır ve en az iki bağımsız puanlayıcı ya da farklı zamanlarda aynı puanlayıcıların puanlandırılmasına dayanır. Yani, PAG analizi, puanlayıcılar arasındaki ölçme hatalardan kaynaklanan varyans çıkartıldığında gözlenen puanda gerçek puandaki varyanstan kaynaklanan ne kadar bir varyans olduğunu belirlemeyi amaçlamaktadır (Novick, 1966).

Puanlayıcı güvenirliği puanlayıcılar arası (interrater) ve puanlayıcılar içi (intrarater) olmak üzere ikiye ayrılmaktadır. Farklı puanlayıcıların birbirinden bağımsız olarak her bir öğrencinin kâğıdına aynı puanı verme tutarlılıklarının derecesi puanlayıcılar arası güvenirlik olarak tanımlanmaktayken, bir puanlayıcının farklı zamanlarda aynı bireylerin kâğıtlarını puanlaması ile elde edilen puanlar arasındaki ilişki ise puanlayıcı içi güvenirliği vermektedir (Güler, 2008).

1.2.2. Puanlama Güvenirliğinde Kullanılan Yöntemler

1.2.2.1. KTK'na Dayalı olan Yöntemler

En uygun puanlama güvenirliği hesaplama yöntemini seçmek için, değişkenin kodlandığı ölçek türü, çalışmanın deseni (tüm kişiler mi yoksa bir kısmı mı birden fazla puanlayıcı tarafından puanlanacağı) ve PAG hesaplamalarının beklenen kullanım amacı (puanlayıcıların bireysel olarak mı yoksa birden fazla puanlayıcıdan alınan ortalamanın mı güvenirliğinin hesaplanacağı) gibi birçok faktör göz önünde bulundurulmalıdır (Hallgren,

2012). Diğer bir ifadeyle, puanlama güvenirliğinde sınıflama ölçeği türündeki bir veri için tavsiye edilen bir yöntem sıralama, aralıklı ve oranlı ölçeğindeki veriler için yetersiz kalabilir. Aynı şekilde oranlı ölçek türünde kullanılan yöntemler de sınıflama ölçeğinde kullanılanlarla benzer olmayabilir. Bu durum, analizde kullanılacak yöntemin seçiminde verinin ölçek türünün doğrudan etkisi olduğunu göstermektedir (Gwet, 2012, s.10). Ayrıca kullanılan puanlayıcı sayısının da yöntemlerin seçimi üzerinde etkisi bulunmaktadır (Hallgren, 2012). İki puanlayıcı için kullanılan bir yöntem ikiden fazla puanlayıcı için uygulanamayabileceği gibi güvenilir olmayan sonuçlar da verebilir. KTK'da bu farklı koşullarda puanlayıcılar arası güvenirliği hesaplamak için Pearson Momentler Çarpımı Korelasyon Katsayısı, Spearman Sıra Farkları Korelasyon Katsayısı, Cohen'in Kappa İstatistiği, Ağırlıklandırılmış Kappa (Weighted Kappa) İstatistiği, Fleiss Kappa İstatistiği, Kendall'ın Uyum Katsayısı, Krippendorff'un Alpha Katsayısı, Kendall τ, Goodman ve Kruskal'ın ϒ, Sınıf İçi Korelasyon (ICC) Katsayısı gibi yöntemler genellikle kullanılmaktadır. Genellikle kullanılan bu yöntemlerin seçilmelerine etki eden yukarıda belirtilen faktörlere göre sınıflandırılması Tablo 1'de verilmiştir.

Tablo 1. Puanlama Güvenirliğinde Kullanılan Yöntemlerin Verinin Ölçek Türüne ve Puanlayıcı Sayısına Göre Sınıflandırılması

Yöntemler	Sınıflandırma Ölçütleri	
	Ölçek Türü	Puanlayıcı Sayısı
Pearson moment çarpım korelasyon katsayısı	EO-EA	2
Spearman sıra korelasyon katsayısı	EO-EA	2
Cohen' in Kappa istatistiği	A	2
Ağırlıklandırılmış Kappa (Weighted Kappa) istatistiği	S	2
Fleiss Kappa istatistiği	A	3+
Kendall'ın uyum katsayısı,	S	3+
Krippendorff'un Alpha katsayısı	EO-EA, A,S	2, 3+
Kendall τ,	S	2
Goodman ve Kruskal's ϒ	S	2
ICC	EO-EA, S	2, 3+
A:Adlandırma (Sınıflama) Ölçeği		S:Sıralama Ölçeği
EA: Eşit Aralıklı Ölçek		EO: Eşit Oranlı Ölçek

Tablo 1'e bakıldığında yöntemlerin hangi veri türlerine daha uygun oldukları görülmektedir. Yöntemleri seçerken belirtilen faktörler dışında sağladıkları bazı avantajlar ve dezavantajlar da göz önünde bulundurularak amaca en uygun olan seçilebilir. Örneğin

bu yöntemlerden Pearson momentler çarpımı korelasyon katsayısı, genellikle kullanılan ve kolay yorumlanabilen bir teknik olması açısından avantaja sahipken, örneklem büyüklüğünden doğrudan etkilenmesi açısından da dezavantaja sahiptir (Goodwin, 2001). Spearman sıra farkları korelasyon katsayısında ise iki değişken (Y ve X) içinde örneklem verilerinin sıralama düzeninde olmaları gereklidir. Genel olarak, örneklem verileri için bu koşul uygun değildir ve veriler sıralama düzeni halinde olmadan oranlı veya aralıklı veya sıralama ölçeği olarak bulunur ve bu halde bir dönüşümle sıralama düzeni haline sokulurlar (Kılıç, 2009). Cohen'in Kappa istatistiğine bakıldığında, bu istatistik uyumluluğun ölçütlerini vermesine rağmen tesadüfen ortaya çıkan uyumluluğu hesaba katmaz. Eğer puanlayıcılar tesadüfen uyumlu oluyorlar ise, hiçbir şekilde gerçekten uyumlu değillerdir; sadece tesadüfen beklenenin dışındaki uyumluluğun "gerçek" uyumluluk olduğu düşünülebilir (Sim ve Wright, 2005). Kappa katsayısı, puanlayıcıların uyumsuzluğunun şansa bağlı olarak mı, yoksa sistematik farklılıklardan dolayı mı olduğunu göstermez (Hartmann, 1977). Uyumsuzluğun derecesini yansıtmak için Kappa ağırlıklandırılabilir, böylece değerlendiriciler arasındaki küçük farklara göre büyük farkların etkisi arttırılır (Sim ve Wright, 2005). Yukarıda belirtilen KTK'nın farklı yöntemleriyle farklı anlamlarda elde edilen güvenirlik katsayılarının birbirinden farklı olmasından hareketle, Cronbach ve arkadaşları, Genellenebilirlik Kuramını (GK) ortaya atmışlardır (Atılgan, 2005).

1.2.2.2. Genellenebilirlik Kuramı

GK, esnek bir alternatif olarak, puanlayıcı, zaman, test formu, madde, görev gibi bir ölçme içinde yer alabilen bütün potansiyel hata kaynaklarını eş zamanlı değerlendirebilen bir yaklaşımdır (Atılgan, 2005). Diğer bir ifadeyle, davranış ölçmede güvenirliğin değerlendirilmesini, güvenilir gözlemlerin tasarlanmasını, araştırılmasını ve kavramsallaştırılmasını sağlayan, istatistiksel bir kuramdır.

Bireylerin yerine getirmeleri gereken görevlerin olduğu ve birden fazla puanlayıcının yer aldığı bir ölçmede görevler ve puanlayıcılar potansiyel değişkenlik kaynaklarını diğer bir ifade ile potansiyel hata kaynaklarını oluşturmaktadır. Söz konusu ölçmede bireylerin değişkenliğe sahip olması ve başarılarının değişkenlik göstermesi beklendiğinden bireyler bir hata kaynağı olarak ele alınmamaktadır. Buna göre bu ölçme deseni iki değişkenlik kaynaklı bir ölçmedir. İki değişkenlik kaynaklı evrenden elde edilen varyans bileşenleri

değişkenlik kaynaklarının her birini ayrı ayrı, ikili ve üçlü etkileşim şeklinde ele almaktadır (Gelbal ve Çakıcı Eser, 2012).

1.2.2.3. Çok Yüzeyli Rasch Ölçme Modeli

Çok Yüzeyli Rasch Ölçme Modeli -ÇYRÖM- (Linacre, 1993) MTK'ya dayalı 1980 yılında Rasch tarafından geliştirilen bir parametreli Rasch ölçme modelinin genişletilmiş halidir. Basit Rasch modelinde, kişiler ve test maddeleri değerlendirilir ve kişilerin yetenek farklılıklarının ya da madde güçlüklerinin ele alındığı eşit aralıklı bir ölçek üzerine yerleştirir ve sonuçlar örneklemden bağımsızdır. Yani kişilerin yetenekleri kullanılan belli madde örnekleminden ve madde güçlükleri maddelerin uygulandığı belli gruplardan bağımsız olarak kestirilir (Sudweeks, Reeve ve Bradshaw, 2005). ÇYRÖM ise kullanıcıların hem görev/madde güçlüğü hem de test edilen kişilerin yetenekleri ile ilgili puanların bulunduğu tek bir eşit aralıklı ölçek oluşturabilmelerini sağlar (Linacre, 1993). Bu puanlar, lojit adı verilen birimle rapor edilir ve lojistik cetvel olarak adlandırılan genelde dikey olan bir cetvel üzerinde yer alırlar. Lojistik cetvelde tüm yüzeyler pozitif tanımlanmışsa, yer alan yüksek pozitif değerlere sahip ölçmeler, bireyler için yüksek yetenek, puanlayıcılar için cömert puanlama ve maddeler içinse kolaylığı ifade eder; yüksek negatif değerler ise düşük yetenekli bireyleri, katı puanlayıcıları ve güç maddeleri ifade etmektedir (Güler, 2008). Lojitlerle toplama, çıkartma, çarpma ve bölme gibi işlemler yapılabilir. Ek olarak, eğitimsel kazanımları değerlendirmede, güçlü ve zayıf yönleri ortaya koymada ve demografik grupların karşılaştırmasında faydalı olabilecek karşılaştırmalar ve istatistiksel çalışmalar bu birimlerle yapılabilir (Engelhard, 1996).

Lojit ölçeğinin, ana etkiler (main effects) için geliştirilen log-doğrusal bir model olduğu da söylenebilir (Hertherman, 2004). ÇYRÖM, tüm değişkenlik kaynaklarını ele alan ve bu değişkenlik kaynaklarının beraber araştırılmasını sağlayan bir modeldir. Regresyon analizi terminolojisine göre değerlendirildiğinde, bağımlı değişken olarak başarılı olma olasılık oranının logaritmik şekli (log odd), bağımsız değişken ise değişkenlik kaynakları olarak tanımlanabilir (Alharby, 2006).

Adından da anlaşılacağı gibi, model ikiden fazla yüzey içermektedir. Bu durum, puanlayıcı katılığı, değerlendirme sayısı gibi daha farklı yüzeylerin modele dâhil edilebilmesini sağlar. ÇYRÖM'de, puanlayıcıların katılık ya da cömertliğiyle ifade edilmek istenilen puanlayıcıların diğer puanlayıcıların puanlarının ortalamasından farklı olarak sistematik bir

şekilde daha yüksek ya da daha düşük puanlamalarıdır (Engelhard ve Myford, 2003). Bu duruma ayrıca puanlayıcı etkisi ya da puanlayıcı hatası da denilebilir (Alharby, 2006). Model genel olarak üç yüzey (birey, görev ve puanlayıcı) içermektedir ve Linacre (1993, s.1) tarafından belirtilen aşağıdaki formül kullanılmaktadır:

$$\log\ (P_{nijk}/P_{nijk-1}) = B_n - D_i - C_j - F_k$$

P_{nijk}: n bireyinin j puanlayıcısından i kategorisi ya da görevinde k puanı alma olasılığı

P_{nijk-1}: n bireyinin j puanlayıcısından i kategorisi ya da görevinde k-1 puanı alma olasılığı

B_n: n bireyinin yeteneği

D_i: i kategorisi, maddesi ya da görevinin güçlük düzeyi

C_j: j puanlayıcısının katılık düzeyi

F_k: j puanlayıcısı için k-1 puanından k puanına geçişin güçlüğü

Bu formülden yola çıkılarak yapılan analiz ilk olarak uyum istatistiklerini vermektedir. Uyum istatistikleri veri matrisindeki uyumsuzlukların derecesini gösterir ve her bir madde için artıkların (residuals) büyüklüğü ve işareti hakkında bilgiler sağlar. Artıklar, gözlenen puanlardan beklenen puanların çıkartılmasıyla hesaplanır. Veriler modelle uyumluysa, her artık değerin sıfır olması gerekir. Veriler modele yaklaştığı ölçüde, standartlaştırılmış artıklar ortalaması 0 ve varyansı 1 olacak şekilde normal ya da normale yakın bir dağılım sergiler. Sonuç olarak, bu ölçütün kullanışlı olmasıyla, standartlaştırılmış artık değerlerin yaklaşık olarak normal bir dağılım sergileyip sergilemediği veri model uyumu hakkında yorum yapmada kullanılabilir (Hertherman, 2004). ÇYRÖM ayrıca analizde oluşturulan model tarafından tahmin edilen beklenen değerlerle her kişinin bireysel performansının, puanlayıcının ya da görevin ne kadar iyi eşleştiği konusunda da bilgi sağlamaktadır. Bu "uyum istatistikleri" Rasch analizinde "uyum içi (infit)" ve "uyum dışı (outfit)" karelerin ortalama değerleri olarak bilinmektedir. Uyum dışı istatistiği, standartlaştırılmış artıkların karelerinin ortalaması olarak tanımlanırken, uyum içi istatistiğinin bu istatistiğin ağırlıklandırılmış hali olduğu belirtilmektedir. Bu uyum istatistikleri her bir yüzeydeki elemanlar için hesaplanabilir. Modele göre, bu istatistiklerin sıfır standart hatayla bir

beklenilen değerini almaları beklenilir. Bazı araştırmacılara göre ise daha kolay yorumlanan uyum istatistikleri olan uyum içi ve uyum dışı değerlerinin makul ranjı 0,6-1,4 arasında olmalıdır (Wright ve Linacre, 1994). Yüzeylerdeki her bir elemana ilişkin uyum içi ve uyum dışı istatistiklerinin 0,5'ten küçük olması, anlamlı olarak uyum üstü bir duruma işarettir. Uyum üstü elemanlar, yüzeyin genel örüntüsü içinde bu elemanlardan bağımsız olarak bir bilgi sağlamadıklarının işaret eder (Hetherman, 2004). Bu düşünce daha sonra genişletilerek, Michael Linacre 1,5'den az olan değerlerin ölçme için istenilen ve 1,5-2,0 arası değerlerin ise istenilmeyen değerler (misfit) olduğunu öne sürmüştür. İstenilmeyen değerleri alan uyumsuz elemanlar, değişkenlik kaynağı içindeki tüm puan örüntüsü içinde beklenmedik durum gösteren elemanlardır. Örnek olarak, uyumsuz maddeler, kötü hazırlanmış olduklarının sinyallerini verebilir. Bu maddeler kendi içlerinde oldukça iyi hazırlanmış olarak değerlendirilebilirken, maddeler örüntüsünde uyumsuzluk gösterirler. 2,0'nin üzerindeki değerler ise yanıltıcı olarak değerlendirilir. (Myford ve Wolfe, 2003). Uyum dışı istatistikler, aykırı puanlara karşı, uyum içi istatistiklerden daha hassastır. Her puanlayıcının puanlamadaki katılığı, modelde yer alan diğer değişkenlik kaynaklarından bağımsız olarak kestirilmektedir. Bu kestirimlerdeki yüksek değerler, puanlayıcıların ortak cetvelinde alttaki puanlayıcıların daha katı olduğunun göstergesidir. Katılık düzeyi kestiriminden sonra, kestirimlerin verilere uyumuna bakılır. Benzer aşamalar diğer yüzeyler (madde güçlüğü, birey yeteneği vb.) için de uygulanabilir (Hetherman, 2004).

GK ile ÇYRÖM karşılaştırıldığında, ilk olarak, iki analizde de kullanılan, fakat farklı anlamlara gelen üç terim anlamlarına bakmak gerekecektir. Bunlar (a) yüzey, (b) etkileşim (c) güvenirlik. Öncelikle GK'da yüzey; sistematik hata varyansının kaynağı olarak tanımlanmaktadır. Ölçülen obje (birey) tarafından oluşturulan varyans doğal karşılandığından, GK'da tanımlanan yüzeyler ölçülen obje (birey) dışındaki model bileşenlerini içermektedir. Bunun aksine, ÇYRÖM yüzeyleri modeldeki tüm sistematik hata kaynakları olarak tanımlamıştır (Güler, 2008). Örneğin; bireyler, görevler ve puanlayıcılar içeren bir model önerilirse, GK'da bireyler ölçme objesi, görevler ve puanlayıcılar yüzey olarak alınmasına karşın ÇYRÖM'de üçü de yüzey olarak alınabilir. Etkileşim terimi de iki modelde biraz farklı şekillerde kullanılır. ÇYRÖM'de yüzeyler arasındaki etkileşimler, değişen yüzey fonksiyonunu olacağı gibi değişen madde ya da kişi fonksiyonunu da gösterebilmekte ve durumdan duruma bireysel olarak yorumlanabilmektedir. GK'da ise etkileşim varyansların faktöriyel analizi olarak

tanımlanmaktadır. Bu analiz, iki ya da daha fazla ana etkinin her etkileşimi için bir varyans bileşeni ileri sürmektedir. Bu etkileşim varyans bileşenlerinin kıyaslamalı boyutu, ölçme nesnesinde ya da bir yüzeydeki elementlerin hangi göreceli durumları ikinci bir yüzeydeki elementlere göre değiştiğinin derecesini gösterir (Alharby, 2006). Örneğin; puanlayıcılar için ana etkiyi anlatan varyans bileşeni örneklemdeki hangi puanlayıcının katılık olarak ne kadar değiştiğini gösterir. Bireyler ve puanlayıcılar arasındaki etkileşimi gösteren varyans bileşeni hangi bireyin farklı puanlayıcılar tarafından ne kadar farklı puanlandığını gösterir (Shalvelson ve Webb, 1991, s. 55). Son olarak, güvenirlik terimi de ÇYRÖM'de GK'dan farklı kullanılmaktadır. Genellenebilirlik analizlerinde g ve phi katsayılarının ikisi de alan örnekleme modeli (domain sampling model) kapsamında güvenirlik katsayısı olarak yorumlanabilir. Bu katsayılar, birey için ortalama puanlamanın onun evrendeki puanını ne kadar iyi yordadığını göstermektedir. G katsayısı tek yüzeyli bir modelde Cronbach alpha katsayısına eşittir ve onun gibi yorumlanabilir. G katsayısı iki ya da daha fazla yüzey içeren G çalışmalarında daha düşük değerler alır. Phi katsayısı, daha fazla kaynak ya da hata varyansı içerdiğinden genel olarak g katsayısından daha küçük değerler alır (Sudweeks, Revee ve Bradshaw, 2005).

ÇYRÖM analizi de her yüzey için iki farklı güvenirlik istatistiği içerir:

1) Ayırma indeksinin güvenirliği (The reliability of separation index)

2) Ayırma oranı (The seperation ratio, G)

Bu iki istatistik farklı birer metrikte olmasına rağmen, aynı bilgilerden hesaplanırlar ve belirli bir yüzey için benzer sonuçlar verirler. İki istatistik de ÇYRÖM ile hesaplanmış ölçümlerin dağılmasının ya da değişkenlik oranını göstermektedir. Ayırma indeksinin güvenirliği, yüzeydeki elemanların her birinden ne kadar güvenilir bir şekilde ayrıldıklarını gösterir. Ayırma oranı ise, yüzeydeki objelerin birbirinden ne kadar ayrıldığını gösteren bir değerdir (Güler, 2008). Her bir yüzey için ayırma indeksinin güvenirliği 0 ile 1 arasında değer alırken, ayırma oranı 1'den sonsuza kadar gider. Bu istatistiklerin yorumları yüzeylere göre değişmektedir. Yani, dağılma indeksinin güvenirliğinin ve ayırma oranının birey yüzeyinde yüksek değerler alması istenirken, diğer yüzeyler için alçak değerler alması genelde istenir. Ayırma indeksinin güvenirliği, Cronbach alpha katsayısına benzerdir (Myford ve Wolfe, 2003). Kişi yeteneği hesaplamalarında bu katsayı gerçek varyansın gözlenen varyansa oranını göstermektedir. Birey, madde ve görev gibi değişkenlik kaynakları için bu istatistiğin 1'e yakın değerler alması istenilir ve bu değerler

yüksek seviyede güvenirliği gösterir. Bireylerin ayırma oranı 1'e eşit olduğunda, birey yeteneklerinin ölçümleri arasındaki değişim bu ölçümlerin hesaplanmasında oluşan belirsizliklerden fazla değildir. Yani, bir bireyin yeteneğini diğer bireyden ayırt edebilmek için güvenilir bir yol yoktur. Yine de dağılma 2'ye eşit ise, o zaman birey yetenek ölçümleri hata kareleri ortalamasının karekökünden (the root mean square error) iki kat daha farklıdır ve ayırma indeksi 4'e eşitse, birey ölçümlerindeki ayırma hesaplamalardaki belirsizliklerden dört kat daha fazladır (Smith, 2001).

Bu iki istatistikteki birey yüzeyi için düşük değerler puanlamalardaki merkezi eğilim hatasına benzer olabilir. Yani puanlayıcılar verilen görevdeki çeşitli bireylerin performanslarını ayırt edememişlerdir (Myford ve Wolfe, 2003). Birey yüzeyinden farklı yüzeyler için bu iki istatistiğin düşük değerleri, belli bir yüzeydeki çeşitli elementlerin ölçümlerindeki yüksek derecede tutarlılığı göstermektedir. Bunlara ek olarak ayırma oranı (G), her bir değişkenlik kaynağında, elemanların kaç farklı düzeye ayrılabildiğini gösteren (4G+1)/3 formülünde kullanılmaktadır (Hetherman, 2004).

Bu analizde yorumlanan diğer bir istatistik de Ki-kare (χ^2)'dir. Bu istatistiğin manidar çıkması yüzeydeki elemanlar arasında anlamlı farklılıkların olduğunu gösterir (Alharby, 2006).

1.3. Araştırmanın Amacı

Bu çalışmada öğrencilerin öz ve akran puanlamalarının öğretmen veya uzmanların puanlarıyla karşılaştırılmasını zaman içerisindeki değişimlerini de dikkate alarak incelenmesi amaçlanmıştır. Bu amaçla, öz, akran ve öğretmen değerlendirme etkinliklerinin kendi içerisinde puanlama özelliklerinin incelenmesi ve yapılan uygulamaların dikey olarak incelenmesi yollarına başvurulmuştur. Bu doğrultuda öğrencilerin yazma becerilerinin değerlendirilmesi amacıyla 5. sınıf Türkçe dersinde beş haftalık bir değerlendirme programı yürütülmüştür. Bu programda öğrencilerin yazma becerilerini sergileyebilmeleri için sınırlandırılmış görevler verilmiştir. Bu görevlerin değerlendirilmesinde öğrencilerin kendi çalışmalarına ilişkin öz puanlamaları, akranlarının çalışmaları için akran puanlamaları ile öğretmenlerin puanlamaları ÇYRÖM ile analiz edilip yorumlanmıştır.

1.4. Problem Cümlesi

Öğrencilerin yazma becerilerinin değerlendirilmesinde öz, akran ve öğretmen puanlarının zaman içerisindeki değişimleri puanlayıcı güvenirliği açısından Çok Yüzeyli Rasch Ölçme Modeli'yle incelenmiştir.

1.4.1. Araştırma Soruları

Bu çalışmada yukarıda belirtilen amaç ve problem cümlesi doğrultusunda aşağıdaki sorulara yanıt aranmıştır:

1) Yazma becerilerinin değerlendirilmesi amacıyla yapılan birinci, ikinci ve üçüncü uygulamaların analizinde;
   a) Kalibrasyon haritasının sonuçları nedir?
   b) Öğrencilerin yazma becerisi düzeyleri nedir?
   c) DPA'daki kategorilerin puanlanma durumları nedir?
   d) Öz, akran ve öğretmen puanlarınn durumuları nedir?
2) Yazma becerilerinin değerlendirilmesi amacıyla yapılan uygulama sürecinin bir bütün olarak alındığı analizde;
   a) Kalibrasyon haritasının sonuçları nedir?
   b) Öğrencilerin yazma becerisi düzeylerinde zaman içinde bir değişme olmuş mudur?
   c) DPA'daki kategorilerin puanlanma durumlarında zaman içinde bir değişme olmuş mudur?
   d) Öz, akran ve öğretmen puanlarının durumunda zaman içerisinde bir değişme olmuş mudur?
3) Uygulamada ayrı ayrı yapılan analizlerdeki puanlama türleri ile ilgili veriler beraber analize alındığında öz, akran ve öğretmen puanlarının durumu nedir?

1.5. Araştırmanın Önemi

İletişim yeteneklerinin ayrılmaz bir parçası olan yazma becerilerinin hayatımızdaki yeri ve önemi gelişen ve değişen dünya şartlarıyla birlikte daha da büyümekte ve yazma becerileri hayatımızı etkileyen mülakat ve sınavları etkileyebilecek düzeye ulaşmıştır. Yazma becerisi; bireyleri araştırmaya, eksiklerini tamamlamaya, yanlışlarını düzeltmeye yönelttiği için insanın bilgi, zihinsel olgunluk ve tutarlı düşünce alışkanlığı kazanması sağlar (Demir,

2003; Göğüş, 1978 s.235). Eleştirel düşünme ve problem çözme becerilerini geliştirmesi (Demirel, 2003, s.7), kişinin kendisini keşfetmesini sağlaması, öğrencilerin kendi bakış açılarını etkin bir biçimde yapılandırabilmeleri yardımcı olur (Demir, 2003). Daha önce belirtilenlere ek olarak, yazma becerisi akademik başarıyı artıran ve bireylerin tüm yaşamları süresince kullanacakları önemli bir beceri olması bakımından eğitimde önemli bir yeri vardır. Bu becerinin kazandırılması kadar izlemeye dayalı durum belirleme (formative assessment) ve düzey belirleyici durum belirleme (summative assessment) süreçleri ile yürütülmesi de önem taşımaktadır. Bu değerlendirme süreçleri arasında öz, akran ve öğretmen değerlendirme çalışmaları dil yeterliliklerini belirleme için en çok kullanılan yöntemlerdendir (Kutlu vd., 2014). Bu durum fark edilmiş olmalıdır ki bu becerinin değerlendirilmesi konusunda da yurtdışında çokça çalışmalar yapılmış ve sonuçlarıyla alan yazına katkı sağlamak amaçlanmıştır. Fakat yurtdışında yapılan bu konudaki çalışmalar kendi ana dilleri olan İngilizce yazma çalışmalarının yabancı dil öğrenenler ya da anadili İngilizce olan kişiler üzerinde uygulanması şeklinde yapılmış ve genel olarak İngilizce yazma becerilerinin değerlendirilmesi konusunda katkı sağlamıştır (Alharby, 2006; Farrokhi ve Esfandiari, 2011; Macmillan, 2000; Matsuno, 2006; Schaefer, 2008; Sudweeks, Reeve ve Bradshaw, 2005). Türkiye'ye baktığımızda Türkçe yazma becerileri üzerinde durulmuş, fakat değerlendirme yönteminin çeşitleri konusunda sınırlı sayıda araştırmaya rastlanmıştır.

Ek olarak, yurtdışında performansların değerlendirilmesinde sıkça ÇYRÖM'nin kullanımı ile alakalı çalışmaların Türkiye'de oldukça kısıtlı olduğu ve bu konuda daha fazla sayıda çalışmaya ihtiyaç duyulduğu görülmektedir. Bu çalışmada Türkçe yazma becerilerinin değerlendirilmesinde öz, akran ve öğretmen puanları ayrıntılı bir şekilde incelenmeye çalışılmıştır.

Okul uygulamalarında daha önce kullanılmış olan sınırlandırılmış yazma görevleriyle yapılacak ölçümlerinde öz, akran ve öğretmen puanlarının yorumlanmaları açısından öğretmenlere realist ve uygulanabilir bir yöntem göstererek, yazma eğitimi veren öğretmenlerin ve bu konuyla ilgili araştırmacıların yazma becerilerinde uygulanan değerlendirme yöntemlerini ileri bir analiz yöntemiyle inceleme imkânı sağlaması açısından önemlidir. Araştırma soruları ve alt problemlerden elde edilen bulguların ışığında bu çalışmanın okullardaki Türkçe öğretmenlerinin ve ilgili araştırmacıların ÇYRÖM'nin ortaokullardaki Türkçe derslerinde yazma becerilerini değerlendirme ve geliştirme açısından ne gibi yararları olacağını görme fırsatı sağlayacağı söylenebilir.

Araştırma sorularının bulgularıyla öğrenci yazma becerileri, puanlama kategorileri ve puanlayıcı türleri hakkında ayrıntılı bilgiler bulabilecekleri gibi, eğitim ve değerlendirme sürecinde bu süreçleri daha kaliteli hale getirmek için neler yapılması gerektiği hakkında da bilgiler edinilebilmeleri açısından önem taşımaktadır.

Puanlamalar için yapılacak birden fazla uygulamanın aynı yapıda ve ayrı ayrı analizlerle incelenmesi açısından puanlama güvenirliği hakkında ayrıntılı bilgiler sağlayacaktır. Ek olarak uygulama yapılan öğrenciler, puanlayıcılar, kullanılan araçlar ve değerlendirme çeşitleri hakkında yüzeyler ve tek tek bireyler hakkında ayrıntılı bilgiler sağlayarak, bu alanda çalışan kişilere bir seçenek sunacağı söylenebilir. Ayrıca zaman içerisinde öz, akran ve öğretmen değerlendirme türlerindeki farklılaşmaları belirlenmeye ve yorumlanmaya çalışılarak metodolojiye katkı sağlayacağı söylenebilir. Bu çalışmanın yazma becerilerinin puanlanmasında ve değerlendirilmesinde öğrenciler hakkında daha doğru kararların verilmesine de katkı sağlayacağı düşünülmektedir.

1.6. Araştırmanın Sınırlılıkları

1. Yazma becerilerini değerlendirme uygulamaları beş hafta boyunca yapılmıştır. Bu uygulamalardan sürecin başındaki, ortasındaki ve sonundaki uygulamalar analizlere dâhil edilmiştir. Uygulamalar bir bütün olduğundan ve değişimler gözlendiğinden, uygulamaların herhangi birinde mazeretleri ya da devamsızlık nedeniyle katılamayan öğrenciler araştırma kapsamı dışında bırakılmıştır.

2. Araştırma başında iki akran değerlendirme bölümü planlanmışken, ders sürelerinin yetersizliğinden dolayı ancak bir akran değerlendirmesi yapılabilmiştir.

1.7. İlgili Araştırmalar

Bu bölümde, alan yazında ÇYRÖM'nin kullanıldığı çalışmalara kronolojik olarak yer verilmiştir. Bu çalışmalardan bazılarının, çalışma konusunun bir parçası olan öz ve akran değerlendirmelerle doğrudan alakalı olmasa da, farklı çalışmalarda çok yüzeyli Rasch ölçme modelinin nasıl kullanıldığını görme açısından faydalı olabileceği düşünülmüştür.

McNamara ve Adams (1991) çalışmalarıyla ÇYRÖM'nin dil becerisinin değerlendirilmesinde puanlayıcı özelliklerini analiz etmek için kullanabileceğini göstermişlerdir. Süreçte 50 bireyin, iki görevden oluşan Uluslararası İngiliz Dili Ölçme

Sistemi testindeki kâğıtları dört profesyonel puanlayıcı tarafından puanlanmıştır. Birinci görev, bireylerin diyagram veya grafiklerle gösterilmiş bilgiler ile ilgili en az 100 kelimelik bir metin oluşturmasını içerirken, ikinci görev, bir ya da birden fazla okuma parçası ve okuduğunun anlama bölümlerine cevap olarak en az 150 kelimelik bir metin oluşturulmasını içermektedir. Analiz sonuçlarında birinci görev için puanlayıcı katılıkları -1.01 ve 1.03 arasında değişmekteyken, ikinci görevde ise, -0,50 ve 0,57 arasında değişmektedir. Yani iki görev arasında puanlayıcı katılıkları konusunda farklılıklar bulunmaktadır. Diğer analizler için, araştırmacılar iki puanlayıcıdan çifte puanlama yapmasını rica etmişlerdir. Dört puanlayıcının olduğu veriyle 2 puanlayıcının çifte puanladığı veriden alınan her bireyin ortalama yetenekleri hesaplanmıştır. Bütün ve kısmi veri setlerinden elde edilen sonuçlar arasındaki korelasyonlar Fisher'ın z dönüşümü ile doğrusal hale getirilmiş ve grafiklerle karşılaştırılmıştır. Puanlama sayısı ve yetenek kestiriminin kalitesi arasındaki bu doğrusal ilişki ne kadar fazla puanlama olursa, o kadar daha doğru yetenek kestirimlerinin olacağına işaret etmektedir. Sonuç olarak, görevlere göre puanlayıcı katılıklarının değiştiği, ne kadar yeniden puanlama olursa o kadar sağlıklı kestirimler yapılabileceği ve ÇYRÖM'nin puanlayıcı özelliklerini araştırmada kullanışlı bir araç olduğu bulunmuştur.

Gyagenda ve Engelhard (1998) puanlayıcı, kompozisyonlar ve öğrencilerin yazma becerileri arasındaki ilişkiyi ÇYRÖM incelemeyi amaçlayan çalışmasında, rastlantısal olarak 87 kişi içerisinden 20 puanlayıcı seçmiş ve 1993 yılı Georgia Lisesi yazma sınavını okumaları için anlaşmıştır. Her puanlayıcı lise öğrencileri tarafından yazılan 366 kompozisyonun hepsini dört kategoride puanlamıştır: (1) içerik ve organizasyon; (2) stil, (3) geçişler ve (4) cümle kurma. Sonuçlara bakıldığında, puanlayıcı ve görev yüzeyindeki elemanların arasında istatistiksel olarak anlamlı değişimler ve puanlayıcılar ile görevler arasında istatistiksel olarak anlamlı bir etkileşim bulunduğu görülmektedir.

Weigle (1998), ÇYRÖM'ni puanlayıcı eğitiminin bireysel olarak puanlayıcı katılıklarını, puanlayıcı katılıklarının dağılımlarını, puanlayıcıların uyum düzeyini nasıl etkilediğini ve deneyimli ve deneyimsiz puanlayıcıların katılıklarının eğitim öncesi ve sonrası nasıl farklılaştığını araştırmak için kullanmıştır. Sekiz deneyimli yerli İngiliz ve sekiz deneyimsiz yerli İngiliz puanlayıcı ikinci dil olarak İngilizce öğretimi gören öğrencilerin kompozisyonlarını eğitim öncesi ve sonrası puanlamışlardır. Çalışma kapsamında, 60 kompozisyon üç kategori içeren 10'lu dereceleme ölçeği kullanılan İkinci Dil Olarak İngilizce Yerleştirme Sınavı (ESLPE) puanlama ölçeği ile puanlanmıştır. Her kompozisyon

zorunlu puanlayıcı eğitimine katılmış iki puanlayıcı tarafından puanlanmış ve iki puanın ortalaması alınmıştır. Eğitim öncesi ve sonrası elde edilen veriler FACETS programı kullanılarak ayrı ayrı analizlere alınmışlardır. Her analizde dört yüzey bulunmaktadır; bireyler, puanlayıcılar, konular ve kategoriler (içerik, söylemsel kontrol ve dil). Eğitim öncesi ve sonrası analiz sonuçlarında deneyimsiz olan sekiz puanlayıcıdan beşinin deneyimlilerden daha katı olduğu bulunmuştur. Geriye kalan üç deneyimsiz puanlayıcının ise puanlamalarında oldukça tutarsız oldukları görülmüştür. Tüm analizler sonucunda, tüm puanlayıcı grupların katılıkları arasındaki değişimler istatistiksel olarak anlamlı bulunmuştur. Ayrıca eğitimler sonrasında puanlayıcı katılıkları azalmış ve tutarsız olan üç puanlayıcılar tutarlılık konusunda gelişme göstermişlerdir.

Weigle (1999), deneyimli ve deneyimsiz puanlayıcıların ikinci dil olarak İngilizce eğitimi alan öğrencilerin iki farklı konudaki kompozisyonlara nasıl puanladıklarını araştırmaktadır. Nicel analizde puanlayıcı katılıklarını ve tutarlılıklarını gösteren ÇYRÖM kullanılmış ve bir konuda deneyimsiz puanlayıcıların deneyimlere göre daha katı oldukları belirlenirken, diğer konuda böyle bir sonuca ulaşılamamıştır. Ayrıca iki puanlayıcı grubu arasındaki farklılıklar eğitimler sonucunda kaybolmuştur. Puanlayıcıların kompozisyonları puanlarken sesli düşünme protokollerini içeren nitel analizi ise farklılıkların sebeplerine ışık tutmuşlardır. İki konu için hangi DPA uygulanacağına dair ve iki grup puanlayıcının konuların uygunluğu hakkındaki farklılıklar olmak üzere iki çeşit farklılık belirlenmiştir.

Nakamura (2002) çalışmasında Japonya'daki İngilizce sözlü sunum becerilerinin değerlendirmesinde kullanılan öğretmen ve akran puanlarının incelenmesinde ÇYRÖM'nin nasıl yardımcı olabileceği konusunu ele almıştır. Bu kapsamda, 12 mezun üniversite öğrencisinin performansları bir öğretmen ve dört akran tarafından değerlendirilmiştir. Analizde üç yüzey (öğrenciler, maddeler ve puanlayıcılar) bulunmaktadır. Araştırma sonucunda genel olarak üç sonuca ulaşılmıştır; akran değerlendirme öğrencilerin sunumlarını geliştirmeleri için motive etmekte başarılı olabilir, öğrenciler akranlarının puanlanmasında güvenilir puanlayıcılar olabilir ve Rasch modeli puanlayıcılar, öğrenciler ve maddeler hakkında ayrıntılı bilgi sağlamak için kullanışlıdır.

Matsuno (2006) çalışmasında, ÇYRÖM'nin üniversite yazma sınıflarında öz, akran ve öğretmen değerlendirmelerinin karşılaştırılmasında nasıl kullanılacağı konusunu ele almıştır. Çalışma kapsamında 91 öğrenci ve dört yerli Japon öğretmen puanlayıcı bulunmaktadır. Öğrenciler bir konuda 300 kelimelik kompozisyonlar yazmış ve bu kompozisyonlar bireyler, akranları ve dört öğretmen tarafından değerlendirilmiştir.

Sonuçlarda, birçok öz puanlayıcı kendini Rasch modelinin hesaplamasından daha düşük olarak değerlendirmiştir. Bu özellikle başarılı öğrenciler için geçerlidir. Akran puanlayıcılar en cömert puanlayıcılardır. Öz değerlendirmenin kompozisyon yazma öğretimi konusunda kullanışlı bir araç olduğu ve ayrıca akran değerlendirmenin resmi değerlendirme süreçlerinin bir parçası olarak faydalı olabileceği sonuçlarına ulaşılmıştır. Çünkü akran puanlayıcılar kendi içlerinde tutarlılık göstermektedir. Ayrıca ÇYRÖM kullanılarak, öğretmenlerin akran puanlayıcıların yanlılık durumları hakkında bilgilendirerek, daha iyi değerlendirme ölçütlerini geliştirmeleri konusunda yardımcı olabilecekleri belirtilmiştir. Bu iki basamağın akran değerlendirmeleri daha da iyi duruma getireceği konularına değinilmiştir.

Eckes (2008) çalışmasıyla, yazma performansının belirlenmesinde puanlayıcı çeşitlerinin etkisinin sınıflama yaklaşımı ile incelenmesini amaçlamıştır. Çalışmada iki yüzeyli Rasch ve kümeleme analizi kullanılmıştır. Bu kapsamda 64 Alman yerli puanlayıcı TestDaf yazma bölümünü puanlamışlardır. Puanlayıcılar üç yıl boyunca (2002-2004) kontrol edilerek, eğitimler almışlardır. 11 puanlama bölümünde 21 ve 32 arasında değişen puanlayıcı sayısı bulunmaktadır. Sonuç olarak bu çalışmayla, puanlayıcı geçerliğinin puanlayıcı özelliklerinden etkilendiği ve bu modellerle hangi özelliğin puanlayıcı geçerliğine etki ettiği hakkında yeterli bilgi sağlayamayacağı ve bu modellerle özelliklerle geçerliğin arasındaki bağlantının niçin ortaya çıktığını açıklayamayacağı belirlenmiştir.

Güler (2008) çalışmasında 2007 yılında matematik başarısını ölçmek için elde edilen puanlara klasik test kuramı, genellenebilirlik kuramı ve çok yüzeyli Rasch ölçme modelini uygulamıştır. Bu kuramlardan hesaplanan puanların güvenirlilikleri karşılaştırılmıştır. Bu bağlamda 24 soru 203 öğrenci tarafından yanıtlanmış ve cevaplar bütüncül dereceli puanlama anahtarı (holistic rubric) ile dört puanlayıcı tarafından puanlanmıştır. Sonuç olarak, 2007 matematik sınavının öğrenci başarısını belirlemede güvenilir sonuçlar verdiği ve puanlayıcıların ortalamalarında farklılıklar olmasına rağmen, birbirleriyle uyumlu puanlama yaptıkları raporlanmıştır. Bu araştırma başarının ölçülmesinde güvenirliğin belirlenmesinde yararlanılacak kuramların hangisinin seçileceği elde edilen puanların hangi amaç için kullanılacağına bağlı olarak değişebileceği gözlenmiştir.

Schaefer (2008) çalışmasında İngilizceyi yabancı dil olarak öğrenen kişilerin kompozisyonları İngilizceyi ana dil olarak konuşan puanlayıcılar tarafından puanlandığında oluşan puanlayıcı yanlılığı yapılarının çok yüzeyli Rasch modeliyle incelemeyi amaçlamıştır. 40 puanlayıcı Japon bayan öğrenciler tarafından TOEFL

sınavından uyarlanmış bir konu hakkında yazılmış 40 kompozisyonu puanlanmıştır. Altı kategorili bir puanlama anahtarı kullanılmıştır. Analiz sonucunda içerik ve düzeni katı puanlayan puanlayıcıların dil kullanımı ve noktalamayı daha esnek puanladıkları ortaya çıkmıştır bunun yanında puanlayıcıların daha yetenekli öğrencileri daha katı ya da esnek puanlama eğiliminde oldukları görülmüştür

Baştürk (2008) çalışmasında çok yüzeyli Rasch ölçme modelinin lise düzeyindeki bireylerin PowerPoint sunumlarına bağlı olarak performanslarının değerlendirilmesindeki kullanışlılığını belirlemeyi amaçlamıştır. Bu çalışmada sunum becerisi görev zorluğu ve puanlayıcı katılık ve cömertliği yüzeyleri kullanılarak ölçülmüştür. Öğrenciler altı alt gruba ayrılmış ve her alt gruba sunum hazırlayacakları kursun amaçları ve içeriği doğrultusunda bir konu verilmiştir. Yedi puanlayıcı ‘‘A+ PowerPoint dereceli puanlama anahtarı’’ ile her grubu puanlamıştır. Sonuç olarak çok yüzeyli Rasch ölçme modelinin çoklu kategorili verilerle performans ve akran değerlendirmede etkili bir araç olduğu bulunmuştur.

Baştürk (2010), öğrencilerin bilimsel araştırma ile ilgili teorik kazanımlarını göstermek için oluşturdukları bilimsel araştırma ödevlerini çok yüzeyli Rasch ölçme yöntemi ile değerlendirme yoluna gitmiştir. Bu çalışmada öğrencilerin bilimsel araştırma becerileri, puanlayıcıların katılık/cömertliklerine ve davranışı sergileme güçlüğü gibi ölçütlere göre değerlendirilmiştir. Araştırmada Pamukkale Üniversitesi Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Rehberlik ve Psikolojik Danışmanlık anabilim dalında “Bilimsel Araştırma Yöntemleri” dersine devam eden 20 öğrenci çalışma grubunu oluşturmaktadır. Verilerin analizi sonucunda, hazırlanan bilimsel araştırma ödevleri içerisinde “iç uyum ” ve “ dış uyum” istatistiklere sahip ödevlerin olmadığı, ödevlerin istatistiksel olarak anlamlı yaklaşık üç ödev kalite tabakasına ayrılabileceği raporlanmıştır. Puanlayıcıların yapmış oldukları puanlamalarda ise istatistiksel olarak katılık/cömertlik bakımından birbirlerinden farklılıklar göstermediği, yalnızca bir puanlayıcı için “dış uyum” kareler toplamının beklenen değerden farklı olduğu gözlenmiştir.

Farrokhi, Esfandiari ve Dalili (2011), çalışmalarında daha yeni bir ölçme modeli olan çok yüzeyli Rasch modeliyle öz, akran ve öğretmen değerlendirmelerindeki merkezi eğilim etkisini belirlemeyi amaçlamışlardır. Bu amaç doğrultusunda, İran Üniversitesi’nde okuyan ve asıl branşı İngilizce olan İranlı öğrencilerin yazdıkları 188 makaleyi toplamda 194 değerlendirici (188’ i öz ve akran değerlendirici ve 6’sı öğretmen olmak üzere) altı dereceli analitik puanlama anahtarı ile değerlendirmişlerdir. Sonuç olarak, ne grup düzeyinde, ne

de bireysel düzeyde üç çeşit değerlendirmede merkeziyetçiliğin izine rastlanmamıştır. Yine Farrokhi ve Esfandiari (2011) bir önceki çalışmalarının devamı nitelikte olan çalışmalarında bu kez ikinci dil makalelerindeki hale etkisini çok yüzeyli Rasch ölçme modeli kullanarak belirlemeyi amaçlamışlardır. Sonucunda ise, grup düzeyinde değerlendiricilerde hale etkisine rastlanmamıştır. Fakat bireysel düzeyde tüm değerlendirici çeşitlerinde belli bir miktarda hale etkisi görülmüştür.

Semerci (2011)'nin, doktora yeterlikler çerçevesinde öğretim üyesi, akran ve öz değerlendirmelerin Rasch ölçme modeliyle analizini amaçlayan çalışmasında, Rasch ölçme modelinde üç yüzey bulunmaktadır. Bu yüzeyler, 11 jüri (beş öğretim üyesi ve altı doktora öğrencisi) olarak belirtilen değerlendirmeci, 15 yeterlik (Bu yeterlikler YÖK tarafından kabul edilen ve halen geliştirilmeye çalışılan ölçütlerdir) ve değerlendirilen altı doktora öğrencisidir. Çalışmalar sonucunda, belirlenen ölçütler çerçevesinde en yüksek performansı gösteren bir başka deyişle diğerlerine göre P1 ve P2 kodlu öğrenciler daha yeterli iken P5 ve P6 kodlu öğrenciler daha az yeterli olarak görülmüştür. J5 numaralı jürinin en cömert ve J4 numaralı jürinin de en katı değerlendirmeleri yaptıkları tespit edilmiştir. En zor gerçekleşen görev B3 (Özgün ve disiplinler arası çalışmalarda liderlik yapar) ve en kolay gerçekleşen görev ise İ3 (Bir yabancı dili en az Avrupa Dil Portföyü C1 Genel Düzeyinde kullanarak ileri düzeyde yazılı, sözlü ve görsel iletişim kurar ve tartışır)'dür.

Baştürk ve Akın (2012) çalışmalarında, Güzel Sanatlar ve Spor Lisesi öğrencilerinin keman eğitiminde kazanmış oldukları temel becerileri çok yüzeyli Rasch ölçme yöntemi ile değerlendirmeyi amaçlamışlardır. Üç farklı Anadolu Güzel Sanatlar Lisesi'nde "Keman eğitimi" dersine devam eden 27 öğrenci örneklem olarak seçilmiş ve verileri toplamak amacı ile "Keman Çalma Becerilerini Değerlendirme Ölçeği" kullanılmıştır. Sonuç olarak müzik eğitiminde özellikle performansa dayalı öğrenmelerin değerlendirilmesinde Çok-Yüzeyli Rasch ölçme modelinin etkin olarak kullanılabileceği gözlenmiştir.

Karakaya, Sarıtaş ve Salmaner (2015) istatistik dersi kapsamında öğrencilerin grup olarak yaptıkları görevlerden oluşan ürünlerini her öğrencinin öz puanlamaları hem de üç öğretim elemanı tarafından puanlanmalarını çok yüzeyli Rasch ölçme modeli ile incelemişlerdir. Araştırmanın çalışma grubunu 2014-2015 eğitim ve öğretim yılı, bahar dönemi Gazi Üniversitesi Eğitim Fakültesi Psikolojik Danışmanlık ve Rehberlik Bölümü 2. sınıf öğrencilerinden istatistik II dersini alan 34 öğrenci oluşturmaktadır. Araştırmaya ait görevler, öncelikle 34 öğrenci tarafından dereceli puanlama anahtarı ile kendileri

puanlanmış daha sonra da aynı görevler üç öğretim elemanı tarafından puanlanarak analize dâhil edilmiştir. Çalışma sonucunda, öğrencilerin genel olarak araştırma görevlisi ve öğretim elemanına oranla kendilerini daha cömert puanladıkları görülmüştür. Ayrıca Dereceli puanlama anahtarı bütün olarak amacına hizmet ettiği, bu nedenle istatistik dersi kapsamında uygulamalı derslerde görevlerin değerlendirilmesinde rahatlıkla kullanılabileceği ve puanlayıcıların bazı kategori ve bireylerde yaptıkları yanlılıkların nitel çalışmalar ile daha ayrıntı olarak incelenebileceği belirlenmiştir. Son olarak da ÇYRÖM'nin derste yapılan görev uygulamaları ve puanlayan puanlayıcılar hakkında detay olarak bilgi edinmek için kullanabilir sonucuna ulaşmışlardır.

Yüzüak, Yüzüak ve Kaptan (2015) çalışmalarında Genetik ve Biyoteknoloji konulu performans görevinin (poster) akran grupları ve öğretmen tarafından çok-yüzeyli Rasch ölçme modeliyle değerlendirmeyi amaçlamaktadırlar. Araştırmada tarama modeli kullanılmıştır. Araştırma 2013-2014 eğitim yılı güz döneminde Bartın ilinde çalışan bir öğretmen ve 7. sınıfta öğrenim gören 50 öğrenci ile yürütülmüştür. Değerlendirme için beşer kişilik 10 öğrenci grubu oluşturulmuştur. Her grup kendi içerisinde birer materyal hazırlamıştır. Çalışma kapsamında, çok-yüzeyli Rasch ölçme modelinin fen eğitiminde performans ve akran grup değerlendirmeleri için etkili bir şekilde kullanılabileceği sonucuna ulaşılmıştır.

Karakaya (2015) çalışmasında portfolyoların değerlendirilmesinde öz, akran ve öğretmen değerlendirmelerinin karşılaştırılmasını amaçlamıştır. Çalışma grubu 74 stajyer öğretmenden oluşmaktadır. Çalışma kapsamında ÇYRÖM kullanılmış ve dört yüzey (öğretmenlerin başarısı, öğretmenlerin cinsiyetleri, öz, akran ve öğretmen puanlayıcı katılığı ve portfolyo değerlendirme kategorileri) oluşturulmuştur. Başarılı stajyer öğretmenlerin daha iyi portfolyo performanslarına sahip oldukları, bayan stajyerlerin erkeklere oranla daha başarılı oldukları bulunmuştur. Öz değerlendirmelerin en esnek ve akran değerlendirmelerin en katı olduğu ve aralarında istatistiksel olarak anlamlı farkın bulunduğu gözlenmiştir. Bu yöntemin üniversitelerde öz, akran ve öğretmen puanları ile uygulamalarda ayrıntılı bilgiler sağlayan ve pratik bir yöntem olduğu sonucuna varılmıştır.

BÖLÜM II

YÖNTEM

Bu bölümde, araştırmanın modeli, çalışma grubu ve veri toplama araçlarına ilişkin bilgiler verilmiştir. Ayrıca bu bölüm verilerin analizi başlığı altında; Çok Yüzeyli Rasch Ölçme Model veri uyumunun kontrolü, birinci, ikinci, üçüncü ve sürecin bir bütün olarak alındığı analizlerin nasıl yapıldığını anlatan kısmı içermektedir.

2.1 Araştırmanın Modeli

Bu çalışma iki temel aşamadan oluşmaktadır. Araştırmanın birinci aşamasında, öz, akran ve öğretmen puanlarının her uygulamadan sonra ayrı ayrı ÇYRÖM analizi yapılmış; ikinci aşamada ise bu uygulamalardan elde edilen veriler ortak bir model oluşturularak beraber olarak genel bir ÇYRÖM analizi yapılarak veriler elde edilmiştir.

Bu çalışma uygulanan ölçme araçlarından elde edilen verilerin analiz edilmesiyle verilen bir durumu olabildiğince tam ve detaylı bir şekilde tanımlayan ya da açıklayan betimsel düzeyde bir araştırmadır. Betimsel araştırmacıların amacı verilen bir durumu mümkün olan en dikkatli ve tam şekilde tanımlamaktır (Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel, 2014, s.15).

2.2. Çalışma Grubu

Çalışma grubunu performans görevlerinin uygulandığı Ankara'nın Keçiören ilçesindeki bir devlet okulunda 5-A ve 5-C şubelerinde okuyan 32 kız ve 38 erkek öğrenciden oluşan toplam 70 ilköğretim 5. sınıf öğrencisi oluşturmaktadır. Fakat yapılan çalışmanın bütünlüğü açısından beş haftalık uygulamalardan herhangi birisine katılamayan öğrenciler

analizden çıkartılarak çalışma grubu 60 kişiye indirilmiştir. Çalışma grubunun cinsiyet ve sınıf değişkenlerine göre dağılımları Tablo 2'de verilmiştir.

Tablo 2. Çalışma Grubunun Cinsiyet ve Sınıf Değişkenlerine Göre Dağılımı

	Çalışma Grubu	
Değişkenler	**N**	**%**
Cinsiyet		
Kız	30	50
Erkek	30	50
Sınıf		
5-A	36	60
5-C	24	40
Toplam	60	100

Tablo 2 incelendiğinde, analize dâhil edilen öğrencilerin 30'unun kız, geri kalan 30'unun ise erkek olduğu ve çalışma grubunun %50'sının erkek ve diğer %50'sinin de kızlardan oluştuğu görülmektedir. Ayrıca öğrencilerin 36'sı (grubun %60'ı) A şubesinde, 24'ü ise (grubun %40'ı) B şubesindendir.

Çalışma grubunun alındığı Keçiören ilçesinin belirlenmesinde Türkiye'nin başkenti olan Ankara'da çok yüksek ve düşük gelir durumuna sahip ailelerden daha çok çoğunlukla orta gelir durumuna sahip ailelerin burada yaşaması göz önünde bulundurulmuştur. Ayrıca orta başarı düzeyinden bir devlet okulu olması nedeniyle de tercih edilmiştir. Öğrencilerin düzeylerinin belirlenmesiyle ilgili olarak 6. 7. ve 8. sınıf öğrencilerinin ve velilerinin lise giriş sınavları konusunda baskı altında olmaları ve ortaokula yeni başlayıp uyum sürecini geçmiş olduklarından dolayı 5. sınıf öğrencileri seçilmişlerdir. Sınıfların seçilmesinde aynı başarı düzeylerine sahip olmaları ve aynı öğretmen tarafından Türkçe dersi almaları göz önünde bulundurulmuştur. Bu iki sınıfın Türkçe öğretmenin sınıflarda daha önce hiç öz ve akran değerlendirme çalışması yaptırmamış olmasının öğrencilerin ilk ölçümü ve son ölçümündeki farkı görmede yarar sağlayacağı düşünülmüştür.

Puanlayıcılara bakılacak olursa, öz ve akran değerlendirme sürecine katılan yukarıda sözü edilen yine aynı 60 öğrenci ve iki kadın ve bir erkek olan üç puanlayıcıdan (R1-R3) oluşturulmuştur. Her öğrenci öz değerlendirme kapsamında kendi çalışmasını değerlendirdikten sonra, akran değerlendirme kapsamında bir arkadaşını değerlendirmiştir.

Puanlayıcıların seçilmesi konusunda ise, yazma uygulamalarını daha önce uygulamış iki Eğitimde Ölçme ve Değerlendirme uzmanları ile bir Türkçe öğretmenin puanlamaya dâhil edilmesi uygun görülürken, daha güvenir ve ayrıntılı bilgi sağlamak amacıyla üç puanlayıcı seçilmiştir.

Çalışmaya katılan tüm öğrenci ve öğretmenler elde edilen verilerin bilimsel bir çalışmada kullanılacağı konusunda bilgilendirilmiştir. Ayrıca öğrencilere motivasyon olması amacıyla teşekkür notları ve betin fırtınası ve kavram haritası içeren çeşitli aktivitelere yer verilmiştir. Buna ek olarak, öğrencilerin yazdıkları yazıların kompozisyon notu olarak az da olsa karne notlarına etkide bulunacağı belirtilmiş ve öğrenciler, eğer yazılarının çalışmada kullanılmasını istemezlerse, öğretmeni bilgilendirerek kompozisyonlarının çalışmaya dâhil edilmeyeceği konusunda uyarılmıştır. Öğretmenler çalışmanın kapsamı ve amacı konusunda bilgilendirilmiş ve gönüllülük esas alınmıştır. Bu uygulamalarla öğretmen ve öğrencilerin çalışmadaki motivasyonu ve ciddiyeti arttırılmaya çalışılmıştır.

2.3 Veri Toplama Araçları

Öğrencilerin yazma becerilerini ölçmek amacıyla araştırmacı tarafından görsel okuma ve yaratıcı yazma konularında yedi tane sınırlandırılmış yazma görevi geliştirilmiştir. Bu görevlerin değerlendirilmesi amacıyla öz, akran ve öğretmen değerlendirmelerinde kullanılmaları için analitik ve göreve özel dereceli puanlama anahtarları (DPA) oluşturulmuştur. Puanlama sürecine katılacak öğretmen ve öğrencilere öncesinde konu hakkında bilgi verilmiştir. Konu hakkındaki soruları araştırmacı tarafından örneklerle açıklanmıştır. Aşağıda performans görevleri ve DPA'ların geliştirilmesi süreci daha ayrıntılı bir şekilde açıklanmıştır.

2.3.1. Yazma Görevlerinin Geliştirilme Süreci

Çalışma kapsamında araştırmacı tarafından öğrencilerin yazma konusunda performanslarının en iyi gözlemlenebileceği ders olan Türkçe seçilmiştir ve araştırmacı tarafından öğrencilerin görsel okuma, betimleyici yazma ve çıkarımlarda bulunma becerilerini ölçen, farklı konularda yedi tane performans görevi geliştirilmiştir. Geliştirme sürecinde öncelikle öğrencilerin ilgisini çekecek, yüksek motivasyonla katılımlarını sağlayacak ya da arttıracak konuların seçilmesine önem verilmiştir. Görevlerin türü ile ilişkili olarak İlköğretim Türkçe Dersi (6, 7, 8. Sınıf) Öğretim Programı incelendiğinde yazma

kazanımlarının aslında metin türleriyle ilgili daha genel bir sınıflamaya bağlı kalınarak düzenlendiği görülür: Olay yazıları, düşünce yazıları, bildirme yazıları, şiir (Arı, 2011). Öğrencinin daha çok yazma imkânı bulabileceği ve görevin konusu ile de ilgili olarak düşünce, olay ve bildirme yazı türleri seçilmiştir. Görevlerin eş değer zorlukta ve konuların birbirine benzer olmasına dikkat edilmiş ve teknik konulardan kaçınılmaya çalışılmıştır. Performans görevleri geliştirilirken, Kutlu vd. (2014) kullandıkları geliştirme aşamaları dikkate alınmıştır. İlk olarak ölçülecek zihinsel süreçler belirlenmiş ve yaşamdaki olaylarla ilişkilendirilmiştir. Daha sonra görev kısımları oluşturulmuştur. Performans görevleri, öğrencilerin yazma becerisini ortaya çıkaracak şekilde ayrıntılı ve görsellerden faydalanarak oluşturması gereken durumlar olarak tasarlanmıştır. Üçüncü olarak yönergeler hazırlanmıştır. Her performans görevine ait gerekli açıklamalar da görevin altında verilmiştir (EK 2, 3, 4, 5, 6). Görevlerin uygulanma sürecinde dikkat edilmesi gereken noktalar sözel olarak açıklanmıştır. Ve en son puanlama yöntemi belirlenmiştir.

2.3.2. Dereceli Puanlama Anahtarlarının Geliştirilme Süreci

DPA geliştirilirken Goodrich (2001)'in önerdiği basamaklar kullanılmıştır:

1. *Performansı belirlemede kullanılacak ölçütlerin listelenmesi:* Yazma becerilerinin belirlenmesinde yedi görevin farklı görev içinde altı ölçüt belirlenmiştir. Bu ölçütler görevlerde öğrencinin yerine getirmesi istenen performanslar dikkate alınarak belirlenmiştir.
2. *Kullanılacak dereceli puanlama anahtarına karar verilmesi:* Performans görevlerinin yapısına uygun olması ve kompozisyonların ayrıntılı puanlanması istendiğinden araştırmada "analitik ve göreve özel puanlama anahtarları" geliştirilmiştir. DPA'ların tümünde, görev türleri benzer olduğundan genel ölçütler (yazım, anlatım, düzen, çıkarımda bulunma, görsellerden faydalanma ve bütünsellik) aynıdır.
3. *Performans düzeylerinin belirlenmesi ve düzey tanımlamalarının yapılması:* En yüksek ve en düşük performans düzeyleri belirlenmiştir. Başarı düzeylerinin tanımları, görevlere uygun olarak yazılmıştır. Bu düzeyler öğrencilerin not kaygısı duymamaları ve arkadaşlarıyla rekabet ortamı oluşturmamaları için düzeylere puanlar yerine belli bazı sıfatlarla (geliştirilebilir-G-, iyi-İ- ve

mükemmel-M-) belirtilmiştir. Buna göre her ölçüt G ile M arasında derecelendirilmiştir.

4. *Uzman görüşünün alınması:* Dereceli puanlama anahtarı hazırlandıktan sonra ilgili alan öğretmenlerinden, Ölçme ve Değerlendirme uzmanlarından ve Türk dili konusunda çalışan uzmanlardan görüşler alınmıştır. Bu konu "Geliştirilen Ölçme Araçlarının Geçerlik ve Güvenirliğine İlişkin Toplanan Kanıtlar" başlığı altında daha ayrıntılı açıklanmıştır.

2.3.3. Geliştirilen Ölçme Araçlarının Geçerlik Ve Güvenirliğine İlişkin Toplanan Kanıtlar

Geliştirilen ölçme araçlarının geçerlik ve güvenirliğine ilişkin kanıtlar ilk olarak uzman görüşleri alınarak sağlanmaya çalışılmıştır. Daha sonra düzenlenen Görev ve Dereceli Puanlama Anahtarı Değerlendirme Formu (EK 1) ile birlikte farklı dört uzmana gönderilmiştir. Elde edilen verilerle performans görevlerinin kapsam geçerlik indeksleri hesaplanmıştır. Geçerlik ve güvenirliğine ilişkin kanıt toplama aşamaları ve tüm uzmanların uzmanlık alanlarına göre dağılımları Tablo 3'te gösterilmektedir.

Tablo 3. Geçerlik ve Güvenirliğe İlişkin Kanıt Toplama Aşamaları ile Uzmanların Alanlarına Göre Dağılımları

	Uzmanlar
Alanlar	**N**
Aşama 1: Uzman görüşü	
Türkçe Öğretmeni	1
Ölçme ve Değerlendirme Uzmanı	2
Aşama 2: Kapsam Geçerlik İndeksi	
Türk Dili Eğitimi alanında Öğretim Üyesi	1
Ölçme ve Değerlendirme Uzmanı	3
Toplam	7

Tablo 3'te görüldüğü gibi, yedi uzmanın beşi Eğitimde Ölçme ve Değerlendirme Anabilim Dalı'nda uzman iken, biri Türkçe Öğretmeni ve birisi de Türk Dili Eğitimi alanında öğretim üyesidir. Ayrıca, uzmanların beşi doktora düzeyinde, biri yüksek lisans ve birisi de

lisans düzeyindendir. Meslekteki tecrübe yıllarına bakıldığında, uzmanların tecrübelerinin 21 ve 3 yıl arasında değiştiği görülmektedir.

Uzman Görüşü: Geliştirilen yedi performans görevi ve dereceli puanlama anahtarı Türkçe dil kurallarına uygunluk, bilimsel doğruluk, ölçme ve değerlendirme ilkelerine uygunluk, yaş düzeyine uygunluk ve görevlerin eş güçlük düzeyinde olup olmadıklarını konularında bir Türkçe öğretmeni ve iki Ölçme ve Değerlendirme uzmanı olmak üzere üç uzmandan yazılı görüşleri alınmıştır. Alınan görüşler doğrultusunda görevler düzenlenmiş ve görevlerin sayısı yediden beşe indirilmiştir.

Kapsam Geçerlik İndeksi: McGartland, Berg-Weger, Tebb, Lee ve Rauch (2003)'e göre ölçme aracının ölçülmesi amaçlanan özelliği kapsama gücünü (kapsam geçerliği) belirlemek amacıyla önsel çalışmaların yapılmasına ihtiyaç vardır. Düzenlenen görevlerin ve DPA'ların da geçerlik ve güvenirlik kanıtlarını elde etmek amacıyla Kapsam Geçerlik İndeksi (KGİ) hesaplanmıştır. Kapsam geçerlik oranları, Lawshe (1975) tarafından geliştirilmiştir ve bu yöntem uzmanlara 3'lü Likert türde ölçek tipi olan formlar verildiğinde kullanılmaktadır. Kapsam geçerliği konusunda kullanılan en yaygın teknik Lawshe tekniği olmasına karşın başka teknikler de geliştirilmiştir. Bunlardan Davis (1992) tekniğinde uzman görüşlerini (1) "uygun" , (2) "madde gözden geçirilmeli" , (3) " madde ciddi olarak gözden geçirilmeli" ve (4) "madde uygun değil" gibi dörtlü derecelendirmektedir.

Aşağıdaki formülde Kapsam Geçerlik İndeksi'nin nasıl hesaplandığı gösterilmektedir.

$$KG\dot{I} = {}^{N_G}/_{N_U}$$

N_G= Gerekli diyen (1 veya 2 veren) uzman sayısı
N_u= Tüm uzmanların sayısı

Formülde görüldüğü gibi, (1) ve (2) seçeneğini işaretleyen uzmanların sayısı toplam uzman sayısına bölünerek maddeye ilişkin "kapsam geçerlik indeksi" elde edilmektedir (Yurdugül, 2005). KGİ için kabul edilebilir değer 0,80 olarak belirlenmiş ve bu değerin

altındaki maddeler ölçme aracından çıkartılabilir (Davis, 1992). Çıkarılmayan maddelerin ortalaması alınarak ölçme aracı için genel bir KGİ elde edilir. Bunun için de yine kapsam geçerlik ölçütü (KGÖ) 0,80'dir. Bu teknikte en az 3 en fazla 20 uzman önerilmektedir (Yurdugül, 2005). Bu çalışmada uzman görüşleri sonrasında düzenlenen görev ve dereceli puanlama anahtarları "Görev ve Dereceli Puanlama Anahtarı Değerlendirme Formu" ile birlikte farklı dört uzmana gönderilmiştir. Görev ve Dereceli Puanlama Anahtarı Değerlendirme Formu araştırmacı tarafından düzenlenmiş olup, demografik bilgilerin olduğu bölüm bulunmaktadır. Değerlendirme formu, performans görevlerinde ve DPA'da olması gereken uygunluğuna bakılacak 10 durum ve her durumun altında öneriler kısımları içermektedir. Ayrıca formun en altında genel önerilerin yazılabileceği bir kısım bulunmaktadır. Formda, 4'lü Likert tipi ölçek türü kullanılmakta ve (1) çok uygun, (2) uygun, (3) az değişiklik gerekiyor (önerdiğim gibi), son olarak (4) çok değişiklik gerekiyor (önerdiğim gibi) anlamına gelmektedir. Görev ve DPA değerlendirme formunda 4'lü Likert tipi dereceleme kullanıldığından, çalışmada Davis tekniğinden faydalanılmıştır. Beş performans görevi ve DPA'ları için hesaplanan KGİ'ler Tablo 4'te gösterilmiştir.

Tablo 4. Performans Görevlerine Ait ve DPA'lara Ait Kapsam Geçerlik İndeksleri

Görevler	KGİ
Görev 1	0,80*
Görev 2	0,85*
Görev 3	0,80*
Görev 4	0,83*
Görev 5	0,85*

* KGÖ (0,80)'e eşit veya 0,80'den büyük olanlar

Tablo 4'te görüldüğü gibi, beş görevin ve içlerinde yer alan puanlama anahtarlarının kapsam geçerlik indeksleri 0,80 ve 0,85 arasında değişmektedir ve görevlerin tümü KGÖ olan 0,80' e eşit ya da ondan büyüktür. Bu değerlere bakılarak ölçme araçlarının geçerliğine ilişkin yeterli kanıtlar toplandığına karar verilmiş ve veri toplama sürecine geçilmiştir.

2.3.4. Veri toplama süreci

Çalışma, 2014-2015 eğitim ve öğretim yılı bahar döneminin ikinci haftasından yedinci haftasına kadar olan beş hafta boyunca Ankara'nın Keçiören ilçesindeki bir devlet okulunda yürütülmüştür. Çalışmanın uygulanması için alınan izin EK 7'de gösterilmiştir. İlk hafta öğrencilere yazacakları yazılar, yazıların dereceli puanlama anahtarları ile öz ve

akran değerlendirmeleri hakkında bilgilendirmeler yapılmış ve bazı örneklerle ve sorularla anlaşılmayan kısımlar kapatılmıştır. İkinci hafta ve sonrasında öncelikle önceki haftanın tekrarı yapılmış daha sonra ısınma etkinlikleri ve konu hakkında beyin jimnastikleri yapılmıştır. Yapılan egzersizler sonunda öğrencilere verilen görev ve yönergelere uygun olarak konuyla alakalı yazılar yazmaları için 15-20 dakika süre verilmiştir. Verilen görevler birinci haftadan son haftaya kadar sıralanacak şekilde EK 2, 3, 4, 5 ve 6' da gösterilmiştir. Yazmak için verilen süre bittiğinde öğrencilerden öz değerlendirme amacıyla kullanıcakları dereceli puanlama anahtarındaki kategorileri ve kategorilerin dereceleri altındaki tanımları okuyup iyice anlamaları daha sonra kendi yazılarını okuyarak derecelerden uygun olanları işaretlemeleri istenmiştir. Öz değerlendirme kısmına yaklaşık olarak 10 dakika ayrılmış ve akran değerlendirme kısmına geçilmiştir. Bu kısımda ise öğrencilerin kâğıtları toplanmış ve karıştırılarak sınıftaki başka bir arkadaşına gelecek şekilde tekrar dağıtılmıştır. Öğrencilerin değerlendirdikleri arkadaşlarını tanımamaları ve yanlı davranmamaları için kâğıtlara isimleri yazılmamış, araştırmacının ayırt edebilmesi amacıyla sadece okul numaraları yazılmıştır. Akran değerlendirme kısmına da yaklaşık olarak 15 dakika verilmiştir. İlk haftalarda bu süreçler sona erdiğinde kalan süreye göre öğrencilerin kendi ve arkadaşlarını değerlendirmelerinde sıkça karşılaştıkları problemler ve bunların nasıl giderilebilecekleri hakkında konuşulmuştur. Zaman zaman bu konular hakkında örnekler verilmiş ve toplanan görüşler şema haline getirilerek kavram haritaları oluşturulmuştur. Son haftada görev olarak öğrencilerin "uygulamanın ilk görevinden son görevine kadar olan süreçte yazma ve değerlendirme konularında kendilerinde meydana gelen değişiklikleri" evde kompozisyon şeklinde yazmaları istenmiştir. Son haftalarda ise yazılan kompozisyonlar toplanmış ve isimsiz olarak bazı yazılar okunmuştur. Yazılar okunduktan sonra öğrencilerin puanlama anahtarına uygun olarak kompozisyonların olumlu ve olumsuz yönlerini, bunların sebepleri ve nasıl giderilebileceği hakkındaki düşüncelerini paylaşmaları istenen bölümler düzenlenmiştir. Her hafta öğretmen puanlamasına göre belirlenen objektif puanlayıcı öğrencilere motive edici notlar verilerek, öğrencileri motive edici etkinlikler düzenlenmiştir. Öğrencilerden toplanan görevler her hafta uygulamayı yapan öğretmen dışındaki iki öğretmen tarafından da puanlanmıştır. Kompozisyon örneği EK 12'de ve Puanlama örnekleri EK 13'te gösterilmiştir. Ayrıca öğrencilerin uygulama ile ilgili görüş örnekleri EK 14'te sunulmuştur.

2.4. Verilerin Analizi

Verilerin analizi üç yardımcı analizden ve bir genel analizden oluşmaktadır. Bu süreç aşağıdaki aşamalar doğrultusunda ilerlemiştir:

1. Toplanan verilerin ilk analizleri için üç yüzey (birey, kategori ve puanlayıcı türü) belirlenmiş ve veriler yüzeylere göre düzenlenmiştir. Puanlayıcı yüzeyinde öğretmenlerin yanında öğrenciler de kodlanmıştır. Ayrıca, puanlayıcı türü yüzeyinde öz, akran ve öğretmen puanları ayrı ayrı kodlanmıştır.
2. Bir bütün olarak alınan analiz için üç yüzeye ek olarak "görev" yüzeyi de eklenmiştir ve veriler bu yüzeylere göre tekrar düzenlenerek üç analizin verileri birleştirilmiştir. Analizde kullanılan yüzeyler ve bu yüzeylerin içlerinde yer alan elemanlar Tablo 5'te gösterilmiştir.

Tablo 5. Analizlerde Kullanılmak İçin Tanımlanan Yüzeyler, Elemanları ve Sayıları

Yüzeyler	Elemanları
Birey (60)	1, 2, 3, ..., 60.
Puanlama Kategorisi (6)	Yazının Bütünselliği, Anlatım, Yazım ve Dil Bilgisi, Düzen, Görsellerden Faydalanma, Çıkarımda Bulunma.
Puanlayıcı Türü (3)	Öz, Akran, Öğretmen.
Görev (3)	İlk, Orta, Son.

3. Model veri uyumuna dikkat edilerek, tüm araştırma sorularındaki veriler Çok Yüzeyli Rasch Ölçme Modeli'yle analiz edilmiştir ve analizler için FACETS (Linacre, 1989-1998) paket programı kullanılmıştır. Son olarak, FACETS programının çıktıları yorumlanmıştır: (a) kalibrasyon haritası, (b) öğrencilerin yetenek ölçümleri ve uyum istatistikleri, (c) DPA'daki kategorilerin puanlanma katılık hesaplamaları ve uyum istatistikleri, (d) puanlayıcıların katılık hesaplamaları ve uyum istatistikleri ve bir bütün olarak alınan analizde ek olarak, (e) görevlerin kolaylık zorlukları hesaplamaları ve uyum istatistikleri. Kalibrasyon haritası yüzeylerdeki farklı elemanların arasında ortaya çıkabilecek farklılıklar hakkında görsel bilgi etmeyi sağlayabilir. Öğrenci

kategorilerinin güçlüğünün lojit ölçümleri ve puanlayıcı türünün katılığı lojit ölçümleri ile eş zamanlı olarak hesaplamaktadır ve bu ölçümleri aynı doğrusal ölçek üzerine yerleştirerek, sonuçların kolayca karşılaştırabilmesi sağlanabilir. Rasch modeli, ayrıca puanlayıcıların puanlama şekillerinin tutarlılıklarının hesaplanmasını sağlayan uyum istatistikleri içermektedir (Lunz ve Linacre, 1998).

2.4.1. Model Veri Uyumu

Model değişken kaynakları arasında bir etkileşim olmadığını varsaymaktadır (Alharby, 2006). Ayrıca bölüm başında bahsedildiği gibi model ve verinin uyum sağlaması gerekir. Veri modelle uyumlu olduğunda, her bir artık değerin (residual) sıfır olması beklenir. Veri modele yaklaştığı derecede, standartlaştırılmış artığın kestirimi, ortalaması 0 ve varyansı yaklaşık 1 olacak şekilde az ya da çok normal dağılım gösterir. Bu sebeple, standartlaştırılmış artıkların yaklaşık olarak normal dağılım sergileyip sergilemediğine bakılarak verinin modele uyumu hakkında yorum yapılabilir (Hertherman, 2004). Başka bir deyişle, verilerin modelle olan uyumu için standartlaştırılmış artık değerlerin yaklaşık en fazla %5'i ±2'nin dışında ve yaklaşık en fazla %1'inin de ±3'ün dışında olması gerekmektedir (Linacre, 2002). Bu çalışmada birinci analiz için yazma becerilerinin ölçülmesinden elde edilen 1636 verinin 47'si (%2,90) ±2'nin dışında ve 2'si (%0,12) ±3'ün dışında yer almıştır. İkinci analiz için yazma becerilerinin ölçülmesinden elde edilen 1636 verinin 44'ü (%2,70) ±2'nin dışında ve 2'si (%0,12) ±3'ün dışında yer almıştır. Üçüncü analiz için yazma becerilerinin ölçülmesinden elde edilen 1636 verinin 47'ü (%2,90) ±2'nin dışında ve 4'ü (%0,24) ±3'ün dışında yer almıştır. Dördüncü analizde yazma becerilerinin ölçülmesinden elde edilen 5168 verinin 94'ü (%1,81) ±2'nin dışında ve 6'si (%0,11) ±3'ün dışında yer almıştır. Son olarak beşinci analizde ise 5168 verinin 92'si (%1,78)±2'nin dışında ve 8'i (%0,15) ±3'ün dışında yer almıştır.

BÖLÜM III

BULGULAR

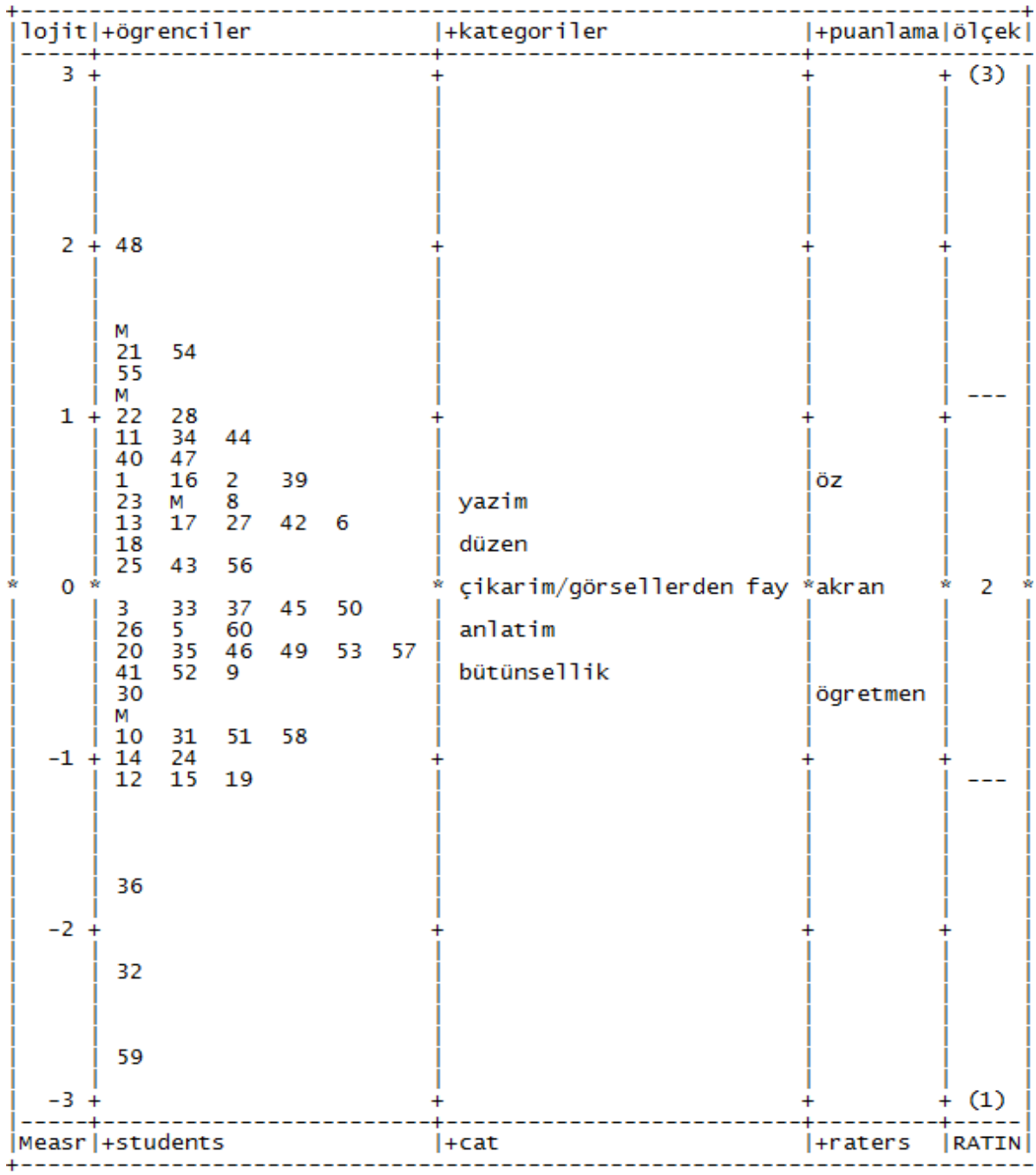
Bu bölümde her bir araştırma sorusu ve alt probleme ilişkin bulgulara yer verilmiştir.

3.1. Birinci Araştırma Sorusuna İlişkin Bulgular

3.1.1. Alt Problem 1.a'ya İlişkin Bulgular

Birinci analizinde kalibrasyon haritası sonuçları nedir?

Puanlayıcı katılıkları, öğrenci yazma becerileri ve kategorilerin puanlanma katılıklarını aynı ölçek üzerinde gösteren kalibrasyon haritası Şekil 1'de gösterilmektedir. Şekil 1'de birinci sütun eşit aralıklı yapıya sahip olan Rasch lojit ölçeğini göstermektedir. İkinci sütunda ise öğrenci beceri düzeyleri görülmektedir. İkinci sütunda her sayı bir öğrenciyi sembolize ederken, "M"ler ise 4, 7, 29 ve 38 numaralı uyumsuz (misfit) öğrencileri göstermektedir. Öğrencilerin beceri düzeyleri şeklin en üstünden altına doğru azalmaktadır. Bu açıklamaya göre Şekil 1'de en yüksek yazma becerisi 48 numaralı öğrencideyken, en düşük yazma becerisinin ise 59 numaralı öğrencide olduğu görülmektedir. Öğrencilerin çoğunun orta başarı düzeyinde toplandığı ve uyumsuz öğrencilerin genelde yüksek başarı düzeyine sahip olduğu söylenebilir.



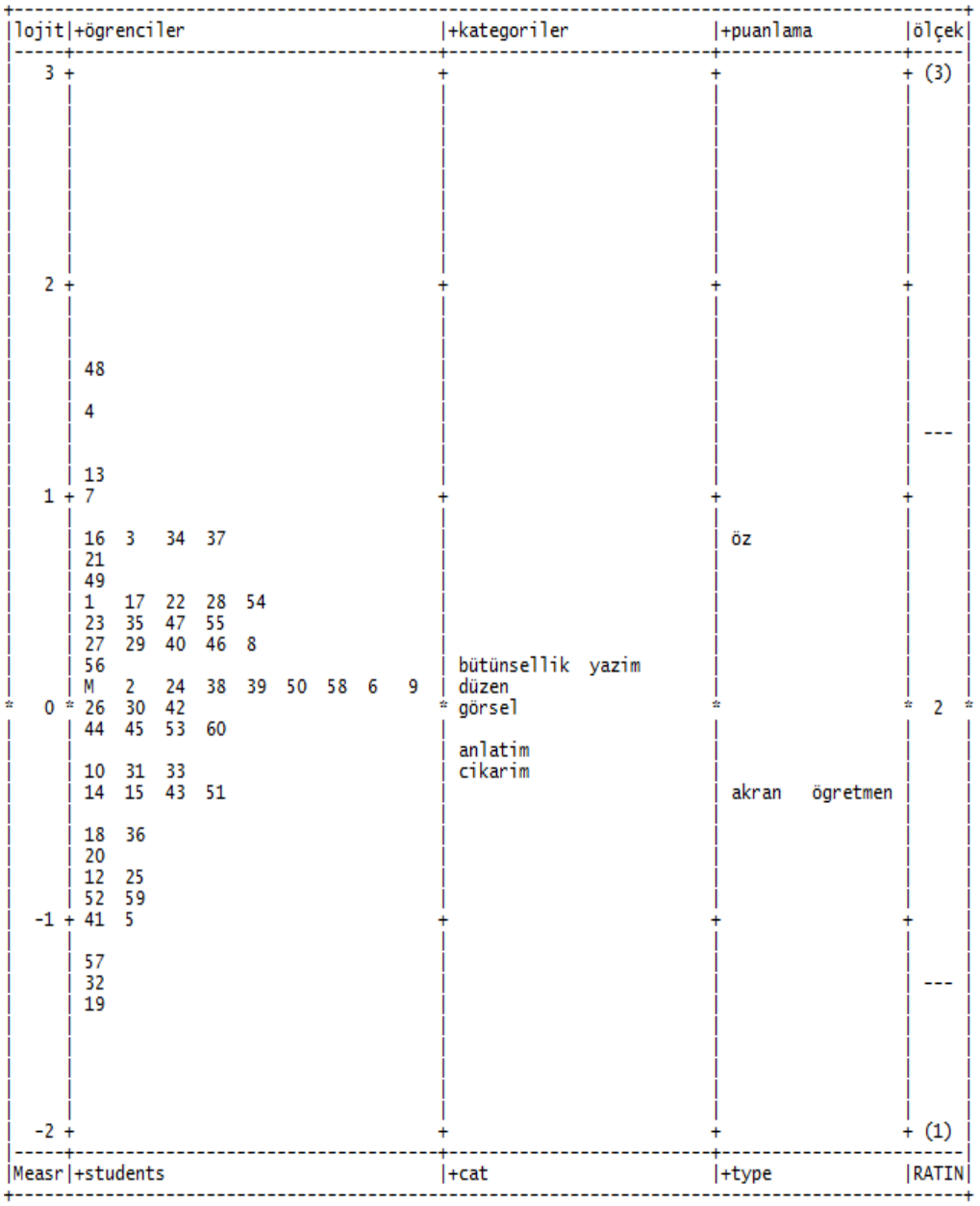
Şekil 1. Birinci analizdeki öğrenci becerisi, kategori puanlama katılığı, puanlayıcı türü katılığı ve Likert ölçek işlevinin kalibrasyon haritası

Şekil 1'de üçüncü sütun, kategorilerin puanlanma katılıklarını göstermektedir ve en cömert puanlanan kategori "yazım" en üstte, en katı puanlanan "bütünsellik" ise en altta yer almaktadır. "Çıkarımda bulunma" ve "görsellerden faydalanma" gibi kategorilerden bazıları da aynı puanlanma katılık düzeyine sahiptir. Dördüncü sütun puanlayıcı türünün katılığını göstermektedir ve öz değerlendirmelerin en cömert düzeye ve öğretmen değerlendirmelerin en katı düzeye sahip olduğu görülmektedir. Akranlar ise orta düzeyde

katılığa sahiptir. Son sütunda ise analizde kullanılan DPA'daki 3'lü Likert tipi ölçek türünde yer alan her bir noktayı göstermektedir.

İkinci analizinde kalibrasyon haritasının sonuçları nedir?

Yazma becerilerinin değerlendirilmesi amacıyla yapılan uygulamaların ikinci analiz sonuçları Şekil 2'de gösterilmiştir.

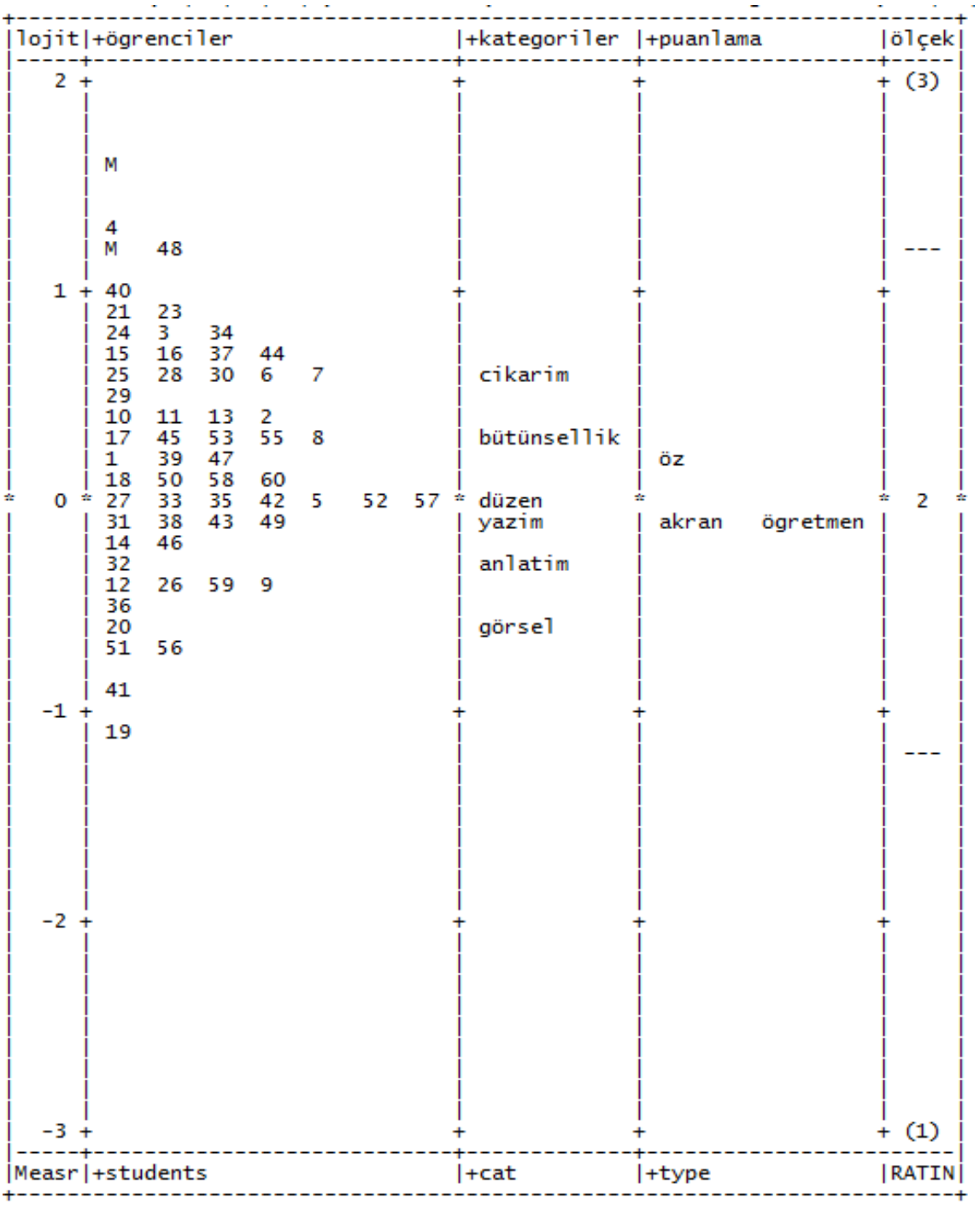


Şekil 2. İkinci analizindeki öğrenci becerisi, kategori puanlama katılığı, puanlayıcı türü katılığı ve Likert ölçek işlevinin kalibrasyon haritası

Şekil 2'de 11 numaralı öğrenci uyumsuz olarak belirlenmiştir. Bu durum ise bu öğrencinin diğer öğrencilerden farklı bir şekilde davrandığını göstermektedir. Şekil 2 incelendiğinde, yazma konusunda en yüksek beceri düzeyinin 48 numaralı öğrencide olduğu ve yine bu konuda en düşük beceri düzeyinin ise 19 numaralı öğrencide olduğu görülmektedir. Öğrencilerin çoğunun orta başarı düzeyinde toplandığı söylenebilir. Üçüncü sütundaki kategorilerin puanlanma katılıkları incelendiğinde, en cömert puanlanan kategorinin "yazım" ve "bütünsellik" en üstte, en katı puanlananın "çıkarımda bulunma" ise en altta yer aldığı söylenebilir. "Bütünsellik" ve "yazım" gibi kategoriler de yaklaşık olarak aynı puanlanma katılık düzeyine sahiptir. Dördüncü sütun ise puanlayıcı türünün katılığını göstermektedir Şekilde 2'de öz değerlendirmelerin en cömert, öğretmen ve akran değerlendirmelerin ise yaklaşık olarak eşit düzeyde ve daha katı olduğu görülmektedir. Şekil 2'de, öğrenci başarı kestirimlerinin (-1,42;1,60) önemli bir şekilde farklılaştığı görülebilir.

Üçüncü analizinde kalibrasyon haritasının sonuçları nedir?

Yazma becerilerinin değerlendirilmesi amacıyla yapılan uygulamaların sonunda yapılan analizlerin sonuçları Şekil 3'te gösterilmiştir. Her sayı bir öğrenciyi sembolize etmektedir.



Şekil 3. Uygulama sonundaki analizin öğrenci yeteneği, kategori puanlama katılığı, puanlayıcı türü katılığı ve Likert ölçek işlevinin kalibrasyon haritası

Şekil 3 incelendiğinde, yazma konusunda en yüksek beceriye uyumsuz olan 54 numaralı öğrencinin sahip olduğu ve yine bu konuda en düşük beceriye ise 19 numaralı öğrencinin sahip olduğu görülmektedir. Öğrencilerin çoğunun orta beceri düzeyinde toplandığı söylenebilir. Uyumsuz öğrenci olarak 54 ve 22 numaralı öğrenciler belirlenmiştir ve "M"

ile gösterilmişlerdir. Üçüncü sütundaki kategorilerin puanlanma katılıkları incelendiğinde, en cömert puanlanan kategorinin "çıkarımda bulunma" en üstte, en katı puanlananın "görsellerden faydalanma" ise en altta yer aldığı söylenebilmektedir. Tüm kategoriler farklı puanlanma katılık düzeyine sahiptir. Dördüncü sütun ise puanlayıcı türünün katılığını göstermektedir Şekil 3'te öz değerlendirmeler en cömertken, öğretmen ve akran değerlendirmelerin ise eşit düzeyde ve daha katı olduğu görülmektedir. Ayrıca öğrenci başarı kestirimlerinin (-1,13;1,61) önemli bir şekilde farklılaştığı görülebilir.

3.1.2. Alt Problem 1.b 'ye İlişkin Bulgular

Birinci analizde öğrencilerin yazma becerisi düzeyleri nedir?

Öğrencilerin yazma becerileri ile ilgili ölçümlerinin raporu ayrıntılı şekilde Tablo 6'da verilmiştir. Tablo 6'da, soldan sağa doğru olarak sütunlarda; öğrenci numaraları, öğrencilerin lojit ölçeğindeki başarı puanları, başarı kestiriminin hatası, uyum içi ortalama kare değerleri ve uyum dışı ortalama kare değerleri gösterilmiştir.

Tablo 6. Birinci Analizdeki Öğrencilerin Beceri Ölçüm Raporu

Öğrenci	Başarı (lojit)	Hata	Uyum içi ortalama kare değerleri (Infit Mean Square)	Uyum dışı ortalama kare değeri (Outfit Mean Square)
1	0,64	0,29	0,86	0,82
2	0,68	0,32	1,11	1,06
3	-0,14	0,32.	0,90	0,87
4	1,07	0,33	**0,35**	**0,40**
5	-0,31	0,34	1,20	1,21
6	0,38	0,28	1,30	1,25
7	0,46	0,33	**1,66**	**1,51**
8	0,49	0,29	0,74	0,74
9	-0,44	0,29	1,13	1,09
10	-0,91	0,31	1,08	1,03
11	0,90	0,29	0,83	0,76
12	-1,11	0,31	1,43	1,41
13	0,35	0,33	1,42	1,29
14	-1,04	0,31	0,71	0,75
15	-1,15	0,31	0,65	0,62
16	0,66	0,32	0,95	1.01
17	0,39	0,29	1,04	1,06
18	0,20	0,31	0,65	0,66
19	-1,12	0,33	1,07	1,05
20	-0,41	0,28	0,60	0,60
21	1,35	0,31	1,10	0,99
22	1,01	0,30	0,99	0,99
23	0,56	0,27	0,91	0,91
24	-0,94	0,31	1,36	1,28
25	0,12	0,28	1,39	1,40
26	-0,24	0,29	1,03	1,02

27	0,32	0,30	1,14	1,07
28	1,04	0,28	1,39	1,40
29	1,47	0,38	**0,44**	**1,85**
30	-0,66	0,30	0,83	0,83
31	-0,88	0,30	0,72	0,70
32	-2,25	0,42	0,78	0,92
33	-0,10	0,31	1,27	1,26
34	0,93	0,34	0,91	0,89
35	-0,32	0,32	0,60	0,60
36	-1,80	0,36	0,98	1,25
37	-0,18	0,28	0,94	0,94
38	-0,70	0,34	**1,51**	1,44
39	0,66	0,29	1,03	1,19
40	0,76	0,43	0,74	0,74
41	-0,54	0,30	0,80	0,79
42	0,40	0,28	0,80	0,80
43	0,15	0,28	0,83	0,83
44	0,92	0,28	1,04	1,04
45	-0,08	0,28	0,83	0,82
46	-0,36	0,29	1,06	1,04
47	0,75	0,29	1,12	1,19
48	2,01	0,35	0,68	0,68
49	-0,39	0,47	1,15	1,13
50	-0,12	0,28	1,39	1,36
51	-0,86	0,35	1,09	1,03
52	-0,54	0,30	0,83	0,84
53	-0,41	0,30	1,12	1,06
54	1,39	0,33	0,85	0,70
55	1,20	0,29	1,10	1,08
56	0,14	0,31	0,99	0,98
57	-0,33	0,32	1,05	1,07
58	-0,87	0,52	1,09	0,98
59	-2,73	1,03	0,93	0,75
60	-0,31	0,34	1,12	1,05
Ortalama	-0,01	0,33	0,99	0,98
Standart sapma	0,90	0,10	0,26	0,25
Ayırma indeksi: 2,45 Güvenirlik: 0,86 Ki-kare: 384,7 sd: 59 p=0,00				

Verilerin Çok Yüzeyli Rasch Ölçme Modeline uyum sağlayıp sağlamadıklarını anlamak için uyum içi ve dışı değerlerin 0,50 ve 1,50 arasında olup olmadığına bakmak gerekmektedir (Alharby, 2006). Buna göre daha önce Şekil 1'de de "M" ile gösterilen dört kişinin (4, 7, 29 ve 38) uyumsuz eleman oldukları belirlenmiş ve bu kişilerin değerleri Tablo 6'da koyu olarak gösterilmiştir.

Analiz sonucunda ayırma indeksi 2,45'tir ve bu indekse bakarak öğrencilerin beceri düzeylerinin birbirinden büyük ölçüde ayrıldığı söylenebilir. Ayrıca, öğrencilerin yaklaşık olarak 4 farklı beceri düzeyine ayrıldığı gözlemlenmiştir. Güvenirlik katsayısı 0,86'dır. Bu güvenirlik katsayısı analizin öğrencileri güvenilir bir şekilde farklı beceri düzeylerine ayırdığını göstermektedir. Ayrıca, ayırma indeksi 2,45 ve güvenirlik katsayısı 0,86 ile sabit etkiye ait "Öğrencilerin yazma becerileri arasında anlamlı bir fark vardır." hipotezi ki-kare

ile test edildiğinde yokluk hipotezi reddedilmiştir (χ^2=384,7, sd=59, p=0,00). Diğer bir ifadeyle öğrencilerin yazma becerileri arasında istatistiksel olarak anlamlı farklılıklar bulunmaktadır.

İkinci analizde öğrencilerin yazma becerisi düzeyleri nedir?

Öğrencilerin yazma beceri ölçümlerinin ölçüm raporu detaylı şekilde Tablo 7'de verilmiştir. Verilerin çok yüzeyli Rasch ölçme modeline uyum sağlayıp sağlamadıklarını anlamak için uyum içi ve dışı değerlerin 0,50 ve 1,50 arasında olup olmadığına bakılmıştır. Buna göre daha önce kalibrasyon haritasında "M" ile gösterilen 11 numaralı öğrencinin uyumsuz eleman olduğu belirlenmiş ve değerleri Tablo 7'de koyu hale getirilmiştir.

Tablo 7. İkinci Analizdeki Öğrencilerin Beceri Ölçüm Raporu

Öğrenci	Başarı (lojit)	Hata	Uyum içi ortalama kare değerleri	Uyum dışı ortalama kare değeri
1	0,53	0,30	0,81	0,82
2	0,11	0,30	0,90	0,88
3	0,82	0,33	0,96	0,95
4	1,44	0,31	1,14	1,11
5	-1,02	0,32	0,88	0,87
6	0,10	0,29	0,78	0,78
7	1,04	0,33	0,88	0,89
8	0,35	0,30	1,04	1,02
9	0,06	0,29	1,15	1,14
10	-0,29	0,29	0,77	0,76
11	0,11	0,29	**1,56**	**1,54**
12	-0,76	0,31	1,35	1,35
13	1,05	0,31	0,96	0,94
14	-0,38	0,30	0,74	0,74
15	-0,44	0,30	1,44	1,45
16	0,76	0,30	0,60	0,64
17	0,54	0,30	0,81	0,82
18	-0,56	0,31	1,09	1,08
19	-1,43	0,34	1,02	1,03
20	-0,73	0,32	0,86	0,84
21	0,75	0,32	1,35	1,23
22	0,48	0,31	0,87	0,96
23	0,39	0,29	1,13	1,13
24	0,09	0,29	1,10	1,10
25	-0,76	0,32	1,14	1,12
26	0,02	0,31	0,79	0,77
27	0,26	0,32	1,08	0,94
28	0,54	0,31	0,70	0,85
29	0,30	0,29	0,53	0,54
30	-0,03	0,29	0,94	0,94
31	-0,30	0,29	1,35	1,34
32	-1,27	0,32	1,44	1,48
33	-0,33	0,30	1,18	1,17
34	0,76	0,36	0,94	1,08
35	0,44	0,30	0,91	0,91
36	-0,61	0,34	0,88	0,87

37	0,81	0,31	0,88	0,91
38	0,08	0,30	1,41	1,44
39	0,11	0,30	0,89	0,88
40	0,28	0,34	1,44	1,40
41	-1,02	0,32	0,80	0,80
42	0,01	0,29	1,16	1,16
43	-0,45	0,35	0,84	0,84
44	-0,06	0,30	1,12	1,11
45	-0,12	0,30	0,87	0,87
46	0,33	0,29	0,83	0,84
47	0,40	0,29	1,07	1,07
48	1,60	0,35	0,94	0,78
49	0,58	0,30	1,06	1,03
50	0,11	0,30	0,84	0,84
51	-0,42	0,30	1,00	0,99
52	-0,88	0,31	1,08	1,06
53	-0,05	0,29	0,90	0,89
54	0,51	0,31	0,86	0,84
55	0,41	0,29	0,98	0,98
56	0,23	0,29	0,99	1,00
57	-1,18	0,32	0,88	0,87
58	0,12	0,30	0,69	0,69
59	-0,85	0,31	0,96	0,99
60	-0,14	0,30	0,99	0,98
Ortalama	0,04	0,31	1,00	0,99
Standart sapma	0,64	0,02	0,22	0,21
Ayırma indeksi: 1,85 Güvenirlik: 0,77 Ki-kare:243,6 sd: 59 p=0,00				

Ayırma indeksi 1,85'tir. Ayrıca öğrenciler yaklaşık olarak 3 farklı başarı düzeyine ayrıldığı gözlemlenmiştir. Güvenirlik katsayısı 0,77'dir ve öğrencilerin başarı düzeylerinin birbirinden büyük ölçüde ayrıldığı söylenebilir. Ek olarak, ayırma indeksi 1,85 ve güvenirlik katsayısı 0,77 ile sabit etkiye ait "Öğrencilerin yazma başarıları arasında anlamlı bir fark vardır." hipotezi ki-kare ile test edildiğinde (X^2=243,6, sd=59, p=0,00) yokluk hipotezi reddedilmiştir. Diğer bir deyişle, öğrencilerin yazma başarıları arasında istatistiksel açıdan anlamlı farklılıklar bulunmaktadır.

Üçüncü analizde öğrencilerin yazma becerisi düzeyleri nedir?

Öğrencilerin yazma beceri ölçümlerinin ölçüm raporu detaylı şekilde Tablo 8'de verilmiştir.

Tablo 8. Üçüncü Analizdeki Öğrencilerin Beceri Ölçüm Raporu

Öğrenci	Başarı (lojit)	Hata	Uyum içi ortalama kare değerleri	Uyum dışı ortalama kare değeri
1	0,16	0,28	0,88	0,87
2	0,38	0,28	1,05	1,06
3	0,75	0,30	0,79	0,76
4	1,28	0,32	0,65	0,69
5	-0,04	0,28	0,90	0,90
6	0,56	0,31	1,21	1,23

7	0,60	0,29	0,74	0,77
8	0,32	0,28	1,03	1,03
9	-0,37	0,32	0,92	0,90
10	0,44	0,30	0,79	0,80
11	0,36	0,28	1,21	1,21
12	-0,44	0,29	1,17	1,15
13	0,41	0,28	0,74	0,76
14	-0,19	0,28	1,04	1,05
15	0,67	0,29	1,06	1,07
16	0,65	0,29	1,22	1,20
17	0,34	0,28	1,22	1,27
18	0,11	0,28	1,01	1,03
19	-1,13	0,31	1,05	1,03
20	-0,61	0,29	1,28	1,26
21	0,95	0,30	0,97	0,98
22	1,21	0,31	1,42	**1,60**
23	0,88	0,29	0,61	0,62
24	0,77	0,29	1,07	1,03
25	0,61	0,28	0,96	0,95
26	-0,40	0,32	1,18	1,18
27	-0,04	0,28	0,81	0,81
28	0,56	0,29	1,03	0,99
29	0,55	0,29	0,91	1,13
30	0,58	0,29	1,16	1,17
31	-0,06	0,28	1,15	1,14
32	-0,31	0,28	1,19	1,20
33	0,03	0,28	0,69	0,69
34	0,84	0,33	0,63	0,62
35	-0,05	0,29	0,62	0,61
36	-0,51	0,32	0,98	0,97
37	0,71	0,29	1,26	1,24
38	-0,07	0,28	0,79	0,79
39	0,24	0,28	1,49	1,46
40	0,98	0,30	0,84	0,82
41	-0,93	0,30	1,03	1,00
42	0,03	0,28	0,69	0,69
43	-0,07	0,28	0,75	0,75
44	0,74	0,29	1,35	1,40
45	0,35	0,28	0,88	0,88
46	-0,22	0,28	0,92	0,91
47	0,18	0,29	1,06	1,07
48	1,19	0,33	0,97	0,97
49	-0,05	0,28	1,06	1,06
50	0,15	0,28	0,90	0,90
51	-0,65	0,29	0,88	0,95
52	0,05	0,28	0,89	0,89
53	0,29	0,28	0,80	0,82
54	1,61	0,32	1,27	**1,56**
55	0,29	0,28	1,19	1,19
56	-0,72	0,29	0,72	0,72
57	0,00	0,28	1,49	1,48
58	0,12	0,28	1,01	1,03
59	-0,42	0,28	1,15	1,14
60	0,14	0,28	0,65	0,66
Ortalama	0,23	0,29	1,00	1,01
Standart sapma	0,55	0,01	0,22	0,23
Ayırma indeksi:1,60 Güvenirlik: 0,72 Ki-kare:198,7 sd: 59 p=0,00				

Tablo 8'deki verilerin çok yüzeyli Rasch ölçme modeline uyum sağlayıp sağlamadıklarını anlamak için uyum içi ve dışı değerlerin 0,50 ve 1,50 arasında olup olmadığına bakıldığında, 54 ve 22 numaralı öğrencinin uyumsuz eleman olduğu belirlenmiş ve değerleri Tablo 8'de koyu hâle getirilmiştir.

Analiz sonucunda elde edilen ayırma indeksi 1,60'tır. Ayrıca öğrencilerin yaklaşık olarak 3 farklı beceri düzeyine ayrıldığı gözlemlenmiştir. Güvenirlik katsayısı 0,72'dir. Öğrencilerin başarı düzeylerinin birbirinden büyük ölçüde ayrıldığı söylenebilir. Ek olarak ayırma indeksi 1,60 ve güvenirlik katsayısı 0,72 ile sabit etkiye ait "Öğrencilerin yazma becerileri arasında anlamlı bir fark yoktur." hipotezi ki-kare ile test edildiğinde (X^2=198,7, sd=59, p=0,00) yokluk hipotezi reddedilmiştir. Diğer bir deyişle, öğrencilerin yazma becerileri arasında istatistiksel açıdan anlamlı farklılıklar bulunmaktadır.

3.1.3. Alt Problem 1.c 'ye İlişkin Bulgular

Birinci analizde DPA'daki kategorilerin puanlanma durumları nedir?

Birinci analizdeki DPA'daki kategorilerin puanlanma katılıkları ölçüm raporu Tablo 9'da verilmiştir.

Tablo 9. Birinci Analizdeki DPA'daki Kategorilerin Puanlanma Katılıkları Ölçüm Raporu

Değerlendirme Kategorisi	Zorluk düzeyi	Hata	Uyum içi ortalama kare değeri (Infit Mean Square)	Uyum dışı ortalama kare değeri (Outfit Mean Square)
Bütünsellik	-0,51	0,10	0,88	0,90
Anlatım	-0,24	0,10	0,94	0,95
Yazım	0,45	0,10	1,10	1,06
Düzen	0,30	0,10	0,97	0,96
Görsellerden Faydalanma	-0,01	0,10	0,92	0,88
Çıkarımlarda Bulunma	0,02	0,10	1,18	1,15
Aritmetik ortalama	0,00	0,10	1,00	0,98
Standart sapma	0,32	0,00	0,11	0,09
Ayırma indeksi: 3,40	Güvenirlik: 0,90		Ki-kare: 61,90 sd: 4	p=0,00

Tablo 9'a bakıldığında en katı puanlanan kategorinin "bütünsellik" (-0,51) ve en cömert puanlananın ise "yazım" (0,45) olduğu görülmektedir. Analiz sonucunda ayırma indeksi 3,40'tır. Güvenirlik katsayısının 0,90 olması, değerlendirme kategorilerinin güçlük bakımından istatistiksel olarak birbirlerinden farklı olduklarını göstermektedir. Bu durum istatistiksel olarak anlamlı (p=0,00) olan ki kare değerine (X^2= 61,90) ve serbestlik

derecesine (sd=4) bakıldığında da doğrulanmaktadır. "Çıkarımda bulunma" ve" görsellerden faydalanma" kategorileri yaklaşık olarak aynı güçlük düzeyine sahip olmalarına rağmen uyum içi ve uyum dışı ortalama kare değerleri 0,50 ile 1,50 arasında olduğundan dolayı çıkartılmalarına gerek duyulmamıştır. Genele bakıldığında, her bir kategorinin mantıksal olarak farklı bilgiler sağladıkları söylenebilir.

İkinci analizde DPA'daki kategorilerin puanlanma durumları nedir?

İkinci analizdeki DPA'daki kategorilerin puanlanma katılıkları ölçüm raporu Tablo 10'da verilmiştir.

Tablo 10. İkinci Analizdeki DPA'daki Kategorilerin Puanlanma Katılıkları Ölçüm Raporu

Değerlendirme Kategorisi	Zorluk düzeyi	Hata	Uyum içi ortalama kare değeri	Uyum dışı ortalama kare değeri
Bütünsellik	0,21	0,10	1,07	1,05
Anlatım	-0,17	0,10	0,84	0,84
Yazım	0,16	0,10	1,02	1,00
Düzen	0,13	0,10	1,05	1,03
Görsellerden Faydalanma	-0,03	0,10	0,96	1,00
Çıkarımlarda Bulunma	-0,30	0,10	1,05	1,06
Ortalama	0,00	0,10	1,00	0,99
Standart sapma	0,19	0,00	0,08	0,07
Ayırma indeksi: 1,85 Güvenirlik: 0,77 Ki-kare: 22,10 sd: 4 p=0,00				

Tablo 10'da en cömert puanlanan kategorinin "bütünsellik" (0,21) ve "yazım" (0,16), en katı puanlananın ise "çıkarımda bulunma" (-0,30) olduğu görülmektedir. Analiz sonucunda ayrıma indeksi 1,85'tir ve güvenirlik katsayısının 0,77'dir. Bu da değerlendirme kategorilerinin güçlük bakımından istatistiksel olarak birbirlerinden farklı olduklarını göstermektedir. Bu durum istatistiksel olarak anlamlı (p=0,00) olan ki kare değeri (χ^2= 22,10) ile serbestlik derecesine (sd= 4) bakıldığında da doğrulanmaktadır. "Bütünsellik" ve "yazım" kategorileri yaklaşık olarak aynı güçlük düzeyine sahip olmalarına rağmen uyum içi ve uyum dışı ortalama kare değerleri 0,50 ve 1,50 arasında olduğundan dolayı çıkartılmalarına gerek duyulmamıştır. Genele bakıldığında, her bir kategorinin mantıksal olarak farklı bilgiler sağladıkları söylenebilir.

Üçüncü analizde DPA'daki kategorilerin puanlanma durumları nedir?

Üçüncü analizdeki DPA'daki kategorilerin puanlanma katılıkları ölçüm raporu Tablo 11'de gösterilmiştir.

Tablo 11. Üçüncü Analizdeki DPA'daki Kategorilerin Puanlanma Katılıkları Ölçüm Raporu

Değerlendirme Kategorisi	Zorluk düzeyi	Hata	Uyum içi ortalama kare değeri	Uyum dışı ortalama kare değeri
Bütünsellik	0,29	0,10	0,99	1,01
Anlatım	-0,26	0,09	0,97	0,96
Yazım	-0,09	0,09	1,11	1,13
Düzen	0,01	0,09	0,82	0,82
Görsellerden Faydalanma	-0,57	0,09	0,78	0,78
Çıkarımlarda Bulunma	0,62	0,10	1,33	1,35
Ortalama	0,00	0,09	1,00	1,01
Standart sapma	0,38	0,00	0,19	0,19
Ayırma indeksi: 4,44 Güvenirlik: 0,95 Ki-kare: 100,9 sd: 4 p=0,00				

Tablo 11'de en cömert puanlanan kategorinin "çıkarımda bulunma" ve en katı puanlanan kategorinin ise "görsellerden faydalanma" olduğu görülmektedir. Yapılan analizler sonucunda ayrıma indeksi 4,44'tür. Güvenirlik katsayısının 0,95 olması, değerlendirme kategorilerinin güçlük bakımından istatistiksel olarak birbirlerinden farklı olduklarını göstermektedir. Bu durum istatistiksel olarak anlamlı (p=0,00) olan ki kare değerine (x^2=100,9) ve serbestlik derecesine (sd=4) bakıldığında da doğrulanmaktadır. Genele bakıldığında, her bir kategorinin farklı bilgiler sağladıkları söylenebilir.

3.1.4. Alt Problem 1.d'ye İlişkin Bulgular

Birinci analizde öz, akran ve öğretmen puanlarınn durumu nedir?

Birinci analizdeki puanlayıcı türü katılık ölçüm raporu Tablo 12'de verilmiştir.

Tablo 12. Birinci Analizdeki Puanlayıcı Türü Katılık Ölçüm Raporu

Puanlayıcı Türü	Katılık (lojit)	Hata	Uyum içi ortalama kare değeri	Uyum dışı ortalama kare değeri
Öz	0,62	0,09	1,00	0,99
Akran	-0,04	0,10	1,10	1,04
Öğretmen	-0,58	0,05	0,97	0,97
Ortalama	0,00	0,08	1,02	1,00
Standart sapma	0,49	0,02	0,06	0,03
Ayırma indeksi: 6,88 Güvenirlik: 0,98 Ki-kare: 131,1 sd: 2 p=0,00				

Tablo 12'de en cömert puan veren puanlayıcıların öz puanlayıcılar (0,62) ve en katı puanlayıcıların ise öğretmenlerin (-0,58) olduğu görülmektedir. Akran puanlayıcılar ise

orta düzey katılıkta puanlama yapmışlardır. Ayrıma indeksi 6,88'dir. Güvenirlik katsayısının yüksek (0,98) olması, puanlayıcı türünün katılık bakımından istatistiksel olarak farklılaştığını göstermektedir. Bu durum istatistiksel olarak anlamlı (p=0,00) olan ki kare değerine (χ^2=131,1) ve serbestlik derecesine (sd=2) bakıldığında da doğrulanmaktadır.

İkinci analizde öz, akran ve öğretmen puanlarınn durumu nedir?

İkinci analizdeki puanlayıcı türü katılık ölçüm raporu Tablo 13'te verilmiştir.

Tablo 13. İkinci Analizdeki Puanlayıcı Türü Katılık Ölçüm Raporu

Puanlayıcı Türü	Katılık (lojit)	Hata	Uyum içi ortalama kare değeri	Uyum dışı ortalama kare değeri
Öz	0,77	0,09	0,90	0,90
Akran	-0,37	0,09	1,12	1,10
Öğretmen	-0,40	0,05	0,99	1,00
Ortalama	0,00	0,08	1,00	1,00
Standart sapma	0,54	0,02	0,09	0,08
Ayırma indeksi: 8,18 Güvenirlik: 0,99 Ki-kare: 128 sd: 2 p=0,00				

Tablo 13'te en cömert puanlayıcıların öz puanlayıcılar (0,77) ve en katı puanlayıcıların ise öğretmen ve akran puanlayıcıları (-0,40; -0,37) olduğu görülmektedir. Analiz sonucunda elde edilen ayrıma indeksi 8,18'tir. Güvenirlik katsayısının yüksek (0,99) olması, değerlendirme puanlayıcı türünün katılık bakımından istatistiksel olarak farklılaştığını göstermektedir. Bu durum istatistiksel olarak anlamlı (p=0,00) olan ki kare değeri (χ^2=128,0) ile serbestlik derecesine (sd=2) bakıldığında da doğrulanmaktadır. Tablo 11 incelendiğinde, öz puanlamalarının öğretmen puanlamalarına göre daha cömert ve akran puanlamalarının aynı düzeyde katılıkta olduğu görülmektedir.

Üçüncü analizde öz, akran ve öğretmen puanlarınn durumu nedir?

Üçüncü analizdeki puanlayıcı türü katılık ölçüm raporu Tablo 14'te verilmiştir.

Tablo 14. Üçüncü Analizdeki Puanlayıcı Türü Katılık Ölçüm Raporu

Puanlayıcı Türü	Katılık (lojit)	Hata	Uyum içi ortalama kare değeri	Uyum dışı ortalama kare değeri
Öz	0,24	0,08	1,02	1,06
Akran	-0,14	0,09	1,11	1,11
Öğretmen	-0,11	0,05	0,95	0,96
Ortalama	0,00	0,07	1,03	1,04
Standart sapma	0,17	0,02	0,06	0,06
Ayırma indeksi: 2,56 Güvenirlik: 0,87 Ki-kare:14 sd: 2 p=0,00				

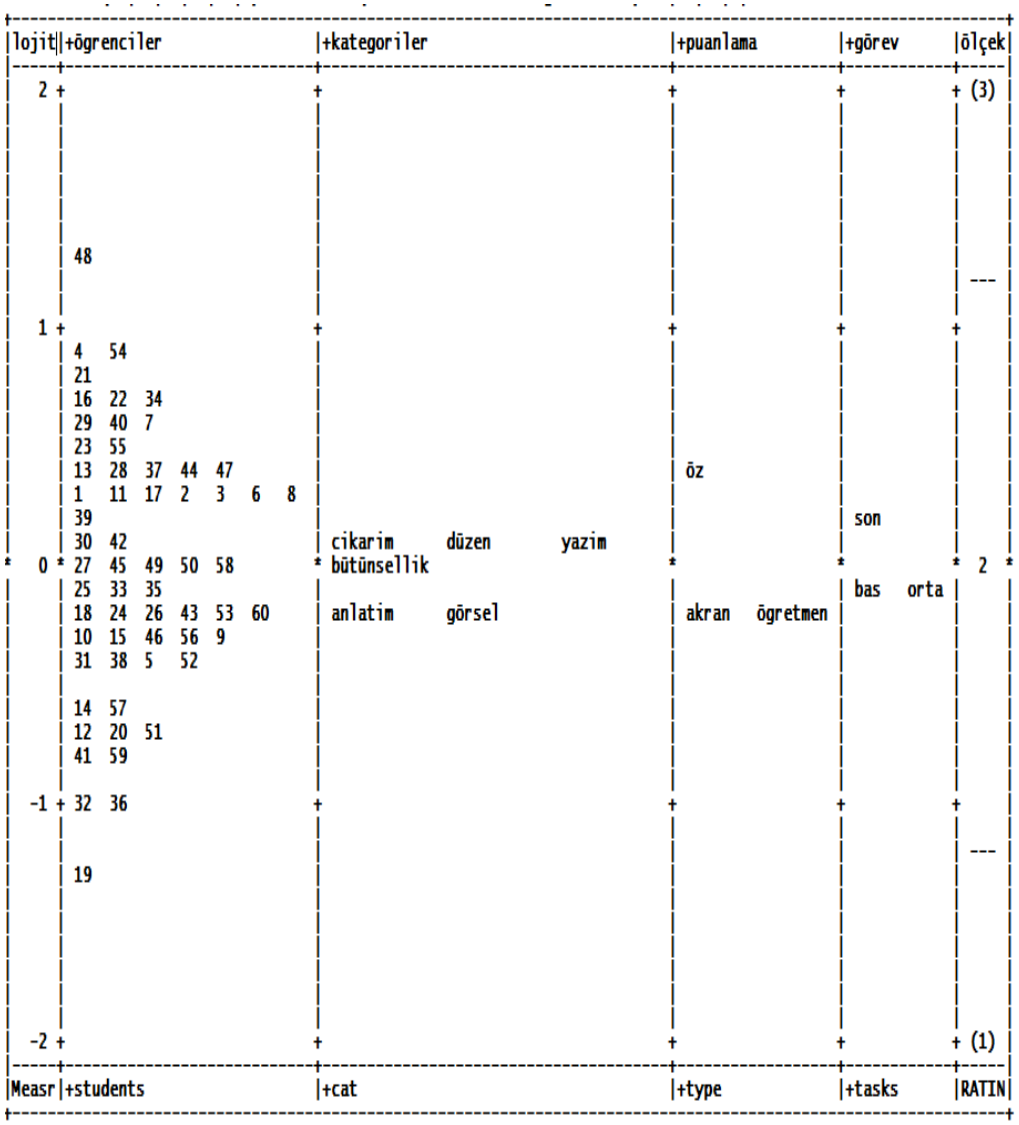
Tablo 14'te en cömert puanlayıcıların öz puanlayıcılar (0,24) oldukları, öğretmen ve akran puanlayıcıların ise (-0,11;-0,14) hemen hemen aynı düzey katılıkta olduğu görülmektedir. Yapılan analizlerde ayrıma indeksi 2,56'dır ve güvenirlik katsayısı 0,87'dir. Bu da puanlayıcı türünün katılık bakımından 0,87 güvenirlik düzeyiyle istatistiksel olarak farklılaştığını göstermektedir. Bu durum istatistiksel olarak anlamlı (p=0,00) olan ki kare değerine (x^2=14,0) ve serbestlik derecesine (sd=2) bakıldığında da doğrulanmaktadır. Tablo 14 incelendiğinde, öz puanlamalarının öğretmen puanlamalarına göre daha cömert ve akran puanlamalarının yaklaşık olarak aynı düzey katılıkta olduğu görülmektedir.

3.2. İkinci Araştırma Sorusuna İlişkin Bulgular

3.2.1 Alt Problem 2.a'ya İlişkin Bulgular

Yapılan uygulama sürecinin bir bütün olarak alındığı analizde kalibrasyon haritası sonuçları nedir?

Yazma becerilerinin değerlendirilmesi amacıyla yapılan uygulamaların sonunda bir bütün olarak yapılan analizlerin sonuçları Şekil 4'te gösterilmiştir. Şekil 4'teki her sayı bir öğrenciyi sembolize etmektedir.



Şekil 4. Bir bütün olarak alınan analizlerdeki öğrenci becerisi, kategori puanlama katılığı, puanlayıcı türü katılığı ve Likert ölçek işlevinin kalibrasyon haritası

Şekil 4 incelendiğinde, yazma konusunda en yüksek beceri düzeyine 48 numaralı öğrencinin sahip olduğu ve yine bu konuda en düşük beceri düzeye ise 19 numaralı öğrencinin sahip olduğu görülmektedir. Öğrencilerin çoğunun orta başarı düzeyinde toplandığı söylenebilir. Uyumsuz öğrenci belirlenememiştir ve bu durumun üç katına çıkarılan veri sayısından kaynaklandığı düşünülmektedir. Üçüncü sütundaki kategorilerin puanlanma katılıkları incelendiğinde, en cömert puanlanan kategorinin "çıkarımda bulunma", "düzen" ve" yazım" en üstte, en katı puanlanan kategorilerin "görsellerden faydalanma" ve "anlatım" ise en altta yer aldığı söylenebilmektedir. "Çıkarımda bulunma", "düzen" ve "yazım" kategorileri aynı düzeyde puanlanmışken, "görsellerden faydalanma" ve "anlatım" kategorileri de aynı katılıkta puanlanmışlardır. Dördüncü sütun ise puanlayıcı

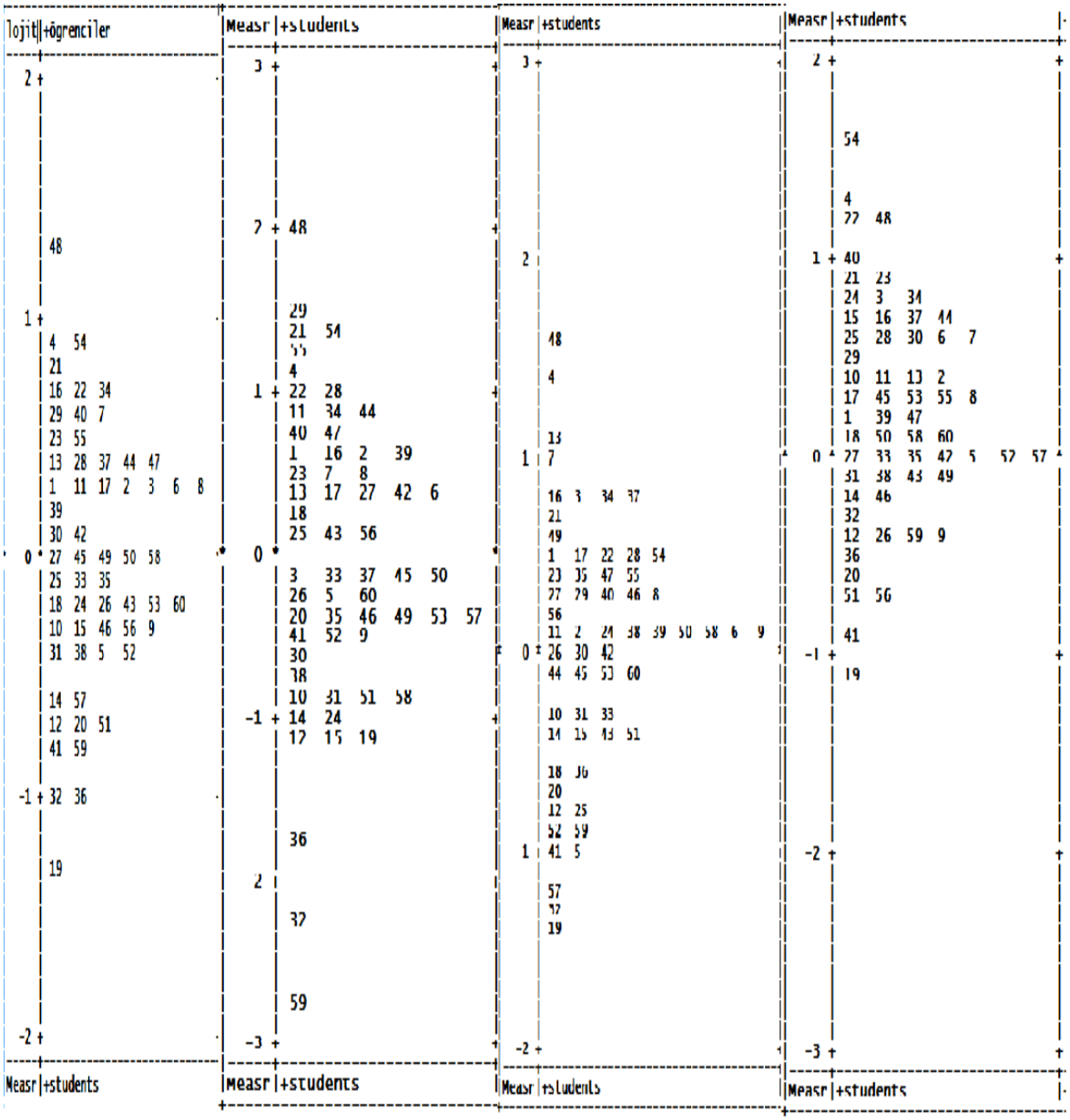
türünün katılığını göstermektedir Şekil 4'te öz değerlendirmelerin en cömert olduğu, öğretmen ve akran değerlendirmelerin eşit düzeyde ve daha katı olduğu görülmektedir. Son sütunda beşinci yüzey olan analizlerin uygulama başında, ortasında ve sonunda olması görülmektedir ve en kolay görevin son görev olduğu ve baştaki ve ortadaki görevlerin son göreve göre daha zor ve aynı düzeyde zorluktaki görevler oldukları söylenebilir. Bu da araştırmacının görevleri geliştirirken yaklaşık olarak aynı güçlük düzeyinde olmalarına dikkat etmesinden kaynaklanabilir. En son görevin daha kolay görülmesi ise öğrencilerin süreci öğrenerek sonuncu görevde daha başarılı olmalarından kaynaklandığı düşünülmektedir. Analizin yüzeylerine ait uyum içi ve uyum dışı değerleriyle uyum indeksleri Ek 8, 9, 10 ve11'de gösterilmiştir.

Şekil 4'te, öğrenci beceri düzey kestirimlerinin önemli bir şekilde farklılaştığı görülmektedir.

3.2.2. Alt Problem 2.b'ye İlişkin Bulgular

Yapılan uygulama sürecinin bir bütün olarak alındığı analizde öğrencilerin yazma becerisi düzeylerinde zaman içinde bir değişme olmuş mudur?

Şekil 5'te uygulama esnasında yapılan analizlerdeki öğrenci yüzeylerinin kalibrasyon haritaları verilmektedir. Birinci sütunda bir bütün olarak alınan analiz olmak üzere, sırasıyla birinci, ikinci ve üçüncü analize ait haritalar vardır.



Şekil 5. Uygulamanın sırasıyla bir bütün olarak alınan analizindeki, birinci, ikinci ve üçüncü analizlerindeki öğrenci yüzeylerinin kalibrasyon haritası

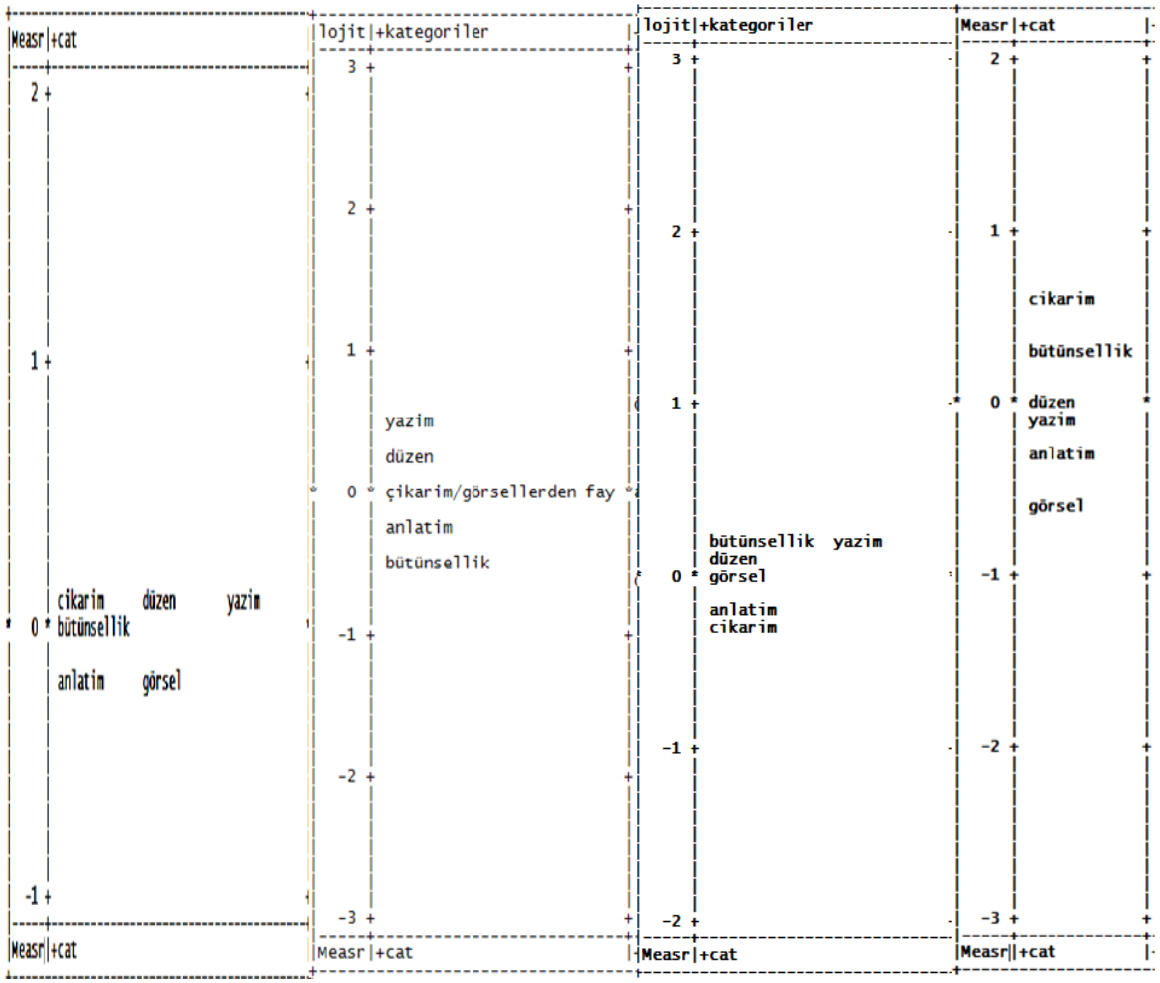
Şekil 5 incelendiğinde, her yüzeyin yanında analizdeki yapının lojit ölçeği bulunmaktadır. Verilen lojit ölçeklerinin hiçbiri aynı olmadığından da anlaşılacağı gibi, bu analizleri istatistiksel olarak karşılaştırmak çok anlamlı olmayacaktır. Şekilsel olarak gruplar içerisinde bakıldığında, 48 numaralı öğrencinin bir bütün olarak alınan analizde, birinci ve ikinci analizlerde grup içerisindeki en yüksek beceriye sahip kişi olduğu, 19 numaralı öğrencinin ise bir bütün olarak alınan analizde, ikinci ve üçüncü analizlerde ise grubun en düşük beceriye sahip olduğu görülmektedir. 1, 2, 4, 6, 8, 11, 17, 21, 39, 40, 47 ve 50 numaralı öğrenciler tüm analizlerde yüksek başarı düzeyinde (0,00 ve 1,00 lojit arasında)

yer almışlardır. 44 numaralı öğrenci genelde yüksek başarı düzeyindedir ve sadece ikinci analizde 0,00 lojit altında yer almaktadır. 3 numaralı öğrenci aynı şekilde genelde yüksek başarı düzeyindeyken, birinci analizde 0,00 lojit altında görülmektedir. 30 ve 56 numaralı öğrenciler birinci ve ikinci analizde daha düşük bir lojit değerinde (0,00 lojit ve altı) iken, bir bütün olarak alınan ve üçüncü analizde daha yüksek değerlerde ve grubun üst kısmındadır. Aynı şekilde 27 numaralı öğrenci genel ve üçüncü analizde sıfıra yakın, birinci ve ikinci de ise 0,00 lojit üstünde ve 42 numaralı öğrenci ise ikinci ve üçüncü analizde 0 ve bir bütün olarak alınan ve birinci analizde ise sıfırın üstündedir. 45 numaralı öğrenci hepsinde sıfıra yakın iken, üçüncü analizde sıfır üstündedir. 57, 52 ve 5 numaralı öğrenciler tüm analizlerde sıfırın altında iken, üçüncü analizde sıfır düzeyinde bulunmaktadır. 10 ve 15 numaralı öğrenciler tüm analizlerde sıfırın altında yer alırken, üçüncü analizde sıfır üzerinde yer almaktadır. 9 numaralı öğrenci tüm analizlerde sıfırın altında iken, ikinci analizde sıfırın üstünde yer almaktadır. 18 numaralı öğrenci genel ve ikinci analizde sıfırın altındaki düzeyde iken, birinci ve üçüncü analizde sıfırın üstündeki düzeydedir. 24. öğrenci bir bütün olarak alınan analizde ve birinci analizde sıfırın altında iken, ikinci analizde sıfırın çok az üstünde ve üçüncü analizde ise sıfırın çok üstündedir. 53., 25. ve 60. öğrenciler ise bir bütün olarak alınan analizde, birinci ve ikinci analizde sıfırın altında iken, üçüncü analizde ise sıfırın üstündedir. 12., 14., 20., 26., 31., 32., 33., 36., 38., 41., 43., 46., 51. ve 59. öğrenciler genel olarak tüm analizlerde düşük başarı düzeyinde (0,00 ve -1,00 lojit arasında) yer almaktadırlar.

3.2.3. Alt Problem 2.c'ye İlişkin Bulgular

Yapılan uygulama sürecinin bir bütün olarak alındığı analizde DPA'daki kategorilerin puanlanma durumlarında zaman içinde bir değişme olmuş mudur?

Şekil 6'da uygulama esnasında yapılan analizlerdeki dereceli puanlama anahtarında yer alan kategorilerin puanlanma durumlarına ait kalibrasyon haritaları verilmektedir. Birinci sütunda bir bütün olarak alınan analiz olmak üzere, daha sonra sırasıyla birinci, ikinci ve üçüncü analize ait haritalara yer verilmiştir.



Şekil 6. Uygulamanın sırasıyla bir bütün olarak alınan analizinde, birinci, ikinci ve üçüncü analizlerinde dereceli puanlama anahtarındaki kategorilerin puanlanma durumlarının kalibrasyon haritası

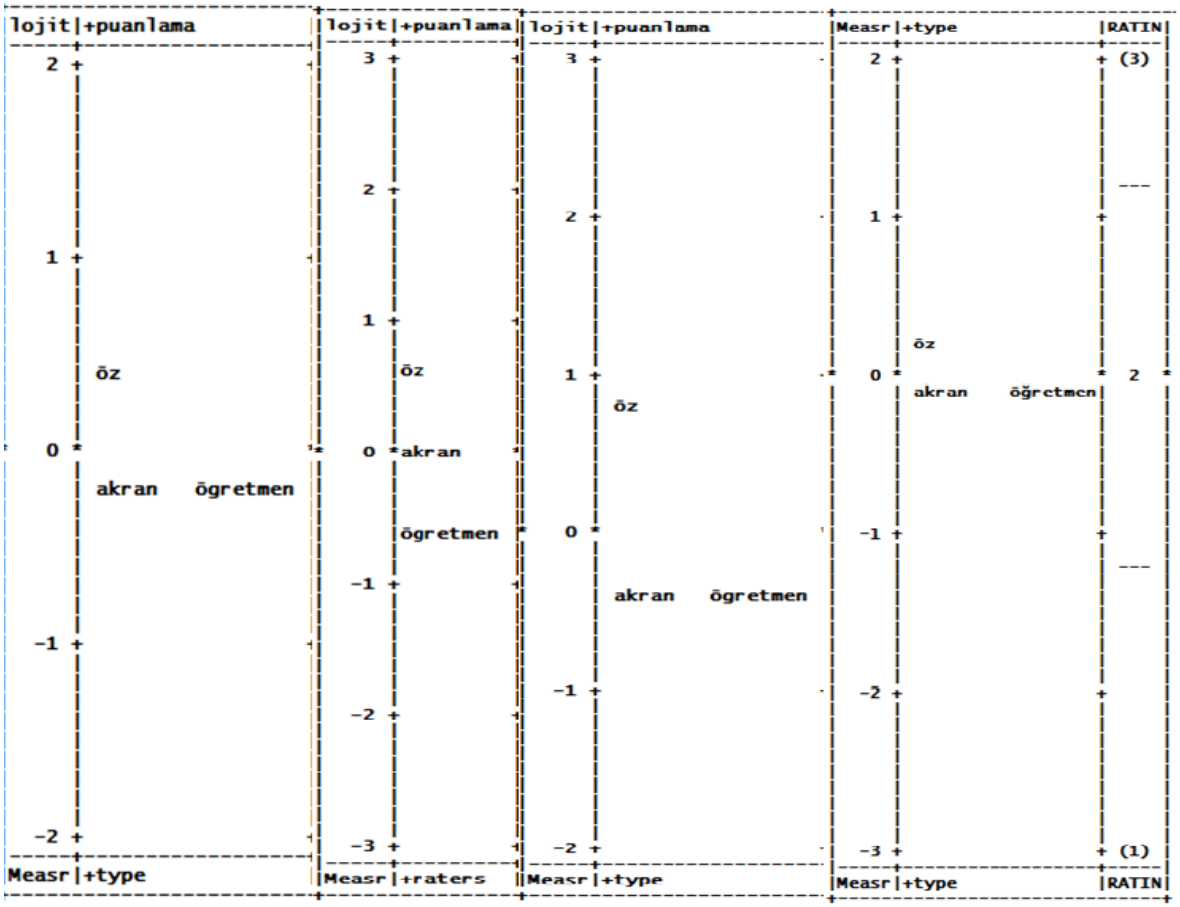
Şekil 6 incelendiğinde, her yüzeyin yanında analizdeki yapının lojit ölçeği bulunmaktadır. Verilen lojit ölçeklerinin hiçbiri birbiri ile aynı olmadığından da anlaşılacağı gibi, bu analizleri istatistiksel olarak karşılaştırmak çok anlamlı olmayacaktır. Şekilsel olarak incelendiğinde ise, birinci, ikinci ve bir bütün olarak alınan analizde 0,00 lojit üzerinde olan "yazım", "düzen" gibi kategorilerin, üçüncü analizde sıfırın altında yer aldığı görülmektedir. "Çıkarımda bulunma" kategorisinin ise bir bütün olarak alınan analizde ve üçüncü analizde en cömert puanlanan olduğu görülmektedir. Ayrıca "çıkarımda bulunma" kategorisinin birinci analizde orta güçlükte ve ikinci analizde en katı puanlanan kategori olduğu söylenebilir. "Bütünsellik" birinci analizde en katı puanlananken, üçüncü ve ikincide en cömert puanlananlar arasındadır. Ayrıca bir bütün olarak alınan analizde orta katılık düzeyinde görülmektedir. "Anlatım" tüm analizlerde katı puanlanırken,

"görsellerden faydalanma" ise, birinci ve ikincide sıfıra eşit iken bir bütün olarak alınan analizde ve üçüncüde en katı puanlananlar arasındadır.

3.2.4. Alt Problem 2.d 'ye İlişkin Bulgular

Yapılan uygulama sürecinin bir bütün olarak alındığı analizde öz, akran ve öğretmen puanlarının durumunda zaman içerisinde bir değişme olmuş mudur?

Şekil 7'de uygulama esnasında yapılan analizlerdeki puanlayıcı türünün katılığına ait kalibrasyon haritaları verilmektedir. Birinci sütunda bir bütün olarak alınan analiz olmak üzere, daha sonra sırasıyla birinci, ikinci ve üçüncü analize ait haritalar yer almaktadır.



Şekil 7. Uygulamanın sırasıyla bir bütün olarak alınan analizindeki, birinci, ikinci ve üçüncü analizlerindeki puanlayıcı türünün katılığının kalibrasyon haritası

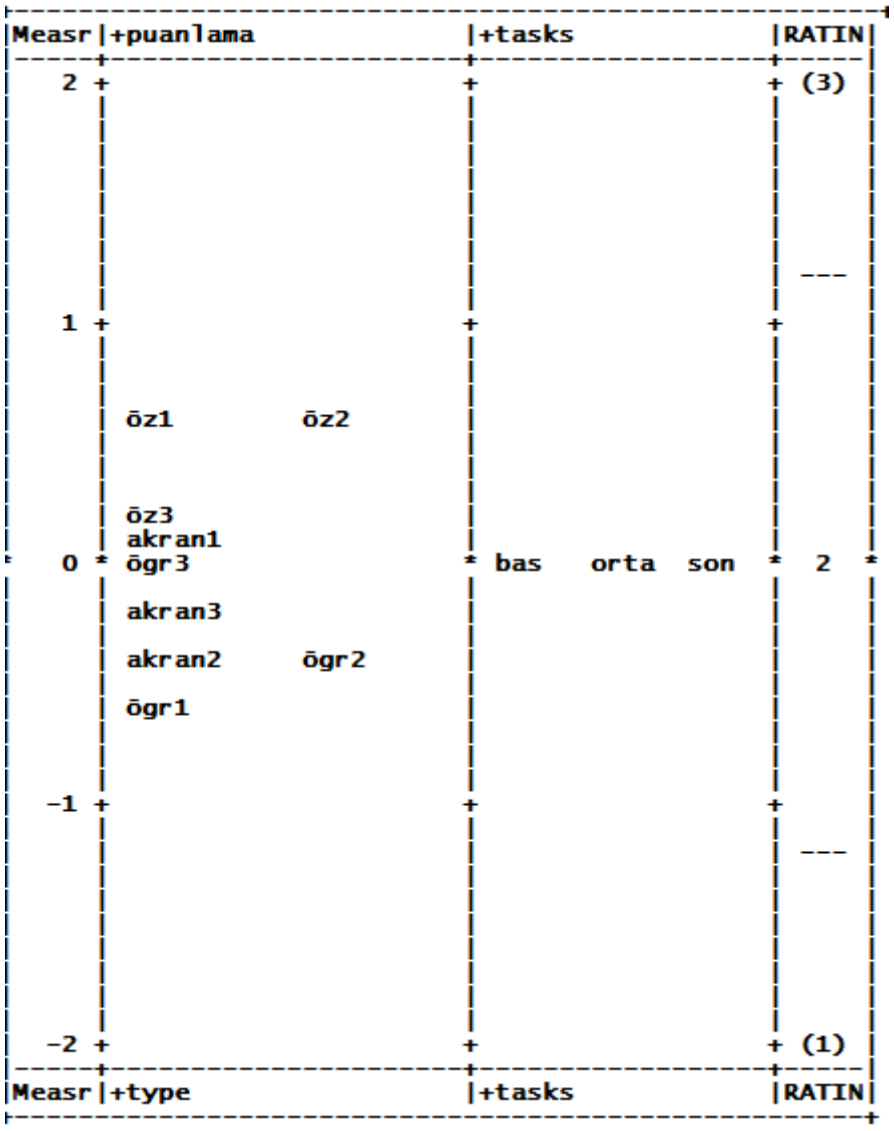
Şekil 7'de görüldüğü gibi her yüzeyin yanında analizdeki yapının lojit ölçeği bulunmaktadır. Verilen lojit ölçeklerinin hiçbiri birbiri ile aynı olmadığından da

anlaşılacağı gibi, bu analizleri istatistiksel olarak karşılaştırmak çok anlamlı olmayacaktır. Şekilsel olarak incelendiğinde ise, tüm analizlerde öz puanlayıcıların daha esnek oldukları görülmektedir. Bir bütün olarak alınan analizde, ikinci ve üçüncü analizlerde akran ve öğretmen aynı düzeyde katılığa sahipken, birinci analizde akranların öğretmenlere göre daha esnek puanlama yaptıkları söylenebilir. Analizler detaylı olarak incelendiğinde, öz, akran ve öğretmen puanları arasında birinci analizde geniş olan aralıklar uygulamalar ilerledikçe daralmış ve tüm puanlayıcı türleri orta düzeyde ve birbirlerine yakın hâle gelmiştir.

3.3. Üçüncü Probleme İlişkin Bulgular

Uygulamada ayrı ayrı yapılan analizlerdeki puanlama türleri ile ilgili veriler beraber analize alındığında öz, akran ve öğretmen puanlarının durumu nedir?

Şekil 8'de uygulama esnasında yapılan analizlerdeki puanlama türleri ile ilgili verilerin ayrı ayrı kodlanarak yapılmış olduğu bütün analizin kalibrasyon haritası verilmektedir.



Şekil 8. Uygulamalar boyunca yapılan analizlerdeki puanlama türleri ile ilgili verilerin ayrı kodlanarak yapılmış olduğu analizinin kalibrasyon haritası

Probleme cevap bulmak için puanlama yüzeyi incelenmiştir. Bu doğrultuda Şekil 8 incelendiğinde, birinci öz puanlama ile ikinci öz puanlamanın aynı katılık düzeyinde olduğu, fakat üçüncü öz puanlamanın daha katı olduğu görülmektedir. Akranlarda birinci puanlama en cömertken, ikinci puanlama en katı ve üçüncü ise orta katılıktadır. Öğretmen puanlamaları incelendiğinde, en katı puanlamanın birinci uygulama olduğu görülmektedir. Birinci uygulamadan sonra ikinci geldiği ve en cömerdin üçüncü uygulamanın olduğu görülmektedir. Ayrıca üçüncü uygulama orta katılığa en yakındır. Tüm puanlayıcı türleri görüldüğü gibi başta daha katı ya da cömert iken, sürecin sonunda orta katılığa daha çok

yaklaşmışlardır. Başka bir deyişle, öz, akran ve öğretmen puanları katılık düzeyi olarak birbirlerine yaklaşmışlardır. Bu da puanlayıcıların zaman içerisinde öz, akran ve öğretmen puanlamalarında daha uyumlu bir yapıya ve aynı katılık düzeyine ulaştıklarını göstermektedir.

Tablo 15'te analizin uyum içi ve dışı değerleri gösterilmiştir.

Tablo 15. Bütün Analizin Puanlayıcı Türleri Katılıklarının Ölçüm Raporu

Puanlayıcı Türü	Katılık (lojit)	Hata	Uyum içi ortalama kare değeri	Uyum dışı ortalama kare değeri
Öz1	0,61	0,09	1,03	1,04
Öz2	0,59	0,09	0,85	0,83
Öz3	0,16	0,08	1,05	1,09
Akran1	0,08	0,10	1,18	1,17
Akran2	-0,36	0,09	1,07	1,07
Akran3	-0,17	0,09	1,20	1,19
Öğretmen1	-0,59	0,05	1,01	1,00
Öğretmen2	-0,38	0,05	0,92	0,93
Öğretmen3	0,05	0,05	0,96	0,97
Ortalama	0,00	0,07	1,03	1,03
Standart sapma	0,39	0,02	0,11	0,11
Ayırma indeksi: 7,44 Güvenirlik: 0,97 Ki-kare:295,4 sd: 5 p=0,00				

Tablo 15'teki tüm değerler 0,50 ile 1,50 arasında yer aldığından uyumsuz eleman belirlenememiştir. Analiz sonundaki ayırma indeksine bakıldığında, 7,44 olduğu görülmektedir. Güvenirlik katsayısı ise 0,97'dir. Bu da değerlendirme puanlayıcı türünün katılık bakımından istatistiksel olarak farklılaştığını göstermektedir. Bu durum istatistiksel olarak anlamlı (p=0,00) olan ki kare değerine (X^2=295,4) ve serbestlik derecesine (sd=5) bakıldığında da doğrulanmaktadır. Şekil 8 incelendiğinde, öz puanlamalarının genel olarak öğretmen ve akran puanlamalarına göre daha cömert ve akran puanlamalarının ise öğretmen puanlamalarından daha cömert olduğu görülmektedir.

BÖLÜM IV

SONUÇ VE TARTIŞMA

Bu bölümde bulgularda ön plana çıkan durumlar alan yazın yardımı ile tartışılmış ve araştırma sonuçları belirtilmiştir. Ayrıca son olarak araştırmayla ilgili önerilere yer verilmiştir.

4.1. Tartışma

Bulguları genel olarak özetlemek gerekirse, ayrı ayrı yapılan analizde en fazla dört olmak üzere belirlenen uyumsuz elemanların olduğu ve bir bütün olarak alınan analizde bu durum düzelmiş ve hiç uyumsuz eleman belirlenmemiştir. Bu konu ile ilgili olarak Luzn ve Linacre (1998), puanlayıcılar arasında uyumsuz öğrencilerin performansı konusunda uyumsuzlukları gösterebileceğini belirtmişlerdir. Başka bir deyişle, analizlerde belirlenen uyumsuz kişilerin gözlenen yazma başarısı değerlendiren kişiler bağlı olarak büyük ölçüde farklılaşabilir.

Puanlama kategorilerinde ise "yazım" kategorisinin en cömert puanlanan kategori olması Matsuno (2006) tarafından çalışmasıyla desteklenmektedir. Karakaya (2015) çalışmasında bu durumu yazıların dışsal özelliklerine değerlendirme sırasında çok önem verilmemesiyle açıklamıştır. Aynı şekilde "anlatım (word choice)" kategorisinin en katı puanlanan kategori olmasına bazı çalışmalarda da rastlanılmıştır (Alharby, 2006; Matsuno, 2006 ve Schaefer, 2008; Lumley ve Mcnamara, 1995). Bunun nedeni "anlatım" kategorisinin, "yazım" kategorisinden daha soyut olması olabilir. Ayrıca anlatım kategorisi kişinin bilgi birikimi ve dildeki okuma düzeyine göre değişebilir. Farklı puanlayıcılar çalışmanın yapıldığı dilde farklı anlatım tarzlarına sahip olabilirler ve bir puanlayıcı için güzel olabilecek bir anlatım diğer bir puanlayıcı için sığ ve yetersiz kalabilir. Ayrıca Matsuno (2006) çalışmasında bu yüzeyle ilgili olarak farklı bir boyut olması açısından aynı katılık düzeyinde olan

kategorilerin istatistiksel olarak gereksiz olabileceğini belirtmiştir. Fakat yaptığı çalışmada aynı değerde çıkan kategorilerin uyum içi ve uyum dışı istatistikleri istenilen aralıkta olduğundan ve her kategorinin ölçmesi gereken farklı bir bilgi olduğundan bu kategorilerin gerekliliğine karar verilmiştir. Aynı şekilde çalışmadaki kategorilerden bazıları aynı düzeyde zorlukta olmasına rağmen yüzeydeki diğer elemanlarla uyum içinde olması ve mantıksal olarak farklı bilgiler ölçmelerinden dolayı çıkarılmamıştır.

Puanlayıcı türü katılıkları hakkında ayrı ayrı yapılan analizlerde tüm analizlerde öz puanlayıcıların daha cömert oldukları görülmektedir. Öz değerlendirme puanlarının akranlara göre daha cömert olduğu sonucu bazı araştırmalar (Farrokhi, Esfandiari ve Dalili, 2011; Farrokhi, Esfandiari ve Schaefer, 2012; Karakaya, 2015; Lejk ve Wyvill, 2001; Topping, 2003) tarafından desteklenmektedir. Farrokhi, Esfandiari ve Schaefer (2012) çalışmalarında öz, akran ve öğretmen puanlamaları karşılaştırıldığında, öğrencilerin akranlarını kendilerini puanlamalarına göre daha yanlı ve katı puanladıklarını belirtmişlerdir. Bir bütün olarak alınan, ikinci ve üçüncü analizlerde akran ve öğretmen yaklaşık olarak aynı düzeyde katılığa sahipken, birinci analizde akranların öğretmenlere göre daha esnek puanlama yaptıkları söylenebilir. Akran değerlendirmelerinin öğretmen değerlendirmeler göre daha esnek olduğu sonucu alan yazında Farrokhi, Esfandiari ve Dalili (2011), Farrokhi, Esfandiari ve Schaefer (2012) tarafından desteklenirken, Karakaya (2015) çalışmasında akran değerlendirmelerinin öğretmenlere göre daha katı olduğunu belirtmiştir. Öz, akran ve öğretmen puanları arasında birinci analizde geniş olan aralıkların uygulamalar ilerledikçe daraldığı ve tüm puanlayıcı türleri orta düzeyde ve birbirlerine yakın hale geldiği gözlemlenmiştir. Bağcı-Kılıç ve Çakan (2007) çalışmalarında benzer bir sonuca rastlamışlardır. Bu da puanlayıcıların zamanla sürece alıştıkları ve daha tutarlı puanlamalara ulaştıkları şeklinde yorumlanabilir. Yani yapılan bilgilendirmelerin ve pratiklerin puanlama türlerinde bir tutarlığa ve olumlu etkiye yol açtığı kanısına varılabilir.

Analizlerde ki-kare değerleri yüzeyler içinde yer alan elemanlar arasındaki değişimlerin istatistiksel olarak anlamlı olup olmadığını göstermekte ve genel olarak her yüzeyde ki-kare değeri anlamlı bulunmaktadır. Bu da öğrenci, kategori ve görev yüzeyleri için istenilen bir durumken, puanlayıcı türünde öz, akran ve öğretmen puanları arasında istatistiksel olarak anlamlı olan farklılıklar olduğunu, bunun da değerlendirmenin güvenirliğini negatif yönde etkileyebileceği şeklinde yorumlanabilir. Güvenirlik katsayı Rasch ölçme modelinin öğrenci performansını, kategorilerin puanlanma durumlarını ve puanlayıcı türlerinin katılıklarını ne kadar güvenilir olarak ayırdığını göstermektedir (Akın

ve Baştürk, 2012). Fakat analizlere bakıldığında yine önemli bir indeks olan ve Cronbach alpha katsayısına eş değer nitelikteki güvenirlik değerlerinin puanlayıcı türü yüzeyinde gayet yüksek olduğu görülmektedir. Yapılan bu çalışmada güvenirlik değerlerinin yüksek olması analizin ortaokullardaki yazma becerilerinin değerlendirilmesinde öz, akran ve öğretmen puanları hakkında önemli bilgiler sağlayabileceği şeklinde yorumlanabilir. Güvenirlik katsayısının yüksek olmasına rağmen ki-kare değeri anlamlı olduğundan değerlendirmelerin istatistiksel olarak anlamlı farklılaştığı durumlarda Karakaya (2015) puanlayıcıların bu uygulamalar ve DPA'lar konusunda eğitilebileceklerini ve öz, akran değerlendirme uygulamalarına sınıflarda daha fazla önem verilebileceğini önermiştir. Buna ek olarak Sluijsman ve Moerkerke (1999) de öz ve akran değerlendirmelerinde yüksek bir performans elde etmek için eğitimlerin gerekli olduğunu belirtmiştir. Bu çalışmada da her hafta yapılan eğitimler ve uygulamalarla bu durum kapatılmaya çalışılmıştır. Fakat ki kare değerlerinin anlamlı olması uygulama süresinin bir eğitim dönemiyle kısıtlı olmasından kaynaklanıyor olabilir.

Ayrı yapılan analizler ve bir bütün olarak alınan analizde, öz ve akran değerlendirmelerin süreci başına göre daha katı düzeye gelerek, öğretmenlere ve dolayısıyla orta düzeye yaklaşmaları puanlayıcılar arası uyum ve güvenirliği arttırmasını sağlamıştır. Verilen eğitimler ve sınıfta toplu yapılarak uygulamalar da bu duruma katkı sağlamış olabilir yorumu yapılabilirken, süreç içerisinde her hafta yapılan tekrarların ve öğrencilerin değerlendirme ölçütlerine daha dikkat ederek yaptıkları çalışmaların yol açmış olabileceği söylenilebilir. Ayrıca yapılan analizlerde birinci ve ikinci hafta uygulanan görevin aynı zorlukta olduğu, fakat üçüncü haftaki görevin daha kolay olduğu görülmektedir. Bu da aynı şekilde öğrencilerin süreçteki gelişimleri ile yorumlanabileceği gibi yazma becerisi ne kadar karmaşık bir beceri de olsa, dereceli puanlama anahtarına ve görevlere alışma ve fazla pratik yapma ile de açıklanabilir. Bu sonucu, Eckes (2008) "Puanlayıcılar ne kadar iyi eğitilirse, puanlayıcılar arası güvenirlik o kadar yükselecektir." görüşü ile doğrulamaktadır. Bu sonuçlardan yola çıkarak öz ve akran değerlendirme uygulamalarının eğitimler verildiğinde öğretmen puanlamaları kadar etkili olduğu söylenebilir. Baştürk (2008) çalışmasında akran değerlendirmenin üç sebepten daha iyi olduğunu belirtmiştir. Sebeplere bakıldığında birincisi öğrencilerin puanlama ölçütlerinden istenilen seviyeyi kavraması ve öğrenciler ne kadar fazla puanlama yaparlarsa, o kadar fazla doğru değerlendirmelerde bulunacaklarını anlamalarıyken, ikincisi ise sınıfta sürekli bir değerlendirme ve uygulama döngüsü olduğundan, sürekli olarak sınıf içi aktivitelere

adapte olmalarıdır. Yani iki görev (değerlendirme ve bu çalışma için kompozisyon yazma) öğrencilerin sınıf aktivitelerine aktif bir şekilde katılmalarını sağlamıştır. Son olarak, puanlamalar gerçek sınıf ortamlarında gerçekleştiğinden, öğrenciler içeriksel ve konuyla bağlantılı olarak nasıl daha iyi kompozisyon yazabileceklerini öğrenebilirler. Öğrenciler kompozisyonlarını ne kadar iyi organize ederlerse, değerlendirmelerin üzerlerindeki etkilerinin o kadar fazla olacağını bilirler (Baştürk, 2008). Bu şekli ile öz ve akran değerlendirme uygulamaları aynı zamanda öğretim materyali olarak da düşünülebilirler (Lindblom-Ylänne, Pihlajamäki ve Kotkas, 2006).

4.2.Sonuçlar

4.2.1 Birinci Araştırma Sorusuna İlişkin Analiz Sonuçları

Birinci Analiz için sonuçlar: Öğrencilerin yazma becerilerini değerlendirmek amacıyla yapılan çalışmanın birinci analizinin sonucunda, öğrencilerin bazılarının (4, 7, 29 ve 38 numaralı öğrenciler) uyumiçi ve uyumdışı istatistiklerin dışında oldukları belirlenmiştir. Analizin kalibrasyon haritasında öğrencilerinin çoğunun ölçeğin orta kısmında toplandığı ve uyumsuz öğrencilerin genelde üst kısmında olduğu görülmektedir (Şekil 1). Diğer bir ifadeyle öğrencilerin genelinin başarılı olduğu ve genelde üst düzeyde yer alan öğrencilerin kendilerinden beklenenden faklı yanıtlar verdikleri belirlenmiştir. Bu durum da yüksek düzeyde başarı gösteren öğrencilerin en kolay kategorilerde düşük puan alarak, öğrenci cevapları genel örüntüsüyle tutarlı olmadıkları göstermektedir. Ayırma indeksi ve ki-kare değeri öğrencilerin yazma beceri düzeylerinin eşit olmadığını göstermektedir ve öğrencilerin yazma becerilerinin istatistiksel olarak anlamlı yaklaşık 4 beceri düzeyine ayrılabileceği gözlenmiştir. Becerileri değerlendirmek amacıyla geliştirilen öz, akran ve öğretmen DPA' daki puanlama kategorileri incelendiğinde, en cömert puanlanan kategorinin "Yazım" ve en katı puanlanan kategorinin ise "bütünsellik" olduğu görülmüştür. "Çıkarımda bulunma" ve "görsellerden faydalanma" kategorileri aynı güçlük düzeyine sahip olmalarına rağmen uyum içi ve uyum dışı ortalama kare değerleri sınırlar içinde olduğundan genel örüntüye uyum sağladıkları belirlenmiştir. Ayrıca ayrıma indeksine bakıldığında, her bir kategorinin mantıksal olarak farklı bilgiler sağladıkları söylenebilir. Puanlayıcı türlerinden en cömert puanlayan puanlayıcıların öz puanlayıcılar ve en katı puanlayanın ise öğretmen puanlayıcılar olduğu görülmektedir. Puanlayıcı

türlerinin katılık bakımından istatistiksel olarak farklılaştığını gözlenmiştir. Rasch ölçme modelinin, öğrencileri yazma becerilerine göre, kategorileri puanlanma katılıklarına göre ve puanlayıcı türlerini katılık ya da cömertlik seviyelerine göre yüksek güvenirlik değerleriyle ayırdığı belirlenmiştir.

İkinci analiz için sonuçlar: Yapılan ikinci analiz sonucunda ise, sadece bir öğrencilerinin (11 numaralı öğrenci) uyumiçi ve uyumdışı istatistiklerin dışında olduğu ve öğrencilerin yazma becerilerinin istatistiksel olarak anlamlı yaklaşık 3 beceri düzeyine ayrılabileceği gözlenmiştir. Öğrencilerinin çoğunun aynı şekilde ölçeğin orta kısmında yer aldıkları ve yazma görevinde başarılı oldukları söylenebilir. DPA' daki kategorilerden en esnek puanlanan kategoriler "yazım" ve "bütünsellik" iken, en katı puanlanan "çıkarımda bulunma" dır. Ayrıca "bütünsellik" ve "yazım" kategorileri de aynı puanlanma katılık düzeyine sahiptir. Fakat uyumiçi ve uyumdışı değerleri normal sınırlar içinden olduğundan, kategorilerin puanlanma örüntüsüyle uyumlu oldukları belirlenmiştir. Puanlayıcı türünün katılığı ile ilgili olarak, en cömert puanlayan puanlayıcıların öz puanlayıcılar ve en katı puanlayanın ise öğretmen ve akran puanlayıcılar olduğu görülmüştür.

Üçüncü analiz için sonuçlar: Üçüncü analizde elde edilen sonuçlarda ise iki öğrencinin (22 ve 54 numaralı öğrenciler) uyum içi ve uyum dışı değerlerin dışında olduğu ve öğrencilerin yazma becerilerinin istatistiksel olarak anlamlı yaklaşık 3 beceri düzeyine ayrılabileceği gözlenmiştir. DPA' daki kategorilerden en esnek puanlanan kategorinin "çıkarımda bulunma" ve en katı puanlananın "görsellerden faydalanma" olduğu belirlenmiştir. Tüm kategoriler farklı puanlanma katılık düzeyine sahiptir. Puanlayıcı türlerinden en cömert puanlayan puanlayıcıların öz puanlayıcılar olduğu ve öğretmen ve akran puanlayıcılar aynı düzeyde katılıkta olduğu görülmektedir.

4.2.2 İkinci Araştırma Sorusuna İlişkin Analiz Sonuçları

İkinci araştırma problemi kapsamında elde edilen sonuçlar incelendiğinde, uyumsuz öğrenci belirlenmemiştir. Bu durumun veri sayısının arttırılmasından kaynaklanıyor olabileceği düşünülmektedir. Kategorilerin puanlanma katılıkları incelendiğinde, "Çıkarımda bulunma", "Düzen" ve "Yazım" kategorilerinin en cömert puanlanan kategoriler oldukları ve "Görsellerden faydalanma" ve "Anlatım" kategorilerinin ise en katı puanlananlar oldukları gözlenmektedir. Bazı kategorilerin aynı katılıkta oldukları

görülmektedir. Ayrıca uyumiçi ve uyumdışı değerlerin sınırları içerisinde olduğundan kategorilerin puanlanma katılık örüntüsü ile tutarlı oldukları belirlenmiştir. Puanlayıcı türünün katılığı ile ilgili olarak öz değerlendirmelerin en cömert oldukları, öğretmen ve akran değerlendirmelerin ise eşit düzeyde ve daha katı olduğu görülmektedir. Dördüncü araştırma sorusunun birinci alt problemine bakıldığında her analizin farklı lojit ölçeğine sahip olduğundan dolayı istatistiksel olarak karşılaştırılamayacakları sonucuna varılmıştır. Şekilsel olarak gruplar içerisinde bakıldığında, 48 numaralı öğrencinin bir bütün olarak alınan, birinci ve ikinci analizlerde grup içerisindeki en yüksek beceriye sahip kişi olduğu, 19 numaralı öğrencinin ise bir bütün olarak alınan, ikinci ve üçüncü analizlerde ise grubun içinde en düşük beceriye sahip olduğu görülmektedir. Analizlerde genel olarak gruplar içerisinde öğrencilerin yerleri belirgin bir şekilde değişmemektedir. Her analizde en üst kısımda yer alan öğrenciler ve alt kısımda yer alan öğrenciler genel olarak aynıdırlar. Dördüncü araştırma sorusunun ikinci alt probleminden elde edilen sonuçlara bakıldığında "yazım" ve "düzen" kategorileri dört analizde de genel olarak üst kısımda yer almaktadırlar. "Anlatım" kategorisi ise analizlerin dördünde de alt kısımda yer almaktadır. Dört analiz incelendiğinde "bütünsellik" ve "görsellerden faydalanma" kategorileri yerlerinin analizlerde farklılaştıkları görülmektedir. Dördüncü araştırma sorusunun üçüncü alt probleminden elde edilen sonuçlar ise tüm analizlerde puanlayıcı türleri arasında öz değerlendirmecilerin en cömerttir. Ek olarak öğretmen ve akran değerlendirmelerinin birinci analiz dışındaki analizlerde eşit katılıkta en katı oldukları görülmektedir. Birinci analizde ise öğretmen değerlendirme akrandan farklı katılıkta ve en katıdır.

4.2.3. Üçüncü Araştırma Sorusuna İlişkin Analiz Sonuçları

Üçüncü ve son araştırma sorusunun cevabından elde edilen sonuçlarda, öğretmen değerlendirmelerin birinci uygulamada en katı olduğu ve gittikçe cömertleşerek en son uygulamada orta düzeye geldiği görülmektedir. Öz değerlendirmelerde ise birinci uygulamadaki katılık düzeyinin ikinci uygulamada değişmediği ve üçüncü uygulamada ise daha katılaştığı gözlenmiştir. Akran değerlendirmelerde duruma bakılacak olunursa, birinci uygulamada en cömert olduğu ve ikinci uygulamada en katı olduğu görülmektedir. Üçüncü uygulamada orta düzeye yakınlaşarak ikinci uygulamaya göre daha cömertleşmiştir. Sonuç olarak tüm değerlendirme türleri orta düzey katılığa ulaşarak birbirlerine yaklaşmışlardır.

4.2.4. Genel Sonuçlar

Öğrencilerin yazma becerilerini değerlendirmek amacıyla yapılan çalışmanın analizleri sonucunda, kalibrasyon haritalarında öğrencilerinin çoğunun ölçeğin orta kısmında toplandığı ve uyumsuz öğrencilerin genelde üst kısmında olduğu görülmektedir (Şekil 1,2,3,4). Diğer bir ifadeyle öğrencilerin genelinin başarılı olduğu ve genelde üst düzeyde yer alan öğrencilerin kendilerinden beklenenden faklı yanıtlar verdikleri belirlenmiştir. Bu durum da yüksek düzeyde başarı gösteren öğrencilerin en kolay kategorilerde düşük puan alarak, öğrenci cevapları genel örüntüsüyle tutarlı olmadıkları göstermektedir. Analizlerde genel olarak gruplar içerisinde öğrencilerin yerleri belirgin bir şekilde değişmemektedir. Yapılan tüm analizlerde en üst kısımda yer alan öğrenciler ve alt kısımda yer alan öğrenciler genel olarak aynıdırlar. Ayırma indeksi ve ki-kare değeri öğrencilerin yazma beceri düzeylerinin eşit olmadığını göstermektedir ve sürecin başında öğrencilerin yazma becerilerinin istatistiksel olarak anlamlı yaklaşık dört beceri düzeyine ayrılabileceği gözlenmiştir. Bu beceri düzeylerinin ikinci ve üçüncü analizde üç beceri düzeyine inmesi, öğrencilerin yazma becerilerinin sürecin ilerleyen uygulamalarında daha homojen bir yapıya sahip olduklarını göstermektedir.

Becerileri değerlendirmek amacıyla geliştirilen öz, akran ve öğretmen DPA'daki puanlama kategorileri incelendiğinde, en cömert puanlanan ya da öğrencilerin en başarılı oldukları kategorinin "yazım" ve en katı puanlanan veya öğrencilerin en başarısız oldukları kategorinin ise "bütünsellik" olduğu görülmüştür. Ayrıca kategorilerin bazılarının yaklaşık olarak aynı güçlük düzeyine sahip olmalarına rağmen uyum içi ve uyum dışı ortalama kare değerleri sınırları içinde olduğundan genel örüntüyle uyum sağladıkları belirlenmiştir. Ayırma indeksine bakıldığında, her bir kategorinin mantıksal olarak farklı bilgiler sağladıkları söylenebilir.

Tüm analizlerde puanlayıcı türleri arasında öz değerlendirmecilerin en cömert olduğu belirlenmiştir. Ek olarak öğretmen ve akran değerlendirmelerinin birinci analiz dışındaki analizlerde yaklaşık olarak eşit katılıkta en katı oldukları görülmektedir. Birinci analizde ise öğretmen değerlendirme akrandan daha katıdır. Puanlayıcı türlerinin katılık bakımından istatistiksel olarak farklılaştığını gözlenmiştir.

Öğretmen değerlendirmelerin birinci uygulamada en katı olduğu ve gittikçe cömertleşerek en son uygulamada orta düzeye geldiği görülmektedir. Öz değerlendirmelerde ise birinci

uygulamadaki katılık düzeyinin ikinci uygulamada anlamlı bir düzeyde değişmediği ve üçüncü uygulamada ise daha katılaştığı gözlenmiştir. Akran değerlendirmelerde duruma bakıldığında, birinci uygulamada en cömert olduğu ve ikinci uygulamada en katı olduğu görülmektedir. Üçüncü uygulamada orta düzeye yakınlaşarak ikinci uygulamaya göre daha cömertleşmiştir. Tüm değerlendirme türleri orta düzey katılığa ulaşarak birbirlerine yaklaşmışlardır.

Son olarak, ÇYRÖM'nin uygulama sürecindeki yüzeyler ve bunların modele olan uyumları hakkında ayrıntılı bilgiler sağlayan ve yüzeylerden elde edilen verilerle yüzeylerdeki elemanlar (öğrenci, kategori, görev ve puanlayıcı) hakkında geleneksel test analizlerinden farklı olarak daha derinlemesine bilgiler sağlayan kullanışlı bir ölçme modeli olduğu sonucuna varılmıştır. Aynı şekilde ÇYRÖM performans değerlendirme konusunda program geliştiricilere ve sınıflarında performans değerlendirme teknikleri kullanan öğretmenlere derinlemesine bilgiler sağlayabilir.

4.3. Öneriler

Bu çalışma sonucunda öneriler, araştırma sonuçlarından çıkan öneriler ve araştırmacılara yönelik öneriler olmak üzere iki başlık altında verilmiştir.

4.3.1. Araştırma Sonuçlarından Çıkan Öneriler

1. İkinci araştırma sorusunun ikinci alt probleminden elde edilen sonuçlarda birinci analizde lojit ölçeğinin üst kısmında yer alan öğrenci sayısının ikinci ve üçüncü analizlerde arttığı görülmektedir. Bu durumda öz ve akran değerlendirme çalışmalarının yazma becerisinin artmasında olumlu bir etkiye sahip olabileceği söylenebilir. Bu sonuç doğrultusunda öz ve akran değerlendirme uygulamalarının Türkçe derslerinde yazma becerilerini geliştirmek amacıyla kullanılması önerilmektedir.
2. İkinci araştırma sorusunun üçüncü alt probleminden elde edilen sonuçlarda öğrencilerin genel olarak "anlatım" ve "bütünsellik" kategorilerinde daha az başarılı oldukları görülmektedir. Bu kapsamda öğrencilerin "anlatım" kategorisi altında yer alan kelime seçimi, örnek ve betimlemelerden yararlanmalarını sağlayacak çalışmalar

ile bütünsellik kategorisi altında yer alan giriş, gelişme ve sonuç bölümlerinin geliştirilmesine yönelik çalışmalara eğitim süreçlerinde yer verilmesi önerilmektedir.

3. İkinci araştırma sorusunun üçüncü alt probleminden elde edilen sonuçlarda "bütünsellik" kategorisinin diğer kategorilere göre daha az başarılı olunan bir kategori olduğu, fakat sürecin ilerlemesiyle öğrencilerin bu kategoride daha başarılı bir düzeye ulaştığı gözlenmiştir. Bu sonuca dayanarak eğitim süreçlerinde öz ve akran değerlenme çalışmalarına daha sık yer verilmesi önerilmektedir.
4. Üçüncü araştırma sorusunun sonucunda, öz, akran ve öğretmen puanlamaları süreç içerisinde birbirine yaklaşmışlardır. Başka bir ifadeyle, uygulamalar yapıldıkça, öğrenciler kendilerini ve akranlarını daha nesnel değerlendirmeye ve öğretmen puanlarına daha yakın puanlamalar yapmaya başlamışlardır. Buna bağlı olarak Türkçe derslerinde düzenli olarak bu uygulamaların kullanılarak öğrencilerin değerlendirmelerine daha çok yer verilmesi önerilmektedir.

4.3.2. Araştırmacılara Yönelik Öneriler

1. Bu çalışmada öğrencilerin genel ders başarı durumları göz önünde bulundurulmamıştır. Öğrencilerin genel ders başarılarıyla olan ilişkilerine de bakılarak araştırmanın kapsamı genişletilebilir.
2. Benzer bir çalışmanın farklı örneklem büyüklükleri üzerinde yapılması ve örneklem büyüklüğüne bağlı olarak analizlerdeki farklılaşmalar incelenebilir.
3. Belirlenen sonuçları desteklemek amacıyla nitel bir çalışma eklenerek çalışma geliştirilebilir.
4. Çalışma farklı becerilerde ve daha uzun süreli olarak uygulanabilir.

KAYNAKÇA

Akın, Ö. & Baştürk, R. (2012). Keman eğitiminde temel becerilerin Rasch ölçme modeli ile değerlendirilmesi. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi, 31*(1), 175-187.

Alharby, E. R. (2006). *A comparison between two scoring methods, holistic vs. analytic, using two measurement models, the generalizability theory and the many-facet rasch measurement, within the context of performance assessment*. Doktora Tezi. The Pennsylvania State University.

Alıcı, D. (2014). Öğrenci performansının değerlendirilmesinde kullanılan diğer ölçme araç ve yöntemleri. Satılmış Tekindal (Ed.) *Eğitimde ölçme ve değerlendirme* içinde (s. 131-174). Ankara: Pegem Akademi.

Andrade, H., Du, Y. & Mycek, K. (2010). Rubric-referenced self-assessment and middle school students' writing. *Assessment in Education: Principles, Policy & Practice, 17* (2), 199-214.

Arı, G. (2011). Türkçe (6, 7, 8. sınıf) ders kitaplarındaki okuma ve dinleme/izleme metinleri ile yazma görevleri arasındaki tür uyumu. *Turkish studies, 6* (3), 489-511.

Atılgan, H. (2005). Genellenebilirlik kuramı ve puanlayıcılar arası güvenirlik için örnek bir uygulama. *Eğitim Bilimleri ve Uygulama*, *4*(7), 67-79.

Bağcı-Kılıç, G., & Çakan, M. (2007). Peer Assessment of elementary science teaching skills . *Journal of Science Teaching Education, 18*(91-107).

Ballantyne, R., Hughes, K. & Mylonas, A. (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment & Evaluation in Higher Education, 27*(5), 427-441.

Baştürk, R. (2008). Applying the many – facet rasch model to evaluate PowerPoint presentation performance in higher education. *Assesment and Evaluation in Higher Education, 33*(4), 431-444.

Baştürk, R. (2010). Bilimsel araştırma ödevlerinin çok yüzeyli rasch ölçme modeli ile değerlendirilmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 1*(1), 51-57.

Bayat, Ö. (2010). İngilizce yazılı anlatım derslerinde uygulanan akran ve öz değerlendirme etkinliklerine yönelik öğrenci görüşleri. *Dil Dergisi, 150*, 70-81.

Berberoğlu, G. (1988). *Seçme amacıyla kullanılan testlerde Rasch modelinin katkıları.* DoktoraTezi. Hacettepe Üniversitesi, Ankara.

Borman, W. C. (1977). Consistency of rating accuracy and rating errors in the judgment of human performance. *Organizational Behavior and Human Performance, 20*, 238-252.

Boud, D. & Lublin, J. (1983). Self-assessment in professional education. A report to the commonwealth research and development committee (Tertiary Education Research Center, University of New South Wales).

Boud, D., Cohen, R. & Sampson, J. (1999). Peer learning and assessment. *Assessment and Evaluation in Higher Education, 24*(4), 413-426.

Brennan, R. L. (2001). *Generalizability theory.* New York: Springer-Verlag.

Brookhart, S. M. (1999). The art and science of classroom assessment: the missing part of pedagogy. *ASHE-ERIC Higher Education Report* (Vol.27,1). Washington DC: The George Washington University, Graduate School of Education and Human Development. 25.08.2015 tarihinde http://files.eric.ed.gov/fulltext/ED364576.pdf sayfasından erişilmiştir.

Brown, J. D. & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly, 32*(4), 653-675.

Büyüköztürk, Ş., Çakmak, K. E., Akgün, E. Ö., Karadeniz, Ş. & Demirel, F. (2014). *Bilimsel araştırma yöntemleri.* Ankara: Pegem Akademi.

Carter, C., Bishop, J. & Kravits, L. (2002). *Key to effective learning* (3. baskı), New Jersey: Printice Hall.

Çakıcı Eser, D. & Gelbal, S. (2011). Genellenebilirlik kuramı ve lojistik regresyona dayalı hesaplanan puanlayıcılar arası tutarlığın karşılaştırılması. *Kastamonu Eğitim Dergisi, 21*(2), 421-438.

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, 5, 194-197.

Demir, T. (2003). Yazma becerisine yönelik kazanımların ortaokulda gerçekleşme düzeyi. *Okuma Yazma Eğitimi Araştırmaları, 1*(1), *45-56.*

Demirel, Ö. (2003). *Türkçe öğretimi*. Ankara: Pegem

Deniz, K. (2003). Yazılı anlatım becerileri bakımından köy ve kent beşinci sınıf öğrencilerinin durumu. *Türklük Bilimi Araştırmaları* (Türkçenin Öğretimi Özel Sayısı), *13*, 233- 255.

Dochy, E, Segers, M. & Sluusmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education, 24*(3), 331-350.

Eckes, T. (2008). Raters types in writing performance assessment: A classfication approach to rater variability. *Language Testing, 25*(2), 155-185.

Engelhard, G. (1996). Evaluating rater accuracy in performance assessment. *Journal of Education Measurement. 33*(1), 56-70.

Engelhard, Jr. & Myford, M. C. (2003). *Monitoring faculty consultant performance in the advanced placement English literature and composition program with a many-faceted Rasch model.* (College Board Research Rep. No. 2003-1). New York: College Entrance Examination Board. 25.08.2015 tarihinde http://files.eric.ed.gov/fulltext/ED345598.pdf sayfasından erişilmiştir.

Erman Aslanoğlu, A. & Kutlu, Ö. (2003). Öğretimde sunu becerilerinin değerlendirilmesinde dereceli puanlama anahtarı (rubric) kullanılmasına ilişkin bir araştırma. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, 36* (1-2), 25-36.

Evans, J. (2001). Introduction: learning and teaching the complexities of writing. Janet Evans (Ed.), *Writing in the Elementary Classroom: A Reconsideration. Portsmout* içinde. NH: Heinemann.

Farrokhi F. & Esfandiari R. (2011). A many-facet rasch model to detect halo effect in three types of raters. *Theory and Practice in Language Studies, 1* (11), 1531-1540.

Farrokhi F., Esfandiari R. & Dalili, M. V. (2011). Applying the many-facet rasch model to detect centrality in self-assessment, peer-assessment and teacher assessment. *World Applied Sciences Journal, 15 (Innovation and Pedagogy for Lifelong Learning),* 70-77.

Farrokhi, F., Esfandiari, R. & Schaefer, E. (2012). A many-facet Rasch measurement of differential rater severity/leniency in self assessment, peer assessment, and teacher assessment. *Journal of Basic and Applied Scientific Research, 2* (9), 8786-8798.

Goodrich, A. H. (2001). The effects of instructional rubrics on learning to write. *Current Issues in Education, 4* (4), 263-272.

Goodwin, L. D. (2001). Interrater agreement and reliability. *Measurement in Psychical Education and Exercises Science*, *5*(1), 13-14.

Göçer, A. (2010). Türkçe öğretiminde yazma eğitimi. *Uluslararası Sosyal Araştırmalar Dergisi, 3* (12), 178-195.

Göğüş, B. (1978). *Orta dereceli okullarımızda Türkçe ve yazın eğitimi*. Ankara: Kadıoğlu.

Güler, N. (2008). *Klasik test kuramı genellenebilirlik kuramı ve Rasch modeli üzerine bir araştırma*. Doktora tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.

Gwet, K. L. (2012). *Handbook of ınter-rater reliability: the definitive guide to measuring the extent of agreement among raters*. Advance Anaytics LLC, USA.

Gyagenda, İ. S. & Engelhard, G. Jr. (1998). *Applying the rasch model to explore rater ınfluences on the assessed quality of students' writing ability.* Presented at the Annual Meeting of the American Educational Research Association , San Diego, CA,USA.

Haladyna, T. M. (1997). *Writing test ıtem to evaluate higher order thinking*. USA: Allyn & Bacon.

Hallgren K. A. (2012). Computing ınter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol. 8*(1), 23-34.

Hambleton, R. K., Swaminathan, H. & Rogers, H.J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage.

Hargreaves, A., Earl, L. & Schimidt, M. (2001). Perspectives on alternative assessment reform. *American Educational Research Journal, 39*(1), 69-95.

Hartmann, D. P. (1977). Considerations in the choice of interobserver reliability estimates. *Journal of Applied Behavior Analysis, 10*, 103-116.

Hertherman, S. C. (2004). *An application of multi faceted rasch measurement to monitor effectiveness of the written composition in english in the new york city department of education.* Doktora tezi. Teacher College, Colombia University, Colombia.

Karakaya, I., Sarıtaş, S. & Salmaner, R. (2015, Mayıs). *Assessment of performance based tasks within the context of statistics lesson with multi faceted rasch model*, International Congress on Education for the Future: Issues and Challenges'da sunulmuş bildiri. Ankara Üniversitesi, Ankara.

Karakaya, İ. (2015). Comparison of self, peer and ınstruction assessments in the portfolio assessments by using many facet rasch model. *Journal of Education and Human Development, 4*(2), 182-192.

Karakaya, İ. (2015a). Öğrencilerin değerlendirme sürecine katılması öz, akran ve grup değerlendirme. Şener Büyüköztürk (Ed.), *Ölçme ve değerlendirme* içinde (s.130-150). Erzurum : Atatürk Üniversitesi.

Karasar, N. (1998). *Bilimsel araştırma yöntemleri*. Ankara: Nobel.

Kılıç, S. (2009). *Ölçümlerin Uyumluluğu ve Tıptaki Uygulamaları*. Yüksek Lisans Tezi. Çukurova Üniversitesi, Adana.

Kutlu, Ö., Doğan, C. D. & Karakaya İ. (2014). *Eğitimde ölçme ve değerlendirme - performansa ve portfolyoya dayalı durum belirleme.* Ankara: Pegem .

Lane, S. & Stone, C. A. (2006). Educational measurment, dördüncü baskı. Robert Brennan (Ed.) USA: American Council On Education Praeger.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28,* 563–575.

Lejk, M. & Wyvill, M. (2001). The Effect of the inclusion of self-assessment with peer assessment of contributions to a group project: A Quantitative study of secret and agreed assessments, *Assessment and Evaluation in Higher Education, 26* (6), 551–561.

Linacre, J. M. (1993, April). Generalizability theory and many facet rasch measurement. *Annual Meeting Of The American Educational Research Association' de sunulmuş bildiri.*(ED 364 573). Atlanta Georgia. 15.11.2014 tarihinde http://files.eric.ed.gov/fulltext/ED364573.pdf sayfasından erişilmiştir.

Linacre, M. J. (2002). Immediate raw score to logit conversion [Electronic version],*Rasch Measurement Transactions,* 16, 877. 15.11.2014 tarihinde http://www.rasch.org/rmt/rmtl62e.htm sayfasından erişilmiştir

Lindblom-Ylänne, S., Pihlajamäki, H. & Kotkas, T. (2006). Self, peer and teacher assessment of student essays. *Active Learning in Higher Education, 7* (1), 51-62.

Lumley, T. & McNamara, T. F. (1995). Rater characteristics and rater bias: Implications for training. *Language Testing, 12*(1), 54-71.

Lunz, M. J. & Linacre, J. M. (1998). Measurement designs using multifacet Rasch modeling. G A. Marcoulides (Ed.), *Modern methods for business research* içinde (s. 44-47). Mahwah, NJ: Lawrence Erlbaum.

Macmillan, P. D. (2000). Classical generalizability and multifaceted rasch detection ınterrater variability in large, sparse data set. *Journal of Experimental Education. 68*(2), 167-190.

Marzona, R. J. (2001). *Designing a new taxonomy of educational objectives.* Thousand Oak, California: Corwin.

Matsuno, S. (2006). *Self-, peer-, and teacher- assessment in japanese university efl writing classrooms.* Doktora Tezi. Temple University.

McGartland, R. D., Berg-Weger, M., Tebb, S., Lee, E. S., & Rauch, S. (2003).Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research, 27*(2), 94 - 104.

McNamara, T. F., & Adams, R. J. (1991). Exploring rater behaviour with Rasch Techniques. *(ERIC Document Reproduction Service* No. ED345498) 25.08.2015 tarihinde http://des.emory.edu./home/documents/EngArticle.pdf sayfasından erişilmiştir.

Moskal, B. M. (2000). Scoring rubrics: what, when and how? *Practical Assessment, Research and Evaluation, 7*(3), 70-80.

Myford, C. M. & Wolfe, E. W. (2003). Detecting and measuring rater effects using many-facet Rasch measurement: Part II. *Journal o f Applied Measurement, 5(2),* 189-227.

Nakamura, Y. (2002). Teacher assessment and peer assessment in practice. *Educational Studies, 44*, 45-85.

Noonnan, B. & Randy, D. (2005). Peer and self assessment in high schools. *Pratical Assessment Research & Evaluation,10*(17), 1-8.

Novick, M. R. (1966). The axioms and principle results of classical test theory. *Journal of Mathematical Psychology*. *3*,1–18.

Rasch, G. (1980). *Probabilistic models for some intelligence and achievement tests.* Copenhagen: Danish Institute for Educational Research. Chicago: The University of Chicago.

Schaefer, E. (2008). Rater bias patterns in an EFL writing assessment. *Language Testing, 25*, 465-493.

Semerci, Ç. (2011). Doktora yeterlikler çerçevesinde öğretim üyesi, akran ve öz değerlendirmelerin Rasch ölçme modeliyle analizi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 2*(2), 164-17.

Shalvelson, R. J. & Webb, N. M. (1991). *Generalizability theory: a primer*. USA: Sage.

Sluijsman, D. & Moerkerke, G. (1999). Student involvement in performance assessment: A research Project. *Europen Journal of Open and Distance Learning.* 15.11.2014 tarihinde http://www.eurodl.org/materials/contrib/1999/assessment_issue/sluijsmans/ sayfasından erişilmiştir.

Sim, J. & Wright, C. C. (2005). The kappa statistic in reliability studies: use, ınterpretation, and sample size requirements. *Physical Therapy, 85*(3):257-268

Smith, Jr., E.V. (2001). Evidence for the reliability of measures and validity of measure interpretation: A Rasch measurement perspective. *Journal of Applied Measurement, 2*, 281-311.

Sperling, M. & Freedman, S.W. (2001) Review of writing research. V.Richardson (Ed.), *Handbook of research on teaching* içinde (s. 370-389). Washington. DC: AMerican Educational Research. 15.11.2014 tarihinde

http://gse.berkeley.edu/sites/default/files/users/sarah-freedman/01chResWrtngHndbk.pdf sayfasından erişilmiştir.

Sudweeks, R. R., Reeve, S. & Bradshaw, W.S. (2005). A comparison of generalizability theory and many-facet Rasch measurement in an analysis of college sophomore writing. *Assessing Writing. 9*, 239-261.

Temizkan, M. (2009). Akran değerlendirmenin konuşma becerisinin geliştirilmesi üzerindeki etkisi. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 6*(12), 90-112.

Topping, K. J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*, 249-276.

Topping, K. (2003). Self and peer assessment in school and university: Reliability, validity and utility. *Optimising New Modes of Assessment: In Search of Qualities and Standards Innovation and Change in Professional Education*, *1*, 55-87.

Turgut, M. F. & Baykul, Y. (2010). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem.

Ungan, S. (2007). Yazma becerisinin geliştirilmesi ve önemi. *Sosyal Bilimler Enstitüsü Dergisi, 23* (2), 461-472.

Weigle, S. C. (1998). Using FACETS to model rater training effects. *Language Testing,15*(2), 263-287.

Weigle, S. C. (1999). Investigating rater/prompt ınteractions in writing assessment: quantitative and qualitative approaches. *Assessing Writing, 6*(2),145-178.

White, J. (1999). Thinking about assessment. *Journal of Philosophy of Education*, *33*(2), 201.

Wiliam, D. (2003). National curriculum assessment: how to make it better. *Research Papers in Education*, *18*(2), 129-136.

Wright, B. D. & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8* (30), 370.

Yangın, B. (2002). *Kuramdan uygulamaya Türkçe öğretimi*. Mersin: Dersal.

Yurdugül, H. (2005). Ölçek geliştirme çalışmalarında kapsam geçerliği için kapsam geçerlik indekslerinin kullanılması. XIV. Ulusal Eğitim Bilimleri Kongresi'nde sunulmuş bildiri, Pamukkale Üniversitesi Eğitim Fakültesi, Denizli.

Yüzüak, A. V., Yüzüak, B. & Kaptan, F. (2015). Performans görevinin akran gruplar ve öğretmen yaklaşımları doğrultusunda çok-yüzeyli rasch ölçme modeli ile analizi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 6*(1),1-11.

EKLER

EK 1: Görevleri ve Dereceli Puanlama Anahtarlarını Değerlendirme Formu

PERFORMANS GÖREVLERİ İÇİN DEĞERLENDİRME FORMU

Değerli Uzman;

Yazma becerilerinin değerlendirilmesinde öz, akran ve öğretmen puanlarını çok yüzeyli Rasch ölçme modeli ile karşılaştırmak amacıyla yüksek lisans tez çalışmasını yürütmekteyim. Tez çalışması kapsamında uygulanacak ölçme araçlarının (performans görevleri ve performans görevlerini değerlendirme formu) uzman görüşleri doğrultusunda geliştirilmesi gerekmektedir.

Bu amaç doğrultusunda; öğrencilere yazma becerilerinin geliştirilmesi ve puanlanması amacıyla hazırlanan performans görevleri ile bu performans görevlerinin uzman görüşleri doğrultusunda değerlendirmesi için uzman görüş formu geliştirilmiştir. Geliştirilen bu formların değerli uzmanların görüşleri doğrultusunda son şeklini alacak ve hazırlanan araçlarının psikometrik özelliklerinin gelişmesine önemli katkı sağlayacaktır.

Form içerisindeki ölçütlerin performans görevleri ile dereceli puanlama anahtarlarının geliştirilme sürecini dikkate alarak ölçütlerin amaç doğrultusunda ne derece etkili olduğuna dair görüşlerinizi tamamen uygunsa (4), uygun ise (3), bazı değişiklikler gerekiyorsa(2) ve ölçüt tamamen veya önemli bir kısmı değişmesi gerekiyorsa (1) işaretleyerek görüşlerinizi belirtiniz.

Yapmış olduğunuz değerlendirmeler ve öneriler için teşekkür ederiz.

Not: Öneriler kısmında kaç numaralı performans görevi için öneride bulunduğunuzu ***ÖZELLİKLE*** belirtiniz.

Doç. Dr. İsmail KARAKAYA

Gazi Üniversitesi

Gazi Eğitim Fakültesi
Eğitim Bilimleri Bölümü

Eğitimde Ölçme ve Değerlendirme AD

Arş. Gör. Rana SALMANER

Gazi Üniversitesi

Gazi Eğitim Fakültesi
Eğitim Bilimleri Bölümü

Eğitimde Ölçme ve Değerlendirme AD

1. Adı ve Soyadı: ……. 2.Unvan: …….

3. Deneyim :…….. 4. Branş …….

5. Eğitim düzeyi ( ) Lisans ( ) Yüksek lisans ( ) Doktora

UYGUNLUĞUNA BAKILCAK ÖLÇÜTLER		**DERECELER**			
	GÖREV	1	2	3	4
Temel dil becerilerinin öğretilmesine ve değerlendirilmesine uygundur. **Öneri:**	1				
	2				
	3				
	4				
	5				
	6				
	7				
Yaratıcılık, problem çözme, iletişim kurma, araştırma yapma, analitik düşünme ve eleştirel düşünme gibi üst düzey zihinsel süreçlerden bir ya da bir kaçına uygundur. **Öneri:**	1				
	2				
	3				
	4				
	5				
	6				
	7				
Görev kısımda verilen soru öğrencilerin dil düzeyine uygun, anlaşılır ve ilgi çekicidir. **Öneri:**	1				
	2				
	3				
	4				
	5				
	6				
	7				
Görevde verilen problem durumu belirlenen zihinsel süreçlere (beklenen performansa ve kazanımlara) uygundur. **Öneri:**	1				
	2				
	3				
	4				
	5				
	6				

	7				
Yönerge kısmında öğrencilerin görevi yaparken dikkat etmesi gereken noktaların hepsi verilmiştir. **Öneri:**	1				
	2				
	3				
	4				
	5				
	6				
	7				
Yönergede yer alan açıklamaların ayrıntısı ve anlatım biçimi öğrencinin sınıf, dil düzeyine ve donanımına uygundur. **Öneri:**	1				
	2				
	3				
	4				
	5				
	6				
	7				

UYGUNLUĞUNA BAKILCAK ÖLÇÜTLER		**DERECELER**			
Performans görevinin değerlendirilmesi amacıyla hazırlanan dereceli puanlama anahtarı (DPA) hazırlanırken olabildiğince az sayıda göreli ifade kullanılmıştır. **Öneri:**	**GÖREV**	1	2	3	4
	1				
	2				
	3				
	4				
	5				
	6				
	7				
DPA içerisindeki ölçütler ayrı ayrı tanımlanmış ve her ölçüt öğrencinin kullandığı bir beceri hakkında bilgi vermiştir. **Öneri:**	1				
	2				
	3				
	4				
	5				
	6				
	7				

DPA içerisindeki ölçtüler beklenen performansların tamamının puanlanması için yeterlidir. **Öneri:**	1				
	2				
	3				
	4				
	5				
	6				
	7				
Görev ve puanlama için verilen süreler kişilerin düzeyine uygundur ve verilen görev için yeterlidir. **Öneri:**	1				
	2				
	3				
	4				
	5				
	6				
	7				
Tanımlama, görev, yönerge ve puanlanma kısımlarının tamamında gereksiz kavram ve şekiller kullanılmamıştır. **Öneri:**	1				
	2				
	3				
	4				
	5				
	6				
	7				

Genel Öneriler:

EK 2: 1. Hafta Kullanılan Performans Görevi ve Dereceli Puanlama Anahtarı

Okul numarası:.....

İçerik Düzeyi	Sınıf Düzeyi	Beklenen Performans	Puanlama Yöntemi	Performans görevinin Süresi
TÜRKÇE **Tema: DOĞA VE EVREN** **Kazanımlar:** **Tür, Yöntem ve Tekniklere Uygun Yazma** 1. Açıklayıcı ve bilgilendirici yazılar yazar. 2. Yazılarında betimlemeler yapar. **Görsel Okuma** **1.** Resim ve fotoğrafları yorumlar.	5.sınıf	*Etkili Yazma *Görsel okuma *Görsellerden çıkarımlarda bulunma	Öz, Akran ve Öğretmen Puanlaması	1 Ders saati



Merhaba arkadaşlar...

Ben sizlere ve ailelerinize sağlıklı bir çevre oluşturmak için çalışan temizlik görevlilerinden yalnızca biriyim. Biz sizler için sağlıklı bir yaşam ve ortam oluşturmaya çalışırken çevremizin zaman içinde ne kadar farklılaştığını fark ettik.

Arka sayfada yaşadığımız kasabanın 20 yıl önceki hâli ve bu günkü hâlini gösteren resimler var. Resimlerdeki değişimi bulmaya çalışalım ve bu değişimlerin sebeplerinin ne olduğunu yazalım. Sonra kasabamın bundan bir 20 yıl sonra ne hâle gelebileceğini ve bu değişimlerin sebeplerinin yazarak yazımızı devam ettirelim.

YÖNERGE

1- Yazdığın metnine bulacağın başlık, metnin içeriği ile ilişkili olmalıdır.

2- Çalışmanın sana ait olmasına özen göstermelisin.

3- Dersin ilk 15 dakikasında(5 dakika ön hazırlık+10 dakika metnin yazımı) çalışmalar tamamlanır.

4- Dersin son 25 dakikasında(5 dakika öz değerlendirme+10 dakika akran + 10 dakika akran) çalışmalar sınıfta değerlendirilir. Öğretmenler çalışmalar hakkında dersten sonra dönüt verir.

5- Çalışma yazının bütünselliği, anlatım, yazım ve dilbilgisi, düzen, görsellerden faydalanma ve çıkarımda bulunma gibi ölçütler açısından değerlendirilecektir.



Kasabanın 20 yıl önceki hâli



Kasabanın bugünkü hâli

.....

1. GÖREVİN DERECELİ PUANLAMA ANAHTARI

Görev no:.....

Ölçütler	Başarı düzeyleri			Oz puanı	Akran puanı	Başarı Puanı
	(G) Geliştirilebilir	(İ) İyi	(M) Mükemmel			
Yazının Bütünselliği	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmamıştır.	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmuş, fakat bölümlerde karışıklıklar yaşanmıştır.	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmuş ve bölümler kurallarına göre yazılmıştır.			
Anlatım	Yazının anlatımını kuvvetlendirecek örneklere ya da betimlerden 1-2 yerde faydalanılmıştır.	Yazının anlatımını kuvvetlendirecek örneklere ya da betimlerden 3-4 yerde faydalanılmıştır.	Yazının anlatımını kuvvetlendirecek örneklere ya da betimlerden 5 ve daha fazla yerde faydalanılmıştır.			
Yazım ve Dilbilgisi	Çalışmada fazlaca yazım ve dilbilgisi hatası vardır.	Çalışmada biraz yazım ve dilbilgisi hatası vardır.	Çalışmada hemen hemen hiç yazım ve dilbilgisi hatası yoktur.			
Düzen	Kâğıt kenarlarına, paragraf ve satır aralıklarına neredeyse hiç dikkat edilmemiştir.	Yazının bir kısmında kenar, paragraf ve satır aralıklarına dikkat edilmiştir.	Yazının çoğunda kenar, paragraf ve satır aralıklarına dikkat edilmiştir.			
Görsellerden Faydalanma	Resimlerdeki değişimlerin 1-2 tanesinden faydalanılmıştır.	Resimlerdeki değişimlerin 3-4 tanesinden faydalanılmıştır.	Resimlerdeki değişimlerin 5 ve daha fazlasından faydalanılmıştır.			
Çıkarımda Bulunma	Resimdeki değişimlerden faydalanarak 20 yıl sonraki kasaba hakkında 1-2 tahminde bulunulmuştur.	Resimdeki değişimlerden faydalanarak 20 yıl sonraki kasaba hakkında 3-4 tahminde bulunulmuştur.	Resimdeki değişimlerden faydalanarak 20 yıl sonraki kasaba hakkında 5 ve daha fazla tahminde bulunulmuştur.			

EK 3: 2. Hafta Kullanılan Performans Görevi ve Dereceli Puanlama Anahtarı

Okul Numarası: .....

İçerik Düzeyi	Sınıf Düzeyi	Beklenen Performans	Puanlama Yöntemi	Performans görevinin Süresi
TÜRKÇE **Yazma Çalışmaları Kazanımlar:** **Tür, Yöntem ve Tekniklere Uygun Yazma** 1. Açıklayıcı ve bilgilendirici yazılar yazar. 2. Hikâye yazar. 3. Yazılarında betimlemeler yapar. **Görsel Okuma** 1. Grafik ve tablo ile verilenleri yorumlar.	5.sınıf	*Etkili Yazma *Görsel okuma *Görsellerden çıkarımlarda bulunma	Öz, Akran ve Öğretmen Puanlama	1 Ders saati



Uyumak ne güzel değil mi? Siz de benim gibi uyumayı çok seviyorsunuzdur. Aşağıdaki araştırma oldukça dikkat çekici. Gerektiğinde şemadaki yerleri karşılaştırarak beş meslek kolunun uyuma saatlerini ve nedenlerini yazınız.

Bu saatlerin sebep olabileceği sonuçlar nelerdir değininiz.

YÖNERGE

1- Yazdığın metnine bulacağın başlık, metnin içeriği ile ilişkili olmalıdır.

2- Çalışmanın sana ait olmasına özen göstermelisin.

3- Dersin ilk 15 dakikasında(5 dakika ön hazırlık+10 dakika metnin yazımı) çalışmalar tamamlanır.

4- Dersin son 25 dakikasında(5 dakika öz değerlendirme+10 dakika akran + 10 dakika akran) çalışmalar sınıfta değerlendirilir. Öğretmenler çalışmalar hakkında dersten sonra dönüt verir.

5- Çalışma yazının bütünselliği, anlatım, yazım ve dilbilgisi, düzen, görsellerden faydalanma ve çıkarımda bulunma gibi ölçütler açısından değerlendirilecektir.







ZAMAN DİLİMLERİ	SAATLER / Meslekler	Öğrenci	Kamyon şoförü	Çalışan anne	İş adamı	Doktor
GECE	Saat: 23.00- 00.00	■		■	■	
	Saat: 00.00- 01.00	■		■	■	
	Saat: 01.00- 02.00	■			■	■
	Saat: 02.00- 03.00	■			■	■
	Saat: 03.00-04.00	■		■	■	
SABAH	Saat: 04.00- 05.00	■		■	■	■
	Saat: 05.00- 06.00	■		■		■
	Saat: 06.00-07.00	■				■
	Saat: 07.00-08.00		■			
	Saat: 08.00-09.00		■			
	Saat: 09.00- 10.00		■			
	Saat: 10.00- 11.00					
	Saat: 11.00- 12.00					
ÖĞLEN	Saat: 12.00- 13.00					
ÖĞLEDEN SONRA	Saat: 13.00-14.00			■		
	Saat: 14.00- 15.00			■		
	Saat: 15.00- 16.00					
	Saat: 16.00- 17.00		■			
AKŞAM	Saat: 17.00- 18.00		■			
	Saat: 18.00-19.00					
	Saat: 19.00-20.00					
	Saat: 20.00-21.00					
	Saat: 21.00-22.00					
	Saat: 22.00-23.00					

■ : Uyuma saatleri □ : Uyanık olunan saatler

DERECELİ PUANLAMA ANAHTARI

Görev no: .....

Ölçütler	Başarı düzeyleri			Öz Puanı	Akran puanı	Öğretmen Puanı
	(G) Geliştirilebilir	(İ) İyi	(M) Mükemmel			
Yazının Bütünselliği	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmamıştır.	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmuş, fakat bölümlerde karışıklıklar yaşanmıştır.	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmuş ve bölümler kurallarına göre yazılmıştır.			
Anlatım	Anlatımı kuvvetlendirecek 1-2 tane örnek ya da betimlemeye yer verilmiştir.	Anlatımı kuvvetlendirecek 3-4 tane örnek ya da betimlemeye yer verilmiştir.	Anlatımı kuvvetlendirecek 5 ya da daha fazla örnek ya da betimlemeye yer verilmiştir.			
Yazım ve Dilbilgisi	Çalışmada fazlaca yazım ve dilbilgisi hatası vardır.	Çalışmada biraz yazım ve dilbilgisi hatası vardır.	Çalışmada hemen hemen hiç yazım ve dilbilgisi hatası yoktur.			
Düzen	Kâğıt kenarlarına, paragraf ve satır aralıklarında boşluklara neredeyse hiç dikkat edilmemiştir.	Yazının bir kısmında kenar, paragraf ve satır aralıklarındaki boşluklara dikkat edilmiştir.	Yazının çoğunda kenar, paragraf ve satır aralıklarındaki boşluklara dikkat edilmiştir.			
Görsellerden faydalanma	Verilen ipuçlarının 1-3 tanesine sebep belirtmiştir.	Verilen ipuçlarının 4-7 tanesine sebep belirtmiştir.	Verilen ipuçlarının 7-9 tanesine sebep belirtmiştir.			
Çıkarımlarda bulunma	Tablodaki değişimlerden faydalanarak 1-2 çıkarımda bulunmuştur.	Tablodaki değişimlerden faydalanarak 3-4 çıkarımda bulunmuştur.	Tablodaki değişimlerden yola çıkarak 5 ve daha fazla çıkarımda bulunmuştur.			

EK 4: 3. Hafta Kullanılan Performans Görevi ve Dereceli Puanlama Anahtarı

Okul numarası: .....

İçerik Düzeyi	Sınıf Düzeyi	Beklenen Performans	Puanlama Yöntemi	Performans görevinin süresi
TÜRKÇE **Yazma Çalışmaları** **Kazanımlar:** **Tür, Yöntem ve Tekniklere Uygun Yazma** 1. Açıklayıcı ve bilgilendirici yazılar yazar. 2. Hikâye yazar. 3. Yazılarında betimlemeler yapar. **Görsel Okuma** 1. Resim ve fotoğrafları yorumlar.	5.sınıf	*Etkili Yazma *Görsel okuma *Görsellerden çıkarımlarda bulunma	Öz, Akran ve Öğretmen Puanlama	1 Ders saati



Selam Dostlar

Beni mutlaka tanıyorsunuzdur. Evet, ben Kaptan Jack. Gemim Kara İnci kaçırıldı. Bu yüzden kendime yeni bir gemi buldum. Altta resmi var. Güzel değil mi? Gemimizle maceraya atılacağız hazır mısın? Öncelikle maceramıza bizimle birlikte kimlerin katılacağını sebepleriyle yazalım ve alttaki ipuçlarından istediklerini seçerek maceramızı etkili bir şekilde yazalım. Bol şans!

YÖNERGE

1- Yazdığın metnine bulacağın başlık, metnin içeriği ile ilişkili olmalıdır.

2- Çalışmanın sana ait olmasına özen göstermelisin.

3- Dersin ilk 15 dakikasında(5 dakika ön hazırlık+10 dakika metnin yazımı) çalışmalar tamamlanır.

4- Dersin son 25 dakikasında(5 dakika öz değerlendirme+10 dakika akran) çalışmalar sınıfta değerlendirilir. Öğretmenler çalışmalar hakkında dersten sonra dönüt verir.

5- Çalışma yazının bütünselliği, anlatım, yazım ve dilbilgisi, düzen, görsellerden faydalanma ve çıkarımda bulunma gibi ölçütler açısından değerlendirilecektir.



1.Korsanlar



2.Denizkızı /
3. Sahil



4. Deniz canavarı/
5. Fırtına



6. Deniz feneri



7. Ada / 8. Baraka/
9. Yerliler

DERECELİ PUANLAMA ANAHTARI

Görev no: .....

Ölçütler	Başarı düzeyleri			Öz Puanı	Akran Puanı	Başarı Puanı
	(G) Geliştirilebilir	(İ) İyi	(M) Mükemmel			
Yazının Bütünselliği	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmamıştır.	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmuş, fakat bölümlerde karışıklıklar yaşanmıştır.	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmuş ve bölümler kurallarına göre yazılmıştır.			
Anlatım	Anlatımı kuvvetlendirecek 1-2 tane örnek ya da betimlemeye yer verilmiştir.	Anlatımı kuvvetlendirecek 3-4 tane örnek ya da betimlemeye yer verilmiştir.	Anlatımı kuvvetlendirecek 5 ya da daha fazla örnek ya da betimlemeden yararlanılmıştır.			
Yazım ve Dilbilgisi	Çalışmada fazlaca yazım ve dilbilgisi hatası vardır.	Çalışmada biraz yazım ve dilbilgisi hatası vardır.	Çalışmada hemen hemen hiç yazım ve dilbilgisi hatası yoktur.			
Düzen	Kâğıt kenarlarına, paragraf ve satır aralıklarında boşluklara neredeyse hiç dikkat edilmemiştir.	Yazının bir kısmında kenar, paragraf ve satır aralıklarındaki boşluklara dikkat edilmiştir.	Yazının çoğunda kenar, paragraf ve satır aralıklarındaki boşluklara dikkat edilmiştir.			
Görsellerden faydalanma	Verilen ipuçlarının 1-3 tanesinden faydalanılmıştır.	Verilen ipuçlarının 4-6 tanesinden faydalanılmıştır.	Verilen ipuçlarının 7-9 tanesinden faydalanılmıştır.			
Çıkarımda Bulunma	Resimden faydalanarak 1-2 çıkarımda bulunulmuştur.	Resimden faydalanarak 3-4 çıkarımda bulunulmuştur.	Resimden faydalanarak 5 ve daha fazla çıkarımda bulunulmuştur.			

EK 5: 4. Hafta Kullanılan Performans Görevi ve Dereceli Puanlama Anahtarı

Okul numarası: ……..

İçerik Düzeyi	Sınıf Düzeyi	Beklenen Performans	Puanlama Yöntemi	Performans görevinin süresi
TÜRKÇE **Yazma Çalışmaları Kazanımlar:** **Tür, Yöntem ve Tekniklere Uygun Yazma** 1. Açıklayıcı ve bilgilendirici yazılar yazar. 2. Yazılarında betimlemeler yapar. 3. Yazılarında önem belirten ifadeleri kullanır.	5.sınıf	*Etkili Yazma *Görsel okuma *Görsellerden çıkarımlarda bulunma	Öz, Akran ve Öğretmen Puanlama	1 Ders saati



Merhaba Arkadaşlar;
Biliyorsunuz ki meslekler insanların ihtiyaçları doğrultusunda ortaya çıkmıştır. Bu ihtiyaçlar da çeşitli yenilik ve gelişmelere sebep olmuştur.

Bu doğrultuda öncelikle hayalinizdeki mesleğin ne olduğuna karar vermelisiniz. Sonra o mesleğin görevleri ve ihtiyaçları ile ilgili düşünerek görevi yerine getirelim.

Düşünün ki hayalinizdeki mesleğe sahipsiniz. 'Hayalindeki mesleğe sahip olsan ne icat ederdin?' sorusunu düşünerek mesleğin ne olduğunun sebep ve sonuçlarını içeren bir yazı yazınız.



YÖNERGE

1- Yazdığın metnine bulacağın başlık, metnin içeriği ile ilişkili olmalıdır.
2- Çalışmanın sana ait olmasına özen göstermelisin.
3- Dersin ilk 15 dakikasında(5 dakika ön hazırlık+10 dakika metnin yazımı) çalışmalar tamamlanır.
4- Dersin son 25 dakikasında(5 dakika öz değerlendirme+10 dakika akran) çalışmalar sınıfta değerlendirilir. Öğretmenler çalışmalar hakkında dersten sonra dönüt verir.
5- Çalışma yazının bütünselliği, anlatım, yazım ve dilbilgisi, düzen, sebep ve sonuçlardan faydalanma ve ilgi çekicilik gibi ölçütler açısından değerlendirilecektir.

.....

Görev no: ……..

Okul no:……

DERECELİ PUANLAMA ANAHTARI

Ölçütler	Başarı düzeyleri			Öz Puanı	Akran puanı	Başarı puanı
	(G) Geliştirilebilir	(İ) İyi	(M) Mükemmel			
Yazının Bütünselliği	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmamıştır.	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmuş, fakat bölümlerde karışıklıklar yaşanmıştır.	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmuş ve bölümler kurallarına göre yazılmıştır.			
Anlatım	Anlatımı kuvvetlendirecek 1-2 tane örnek ya da betimlemeye yer verilmiştir.	Anlatımı kuvvetlendirecek 3-4 tane örnek ya da betimlemeye yer verilmiştir.	Anlatımı kuvvetlendirecek 5 ya da daha fazla örnek ya da betimlemelerden yararlanılmıştır.			
Yazım ve Dilbilgisi	Çalışmada fazlaca yazım ve dilbilgisi hatası vardır.	Çalışmada biraz yazım ve dilbilgisi hatası vardır.	Çalışmada hemen hemen hiç yazım ve dilbilgisi hatası yoktur.			
Düzen	Kâğıt kenarlarına, paragraf ve satır aralıklarında boşluklara neredeyse hiç dikkat edilmemiştir.	Yazının bir kısmında kenar, paragraf ve satır aralıklarındaki boşluklara dikkat edilmiştir.	Yazının çoğunda kenar, paragraf ve satır aralıklarındaki boşluklara dikkat edilmiştir.			
Sebep ve sonuçlardan faydalanma	Konuda sebep ve sonuçların 1-3 tanesinden bahsetmiştir.	Konuda sebep ve sonuçların 4-6 tanesinden bahsetmiştir.	Konuda sebep ve sonuçların 7-9 tanesinden bahsetmiştir.			
İlgi çekicilik	Yazı okuyucunun çok az dikkatini çekecek şekilde hazırlanmıştır.	Yazı okuyucunun belli bir düzeyde dikkatini çekecek şekilde hazırlanmıştır.	Yazı okuyucunun önemli düzeyde dikkatini çekecek şekilde hazırlanmıştır.			

EK 6: 5. Hafta Kullanılan Performans Görevi ve Dereceli Puanlama Anahtarı

Okul numarası: .....

İçerik Düzeyi	Sınıf Düzeyi	Beklenen Performans	Puanlama Yöntemi	Performans görevinin süresi
TÜRKÇE **Yazma Çalışmaları** **Kazanımlar:** **Tür, Yöntem ve Tekniklere Uygun Yazma** 1. Açıklayıcı ve bilgilendirici yazılar yazar. 2. Hikâye yazar. 3. Yazılarında betimlemeler yapar. **Görsel Okuma** 1. Resim ve fotoğrafları yorumlar.	5.sınıf	*Etkili Yazma *Görsel okuma *Görsellerden çıkarımlarda bulunma	Öz, Akran ve Öğretmen Puanlama	1 Ders saati

Dün sizin için çok yorucu bir gündü. Çünkü almanız gereken bazı hediyeler vardı. Aşağıda alabileceğiniz hediyelerden bazıları bulunmaktadır. Bu hediyelerden bazılarını seçerek o gününüzü, kime, niçin aldığınızı ve nasıl vereceğinizi anlatan bir hikaye yazınız.

HEDİYELER



YÖNERGE

1. Metni yazmadan önce herhangi bir kâğıda "Günüm hakkında neler yazabilirim?" konusunda ön bir hazırlık çalışması yapmalısın.

2- Yazdığın metnine bulacağın başlık, metnin içeriği ile ilişkili olmalıdır.

3- Çalışmanın sana ait olmasına özen göstermelisin.

4- Dersin ilk 15 dakikasında(5 dakika ön hazırlık+10 dakika metnin yazımı) çalışmalar tamamlanır.

5- Dersin son 25 dakikasında(5 dakika öz değerlendirme+10 dakika akran + 10 dakika akran) çalışmalar sınıfta değerlendirilir. Öğretmenler çalışmalar hakkında dersten sonra dönüt verir.

6- Çalışma yazının bütünselliği, anlatım, yazım ve dilbilgisi, düzen, görsellerden faydalanma ve çıkarımda bulunma gibi ölçütler açısından değerlendirilecektir.

Dün ……. çok sıkılmıştım ve yapılacak bir şey bulamıyordum. Annemle babam da evde yoktu. Ve sonra

……

……

……

……

……

……

……

……

……

……

……

……

……

……

Günün sonunda ……

……

……

Okul numarası: .....

DERECELİ PUANLAMA ANAHTARI

Görev no: .....

Ölçütler	Başarı düzeyleri			Öz Puanı	Akran puanı	Başarı puanı
	(G) Geliştirilebilir	(İ) İyi	(M) Mükemmel			
Yazının Bütünselliği	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmamıştır.	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmuş, fakat bölümlerde karışıklıklar yaşanmıştır.	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmuş ve bölümler kurallarına göre yazılmıştır.			
Anlatım	Anlatımı kuvvetlendirecek örnek veya betimlerden 1 ya da 2 yerde yararlanılmıştır.	Anlatımı kuvvetlendirecek 3-4 tane örnek ya da betimlemeye yer verilmiştir.	Anlatımı kuvvetlendirecek örnek ya da betimlemelerden 5 ve daha fazla yerde yararlanılmıştır.			
Yazım ve Dilbilgisi	Çalışmada fazlaca yazım ve dilbilgisi hatası vardır.	Çalışmada biraz yazım ve dilbilgisi hatası vardır.	Çalışmada hemen hemen hiç yazım ve dilbilgisi hatası yoktur.			
Düzen	Kâğıt kenarlarına, paragraf ve satır aralıklarında boşluklara neredeyse hiç dikkat edilmemiştir.	Yazının bir kısmında kenar, paragraf ve satır aralıklarındaki boşluklara dikkat edilmiştir.	Yazının çoğunda kenar, paragraf ve satır aralıklarındaki boşluklara dikkat edilmiştir.			
Görsellerden faydalanma	Verilen resimlerin 1-3 tanesinden faydalanılmıştır.	Verilen resimlerin 4-7 tanesinden faydalanılmıştır.	Verilen resimlerin 9+ tanesinden faydalanılmıştır.			
Çıkarımlarda bulunma	Resimlerden faydalanarak hediyeyi nasıl aldığı, kime ve nasıl vereceği konularından yalnızca bir tanesine değinmiştir.	Resimlerden faydalanarak hediyeyi nasıl aldığı, kime ve nasıl vereceği konularından iki tanesine değinmiştir.	Resimlerden faydalanarak hediyeyi nasıl aldığı, kime ve nasıl vereceği konularının hepsine değinmiştir.			

EK 7: Uygulamalar için Milli Eğitim Bakanlığı'ndan Alınan İzin Belgesi

T.C.
ANKARA VALİLİĞİ
Milli Eğitim Müdürlüğü

Sayı : 14588481/605.99/2039143 24/02/2015
Konu: Araştırma izni

GAZİ ÜNİVERSİTESİNE
(Eğitim Bilimleri Enstitüsü)

İlgi: a) MEB Yenilik ve Eğitim Teknolojileri Genel Müdürlüğünün 2012/13 nolu Genelgesi.
b) 05/01/2015 tarihli ve 6401 sayılı yazınız.

Üniversiteniz Eğitim Bilimleri Enstitüsü Yüksek Lisans Öğrencisi Rana SALMANER' in **"Yazma becerilerinin değerlendirmesinde öz, akran ve öğretmen puanlarının çok yüzeyli rasch ölçme modeliyle incelenmesi"** başlıklı tezi kapsamında çalışma yapma talebi Müdürlüğümüzce uygun görülmüş ve araştırmanın yapılacağı İlçe Milli Eğitim Müdürlüğüne bilgi verilmiştir.

Uygulama formunun (6 sayfa) araştırmacı tarafından uygulama yapılacak sayıda çoğaltılması ve çalışmanın bitiminde iki örneğinin (cd ortamında) Müdürlüğümüz Strateji Geliştirme (1) Şubesine gönderilmesini arz ederim.

Ali GÜNGÖR
Müdür a.
Şube Müdürü

Güvenli Elektronik İmzalı
Aslı İle Aynıdır.
24/02/2015

AŞI

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

Ayrıntılı bilgi için: Ad SOYAD Ünvan
Tel: (0 312) XXX XX XX
Faks: (0 312) XXX XX XX

Bu evrak güvenli elektronik imza ile imzalanmıştır. http://evraksorgu.meb.gov.tr adresinden 0e5b-2b9a-30a0-a48f-a04a kodu ile teyit edilebilir.

EK 8: Bir Bütün Olarak Alınan Analizin Öğrenci Yazma Beceri Düzeylerini Gösteren Yüzeyin Uyum İçi, Uyum Dışı Değerleri ve Uyum İstatistikleri Tablosu

Self, Peer and Teacher Assessment in genel 7/28/2015 8:29:21 PM
Table 7.1.1 students Measurement Report (arranged by lmN).

Total Score	Total Count	Obsvd Average	Fair(M) Average	Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Estim. Discrm	Correlation PtMea	PtExp	Nu	students
232	90	2.58	2.66	1.73	.20	.94	-.3	1.14	.7	.94	.28	.37	48	48
219	90	2.43	2.55	1.33	.18	.98	-.1	1.05	.3	1.02	.36	.35	54	54
202	84	2.40	2.54	1.32	.18	.86	-1.0	.87	-.8	1.13	.28	.34	4	4
216	90	2.40	2.50	1.20	.17	1.04	.3	1.00	.0	1.04	.49	.37	21	21
212	90	2.36	2.47	1.12	.17	1.10	.8	1.20	1.4	.83	.30	.38	22	22
205	90	2.28	2.45	1.07	.16	1.00	.0	1.02	.1	.98	.21	.29	16	16
169	72	2.35	2.45	1.06	.20	.89	-.7	.88	-.7	1.11	.45	.46	34	34
192	84	2.29	2.43	1.01	.17	.92	-.6	.89	-.8	1.22	.54	.33	7	7
174	78	2.23	2.43	1.00	.18	.71	-2.4	.73	-2.2	1.45	.25	.31	29	29
148	66	2.24	2.42	.98	.19	.96	-.2	.94	-.4	1.12	.46	.36	40	40
192	90	2.13	2.37	.85	.16	.93	-.5	.91	-.7	1.06	.11	.31	23	23
192	90	2.13	2.36	.84	.16	1.07	.5	1.07	.5	.87	.25	.27	55	55
201	89	2.26	2.36	.82	.17	1.06	.5	1.08	.6	.82	.27	.41	28	28
195	90	2.17	2.35	.81	.16	1.11	.9	1.11	.9	.81	.30	.32	37	37
196	90	2.18	2.33	.76	.16	.99	.0	.98	-.1	.97	.23	.33	47	47
181	84	2.15	2.33	.76	.17	1.13	1.0	1.12	.9	.76	.29	.36	13	13
194	90	2.16	2.32	.74	.16	1.27	2.1	1.27	2.1	.40	.04	.32	44	44
170	78	2.18	2.31	.70	.18	.90	-.7	.89	-.8	1.21	.49	.39	3	3
194	90	2.16	2.31	.70	.16	1.27	2.1	1.27	2.1	.54	.32	.33	11	11
176	84	2.10	2.30	.68	.17	1.00	.0	1.00	.0	1.08	.49	.28	6	6
197	90	2.19	2.29	.66	.16	.85	-1.2	.84	-1.3	1.29	.51	.40	8	8
194	90	2.16	2.28	.63	.16	.91	-.7	.89	-.8	1.29	.61	.40	2	2
194	90	2.16	2.28	.62	.16	1.01	.0	1.00	.0	1.02	.43	.39	1	1
189	89	2.12	2.27	.62	.16	.85	-1.2	.88	-1.0	1.24	.39	.38	17	17
192	90	2.13	2.24	.53	.16	1.02	.2	1.02	.2	.93	.38	.41	39	39
181	90	2.01	2.20	.45	.16	.91	-.7	.90	-.8	1.14	.22	.30	42	42
145	72	2.01	2.20	.44	.18	1.15	1.1	1.15	1.1	.62	.13	.34	49	49
176	90	1.96	2.19	.42	.16	.96	-.2	.96	-.3	1.11	.38	.27	30	30
187	90	2.08	2.18	.39	.17	1.06	.5	1.08	.6	.91	.45	.46	27	27
139	72	1.93	2.17	.37	.18	.87	-1.0	.86	-1.0	1.30	.42	.29	58	58
180	90	2.00	2.16	.36	.16	1.01	.0	1.01	.1	1.02	.39	.33	50	50
180	90	2.00	2.15	.34	.16	.95	-.4	.95	-.3	1.06	.30	.36	45	45
171	84	2.04	2.13	.28	.17	.67	-2.9	.68	-2.8	1.54	.54	.45	35	35
175	90	1.94	2.10	.22	.16	1.29	2.3	1.28	2.1	.42	.23	.33	25	25
150	78	1.92	2.09	.20	.17	1.04	.3	1.03	.2	1.00	.47	.33	33	33
166	84	1.98	2.08	.18	.17	.85	-1.2	.85	-1.2	1.22	.42	.44	43	43
176	90	1.96	2.08	.18	.16	1.24	1.9	1.22	1.7	.60	.39	.38	24	24
163	84	1.94	2.08	.18	.17	1.01	.1	1.03	.2	.92	.31	.42	18	18
175	90	1.94	2.08	.17	.16	.95	-.4	.94	-.4	1.07	.31	.34	53	53
166	84	1.98	2.06	.14	.17	.84	-1.2	.83	-1.3	1.27	.54	.48	60	60
162	84	1.93	2.06	.13	.17	1.12	.9	1.11	.9	.77	.39	.45	26	26
170	90	1.89	2.04	.08	.16	.94	-.5	.93	-.6	1.14	.41	.35	46	46
169	90	1.88	2.04	.08	.16	1.07	.5	1.05	.4	.91	.38	.33	15	15
157	84	1.87	2.01	.02	.17	.96	-.2	.95	-.3	1.12	.44	.36	9	9
167	90	1.86	2.00	.00	.16	.93	-.5	.93	-.5	1.16	.43	.34	10	10
156	84	1.86	2.00	-.01	.17	1.04	.3	1.08	.6	.78	.13	.39	56	56
161	84	1.92	1.99	-.02	.17	1.20	1.5	1.18	1.4	.62	.32	.42	38	38
167	90	1.86	1.96	-.10	.16	1.06	.5	1.05	.4	.93	.44	.39	31	31
165	90	1.83	1.95	-.10	.16	.95	-.3	.95	-.3	1.08	.39	.36	52	52
155	84	1.85	1.94	-.12	.17	.89	-.8	.89	-.8	1.21	.52	.45	5	5
152	84	1.81	1.89	-.24	.17	1.10	.8	1.09	.7	.82	.35	.40	57	57
157	90	1.74	1.88	-.26	.16	.83	-1.4	.89	-.8	1.27	.34	.30	14	14
158	90	1.76	1.84	-.35	.17	1.27	2.0	1.24	1.8	.53	.25	.37	12	12
154	90	1.71	1.84	-.36	.17	.88	-.9	.87	-1.0	1.30	.59	.36	20	20
144	84	1.71	1.83	-.38	.17	.87	-1.0	.85	-1.1	1.22	.44	.39	51	51
157	90	1.74	1.78	-.49	.17	1.04	.3	1.03	.2	.90	.46	.51	41	41
123	72	1.71	1.78	-.50	.19	.83	-1.2	.80	-1.4	1.37	.63	.43	59	59
127	78	1.63	1.70	-.68	.19	.87	-.9	.86	-.9	1.16	.44	.43	36	36
140	90	1.56	1.68	-.72	.18	1.26	1.8	1.22	1.5	.68	.21	.28	32	32
136	90	1.51	1.59	-.95	.19	1.03	.2	1.01	.1	.91	.30	.41	19	19
173.9	86.1	2.02	2.16	.36	.17	.99	.0	1.00	.0		.37		Mean (Count: 60)	
22.8	5.9	.22	.24	.56	.01	.14	1.1	.14	1.1		.12		S.D. (Population)	
23.0	5.9	.22	.24	.56	.01	.14	1.1	.14	1.1		.13		S.D. (Sample)	

Model, Popuin: RMSE .17 Adj (True) S.D. .53 Separation 3.12 Strata 4.50 Reliability .91
Model, Sample: RMSE .17 Adj (True) S.D. .54 Separation 3.15 Strata 4.53 Reliability .91
Model, Fixed (all same) chi-square: 597.4 d.f.: 59 significance (probability): .00
Model, Random (normal) chi-square: 53.8 d.f.: 58 significance (probability): .63

EK 9: Bir Bütün Olarak Alınan Analizin DPA'daki Kategorilerin Katılık ve Cömertlik Düzeylerini Gösteren Yüzeyin Uyum İçi, Uyum Dışı Değerleri ve Uyum İstatistikleri Tablosu

Self, Peer and Teacher Assessment in genel 7/28/2015 8:29:21 PM
Table 7.2.1 cat Measurement Report (arranged by 2mN).

Total Score	Total Count	Obsvd Average	Fair(M) Average	Model Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Estim. Discrm	Correlation PtMea	Correlation PtExp	N cat
1787	859	2.08	2.23	.15	.05	.90	-2.5	.90	-2.4	1.26	.60	.46	3 yazim
1787	861	2.08	2.22	.14	.05	.96	-.9	.97	-.7	1.04	.45	.46	4 düzen
1783	861	2.07	2.22	.12	.05	1.18	4.2	1.16	3.7	.76	.47	.46	6 cikarim
1745	861	2.03	2.17	.01	.05	.98	-.5	.98	-.4	1.05	.49	.46	1 bütünsellik
1666	861	1.93	2.07	-.21	.05	1.06	1.6	1.08	1.9	.76	.29	.46	5 görsel
1665	861	1.93	2.07	-.21	.05	.90	-2.4	.91	-2.3	1.13	.45	.46	2 anlatim
1738.8	860.7	2.02	2.16	.00	.05	1.00	-.1	1.00	-.1		.46		Mean (Count: 6)
53.8	.7	.06	.07	.16	.00	.10	2.4	.09	2.2		.09		S.D. (Population)
59.0	.8	.07	.08	.17	.00	.10	2.6	.10	2.5		.10		S.D. (Sample)

Model, Populn: RMSE .05 Adj (True) S.D. .15 Separation 2.73 Strata 3.98 Reliability .88
Model, Sample: RMSE .05 Adj (True) S.D. .16 Separation 3.03 Strata 4.37 Reliability .90
Model, Fixed (all same) chi-square: 50.8 d.f.: 5 significance (probability): .00
Model, Random (normal) chi-square: 4.6 d.f.: 4 significance (probability): .34

EK 10: Bir Bütün Olarak Alınan Analizin Puanlayıcı Türünün Katılık Ve Cömertlik Düzeylerini Gösteren Yüzeyin Uyum İçi, Uyum Dışı Değerleri ve Uyum İstatistikleri Tablosu

Self, Peer and Teacher Assessment in genel 7/28/2015 8:29:21 PM

Table 7.4.1 type Measurement Report (arranged by mN).

Total Score	Total Count	Obsrvd Average	Fair(M) Average	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Estim. Discrm	Correlation PtMea	Exact Agree. PtExp	Obs %	Exp %	Nu	type
2454	1049	2.34	2.36	.46	.05	1.08	.8	1.05	1.3	.94	.33	.36	.0	.0	10 self
1875	899	2.09	2.09	-.18	.05	1.18	4.5	1.18	4.3	.68	.37	.43	.0	.0	11 peer
6104	3216	1.90	2.04	-.28	.03	.94	-3.2	.93	-3.3	1.11	.45	.42	.0	.0	12 teacher
3477.7	1721.3	2.11	2.16	.00	.04	1.05	.7	1.06	.8		.38				Mean (Count: 3)
1872.1	1058.7	.18	.14	.32	.01	.10	3.2	.10	3.2		.05				S.D. (Population)
2292.8	1296.6	.22	.17	.40	.01	.18	3.9	.12	3.9		.06				S.D. (Sample)

Model, Populn: RMSE .04 Adj (True) S.D. .32 Separation 7.22 Strata 9.96 Reliability (not inter-rater) .98

Model, Sample: RMSE .04 Adj (True) S.D. .39 Separation 8.87 Strata 12.16 Reliability (not inter-rater) .99

Model, Fixed (all same) chi-square: 166.5 d.f.: 2 significance (probability): .00

Model, Random (normal) chi-square: 2.0 d.f.: 1 significance (probability): .16

Inter-Rater agreement opportunities: 0

EK 11: Bir Bütün Olarak Alınan Analizin Uygulanan Görevlerin Zorluk Düzeylerini Gösteren Yüzeyin Uyum İçi, Uyum Dışı Değerleri ve Uyum İstatistikleri Tablosu

Self, Peer and Teacher Assessment in genel 7/28/2015 8:29:21 PM

Table 7.5.1 tasks Measurement Report (arranged by 5mN).

Total Score	Total Count	Obsrvd Average	Fair(M) Average	Model S.E.	Infit MnSq	ZStd	Outfit MnSq	ZStd	Estim. Discrm	Correlation PtMea	PtExp	N tasks	
3695	1764	2.09	2.24	.18	.04	1.05	1.9	1.08	2.5	.85	.38	.46	3 son
3526	1764	2.00	2.14	-.06	.04	.93	-2.6	.92	-2.8	1.11	.46	.46	2 orta
3212	1636	1.96	2.11	-.12	.04	1.01	.4	1.00	.0	1.05	.54	.47	1 bas
3477.7	1721.3	2.02	2.16	.00	.04	1.00	-.1	1.00	-.1		.46		Mean (Count: 3)
200.1	60.3	.06	.06	.13	.00	.05	1.9	.06	2.2		.07		S.D. (Population)
245.1	73.9	.07	.07	.16	.00	.06	2.4	.08	2.7		.08		S.D. (Sample)

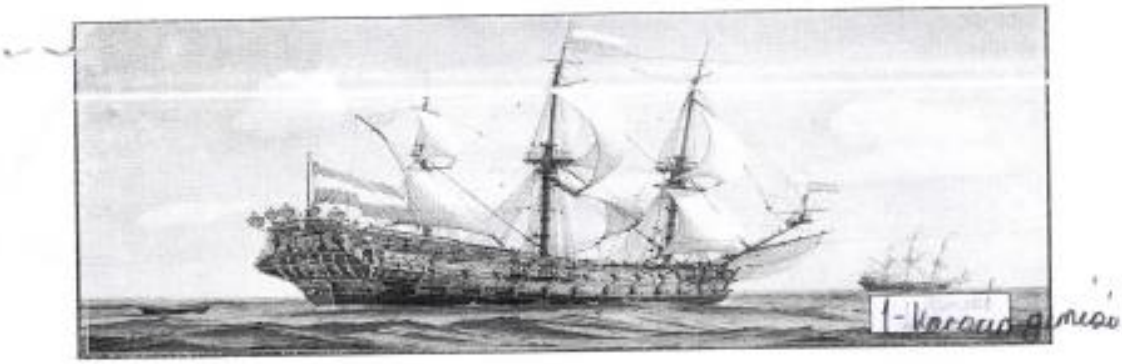
Model, Populn: RMSE .04 Adj (True) S.D. .12 Separation 3.23 Strata 4.64 Reliability .91

Model, Sample: RMSE .04 Adj (True) S.D. .15 Separation 4.01 Strata 5.69 Reliability .94

Model, Fixed (all same) chi-square: 34.3 d.f.: 2 significance (probability): .00

Model, Random (normal) chi-square: 1.9 d.f.: 1 significance (probability): .17

EK 12: Öğrencilerin Kompozisyonlarından Bir Örnek





2.Denizkızı / 3. Sahil | 4. Deniz canavarı/ 5. Fırtına | 6. Deniz feneri | 7. Ada / 8. Baraka/ 9. Yerliler

İP UÇU

Bir zamanlar Ben, Doğan ve Jack adlarında korsanlar varmış bunların içinde Doğan adındaki korsan kaçırılmış ve onu bulmak için yola düşmüşler Doğan adındaki korsan ip uçu bırakmış o ip uçundan 1'si bir sahil resmi çıkmış araya gitmişler ve 2. ip ucunu bulmuşlar ada deniz feneri imiş araya çıkmışlar ve bir ada bulmuşlar 3.'cü ip ucuda ada imiş 4.'cü ip uçu ise yerlilermiş anlardan birini kurtarmışlar korsanlara 5.'cü ip uçu ise gemi imiş binip son ip ucuna deniz canavarı imiş Doğan adındaki korsan deniz canavarının içindeymiş ve kurtarmışlar.

LON

EK 13: Öz, Akran ve Öğretmen Puanlama Örnekleri



ÖZ DEĞERLENDİRME FORMU

Ölçütler	Başarı düzeyleri			Öz Puanı
	(G) Geliştirilebilir	(İ) İyi	(M) Mükemmel	
Yazının Bütünselliği	Yazım da başta konuya genel giriş yapılan bir bölüm, ortada olayı detayları ile anlatan bir bölüm ve sonda olayların özetlendiği bir bölüm yoktur.	Yazımda baştaki giriş bölümü ortadaki gelişme bölümü ve sondaki sonuç bölümü vardır ama bazı özellikleri karışmıştır. X	Yazım giriş, gelişme ve sonuç bölümlerinden oluşmuş ve bölümler kurallarına göre yazılmıştır.	
Anlatım	Yazımda anlatımımı kuvvetlendirecek örnek ya da betimlere çok az yer vermişim.	Yazımda anlatımı kuvvetlendirecek 3-4 tane örnek ya da betimlemeye yer verilmişim. X	Yazımda anlatımı kuvvetlendirecek örnek ya da betimlemelerden fazlaca yararlanılmışım.	
Yazım ve Dilbilgisi	Çalışmamda fazlaca yazım ve dilbilgisi hatası vardır.	Çalışmamda biraz yazım ve dilbilgisi hatası vardır.	Çalışmada hemen hemen hiç yazım ve dilbilgisi hatası yoktur. X	
Düzen	Kâğıt kenarlarına, paragraf ve satır aralıklarında boşluklara neredeyse hiç dikkat etmedim.	Yazının bir kısmında kenar, paragraf ve satır aralıklarındaki boşluklara dikkat ettim.	Yazının çoğunda kenar, paragraf ve satır aralıklarındaki boşluklara dikkat ettim. X	
Görsellerden faydalanma	Resimlerdeki değişimlerin 1-2 tanesinden faydalandım.	Resimlerdeki değişimlerin 3-4 tanesinden faydalandım.	Resimlerdeki değişimlerin tamamından faydalandım. X	
Çıkarımlarda bulunma	Resimlerdeki değişimlerden faydalanarak 20 yıl sonraki kasaba hakkında 1-2 tahminde bulundum.	Resimlerdeki değişimlerden faydalanarak 3-4 tahminde bulundum. X	Resimlerdeki değişimlerden yola çıkarak 5 ve daha fazla tahminde bulundum.	

5

AKRAN DEĞERLENDİRME FORMU

Ölçütler	Başarı düzeyleri			1. Akran puanı
	(G) Geliştirilebilir	(İ) İyi	(M) Mükemmel	
Yazının Bütünselliği	Arkadaşımın yazısında başta konuya genel giriş yapılan bir bölüm, ortada olayı detayları ile anlatan bir bölüm ve sonda olayların özetlendiği bir bölüm yoktur.	Arkadaşımın yazısında baştaki giriş bölümü ortadaki gelişme bölümü ve sondaki sonuç bölümü vardır ama bazı özellikleri karışmıştır.	Arkadaşımın yazısı giriş, gelişme ve sonuç bölümlerinden oluşmuş ve bölümler kurallarına göre yazılmıştır. X	
Anlatım	Arkadaşımın yazısında anlatımını kuvvetlendirecek örneklere ya da betimlemelere çok az yer vermiş.	Arkadaşımın yazısında anlatımını kuvvetlendirecek 3-4 tane örnek ya da betimlemelere X yer verilmiş.	Arkadaşım yazısında anlatımını kuvvetlendirecek örnekler ya da betimlemelerden fazlaca yararlanmış.	
Yazım ve Dilbilgisi	Arkadaşımın çalışmasında fazlaca yazım ve dilbilgisi hatası vardır.	Arkadaşımın çalışmasında biraz yazım ve dilbilgisi hatası vardır.	Arkadaşımın çalışmasında hemen hemen hiç yazım ve X dilbilgisi hatası yoktur.	
Düzen	Arkadaşım kâğıt kenarlarına, paragraf ve satır aralıklarında boşluklara neredeyse hiç dikkat etmemiş.	Arkadaşım yazısının bir kısmında kenar, paragraf ve satır aralıklarındaki boşluklara dikkat etmiş.	Arkadaşım yazısının çoğunda kenar, paragraf ve satır X aralıklarındaki boşluklara dikkat etmiştir.	
Görsellerden faydalanma	Resimlerdeki değişimlerin 1-2 tanesinden faydalanmış.	Resimlerdeki değişimlerin 3-4 tanesinden faydalanmış.	Resimlerdeki değişimlerin X tamamından faydalanmış.	
Çıkarımlarda bulunma	Resimlerdeki değişimlerden faydalanarak 20 yıl sonraki kasaba hakkında 1-2 tahminde bulunmuş.	Resimlerdeki değişimlerden faydalanarak 3-4 tahminde X bulunmuş.	Resimlerdeki değişimlerden yola çıkarak 5 ve daha fazla tahminde bulunmuş.	

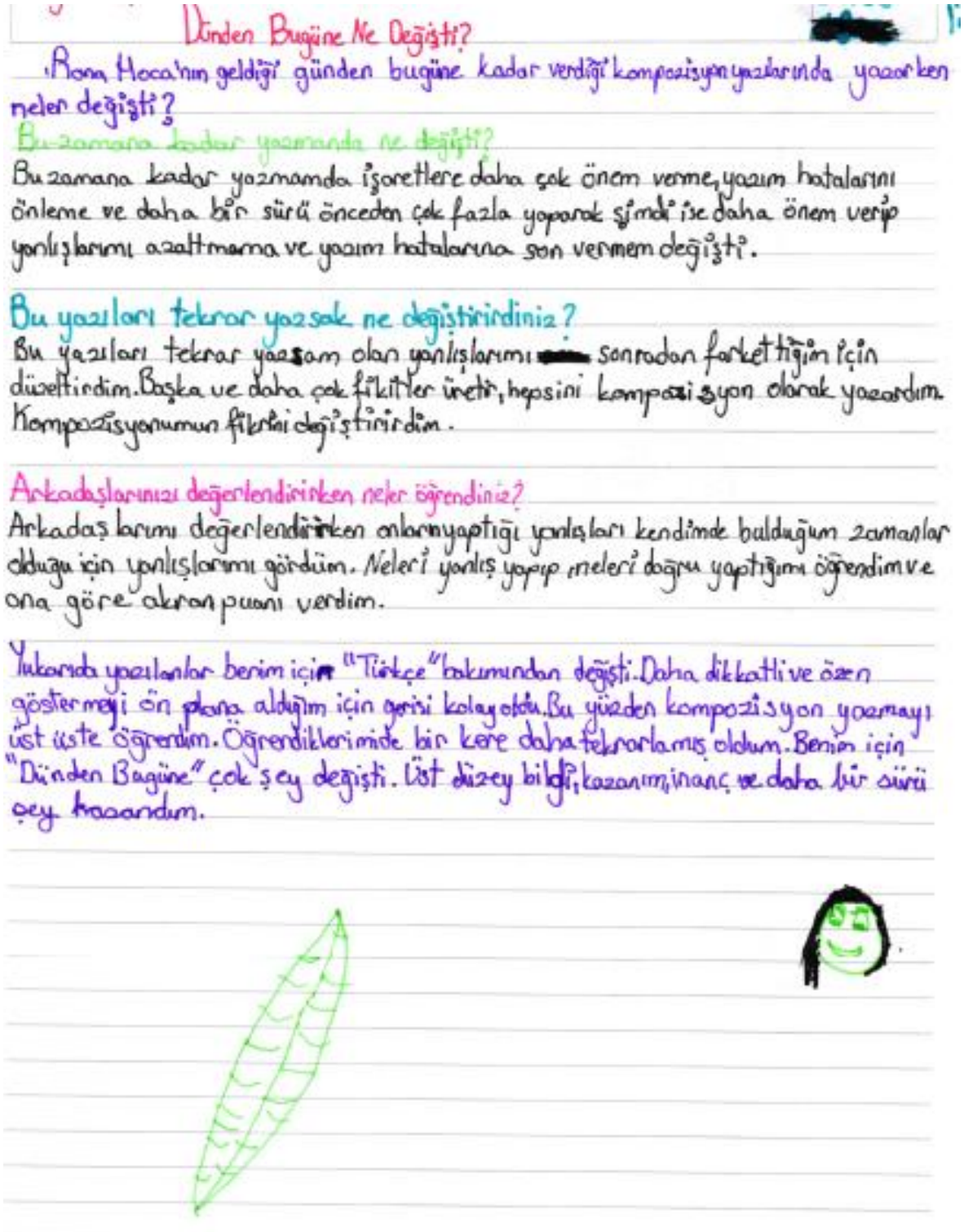
16

DERECELİ PUANLAMA ANAHTARI

14 | 20

Ölçütler	Başarı düzeyleri			Başarı Puanı
	(G) Geliştirilebilir	(İ) İyi	(M) Mükemmel	
Yazının Bütünselliği	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmamıştır.	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmuş, fakat bölümlerde karışıklıklar yaşanmıştır.	Yazı giriş, gelişme ve sonuç bölümlerinden oluşmuş ve bölümler kurallarına göre yazılmıştır.	5 / i
Anlatım	Anlatımı kuvvetlendirecek örnek ya da betimlere çok az yer verilmiştir.	Anlatımı kuvvetlendirecek 3-4 tane örnek ya da betimlemeye yer verilmiştir.	Anlatımı kuvvetlendirecek örnek ya da betimlemelerden fazlaca yararlanılmıştır.	5 / 5
Yazım ve Dilbilgisi	Çalışmada fazlaca yazım ve dilbilgisi hatası vardır.	Çalışmada biraz yazım ve dilbilgisi hatası vardır.	Çalışmada hemen hemen hiç yazım ve dilbilgisi hatası yoktur.	5 / i
Düzen	Kâğıt kenarlarına, paragraf ve satır aralıklarında boşluklara neredeyse hiç dikkat edilmemiştir.	Yazının bir kısmında kenar, paragraf ve satır aralıklarındaki boşluklara dikkat edilmiştir.	Yazının çoğunda kenar, paragraf ve satır aralıklarındaki boşluklara dikkat edilmiştir.	5 / 5
Görsellerden faydalanma	Resimlerdeki değişimlerin 1-2 tanesinden faydalanılmıştır.	Resimlerdeki değişimlerin 3-4 tanesinden faydalanılmıştır.	Resimlerdeki değişimlerin tamamından faydalanılmıştır.	5 / i
Çıkarımlarda bulunma	Resimlerdeki değişimlerden faydalanarak 20 yıl sonraki kasaba hakkında 1-2 çıkarımda bulunmuştur.	Resimlerdeki değişimlerden faydalanarak 3-4 çıkarımda bulunmuştur.	Resimlerdeki değişimlerden yola çıkarak 5 ve daha fazla çıkarımda bulunmuştur.	5 / 5

EK 14: Öğrencilerin Uygulama ile İlgili Görüş Örnekleri



Dünden Bugüne Ne Değişti?

Rana Hoca'nın geldiği günden bugüne kadar verdiği kompozisyon yazılarında yazarken neler değişti?

Bu zamana kadar yazmamda ne değişti?

Bu zamana kadar yazmamda işaretlere daha çok önem verme, yazım hatalarını önleme ve daha bir sürü önceden çok fazla yaparak şimdi ise daha önem verip yanlışlarımı azaltmama ve yazım hatalarına son vermem değişti.

Bu yazıları tekrar yazsak ne değiştirirdiniz?

Bu yazıları tekrar yazsam olan yanlışlarımı [redacted] sonradan farket tiğim için düzeltirdim. Başka ve daha çok fikirler üretir, hepsini kompozisyon olarak yazardım. Kompozisyonumun fikrini değiştirirdim.

Arkadaşlarınızı değerlendirirken neler öğrendiniz?

Arkadaşlarımı değerlendirirken onlarınyaptığı yanlışları kendimde bulduğum zamanlar olduğu için yanlışlarımı gördüm. Neleri yanlış yapıp neleri doğru yaptığımı öğrendim ve ona göre akran puanı verdim.

Yukarıda yazılanlar benim için "Türkçe" bakımından değişti. Daha dikkatli ve özen göstermeyi ön plana aldığım için geri si kolay oldu. Bu yüzden kompozisyon yazmayı üst üste öğrendim. Öğrendiklerimide bir kere daha tekrarlamış oldum. Benim için "Dünden Bugüne" çok şey değişti. Üst düzey bilgi, kazanım, inanç ve daha bir sürü şey kazandım.

KOMPOZİSYONUMDA DEĞİŞENLER

Yaklaşık 1 aydır okulumuza, Türkçe dersinde kompozisyon yazdıran Rana adında bir öğretmen geliyor. Onun sayesinde daha düzgün yazmaya başladım bunun için ona çok teşekkür ediyorum.

Bize her kompozisyon yazdığımda yazım daha da çok düzeliyor, noktalama işaretlerine, büyük harfin kullanıldığı yerlere, yazdığım yazıda betimleme ve örneklendirmenin sık kullanılmasına daha çok dikkat ediyorum. Fakat bu yazıları tekrar yazsam anahtar kelimeleri, örneklendirmeleri daha çok kullanırdım. Bu uygulamadan sadece dilbilgisi kurallarına daha çok özen göstermeyi öğrenmedim, kendimi ve başka kişileri objektif değerlendirmeyi öğrendim. Eskiden kendi eksikliklerimi çok fazla göz önünde bulunduran biri değildim. Bu değerlendirme sonucunda kendi eksikliklerimi görmeye başladım.

Bu uygulamadan sonra insanın eksikliklerini görmesinin kendisini geliştirme yolunda en büyük adım olduğunu gördüm.