

MEÜ EĞİTİM BİLİMLERİ ENSTİTÜSÜ



**T.C.
Mersin Üniversitesi
Eğitim Bilimleri Enstitüsü
Eğitim Bilimleri Ana Bilim Dalı
Eğitimde Ölçme ve Değerlendirme Bilim Dalı**

**AYNI PERFORMANS GÖREVİNİN FARKLI SAYIDA
PUANLAYICILAR TARAFINDAN ÜÇ FARKLI TEKNİKLE
PUANLANMASINDAN ELDE EDİLEN PUANLARIN
GÜVENİRLİKLERİNİN GENELLENEBİLİRLİK KURAMINA GÖRE
İNCELENMESİ**

YÜKSEK LİSANS TEZİ

Mehtap AKTAŞ

Mersin, 2013

T.C.
Mersin Üniversitesi
Eđitim Bilimleri Enstitüsü
Eđitim Bilimleri Ana Bilim Dalı
Eđitimde Ölçme ve Deđerlendirme Bilim Dalı

AYNI PERFORMANS GÖREVİNİN FARKLI SAYIDA
PUANLAYICILAR TARAFINDAN ÜÇ FARKLI TEKNİKLE
PUANLANMASINDAN ELDE EDİLEN PUANLARIN
GÜVENİRLİKLERİNİN GENELLENEBİLİRLİK KURAMINA GÖRE
İNCELENMESİ

YÜKSEK LİSANS TEZİ

Mehtap AKTAŞ

Danışman

Yrd. Doç. Dr. Devrim ALICI

Mersin, 2013

KABUL VE ONAY

Mersin Üniversitesi Eğitim Bilimleri Enstitüsü Müdürlüğü'ne,

Mehtap AKTAŞ tarafından hazırlanan "Aynı Performans Görevinin Farklı Sayıda Puanlayıcılar Tarafından Üç Farklı Teknikle Puanlanmasından Elde Edilen Puanların Güvenirliklerinin Genellenebilirlik Kuramına Göre İncelenmesi" başlıklı bu çalışma, jürimiz tarafından Eğitimde Ölçme ve Değerlendirme Anabilim Dalı'nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Başkan..........(imza)

Doç. Dr. Adnan KAN

Üye..........(imza)

Yrd. Doç. Dr. Devrim ALICI

(Danışman)


Üye..........(imza)

Yrd. Doç. Dr. Önder SÜNBÜL

Onay

Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylıyorum.

2./9./2013


Prof. Dr. Yüksel KELES
Enstitü Müdürü

TEŞEKKÜR

Tezimin bütün aşamalarında bana yol gösteren, hiçbir desteğini esirgemeyen, benimle birlikte emek sarf eden değerli hocam ve tez danışmanım Yrd. Doç. Dr. Devrim ALICI'ya çok teşekkür ederim.

Tez çalışmama geri bildirimleriyle katkı sağlayan tez savunma komitesinin değerli jüri üyesi Sayın Doç. Dr. Adnan Kan'a, ve ayrıca süreç boyunca da değerli desteğini ve katkılarını hiç esirgemeyen, çalışmamda büyük emeği olan jüri üyesi Sayın Yrd. Doç. Dr. Önder SÜNBÜL'e teşekkür ederim.

Bu çalışmanın ortaya çıkış aşamasında, değerli katkılarıyla çalışmamın şekillenmesine yardımcı olan, kendisinden çok şey öğrendiğim Sayın Hocam Prof. Dr. Adnan ERKUŞ'a teşekkür ediyorum.

Çalışmama tez önerisi aşamasında, zaman ayırıp katkı sunan Sayın Prof. Dr. A. Ata TEZBAŞARAN'a teşekkür ediyorum.

Ayrıca, uygulamalarım sırasında bana çok yardımcı olan uygulama yaptığım okullardaki tüm öğretmenlere ve çalışma arkadaşlarıma teşekkür ediyorum.

Tüm hayatım boyunca her anımda beni destekleyen annem, babam ve abime, ayrıca her koşulda yanımda olan, varlığıyla hayatımı anlamlandıran canım eşime çok teşekkür ediyorum.

Mehtap AKTAŞ

ÖZET

AYNI PERFORMANS GÖREVİNİN FARKLI SAYIDA PUANLAYICILAR TARAFINDAN ÜÇ FARKLI TEKNİKLE PUANLANMASINDAN ELDE EDİLEN PUANLARIN GÜVENİRLİKLERİNİN GENELLENEBİLİRLİK KURAMINA GÖRE İNCELENMESİ

Mehtap AKTAŞ

Yüksek Lisans Tezi

Eğitim Bilimleri Enstitüsü

Yrd. Doç. Dr. Devrim ALICI (Danışman)

Mersin, 2013

89 Sayfa

Bu araştırmanın amacı, aynı performans görevlerinin farklı sayıda puanlayıcı tarafından kontrol listesi, dereceleme ölçeği ve analitik rubrik yardımıyla puanlanması durumunda, puanlayıcılar arası güvenilirliklerinin Genellenebilirlik Kuramı çerçevesinde incelenmesidir.

Bu amaç doğrultusunda, 5. sınıf düzeyindeki öğrencilerin hikâye yazma becerilerini puanlamak amacıyla, kontrol listesi, dereceleme ölçeği ve analitik rubrik hazırlanmıştır. İlköğretim 5. sınıf öğrencilerine yazdırılan hikâyeler arasından seçilen 6 hikâye 45 puanlayıcıya üç farklı puanlama anahtarı ile 10-15 gün aralıklarla puanlattırılmıştır. Araştırmaya katılan 45 puanlayıcı içerisinde 2, 3, 5 ve 10 puanlayıcılı 100'er örneklem çekilmiştir. Elde edilen 400 örneklem için Genellenebilirlik Kuramı'na göre puanlayıcılar arası güvenilirlikleri hesaplanmıştır. Elde edilen 1200 hesaplamanın her bir durum için elde edilen 100 örnekleme için ortancaları ve standart hataları hesaplanmıştır.

Genellenebilirlik Kuramı güvenilirlik kestirimlerinin ortanca deęerleri incelendięinde, 5 puanlayıcının kontrol listesi kullanarak yaptıkları puanlamaların güvenilirliklerinin ortanca deęeri hariç olmak üzere, puanlayıcı sayısı ve aynı zamanda kullanılan ölçeęin kategori sayısı arttıkça ortanca deęerlerinin de arttığı gözlenmiştir.

Genellenebilirlik Kuramı'ndan elde edilen standart hataların, puanlayıcı sayısı arttıkça azaldığı gözlenmiştir. En düşük standart hata deęerlerinin, 10 puanlayıcı olması durumunda elde edildięi saptanmıştır.

Puanlayıcı sayısı 5 ve kategori sayısı 2 olduęunda, Genellenebilirlik Kuramı'nda güvenilirlik kestiriminin en yüksek deęeri verdięi belirlenmiştir.

Anahtar Kelimeler: Genellenebilirlik Kuramı, puanlayıcılar arası güvenilirlik, kontrol listesi, dereceleme ölçeęi, analitik rubrik, performans görevi.

ABSTRACT

AN INVESTIGATION OF THE RELIABILITY OF THE SCORES OBTAINED THROUGH RATING THE SAME PERFORMANCE TASK WITH THREE DIFFERENT TECHNIQUES BY DIFFERENT NUMBERS OF RATERS ACCORDING TO GENERALIZABILITY THEORY

Mehtap AKTAŞ

Master's Thesis

Institute of Education Sciences

Asst. Prof. Dr. Devrim ALICI (Supervisor)

Mersin, 2013

89 Pages

The aim of this study is to investigate inter-rater reliability in a condition that the same performance tasks are scored by different number of raters through check list, rating scale, and analytic rubric within the frame of Generalizability Theory.

In line with the aim of the study, check list, rating scale, and analytic rubric were prepared in order to rate 5th grade students' story writing skills. Six stories chosen among the stories which had been written by 5th grade students were scored by 45 raters with three different rating keys between 10-15 days. Among the 45 raters who participated into the study, 100 samples with 2, 3, 5, and 10 rating were drawn. Inter-rater reliability was calculated for the obtained 400 samples according to Generalizability Theory. Median and standard error of 100 samples obtained for each situation among 1200 calculations were calculated.

When the median value of Generalizability Theory reliability estimate was illustrated, it was observed that when the number of raters and at the same time category labels were increased, the median value also increased except the median value of reliability of rating done through using control list by 5 raters.

It was found out that standard errors obtained from Generalizability Theory decreased as long as the numbers of raters were increased. It was ascertained that the lowest standard error value was obtained in case there were 10 raters.

When there were 5 raters and category number was 2, reliability estimate presented maximum value within Generalizability Theory.

Keywords: Generalizability Theory, inter-rater reliability, check list, rating scale, analytic rubric, performance tasks.

İÇİNDEKİLER

KABUL VE ONAY	i
TEŞEKKÜR	ii
ÖZET	iii
ABSTRACT	v
TABLolar DİZİNİ.....	ix
KISALTMALAR.....	x
BÖLÜM I	1
GİRİŞ	1
1.1. Problem	1
1.1.1. Performans Görevi	1
1.1.2. Performans Görevinin Puanlanması	2
1.1.3. Puanlama Güvenirliği	7
1.1.3.1. Genellenebilirlik Kuramında Güvenirlik	9
1.1.4. İlgili Araştırmalar	12
1.2. Amaç	17
1.3. Önem	18
1.4. Sayıtlar	20
1.5. Sınırlılıklar	20
BÖLÜM II	21
GEREÇ VE YÖNTEM	21
2.1. Araştırmanın Modeli	21
2.2. Araştırma Grubu	21

2.3. Veri Toplama Araç ve/veya Teknikleri	21
2.4. Verilerin Analizi	26
2.5. İşlem	26
BÖLÜM III	28
BULGULAR	28
3.1. “Genellenebilirlik Kuramı’na göre, aynı performans görevlerinin; kontrol listesi, dereceleme ölçeği, analitik rubrik kullanılarak ve 2, 3, 5, 10 puanlayıcıdan oluşan 100’er örneklemin puanlama güvenilirlikleri nasıldır?” araştırma sorusuna ilişkin bulgular.....	28
3.2. “Genellenebilirlik Kuramı’na göre, aynı performans görevlerinin; kontrol listesi, dereceleme ölçeği, analitik rubrik kullanılarak ve 2, 3, 5, 10 puanlayıcıdan oluşan 100’er örneklemin puanlama güvenilirliklerinin ortancaları ve standart hataları nasıldır?” araştırma sorusuna ilişkin bulgular	33
BÖLÜM IV	35
TARTIŞMA VE YORUM	35
SONUÇ VE ÖNERİLER	37
KAYNAKLAR	39
EKLER	43
ÖZGEÇMİŞ	89

TABLolar DİZİNİ

Tablo 1. Hoşgörü gösterebilme konusunda gerçekleştirmesi beklenen kazanımları gözlemek amacıyla 4 farklı puanlama anahtarı.....	6
Tablo 2. Çalışma grubunda yer alan öğretmenlerin kıdem ve cinsiyet bilgileri	21
Tablo 3. Performans görevi yardımıyla gözlenebileceği belirlenen ve uzmanlar arasında uzlaşmanın sağlandığı kazanımlar.....	23
Tablo 4. Genellenebilirlik Kuramına göre, üç farklı puanlama anahtarı kullanılarak ve 2 puanlayıcıdan oluşan 100'er örneklemin puanlama güvenilirlikleri	29
Tablo 5. Genellenebilirlik Kuramına göre, üç farklı puanlama anahtarı kullanılarak ve 3 puanlayıcıdan oluşan 100'er örneklemin puanlama güvenilirlikleri.....	30
Tablo 6. Genellenebilirlik Kuramına göre, üç farklı puanlama anahtarı kullanılarak ve 5 puanlayıcıdan oluşan 100'er örneklemin puanlama güvenilirlikleri	31
Tablo 7. Genellenebilirlik Kuramına göre, üç farklı puanlama anahtarı kullanılarak ve 10 puanlayıcıdan oluşan 100'er örneklemin puanlama güvenilirlikleri	32
Tablo 8. Genellenebilirlik Kuram'ına ilişkin güvenilirliklerinin ortancaları ve standart hataları.....	33

KISALTMALAR

GK	: (Generalizability Theory) Genellebilirlik Kuramı
G Çalışması	: (Generalizability Study) Genellebilirlik Çalışması
K Çalışması	: (Decision Study) Karar Çalışması
G Katsayısı	: Genellebilirlik katsayısı
KTK	: Klasik Test Kuramı

BÖLÜM I

GİRİŞ

1.1. Problem

Eğitimde ölçme ve değerlendirme çalışmalarının önemli bir bölümünü öğrenciler hakkında karar verebilmek amacıyla yapılan çalışmalar oluşturur. Öğrencilerin bir derste başarılı olmak için gerekli olan ön koşul düzeydeki bilgi ve becerilere sahip olup olmadıklarını, bir dersteki öğrenme eksikliklerini ve yanlışlıklarını ve öğrenme düzeylerini belirlemek ve çeşitli konularda öğrencilere rehberlik çalışmaları yapmak vb. amacıyla ölçme değerlendirme çalışmalarına başvurulur.

Öğrenciler hakkında bilgi toplayabilmek ve çeşitli kararlar alabilmek için farklı ölçme ve değerlendirme araç ve yöntemlerinden yararlanılır. Bu araçlar, uzun cevaplı, kısa cevaplı, doğru yanlış ve çoktan seçmeli maddelerden oluşan ölçme araçları, performans görevleri, projeler, gelişim dosyaları gibi etkinlikler ve kontrol listeleri, dereceleme ölçekleri, dereceli puanlama anahtarları şeklinde sıralanabilir.

Her ölçme aracı farklı durumlarda, farklı düzeydeki kazanımlarda diğerlerine göre daha uygun olabilir. Ölçme aracı hazırlayacak olan eğitimcilerin hangi madde türlerinden yararlanacaklarına öncelikle amaca ve ölçülecek özelliğe bağlı olmak üzere, öğrenci özellikleri, süre vb. ölçütleri dikkate alarak karar vermeleri beklenir.

1.1.1. Performans Görevi

Performans, bireyin sahip olduğu kapasiteyi belli zaman dilimi içinde bir işi başarıyla tamamlamak için kullanabilme yüzdesi olarak tanımlanabilir (Deliceoğlu, 2009). Performans değerlendirme ise bireylerin aktif öğrenme yoluyla süreç içerisinde gerçekleştirdikleri çalışmaların ve süreç sonunda ortaya çıkardıkları ürünlerin değerlendirilmesi şeklinde ifade edilebilir (Tekindal, 2008).

Öğrencilerin performanslarını gözlemlemenin bir yolu, performans görevi çalışmaları yaptırmaktır. Performans görevleri, öğrencilere gerçek dünyada karşılaşılabilecekleri problem durumlarını sunan ve öğrencilerin üst düzey zihinsel becerilerini geliştirilmesini ve ölçülmesini amaçlayan etkinliklerdir (Kutlu, Doğan ve Karakaya, 2009). Bu etkinlikler, öğrencilerin bilgiyi edinme, düzenleme, kritik etme, kendini ifade etme,

yaratıcılığı ve vücut dilini kullanma gibi hem bilişsel hem duyuşsal hem de psikomotor becerilerini işe koşmalarını sağlar (MEB, 2009).

Performans görevlerinin, okul imkânlarının kısıtlı olduğu durumlarda ve eğer istenilen görevde araştırma yapma gerekliliği varsa bu aşamalarda okul dışı bir etkinlik olarak gerçekleştirilmesi öğrenciden istenen diğer görevlerin ise sınıf içi bir etkinlik olarak gerçekleştirilmesi gerekmektedir (MEB, 2009).

Bir performans görevi hazırlanırken; hangi sınıf düzeyinde olması gerektiğine, ders adının, ilgili ders konusunun ve kazanımların belirlenmesine, problem durumunun ortaya koyulmasına, süreç boyunca öğrencinin takip edeceği yönergenin düzenlenmesine ve nasıl değerlendirilecekleri konusunda öğrencilerin bilgilendirilmesine dikkat edilmelidir (Tekindal, 2008).

1.1.2. Performans Görevinin Puanlanması

Öğrencilerin ilgilenilen özelliklerinin gözlenmesi veya programda yer alan kazanımların ne kadarının öğrenciler tarafından gerçekleştirildiğinin ortaya konabilmesi amacıyla verilen performans görevlerinin güvenilir bir biçimde puanlanabilmesi önemlidir. Bir performans görevinin önceden belirlenen ölçütlere göre ve bir puanlama sistemi kullanılarak değerlendirilmesi, hatasız ve amaca uygun değerlendirme yapmaya olanak sağlar.

Kontrol listeleri, dereceleme ölçekleri ve dereceli puanlama anahtarları (holistik rubrik ve analitik rubrik) performans görevinin değerlendirilmesinde puanlama anahtarı olarak kullanılabilirler.

Kontrol listeleri, bir yandan herhangi bir performansın ölçümünde doğrudan kullanılabilir bir ölçme aracı iken, bir yandan da performans görevi, proje, gelişim dosyası gibi etkinliklerin değerlendirilmesinde ya da özdeğerlendirme, akran değerlendirme gibi çalışmalarda puanlama anahtarı olarak yararlanılabilen bir değerlendirme aracıdır.

Kontrol listeleri, içinde bir dizi uyarıcının (kazanım, soru, ölçüt) yer aldığı ve puanlayıcının (cevaplayıcının) bu uyarıcılara var/yok, evet/hayır vb. şeklinde verilen ikili kategorilerden birini seçerek tepkide bulunduğu ölçme araçlarıdır. Diğer ifadeyle kontrol listeleri, gösterilen performansın ya da değerlendirilen kişinin sahip olması gereken özelliklere sahip olup olmadığını ortaya koyar (Aiken, 2000). Ancak kontrol listeleri, performansın düzeylerinin ya da ölçütlerin ne dereceye kadar karşılandığı hakkında bilgi

vermez (Moskal, 2000). Kontrol listeleri özdeğerlendirme, akran değerlendirme ve grup değerlendirme gibi çalışmalarda da yaygın olarak kullanılmaktadır. Bu araç, eğitim, sanayi, tıp gibi birçok alanda yaygın olarak kullanılan (Aiken, 2000), istenilen bilgiye çabuk ulaşılması, kolay uygulanması, sonuçların daha anlaşılabilir olması açısından avantajlı bir ölçme aracıdır (Hobart ve Frankel, 1999).

Kontrol listesi hazırlanırken dikkat edilmesi gereken kurallar şu şekilde sıralanabilir: (1) ölçülecek davranış açık bir biçimde tanımlanmalı ve bu davranışın göstergeleri eksiksiz bir biçimde listede yer almalıdır, (2) davranış göstergelerinin hangi kategorilerle (evet/hayır, var/yok vb.) gözlemleneceği kararlaştırılmalıdır, (3) davranış göstergelerinin puanlanmasında her göstergenin toplam puan içindeki katkısı kararlaştırılmalıdır (Tekindal, 2008).

Eğitim sisteminde ölçme ve/veya değerlendirme amacıyla yararlanılan bir diğer araç, dereceleme ölçekleridir. Dereceleme ölçekleri hem ilgilenilen özelliğin bireyde bulunma düzeyini ortaya koymak amacıyla ölçme aracı olarak kullanılabilirken hem de performans görevlerinin puanlanmasında bir puanlama anahtarı olarak kullanılabilir.

Dereceleme ölçekleri, birey ya da nesnelere belli bir özelliğini doğrudan ölçmek için bir araç bulunmuyorsa, her bir özelliğin bulunma derecesini belirlemek için kullanılır (Keeves, 1988). Bu araçlar, ölçülen özelliğe ilişkin performans ölçütlerinin ne dereceye kadar karşılandığını görebilmek açısından uygundur (Moskal, 2000). Bir başka deyişle kontrol listeleri yalnızca ilgilenilen özelliğin bireyde bulunup bulunmadığını ortaya koyarken, dereceleme ölçekleri bu özelliklerin yalnızca bireyde bulunup bulunmadığıyla değil, aynı zamanda bireyde bulunma düzeyi/sıklığı ile de ilgilenmektedir. Bu durum, dereceleme ölçeklerinin en az üç kategorili olarak (evet/kısmen/hayır gibi) düzenlenebileceğini de ortaya koymaktadır (Tekindal, 2008).

Dereceleme ölçeklerinde, değerlendirme yapan kişiden ilgilenilen özellik ya da davranışa bireyin ne derece sahip olduğunu sıfat, zarf ve bunlara karşılık gelen sayısal değerlere göre derecelemesi beklenir. Dereceleme ölçekleri yardımıyla kişiler kendilerinin ya da bir başkasının sahip olduğu duyuşsal özellikleri betimlemenin yanında, belli bir durumda göstermesi gereken performansı da değerlendirebilir (Erkuş, 2003).

Dereceleme ölçekleri, en az sıralama düzeyinde ölçme yapmaya uygun ölçeklerdir. Ölçülen özelliğin niteliğine bağlı olarak tepki kategorilerinin sayısı değişebilmektedir. Fakat tepki kategorilerinin sayısı azaldıkça ölçme aracının duyarlılığı da azalmaktadır (Aiken, 2000).

Eğitim sürecinde dereceleme ölçekleri kullanmak öğrencilere çeşitli yararlar sağlar. Bunlar sadece durum belirleme süreciyle sınırlı değildir; hem durum belirlerken hem durum belirleme öncesinde ve hem de sonrasında öğrenciler için yararlıdır (Kutlu, Doğan, Karakaya, 2009).

Dereceleme ölçeği hazırlanırken dikkat edilmesi gereken kurallar şu şekilde sıralanabilir: (1) ölçülmek istenen özelliğin davranış göstergesi kabul edilecek ölçütler ve (varsa) alt ölçütler belirlenmelidir, (2) belirlenen ölçütlere karşılık gelen tepki kategorileri (evet/kısmen/ hayır ya da hiç/ara sıra/ sık sık/ her zaman gibi) karşılaştırılmalıdır, (3) ölçmenin amacına uygun planlama sistemi (analitik/holistik vb.) karşılaştırılmalıdır, (4) her ölçütün ve/veya alt ölçütün puanlamaya katkısının/ağırlığının ne olacağı belirlenmelidir.

Dereceleme ölçekleri kullanmanın değerlendiriciler için değerlendirme ölçütlerinin açık ve net olmasını sağlama, öğrencilerden beklentileri açıkça ortaya koyma, nesnellik sağlama, değerlendirme sonuçlarının doğru bir şekilde açıklanmasını sağlama ve öğrencilerin güçlü ve zayıf yönleri konusunda bilgilenmelerini sağlama gibi yararları da vardır (Goodrich, 2005).

Bir ölçme aracında yer alan maddelerin ilgili davranışın gözlenme düzeyine bağlı olarak (1,0) dışında puanlanması durumunda ise çoklu puanlama söz konusudur. Örneğin uzun cevaplı bir maddenin, bir projenin ya da bir performans görevinin bir dereceleme ölçeği yardımıyla puanlanması durumunda çoklu puanlamadan yararlanıldığı söylenilebilir.

Performans görevlerini puanlamanın bir başka yolu, dereceli puanlama anahtarından yararlanmaktır. Dereceli puanlama anahtarı, holistik (bütüncül) ve analitik (ayrıntılı) puanlama anahtarı olmak üzere iki grupta incelenebilir.

Dereceli puanlama anahtarları (rubrikler); değerlendirme ölçütleri, ölçüt tanımlamaları ve puanlama stratejisi olmak üzere üç bölümden oluşur (Popham, 1997). Değerlendirme ölçütleri; kabul edilebilir yanıtları kabul edilemez yanıtlardan ayırmak için kullanılır. Ölçüt tanımlamaları; öğrencilerin değerlendirilmek istenen yanıtlarındaki niteliksel

farklılıkları tanımlama yolunu ifade eder. Puanlama stratejisi; bütünsel (holistic) ya da analitik (analytical) puanlama biçiminde olabilir. Popham'a (1997) göre, bütünsel (holistic) puanlama, performansın daha küçük parçalara ayıramadığı durumlarda kullanılan ve değerlendirmenin genel olarak yapıldığı durumlarda kullanılırken; analitik (analytical) puanlama, sergilenen performansın en küçük alt parçalarının ayrı ayrı puanlandığı durumlarda kullanılır. Dereceli puanlama anahtarlarından hangisinin kullanılacağına değerlendirmenin amacına bağlı olarak karar verilir.

Her bir performans düzeyinde kişiden ne beklenildiğinin tanımlandığı dereceli puanlama anahtarlarının, değerlendirme aracı olarak kullanılmasından çok, öğretim aracı olarak kullanılmasının daha yararlı olduğu söylenebilir. Öğretmenler dereceli puanlama anahtarı kullanarak öğrenme hedefleri doğrultusunda öğrencilerin gelişimini takip edebilir (Arter, 2002) ve ayrıca öğrenciler, puanlama yönergesinden edindikleri bilgiyle bir sonraki adımda hangi performansı göstermeleri gerektiğini öğrenebilirler (Moskal ve Leydens, 2000).

Tablo 1'de, 36-72 aylık çocukların sosyal-duygusal alanda hoşgörü gösterebilme becerisini gözlemek amacıyla hazırlanan 4 farklı puanlama anahtarı (kontrol listesi, dereceleme ölçeği, holistik rubrik ve analitik rubrik) yer almaktadır.

Tablo 1, öğrencilerin aynı performansının farklı yöntemlerle nasıl gözlemlenebileceğine ilişkin bir örnek olarak düşünülebilir. Tablo 1'de görüldüğü gibi aynı davranış göstergesi bir kontrol listesi yardımı ile gözlemlenebileceği gibi, bir analitik rubrik yardımı ile de gözlenebilir. Hangi değerlendirme yöntemlerinden yararlanacağımıza karar verirken amacımız, öğrenci gurubunun düzeyi, davranış göstergelerinin içeriği gibi pek çok ölçütün göz önünde bulundurulması önemlidir. Kimi durumda kontrol listesi oldukça yeterli bir araç konumundayken kimi durumda analitik rubrik çok daha yararlı bir araç konumunda olabilir.

Tablo 1'de kontrol listesinde Evet kategorisinin 1'le, Hayır kategorisinin 0'la puanlandığı gözlenmektedir. Bir ölçme aracında yer alan maddelerin ilgili davranış gözlendiyse 1, gözlenmediyse 0 olacak şekilde puanlanması durumunda ikili puanlamadan bahsetmek mümkündür. Çoktan seçmeli test maddesinin doğru cevaplanması durumunda 1 puan, yanlış, birden fazla cevaplama ya da boş bırakılması durumlarında 0 puanla puanlanması ikili puanlamaya örnektir. Yine kontrol listeleri yardımıyla yapılan ölçme çalışmalarında ikili puanlama yapılır. Uzun cevaplı bir maddenin bir kontrol listesi yardımıyla puanlanması durumunda da ikili puanlamadan söz edilebilir.

Tablo 1. Hoşgörü gösterebilme konusunda gerçekleştirmesi beklenen kazanımları gözlemek amacıyla 4 farklı puanlama anahtarı

Kazanımlar	Kontrol Listesi		Dereceleme Ölçeği				Holistik Rubrik			Analitik Rubrik					
	Evet (1)	Hayır (0)	Kesinlikle katılmıyorum	Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle katılıyorum	Yetersiz	Kısmen yeterli	Yeterli	Hiçbir zaman	Nadiren	Ara sıra	Çoğu zaman	Her zaman
Hata yapabileceğini kabul eder.								Kendine ve diğer insanlara hoşgörülü değildir. Kendisinin ve diğer insanların hatasını fark etmez, dile getirmez.	Kendisinin ve diğer insanların hatasını bazen fark eder, dile getirir.	Kendisinin ve diğer insanların hatasını fark eder, dile getirir ve kabul eder.	Hata yaptığını hiçbir zaman kabul etmez. Hata yapmadığını düşünür.	Hata yaptığını nadiren kabul eder. Seyrek de olsa hata yaptığını düşünür.	Sadece belli durumlarda hata yapabileceğine inanır.	Çoğu zaman hata yapabileceğine inanır.	Her zaman hata yapabileceğini bilir ve bunu kabul eder.
Kendi hatalarını söyler.								Kendine ve diğer insanlara hoşgörülü değildir. Kendisinin ve diğer insanların hatasını fark etmez, dile getirmez.	Kendisinin ve diğer insanların hatasını bazen fark eder, dile getirir.	Kendisinin ve diğer insanların hatasını fark eder, dile getirir ve kabul eder.	Hiçbir zaman hatasını fark etmez ve bunu dile getirmez.	Nadiren hatasını fark eder ve bunu dile getirir.	Bazı durumlarda hata yaptığını fark eder ve bunu söyler.	Çoğunlukla hata yaptığını fark eder ve bunu söyler.	Hatalarının her zaman farkındadır ve bunu çekinmeden dile getirir.
Kendisini başkalarının yerine koyarak duygularını açıklar.								Kendine ve diğer insanlara hoşgörülü değildir. Kendisinin ve diğer insanların hatasını fark etmez, dile getirmez.	Kendisinin ve diğer insanların hatasını bazen fark eder, dile getirir.	Kendisinin ve diğer insanların hatasını fark eder, dile getirir ve kabul eder.	Kendisini başkasının yerine koyamaz. Başkalarının duygu ve düşünceleriyle empati kuramaz.	Kendisini nadiren başkasının yerine koyar. Başkalarının duygu ve düşünceleriyle seyrekle de olsa empati kurar.	Bazı durumlarda başkalarının duygu ve düşüncelerini anlayabilir ve bunu söyleyebilir.	Çoğu zaman başkalarının duygu ve düşüncelerini anlayabilir ve bunu söyleyebilir.	Karşısındakinin duygu ve düşünceleriyle empati kurabilir ve bu duygu ve düşüncelerini açıklayabilir.
Başkalarının hatalarını uygun yollarla ifade eder.								Kendine ve diğer insanlara hoşgörülü değildir. Kendisinin ve diğer insanların hatasını fark etmez, dile getirmez.	Kendisinin ve diğer insanların hatasını bazen fark eder, dile getirir.	Kendisinin ve diğer insanların hatasını fark eder, dile getirir ve kabul eder.	Başkalarının hatalarının farkına varmaz. Bu hataları herhangi bir şekilde ifade etmekten kaçınır.	Başkalarının hatalarının nadiren farkına varır. Bu hataları seyrekle de olsa ifade etmeye çalışır.	Bazı durumlarda başkalarının hatalarını fark eder ve bunları ifade etmeye çalışır.	Çoğu zaman başkalarının hatalarını fark eder ve bunları ifade etmeye çalışır.	Başkalarının hatalarının her zaman farkındadır ve bu hataları onların duygu ve düşüncelerini incitmeden söyleyebilir.
Başkalarının hata yapabileceğini kabul eder.								Kendine ve diğer insanlara hoşgörülü değildir. Kendisinin ve diğer insanların hatasını fark etmez, dile getirmez.	Kendisinin ve diğer insanların hatasını bazen fark eder, dile getirir.	Kendisinin ve diğer insanların hatasını fark eder, dile getirir ve kabul eder.	Hiç kimsenin hata yapmasına tahammül edemez. Kimsenin hata yapmaması gerektiğini düşünür.	Nadiren başkalarının hata yapabileceğini kabullenir.	Başkalarının hata yapabileceğini bazı durumlarda kabul eder.	Başkalarının hata yapabileceğini kabul eder.	Kendi dahil herkesin zaman zaman hata yapabileceğini bilir ve bunu doğal karşılar.

Performans görevinin puanlanmasında ister kontrol listesi, isterse dereceleme ölçekleri ya da dereceli puanlama anahtarı kullanılsın, bu puanlama sistemlerinin güvenilir ve geçerli bilgi taşıması önemlidir. Bu anlamda, performans görevleri değerlendirilirken, yararlanılan puanlama araçlarının güvenilirlik ve geçerliklerinin kanıtlarının elde edilmiş olması gerekir.

1.1.3. Puanlama Güvenirliği

Eğitimde ve psikolojide bireyler hakkında verilen kararların diğer ifadeyle yapılan değerlendirmelerin isabetliliği, bu kararları almada kullanılan ölçme sonuçlarına ölçme sonuçlarının niteliği ise bu sonuçları elde etmemize olanak sağlayan ölçme araçlarının niteliğine bağlıdır.

Bir ölçme aracının seçkisiz hatalardan arınık ölçüm yapabilme yeterliliği olarak adlandırılan güvenilirlik ile kullanılma amacına hizmet etme derecesi olarak adlandırılan geçerlik, ölçme araçlarında bulunması gereken iki önemli psikometrik özelliktir (Turgut ve Baykul, 2010).

Crocker ve Algina (1986, ss105), Klasik Test Kuramı'na göre güvenirligi "aynı bireyler üzerinde yapılan ölçmelerin benzer şartlarda tekrar elde edilebilirliği" şeklinde tanımlamaktadır. Turgut (1997) ise güvenirlilik kavramıyla, ölçme sonuçlarının seçkisiz hatalardan ne derece arınık olduğunu ve bir ölçme sonucunun, içindeki seçkisiz hataların azlığı oranında güvenilir sayıldığını ifade etmektedir. Bir ölçme aracının güvenirligini etkileyen pek çok faktörden bahsedilebilir. Bu faktörler; ölçmenin yapıldığı nesne ya da kişilerden, ölçmeyi yapan kişi (ölçmeci) ya da kişilerden, ölçmenin yapıldığı koşullardan (ortam) ve ölçme aracından kaynaklanan faktörler olarak sınıflanabilir (Erkuş, 2003).

Yapılan bir ölçme işleminde işlem, içerisinde hata barındırdığından ölçmek istediğimiz özelliğin gerçek değeri doğrudan elde edilemez. Bu nedenle ölçme sonucundan elde edilen gözlenen puanlar yardımıyla gerçek değerler kestirilmeye çalışılır. Klasik Test Kuramı'nda, ölçülmek istenen özelliğe ait gerçek değer, gerçek puan olarak adlandırılır. Ölçme yoluyla doğrudan elde edilemeyen gerçek puan, bazı varsayımlarla gözlenen

puanlardan kestirilir. Bu nedenle Klasik Test Kuramı'na *gerçek puan kuramı* da denir (Baykul, 2000).

Ölçmede amaç, içerisinde hata barındırmayan gerçek puanı elde etmektir. Klasik Test Kuramı, ölçme sonucunda elde edilen puanın gerçek puanına eşit olmadığı ve ölçülen puan ile gerçek puan arasındaki farkın hata/hatalardan kaynaklandığını vurgular. Klasik Test Kuramı'nın temel denklemi 1 nolu eşitlikteki gibidir (Crocker ve Algina, 1986; Lord ve Novick, 1968);

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

1 nolu eşitlikte, X, ölçme sonucunu, T, gerçek puanı ve E ise ölçme sonuçlarına karışan hatayı göstermektedir.

Klasik Test Kuramı'nın temel varsayımları aşağıdaki gibi açıklanabilir:

- Gerçek puan ile hata puanının toplamı, gözlenen puanı verir.
- Gerçek puan ve hata puanları arasında ilişki yoktur.
- Her ölçmedeki hata puanları, birbirinden bağımsızdır. Yani, herhangi iki ölçmenin hata puanları arasındaki ilişki sıfır'dır (0,00).
- Sayılabilir sonsuz çoklukta tekrarlanan ölçmelerdeki hata puanlarının ortalaması sıfır'a yaklaşır (Crocker ve Algina, 1986).

Genellenebilirlik kuramı Klasik Test Kuramı'nın sınırlılıklarına tepki olarak Cronbach, Gleser, Nanda ve Rajaratnam tarafından (1972) önerilmiştir (Akt: Brennan 2001). Bu kuram hem sistematik hem de sistematik olmayan hata kaynaklarını ele alır ve aynı zamanda hesaplar.

Genellenebilirlik kuramı (G Kuramı), davranış ölçümlerinde güvenirliliğin değerlendirilmesini, güvenilir gözlemlerin tasarısını, araştırılmasını ve kavramlaştırılmasını sağlayan bir istatistiksel kuramdır (Brennan, 2001).

Genellenebilirlik (G) kuramı temeli varyans analizi (ANOVA) üzerine kurulmuştur. Varyans analiziyle, toplam varyans desendeki varyans bileşenlerine bölünür. Böylece ölçme sonuçları farklı varyans kaynaklarına ayrılarak bireylerin ya da objelerin

gözlenen puanlarının evren puanlarına (gerçek puanlarına) genellenebilmesi sağlanır (Brennan, 2001).

Bu varyans kaynakları genellenebilirlik kuramında değişkenlik kaynağı (facet) olarak adlandırılır. Bu değişkenlik kaynaklarının da koşulları (conditions) ya da seviyeleri (levels) vardır. Değişkenlik kaynağı ve koşul ifadeleri varyans analizi alanyazınında faktör ve düzey kavramlarına karşılık gelmektedir (Brennan, 2001; Atılgan, 2004).

I.1.3.1. Genellenebilirlik Kuramında Güvenirlik

Performans görevlerinde güvenirlilik çalışması yapılırken, puanlayıcı güvenirliliğinin (intrarater reliability) ve/veya puanlayıcılar arası güvenirliliğinin (interrater reliability)) incelenmesine de ihtiyaç duyulabilir.

Puanlayıcı güvenirliliği (intrarater reliability), bir puanlayıcının farklı zamanlarda aynı bireylerin kâğıdını birden fazla puanlamasından elde edilen verilerin tutarlılığı olarak ifade edilebilir. Puanlayıcılar arası güvenirlilik (interrater reliability) ise, iki ya da daha çok puanlayıcının aynı bireyi birbirlerinden bağımsız olarak puanlamalarından elde edilen puanların tutarlılık derecesi olarak tanımlanır (Anastasi ve Urbina, 1997).

Yapılandırmacı yaklaşımın benimsendiği eğitim sistemimizde öğrencilerin bir ürün ortaya çıkarması diğer bir deyişle performans sergilemeleri gerekliliği ön plana çıkmıştır. Sergilenen performans ürünleri de ancak değerlendirme ölçekleriyle puanlanabilir ki burada puanlayıcı önemli bir faktördür. Dolayısıyla eğitimde bireyler hakkında verilen kararlarda puanlayıcıların önemi büyüktür. Puanlayıcıların bireyler hakkında verdikleri kararların niteliği puanlama güvenirliliği incelenerek belirlenebilir. Farklı durumlarda kullanılan birçok farklı puanlayıcılar arası güvenirlilik hesaplama teknikleri vardır. Bireyler hakkında güvenilir sonuçlara ulaşmak adına kullanılacak katsayının duruma uygun olması gerekmektedir.

G kuramı yaklaşımı ile güvenirlilik kestirmenin en önemli avantajı, bütün değişkenlik kaynaklarını dikkate alarak, yapılan hesaplamalarla kapsamlı bir güvenirlilik elde edilmesidir. G kuramında, bir tek çalışmayla potansiyel bütün hata kaynaklarının ayrı ayrı ve birbiriyle etkileşiminden meydana gelen hatalar değerlendirilerek bir katsayı elde edilir (Brennan, 2001).

Genellenebilirlik Kuramı'nda güvenilirliğin araştırılmasında (1) Genellenebilirlik (G) çalışması ve (2) Karar (K) çalışması olmak üzere iki temel çalışma yer almaktadır (Güler, 2011). Genellenebilirlik Kuramı'nda G çalışması ile varyansların bağıl büyüklüklerine dayalı olarak ana ve ortak etkilerin yorumlanması ve aynı zamanda güvenilirlik katsayılarının (G ve Phi) hesaplanması için gerekli değerler sağlanmaktadır. K çalışmasıyla ise madde ve/veya puanlayıcı sayılarının artırılıp azaltılması sonucunda en güvenilir ve en verimli ölçme desenleri belirlenmektedir (Atılgan, 2005).

G kuramı sadece klasik güvenilirlik kuramını yeniden yorumlamaz, ayrıca güvenilirlik ve geçerlik arasındaki geleneksel farkın güvenilir gözlemler düzenlemekle nasıl ortadan kaldırılabileceğini de gösterir (Brennan 2001). Geleneksel olarak yapı geçerliği; gözlenebilir ölçmeler üzerine temellenmiş örtük bir yapı hakkında çıkarım yapılmasıyla ilgilidir. Oysa güvenilirlik, paralel ölçmelerle gözlenen gerçek puan hakkında çıkarım yapmayla ilgilidir (Brennan 2001). G kuramında bir evren, onun değişkenlik kaynakları ve gözlemlerin koşulları geçerlik kuramının geleneksel alanında açıklanan yapıdan tanımlanır. G kuramı, gözlemlerin örtülü yapısı hakkında çıkarımları doğrulukla göstermeyi sağlarsa güvenilir olarak tanımlar. Böyle olduğunda, G kuramı güvenilirlik ve geçerlik arasındaki geleneksel farklılığı da ortadan kaldırır. G kuramı ile "genellenebilirlik katsayısı" olarak adlandırılan bir katsayı hesaplanır (Crocker ve Algina, 1986; Brennan, 2001). Bu katsayı Klasik Test Kuramı'ndaki güvenilirlik katsayısına benzer. G kuramında "Güvenirlik" (dependability) ve "Genellenebilirlik" (generalizability) kavramlarının "güvenirlik" (reliability) yerine kullanımı, geçerlik ve güvenilirliğin birleştirilmesini yansıtır (Brennan, 2001).

G Katsayısı

Genellenebilirlik Kuramı (GK) yoluyla puanlayıcı güvenilirliği üzerinde çalışmalar yapılabilir. Genellenebilirlik Kuramı'na göre değişkenlik kaynakları çapraz (crossed) ya da yuvalanmış (nested) şekilde olabilir. Bir değişkenlik kaynağının koşulları başka bir değişkenlik kaynağının koşullarıyla örtüştüğü duruma çapraz (crossed) desen denilmektedir (Brennan 2001). Çapraz olarak tasarlanmış bir desende değişkenlik kaynakları arasında "x" işareti kullanılır. Örneğin; öğrenciler aynı ürünler doğrultusunda aynı puanlayıcılar tarafından puanlanmış ise öğrenci, ürün ve puanlayıcı değişkenleri çaprazlanarak $\bar{o} \bar{x} \bar{u} \bar{x} \bar{p}$ desenini oluşturur (ö: öğrenci, ü: ürün, p: puanlayıcı).

Değişkenlik kaynaklarının diğer bazı değişkenlik kaynaklarıyla örtüşmediği yani çaprazlanmadığı durumlar da söz konusudur. Bir değişkenlik kaynağının koşulları diğer değişkenin bazı koşullarıyla örtüşüyorsa değişkenlik kaynakları yuvalanmış (nested) olarak tasarlanmalıdır. Yuvalanmış olarak tasarlanmış bir desende değişkenler arasında “:” işareti kullanılır (Brennan 2001).

Öğrencilerin bir performans durumuna ait becerilerinin değerlendirilmesinde öğrenci, ürün ve puanlayıcı değişkenlerinin çapraz tasarlandığı $\bar{o} \times \bar{u} \times \bar{p}$ (\bar{o} : öğrenci, \bar{u} : ürün ve \bar{p} : puanlayıcı olmak üzere) desenine ait bir G çalışması örneği verilecek olursa; $\bar{o} \times \bar{u} \times \bar{p}$ deseninde öğrenciler ölçmenin nesnesi (object of measurement) olmak üzere desene ait gözlenen puan;

$$\begin{aligned}
& \mu \quad \text{[Genel Ortalama]} \\
& + \mu_{\bar{o}} - \mu \quad \text{[Öğrenci Etkisi]} \\
& + \mu_{\bar{u}} - \mu \quad \text{[Görev Etkisi]} \\
& + \mu_{\bar{p}} - \mu \quad \text{[Puanlayıcı Etkisi]} \\
& + \mu_{\bar{o}\bar{p}} - \mu_{\bar{o}} - \mu_{\bar{p}} + \mu \quad \text{[Öğrenci-Puanlayıcı Etkisi]} \\
& + \mu_{\bar{o}\bar{u}} - \mu_{\bar{o}} - \mu_{\bar{u}} + \mu \quad \text{[Öğrenci- Görev Etkisi]} \\
& + \mu_{\bar{p}\bar{u}} - \mu_{\bar{p}} - \mu_{\bar{u}} + \mu \quad \text{[Puanlayıcı- Görev Etkisi]} \\
& + \mu_{\bar{o}\bar{u}\bar{p}} - \mu_{\bar{o}\bar{p}} - \mu_{\bar{o}\bar{u}} - \mu_{\bar{u}\bar{p}} + \mu_{\bar{o}} + \mu_{\bar{p}} + \mu_{\bar{u}} - \mu \quad \text{[Artık Etki]}
\end{aligned}$$

$\bar{o} \times \bar{u} \times \bar{p}$ deseni, yedi varyans bileşenine ayrılmaktadır. Her bir varyans kaynağı için varyans analiziyle hesaplanan değerler araştırmacı için bilgi sağlamaktadır. Yukarıda gösterilen varyans kaynaklarından; \bar{o} öğrencileri, \bar{u} ürünleri, \bar{p} puanlayıcıları göstermektedir. Öğrencilere ait varyans bileşeni ($\sigma^2_{(\bar{o})}$), ölçülen beceri bakımından öğrencilerin birbirlerinden farklılık gösterip göstermediğini, ne kadar ayrıldığını gösterir. Ürünlere ait varyans bileşeni ($\sigma^2_{(\bar{u})}$), bazı ürünlerin diğerlerine göre zorluk-kolaylık bakımından farklılaşp farklılaşmadığını, puanlayıcılara ait varyans bileşeni ($\sigma^2_{(\bar{p})}$) ise, puanlayıcıların öğrencilerin performansını değerlendirirken birbirlerine göre daha cömert (yumuşak) olup olmadığını göstermektedir. Öğrenci x ürün etkisine ait varyans bileşeni ($\sigma^2_{(\bar{o}\bar{u})}$), öğrencilerin bir üründen diğerine durumlarındaki değişikliği, öğrenci x puanlayıcı etkisine ait varyans bileşeni ($\sigma^2_{(\bar{o}\bar{p})}$), öğrencilerin durumlarının bir puanlayıcıdan diğerine farklılık gösterip göstermediğini, yani

puanlayıcıların öğrencileri puanlamada tutarlı olup olmadıklarını gösterir. Ürün x puanlayıcı etkisine ait varyans bileşeni ($\sigma^2_{(pü)}$), puanlayıcıların bir üründen diğerine kararlılıklarını gösterir. Öğrenci x ürün x puanlayıcı artan etkisine ait varyans bileşeni ($\sigma^2_{(öüp)}$) ise öğrenci x ürün x puanlayıcı ortak etkisini ve/veya tesadüfi hata kaynaklarını içerir (Brennan, 2001).

G katsayısı hesaplamaları yapılırken bazen negatif değerlerde varyans bileşenleri elde edilebilir. Bu durum, eğer negatif değerler görece büyük ise ölçme modelinin hatalı olduğunun, değerler görece küçük ise (sıfıra yakın) çok büyük bir evreni temsil eden çok küçük örneklemin varlığının bir göstergesi olarak kabul edilmektedir (Güler ve diğerleri, 2012). Böyle bir durum söz konusu olduğunda negatif çıkan varyans değerinin sıfır olarak değiştirilmesi önerilmektedir (Güler ve diğerleri, 2012; Güler, 2011; Atılğan, 2005; Brennan, 2001).

Elde edilen hata kaynakları kullanılarak G katsayısı aşağıdaki gibi hesaplanmaktadır:

$$G = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_\delta^2} \quad (15)$$

σ_b^2 : Bireylere ait varyans bileşeni

σ_δ^2 : Bağıl hata varyansı

I.1.4. İlgili Araştırmalar

Büyükkıdık (2012) araştırmasında, matematik problemi çözme becerilerinin değerlendirilmesine yönelik hazırlanan iki adet performans görevinin, analitik ve bütünsel dereceli puanlama anahtarıyla, dört puanlayıcı tarafından puanlanmasından elde edilen puanların Klasik Test Kuramı ve genellenebilirlik kuramı çerçevesinde güvenilirlik katsayıları hesaplamaları yapmıştır. Araştırma sonucunda genellenebilirlik kuramından elde edilen katsayıların Klasik Test Kuramından elde edilen katsayılara göre göreceli olarak daha yüksek olduğu ve her iki kuramdan da elde edilen güvenilirlik katsayılarının yüksek olduğu görülmüştür. Dereceli puanlama anahtarı türünün de ölçümlerin güvenilirliğini etkileyebileceği, bu araştırmada analitik ve bütünsel dereceli puanlama anahtarlarından elde edilen puanların benzer güvenilirlik katsayıları verdiği; ancak analitik dereceli puanlama anahtarından elde edilen puanların göreceli olarak daha yüksek güvenilirliğe sahip olduğu belirlenmiştir.

Nalbantoğlu (2009) araştırmasında, tıp fakültesi öğrencilerinin üst düzey düşünme becerilerini ölçmeye yönelik olarak hazırlanan ve önceden belirlenen ölçütlere göre farklı puanlayıcılar tarafından birlikte ve dönüşümlü olarak puanlanan OSCE sınavının (objektif yapılandırılmış klinik sınav) güvenilirliğini Genellenabilirlik Kuramı'na göre incelemiştir. Araştırmada, puanlayıcıların öğrencileri aynı iletişim becerileri değerlendirme formuyla 15 görev doğrultusunda birlikte ve dönüşümlü olarak puanlamasıyla oluşturulan $\bar{o} \times g \times p$ ve $(\bar{o}:p) \times g$ desenleri (\bar{o} : öğrenci, g : görev, p : puanlayıcı) için ayrı ayrı G ve K çalışması yapılmıştır. Araştırma sonucunda, hem $\bar{o} \times g \times p$ hem de $(\bar{o}:p) \times g$ deseninde değişkenler için kestirilen varyans değerlerinin birbirleriyle paralellik gösterdiği; puanlayıcıların öğrencileri puanlamada aralarında farklılıklar olmadığı; her iki desene yapılan karar çalışmaları sonucunda her iki desendeki G ve Phi katsayıları arasında çok fark olmamakla birlikte $(\bar{o}:p) \times g$ deseninde katsayıların daha büyük çıkma eğiliminde olduğu belirlenmiştir.

Atmaz'ın (2009) araştırmasında, ilköğretim 8. Sınıfta okuyan 50 öğrenciye grafik yorumlama becerisine ilişkin dört açık uçlu maddeden oluşan bir ölçme aracı uygulanmış ve bu ölçme aracını puanlamak için bir puanlama yönergesi (rubrik) hazırlanmıştır. Öğrencilerin verdikleri cevaplar 17 matematik öğretmeni tarafından puanlanmıştır. Bu amaçla öğretmenler öncelikle puanlama yönergesi kullanmadan; iki hafta aradan sonra da puanlama yönergesi kullanarak puanlama yapmışlardır. Zamanın etkisini incelemek amacıyla aynı uygulama on hafta sonra tekrarlanmıştır. Her uygulamada aynı öğretmenlerin yaptıkları puanlamalar arasında pozitif yönde, yüksek ve anlamlı bir ilişki bulunurken; puan ortalamaları arasında da farklılıklar gözlenmiştir. Her uygulamada farklı öğretmenlerin yaptıkları puanlamalar arasında da pozitif yönde, yüksek ve anlamlı bir ilişki bulunmuştur. Farklı öğretmenlerin puan ortalamaları arasında ise tutarsızlığa neden olacak kadar manidar farklılıklar gözlenmiştir.

Tobar ve ark., (1999) araştırmalarında, "Duygu Durum Profili" ölçeğinin güvenilirliğini G kuramı ile araştırmışlardır. Ölçeğinin 6 alt boyutu için 0,74'den büyük G katsayıları elde edilirken toplam puanlardan elde edilen G katsayısı 0,90 olarak bulunmuştur. Tesadüfî desende madde sayısı artırıldığında G katsayısının artma eğiliminde olduğu, bir ölçekteki madde sayısı ikiye katlandığında G katsayısının çok az arttığı ve alt ölçeklerin sabit değişken kaynağı olarak tanımlanması durumunda G katsayısının 0,96'ya yükseldiği sonucuna varılmıştır.

Ömür'ün (2009) araştırmasında, değerlendirme işleminin bir aracı olarak yazılı anlatım becerisi ve bunun özel bir uygulaması olan kompozisyon kullanılacağından ilk olarak

“Kompozisyon Yazmaya İlişkin Dereceli Puanlama Anahtarı” geliştirilmiştir. Yargıcıların kompozisyonları önce ikili karşılaştırmalarla, sonra genel izlenimle ve en son olarak dereceli puanlama anahtarıyla değerlendirmeleri sonucunda, hangi değerlendirme yapılırsa yapılsın kompozisyonların sıralamalarında bir değişikliğin olmadığı gözlenmiştir. Kompozisyonlar dereceli puanlama anahtarıyla aynı yargıcılara 7-14 gün zaman aralığıyla iki kez puanlatılmış ve iki değerlendirme arasında pozitif yönde ve anlamlı bir ilişki bulunmuştur. Farklı yargıcıların beş ayrı kompozisyonu, genel izlenimle ve dereceli puanlama anahtarı ile puanlanmasında puanlayıcı güvenilirliğinin belirlenmesinde G katsayıları hesaplanmış ve bu katsayılar yüksek bulunmuştur.

Derkuş'un (2009) araştırmasında, “Problem Çözme Becerisi” ile ilgili bir ölçek kullanılarak puanlayıcılar arasındaki uzlaşma, sınıf içi korelasyon katsayısı ve lojistik regresyon teknikleri kullanılarak incelenmiştir. Sınıf içi korelasyon katsayıları ve lojistik regresyondan elde edilen sıralamalar sonucunda iki teknik arasındaki sıra farkları korelasyon katsayısı -0,23 bulunmuştur. Sonuç olarak her iki teknik ile elde edilen sıralamaların birbiriyle düşük düzeyde ve ters yönlü uyum gözlenmiştir.

Deliceoğlu'nun (2009) araştırmasında, futbolcuların teknik yetilerinin (pas, şut çekme, top sürme ve top kontrolü) belirlenmesinde kullanılan 56 maddelik Futbol Yetilerine İlişkin Dereceleme Ölçeği'nden elde edilen ölçmelerin güvenilirliği Klasik Test Kuramı ve Genellenebilirlik Kuramı'na dayalı olarak incelenmiştir. Ölçekten elde edilen veriler ile ilk puanlama ve ikinci puanlama arasındaki tutarlılık için Pearson korelasyon katsayısı, maddelerin iç tutarlılık güvenilirliği için Cronbach α (alfa), puanlayıcılar arasındaki tutarlılık için Kendall uyum katsayısı hesaplanmıştır. Ölçeğin alt boyutları ve ölçeğin toplamından elde edilen puanların güvenilirliği için puanlayıcı sayılarının 3, 4 ve 5 olduğu ve madde sayılarının bir arttırılıp, bir azaltıldığı koşullarda, Genellenebilirlik Kuramı'nda ana ve ortak etkilerin varyans bileşenlerinin kestirilmesi için G ve Phi katsayıları bulunmuştur. Araştırma sonucunda, alt boyutlar için en yüksek varyans bileşeninin birey temel etkisi için kestirilen varyans bileşeni olduğu; bunu $\beta_1\beta_2$ ortak etkisi varyans bileşeninin takip ettiği; bunların dışındaki temel ve ortak etkilere ait varyans bileşenlerinin ise sıfıra yakın olduğu belirlenmiştir. Ölçeğin güvenilirliği iç ölçütlere göre incelendiğinde G katsayısı ve Cronbach Alfa katsayısının beklenen değerlerinden yüksek olduğu; Phi ve Kendall W katsayısının ise beklenen değerlerinden düşük olduğu ortaya konmuştur.

Güler'in (2008) araştırmasında, TIMMS-1999'da yer alan açık uçlu matematik sorularından 24'ü 2007 yılı bahar döneminde 203 öğrenciye uygulanmış; öğrencilerin verdikleri cevaplar dört puanlayıcı tarafından holistik rubrik yardımıyla puanlanmıştır. Puanlayıcıların yaptıkları puanlamalara ilişkin güvenilirlik Klasik Test Kuramı, Genellenebilirlik Kuramı ve çok değişkenlik kaynaklı Rasch ölçme modeli yardımıyla incelenmiştir. Klasik Test Kuramı'nda Cronbach alfa katsayısı ve Kendall'in konkordans katsayısı; Genellenebilirlik Kuramı'nda $b \times g \times p$ (b: birey; p: puanlayıcı; g: görev) tümüyle çaprazlanmış desen kullanılarak genellenebilirlik ve güvenilirlik katsayıları; Rasch ölçme modelinde ise birey, puanlayıcı ve madde boyutlarına ilişkin ayrı ayrı güvenilirlik hesaplamaları yapılmıştır. Araştırma sonucunda, puanlayıcıların puan ortalamaları arasında fark olmakla birlikte, birbirleriyle uyumlu puanlamalar yaptıkları belirlenmiştir.

Lee ve Kantor (2007) çalışmalarında, farklı yazma görevleri ve puanlayıcıların yazma becerilerine ait puanlarındaki güvenilirliği araştırarak farklı senaryolara bağlı karar çalışması ile en yüksek puan güvenilirliği için puanlayıcı ve görev sayısını belirlemeye çalışmışlardır. Araştırma sonuçlarına göre, görevlerinin sayısını artırmanın puanlayıcı sayısını artırmaktan daha etkili olduğu tespit edilmiştir.

Yelboğa'nın (2007) araştırmasında, ortak performans ölçüt ve standartlarına dayalı yöntem yaklaşımı ile geliştirilmiş İş Performansı Ölçeği'nin güvenilirliği, Klasik Test Kuramı ve Genellenebilirlik Kuramı yardımıyla incelenmiş ve ayrıca Kendall'in konkordans katsayısı yardımıyla değerlendiriciler arası tutarlılığa bakılmıştır. Klasik Test Kuramı'nda, test tekrar test ve Cronbach alfa güvenilirlik katsayıları; Genellenebilirlik Kuramı'nda ise çok değişkenli modelle G ve Phi katsayıları hesaplanmıştır. Araştırma sonucunda, İş Performansı Ölçeği'nde aynı ölçme durumu için Klasik Test Kuramı'nda ve Genellenebilirlik Kuramı'nın çok değişkenli modelle elde edilen güvenilirlik katsayılarının birbirleriyle uyumlu sonuçlar ürettiği belirlenmiştir.

Alharby'nin (2006) araştırmasında, analitik ve holistik puanlama yöntemlerini kullanarak puanlattığı performans görevlerinden elde ettiği verileri Genellenebilirlik Kuramı ve çok değişken kaynaklı Rasch modeli ile analiz etmiştir. Bu çalışmada, hangi puanlama yöntemi ve hangi model birlikte kullanılırsa en yüksek düzeyde güvenilirlik elde edileceği ortaya koyulmaya çalışılmıştır. Araştırmanın sonucunda, olarak analitik puanlama kullanılarak Genellenebilirlik Kuramı ile hesaplamalar yapıldığında, holistik puanlama

kullanılarak yapılan hesaplamalardan elde edilen güvenilirlikten daha yüksek bir güvenilirlik elde edildiği görülmüştür.

Durand ve arkadaşları (2004) araştırmalarında, fiziksel çalışma performansını değerlendirmek için bir derecelendirme ölçeği kullanarak, 40 kişiyi, 3 mesleki ve 2 fiziksel terapistte puanlattırılmış ve bu puanların puanlayıcı güvenilirliğini uyuma yüzdeleri ve Cohen'in kappa indekslerini kullanarak hesaplamıştır. Araştırmanın sonucunda puanlanmalar arasında yüksek oranda bir uyuma gözlenmiştir.

Atılgan'ın (2004) yaptığı çalışmada, 2002 ve 2003 yıllarında Müzik öğretmenliği özel yetenek seçme sınavı verilerini kullanarak Genellenebilirlik Kuramı ve çok değişken kaynaklı Rasch modeli analizlerini yapmıştır. Analizler sonucunda elde edilen verilerle Genellenebilirlik Kuramı ve çok değişken kaynaklı Rasch modeli istatistikleri karşılaştırılmıştır. Elde edilen sonuçlara göre, Genellenebilirlik Kuramı'nın tek değişkenli ve çok değişkenli modellerinin aynı ölçme durumu için alt boyutlardan oluşan testlerde farklı sonuçlar ürettiği, G ve Phi katsayılarının modelden etkilendiği ve alternatif karar çalışmalarıyla farklı değerlendirici sayıları senaryoları karşısında elde edilen G ve Phi katsayılarının gerçek durumda kestirilenlerden farklı olduğu görülmüştür. Ayrıca, hem Genellenebilirlik Kuramı hem de çok değişken kaynaklı Rasch modeli yaklaşımlarının her ikisinin de oldukça kullanışlı sonuçlar ürettiği görülmüştür.

Goodwin (2001) araştırmasını, hipotetik veriler kullanarak yapmıştır. Araştırmada basit yüzde, kappa, basit korelasyon teknikleri ve Genellenebilirlik Kuramı hesaplamaları kullanılmıştır. Araştırmada, basit yüzde ile hesaplanan değerlendiricilerin farklı zamanlarda verdikleri puanların uyumunun ortanca değeri %20 ve değerlendiricilerin uyumunun ortanca değeri %80, Pearson korelasyon katsayısı 0.90, Kappa istatistikleri 0 ile 0.63 arasında hesaplanmıştır. Genellenebilirlik kuramı ile hesaplanan G katsayısı ise 0.88 olarak bulunmuştur. Bir çalışmada birçok hata kaynağını ayırabilmesi ve kapsamlı olması sebebiyle Genellenebilirlik Kuramının bu tür araştırmalarda daha avantajlı olduğu sonucuna varmıştır.

Stuhlmann ve arkadaşlarının (1999) yaptıkları araştırmada, uzun cevaplı testlerin puanlanmasında öğretmenlere verilen rubrik puanlama eğitiminin etkisi genellenebilirlik çalışması yardımıyla incelenmiştir. Rubrik eğitimi alan 23 ve eğitim almayan 17 öğretmen üzerinde yürütülen çalışmada sonuçlar rubrik puanlama eğitimi almanın puanlayıcı

güvenirliğini arttırıp arttırmadığını ortaya koymuştur. Puanlayıcılar eğitim aldığı anda aynı ürün için yapılan puanlamalar arasında varyansın daha az olduğu gözlenmiştir.

Lee ve Frisbie (1999) araştırmada, test puanlarının güvenilirliği için hesaplanan Cronbach alfa güvenilirlik katsayısı ile G katsayısı arasındaki farkı belirlemek amaçlanmıştır. Araştırma sonucunda, kestirilen G katsayısı ve Cronbach alfa katsayısı arasındaki fark 0,04 olarak bulunmuştur. Ayrıca G katsayısının, test içindeki her bir pasaj içindeki madde sayısının değiştirilmesi yerine pasaj sayısının değiştirilmesinin daha etkili olduğu sonucuna varılmıştır.

Lane ve Sabers (1989) araştırmalarında, öğrencilerin yazma becerilerini değerlendirmede puanlama sisteminin güvenilirliğini belirlemek amacıyla genellenebilirlik kuramını kullanmışlardır. Araştırmada 15 öğrenciye bir konu verip bu konu hakkında kompozisyon hazırlamaları istenmiştir. Öğrencilerin kâğıtları 8 öğretmen tarafından ve bir puanlama sistemi oluşturularak değerlendirilmiştir. Öğretmenler puanlama sisteminin dört kategoride hazırlamıştır. Araştırmada öğrenciler (s), puanlayıcılar (r) ve puanlama sistemindeki kategoriler (c) ile gösterilmek üzere kategori değişkeni sabit olarak belirlenip tüm değişkenlerin çapraz tasarlandığı $s \times r \times c$ deseniyle G ve K çalışmaları yapılmıştır. Araştırma sonuçlarına göre puanlayıcı sayısının azaldıkça mutlak ve bağıl hata varyanslarının arttığı, G ve Phi katsayılarının ise azaldığı tespit edilmiştir.

İlgili alanyazın incelendiğinde, Klasik Test Kuramı, Genellenebilirlik Kuramı ve çok değişkenli Rasch ölçme modeli kullanılarak güvenilirlik hesaplamaları yapıldığı gözlenmektedir. Ancak Klasik Test Kuramı'nda yapılan güvenilirlik çalışmalarında Fleiss'in Kappa katsayısının kullanıldığı, her bir ölçme aracı ve puanlayıcı sayısı için 100'er örneklem çekerek bunlardan elde edilen güvenilirlik katsayılarının ortanca ve standart hata değerlerinin incelendiği bir çalışmaya rastlanmamıştır. Bu nedenlerle bu araştırmanın problemini, aynı performans görevinin farklı sayıda puanlayıcılar tarafından üç farklı puanlama anahtarı kullanılarak puanlanması durumunda puanlayıcılar arası güvenilirliğin Genellenebilirlik Kuramına dayalı olarak incelenmesi oluşturmaktadır.

I.2. Amaç

Bu araştırmanın amacı, aynı performans görevlerinin farklı sayıda puanlayıcı tarafından kontrol listesi, dereceleme ölçeği ve analitik rubrik yardımıyla puanlanması

durumunda, puanlayıcı güvenilirliklerinin Genellenebilirlik Kuramı çerçevesinde incelenmesidir. Bu amaç doğrultusunda, aşağıdaki sorulara yanıt aranmıştır.

- 1) Genellenebilirlik kuramına göre, aynı performans görevlerinin
 - a) Kontrol listesi,
 - b) Dereceleme ölçeği,
 - c) Analitik rubrik

kullanılarak ve 45 puanlayıcı arasından seçkisiz olarak seçilen

- i) 2'şer
- ii) 3'er
- iii) 5'er
- iv) 10'ar

puanlayıcıdan oluşan 100'er örneklemin genellenebilirlik kuramına göre kestirilen G katsayılarının derecesi nasıldır?

- 2) Genellenebilirlik Kuramı'na göre aynı performans görevlerinin
 - d) Kontrol listesi,
 - e) Dereceleme ölçeği,
 - f) Analitik rubrik

kullanılarak ve 45 puanlayıcı arasından seçkisiz olarak seçilen

- i) 2'şer
- ii) 3'er
- iii) 5'er
- iv) 10'ar

puanlayıcıdan oluşan 100'er örneklemin puanlama güvenilirliklerinin ortancaları ve standart hataları nasıldır?

I.3. Önem

Performans görevleri, öğrencilerin belirli kazanımları gerçekleştirme düzeylerini belirlemede yararlanılan önemli ölçme araçlarıdır. Öğrencilerin performans görevlerine dayalı olarak verilecek kararların niteliği, bu çalışmaların hatasız ve tutarlı bir biçimde

puanlanabilmesiyle doğrudan ilgilidir. Öğrencilerin performans görevleri ise öğretmenler tarafından puanlanır ve değerlendirilir. Bu nedenle, öğretmenlerin yaptığı bu puanlamaların güvenilirliğinin test edilmesi önem taşımaktadır.

Performans görevlerinin puanlanmasında, puanlama ölçütlerinin önceden belirlenmiş olması gerekir. Performans görevlerinin belirlenen bu ölçütlere uygun olarak kontrol listesi, dereceleme ölçeği ve analitik rubrik yardımıyla puanlanmasının bu araçlardan herhangi biri kullanılmaksızın yapılacak puanlamalardan daha nesnel olması beklenir. Bunun yanında, aynı performans görevinin farklı puanlayıcılar tarafından değerlendirilmesi de söz konusu olabilir. Böyle bir durumda, puanlayıcıların yaptıkları puanlamaların birbiriyle tutarlı olması beklenir.

Farklı sayıda puanlayıcının farklı araçlar yardımıyla yaptıkları puanlamaların güvenilirliği farklı tekniklerle incelenebilir. Bu araştırmanın amacı, aynı performans görevinin puanlanmasında kontrol listesi, dereceleme ölçeği ve analitik rubrik kullanılması durumunda farklı sayıda puanlayıcı tarafından yapılacak puanlamaların güvenilirliğinin çeşitli yöntemlere göre incelenmesi açısından önem taşımaktadır. Yapılan bu araştırmanın, farklı sayıda puanlayıcı tarafından puanlanan aynı performans görevlerine ilişkin puanlama güvenilirliğinin belirlenmesinde üç farklı araçtan faydalanılması açısından ve bir aracın diğerine göre daha tutarlı bilgi sağlayıp sağlamadığını belirlemek açısından önemli olacağı düşünülmektedir. Ayrıca, araştırma sonuçlarının pratikte uygulayıcılara performans görevlerinin puanlanmasında güvenilir ölçme araçlarından yararlanmanın önemi ve performans görevlerini puanlayacak puanlayıcı sayısını belirleme konusunda bir görüş sağlayacağı; performans görevlerini puanlamada tercih edebilecekleri araçlara karar verme konusunda yararlı olacağı düşünülmektedir.

Yapılan alanyazın araştırmasında, performans görevlerinin farklı sayıda puanlayıcılarla ve/veya üç farklı ölçme aracı yardımıyla puanlanmasına ilişkin bir çalışmaya rastlanmamıştır. Bu anlamda, araştırma bulgularının ölçme değerlendirme alanyazınına önemli bir katkı sağlaması beklenmektedir.

I.4. Sayıtlar

1) Her puanlayıcının aynı performans görevlerini kontrol listesi, dereceleme ölçeği ve analitik rubrik yardımıyla yaptığı puanlamalar birbirinden bağımsızdır.

2) Puanlayıcılar her öğrencinin performans görevini aynı titizlikle ve dikkatle puanlamışlardır.

I.5. Sınırlılıklar

1) Bu araştırma, ilköğretim 5. Sınıf Türkçe dersinde bir grup öğrenciye verilecek aynı performans göreviyle,

2) bu performans görevlerinin puanlanmasında kullanılacak ve araştırmacı ve ilköğretim sınıf öğretmenleri tarafından hazırlanacak kontrol listesi, dereceleme ölçeği ve analitik rubrik ve bu araçlardan elde edilecek verilerle,

3) performans görevlerini puanlayacak toplam 45 puanlayıcı ve puanlayıcılar arasından 2'li, 3'lü, 5'li ve 10'lu olarak seçkisiz seçilecek 100'er örnekleme

sınırlıdır.

BÖLÜM II

GEREÇ VE YÖNTEM

2.1. Araştırmanın Modeli

Bu araştırmada Genellenebilirlik Kuramı ile elde edilen puanlayıcı güvenilirlikleri incelendiğinden, bu çalışma bir temel araştırma olarak tanımlanabilir.

2.2. Araştırma Grubu

Araştırmanın verileri, sınıf öğretmenlerinin aynı öğrencilere ait hikayeleri üç farklı değerlendirme formu yardımıyla puanlanması sonucu elde edilmiştir. Araştırmaya Mersin ili merkez ilköğretim okullarında sınıf öğretmeni olarak çalışan öğretmenler katılmıştır.

Araştırma sorularına yanıt aramak amacıyla ilköğretim okullarında sınıf öğretmeni olarak çalışan ve araştırmaya katılmaya gönüllü 104 sınıf öğretmeni ile çalışmaya başlanmıştır. Fakat 3 ayrı ölçme aracının 10-15 gün ara ile uygulanması sürecinde katılımcı kayıpları olmuştur. Nihai olarak 3 ölçme aracı ile 6 hikâyeyi puanlayan 45 sınıf öğretmenin katılımıyla puanlayıcı grubu oluşturulmuştur. 45 kişilik puanlayıcı grubunun cinsiyet ve kıdemlerine ilişkin bilgiler Tablo 2'deki gibidir.

Tablo 2. Çalışma grubunda yer alan öğretmenlerin kıdem ve cinsiyet bilgileri

Cinsiyet		Kıdem			Toplam
		1-10 yıl	11-20 yıl	21 yıl ve üstü	
Kadın	f	1	10	17	28
	%	%2,22	%22,2	%37,74	%62,16
Erkek	f	2	9	6	17
	%	%4,44	%20	%13,2	%37,74
Toplam	f	3	19	23	45
	%	6,67	%42,2	%51,06	%100

2.3. Veri Toplama Araç ve/veya Teknikleri

Araştırma verilerinin elde edilmesi amacıyla, aynı performans görevinin değerlendirilmesinde kullanılacak 3 puanlama anahtarı hazırlanmıştır. Bu puanlama

anahtarları, kontrol listesi dereceleme ölçeği ve analitik rubriktir. Bu araçların hazırlanmasında izlenen yol, aşağıda ayrıntılı olarak açıklanmıştır.

Kazanımların Belirlenmesi

İlk aşamada, geliştirilecek ölçme araçlarının konu alanı ve içeriği kararlaştırılmış ve ilköğretim 5. Sınıf Türkçe dersi öğretim programı incelenmiştir. Bu program okuma, yazma, dinleme ve konuşma olmak üzere 4 temel beceriden ve bu becerilere ilişkin kazanımlardan oluşmaktadır. Ölçme araçlarının oluşturulmasında ve öğretmenler tarafından öğrenci performanslarını değerlendirme amacıyla kullanılmasında yazma becerisi ile ilgili kazanımların uygun olabileceği düşünülmüş ve performans görevinin yazma becerisine yönelik olarak hazırlanması kararlaştırılmıştır.

İkinci aşamada, 5. Sınıf Türkçe dersi öğretim programında yazma ile ilgili 64 kazanımdan hangilerinin bir performans görevi yardımıyla gözlemlenebileceğine karar verilmiştir. Bu amaçla bir kontrol listesi hazırlanmıştır. Bu kontrol listesinde 64 kazanım yer almış ve her kazanımın karşısında da Uygundur/Uygun değildir seçeneklerine yer verilmiştir (Ek 1).

Oluşturulan bu liste, 7 ilköğretim 5. sınıf öğretmenine verilerek, ilgili kazanımların bir performans görevi yardımıyla gözlenmede uygun olup olmadığına karar vermeleri istenmiştir. Öğretmenlerin verdiği yanıtların birbiriyle tutarlı olup olmadığını, diğer bir ifadeyle uzmanlar arası uyumun sağlanıp sağlanmadığını belirlemek amacıyla, Basit Uyum İndeksinden (Erkuş, 2006) yararlanılmıştır.

Yapılan analiz sonrasında, 13 kazanım için öğretmenler arasında %70 ve üzeri uyum sağlandığı (Erkuş, 2006) belirlenmiştir. Diğer ifadeyle öğretmenler, yazma becerisi ile ilgili 64 kazanımdan 13'ünün performans görevi yardımıyla yapılabileceği konusunda uzlaşmışlardır. Öğretmenler arası uyumun sağlandığı bu kazanımlar Tablo 3'te verilmiştir.

Üçüncü aşamada, 6 beşinci sınıf öğretmeninden, seçilen 13 kazanımın gerçekleşip gerçekleşmediğini gözlemlemek üzere performans görevi örnekleri yazmaları istenmiştir. Öğretmenler farklı kazanımlarla ilgili toplam 10 performans görevi örneği (Ek 2) yazmıştır.

Dördüncü aşamada, öğretmenlerin önerdikleri performans görevleri dikkate alınarak, Türkçe Eğitimi alanında öğretim üyesi olarak çalışan bir uzman yardımıyla, 13

kazanım ve öğretmenlerin bu kazanımların gerçekleşmesi için önerdikleri performans görevi örnekleri birlikte incelenmiştir. Bu inceleme sonucunda “Hikâye yazar” kazanımı ve bu kazanıma uygun olarak da performans görevi örnekleri arasından “Bir resme bakar ve bu resmi temel alarak bir hikâye yazar” şeklindeki performans görevinin araştırma kapsamında kullanılmasına karar verilmiştir. Bu belirlemenin ardından 5. sınıf öğrencilerine yönelik olarak Ek 3’te gösterilen performans görevi hazırlanmıştır.

Tablo 3. Performans görevi yardımıyla gözlenebileceği belirlenen ve uzmanlar arasında uzlaşmanın sağlandığı kazanımlar

No	Kazanımlar
1	Konu dışına çıkmadan yazar.
2	İzlenim ve deneyimlerine dayalı yazılar yazar.
3	Duygu, düşünce ve hayallerini anlatan yazılar yazar.
4	Kendisini, ailesini ve çevresini tanıtan yazılar yazar.
5	Özet çıkarır.
6	Kendisine ilginç ve çekici gelen karakter, olay, yer, an vb. ile ilgili görüş ve düşüncelerini yazar.
7	Dilek, istek ve şikâyetlerini yazılı olarak ifade eder.
8	Eksik bırakılan metni yazarak tamamlar.
9	Mektup yazar.
10	Açıklayıcı ve bilgilendirici yazılar yazar.
11	Hikâye yazar.
12	Yazılarında betimlemeler yapar.
13	Şiir yazar.

Performans Görevinin Uygulanması

Beşinci aşamada, hazırlanan performans görevi, 2011-2012 eğitim öğretim yılında Mersin il merkezindeki dört farklı ilköğretim okulunun 5. sınıflarında okuyan 104 öğrenciye uygulanmıştır. Bu amaçla öncelikle ilgili sınıfların öğretmenleriyle görüşülmüş ve seçilen kazanımla ilgili ne tür etkinlikler yapıldığı belirlenmiştir.

Bilindiği gibi performans görevi hem sınıf içi hem de sınıf dışı etkinlikleri birlikte gerçekleştirmeyi gerektirmektedir. Bu nedenle performans görevinin sınıf dışı etkinlik kısmını gerçekleştirmek üzere onlardan 5’er kısa hikâye okumaları ve bunlardan bir tanesinin özetini yazmaları istenmiştir. Öğrencilere bu etkinlik için 2 hafta süre verilmiştir. Öğrencilerin bu sürenin sonunda yazdıkları hikâye özetleri araştırmacı tarafından toplanmıştır. Bu özetlere ilişkin örnekler Ek 4’te verilmiştir.

Bu sürenin sonunda arařtırmacının kendisi performans görevinin sınıf ii etkinlik kısmını gerekleřtirmek üzere sınıflara gitmiřtir. Tm uygulamalar arařtırmacının kendisi tarafından gerekleřtirilmiřtir.

Sınıf ii etkinlik ařamasında bir grsel materyal (Ek 3) A4 boyutunda renkli ođaltılmak suretiyle ğrencilere dađıtılmıř ve performans görevinin sınıf ii etkinlik ařaması arařtırmacı tarafından ğrencilere “elinizdeki tabloya bakın ve hayal gcnz kullanarak bu tabloda grdğnz resmi anlatan bir hikye yazın” řeklinde aıklanmıřtır. ğrencilere performans görevini sonulandırmaları iin bir ders saati sre tanınmıřtır. Srenin bitiminde tm ğrencilerden hazırlamıř oldukları alıřmalar toplanmıřtır.

Arařtırmanın Amacına Uygun Olarak Hikyelerin Belirlenmesi

Altıncı ařamada, ğrenciler tarafından yazılan 104 hikye, arařtırmacı tarafından tek tek okunarak incelenmiř; anlařılamayan, okunamayan, performans görevinin amacının dıřında hazırlanan hikayeler elenmiřtir. Bu eleme sonucunda toplam hikye sayısı 54’e dřmřtr.

54 hikye arasından arařtırmada kullanılmak üzere hikye semek amacıyla, ncelikle bu hikyelerin “ok iyi”, “iyi”, “orta”, “kt” ve “ok kt” olmak üzere 5 kategori altında sınıflanması kararlařtırılmıřtır. Bu amala her hikyeye bir numara verilmiř ve 2’si Trke, 2’si Sınıf ğretmeni olmak üzere 4 ğretmeden belirlenen kategorilere gre sınıflamaları istenmiřtir. Her ğretmen birbirinden bađımsız olarak alıřmıř ve her hikyeyi okuduktan sonra, ilgili kategorilerden uygun olduđunu dřndğ kategoriye yerleřtirmiřtir. Bu alıřmanın yapılabilmesi iin Ek 5’te yer alan form hazırlanmıřtır.

ğretmenlerin yaptığı sınıflamalar tamamlandıđında, “ok kt” kategorisinde hi bir hikyenin olmadığı gzlenmiřtir. Bunun zerine kategori birleřtirme yoluna gidilmiř, “ok iyi” ve “iyi” olarak tanımlanan kategoriler “iyi”; “orta” olarak tanımlanan kategori “orta” ve “kt” ve “ok kt” olarak tanımlanan kategoriler de “kt” olarak yeniden tanımlanarak hikyeler  kategori altına yerleřtirilmiřtir (Ek 6).

Yapılan  kategorili dzenlemeye gre 4 ğretmenin de ortaklařtığı yani aynı hikyeyi aynı kategoriye yerleřtirdiđi (iyi, orta ya da kt) 17 hikye olduđu tespit edilmiřtir.

Bunun sonucunda her kategoriden seçkisiz olarak 2'şer hikâye seçilmiş ve araştırmanın bu 6 hikâye kullanılarak yürütülmesi kararlaştırılmıştır.

Performans Görevini Değerlendirme Ölçütlerinin Belirlenmesi

Hikâye yazmaya ilişkin puanlama anahtarlarını geliştirme çalışmaları için alanında uzman bir öğretim üyesi, bir öğretim elemanı ve 8 öğretmenle görüşülerek onlara yazılan hikâyeyi değerlendirmede göz önüne alınması gereken ölçütlerin (davranış göstergelerinin) neler olduğu sorulmuş ve ilgili alan yazın taraması da gerçekleştirilerek bu ölçütler listelenmiştir (Ek 7).

Listelenen ölçütler ana ve alt ölçütler başlıkları altına yerleştirilmiş ve davranış göstergelerinin ana ve alt ölçütlere uygunluğunun ve puanlama yapılırken ağırlıklandırmanın nasıl olması gerektiğinin kararlaştırılabilmesi amacıyla bir değerlendirme formu hazırlanmıştır (Ek 8). 9 Türkçe öğretmeni bu formu dikkate alarak listelenen ölçüt ve alt ölçütlerin uygun olup olmadığına ve bu ölçüt ve alt ölçütlerin puan ağırlıklarına karar vermişlerdir.

9 öğretmenin her ölçüt/alt ölçüt için yaptığı değerlendirmeler arasındaki uyum basit uyum katsayısı ile incelenmiştir. %77,7 ve üzerinde uyum gösteren ölçütler puanlama anahtarına dâhil edilmiştir. Uyumun nispeten düşük çıktığı ölçütler için araştırmacı aracılığıyla uzmanlar arası uzlaşma sağlanmış ve uzlaşma sağlanan ölçütler puanlama anahtarına dâhil edilmiştir. Basit uyum katsayısı ile elde edilen bulgular ve bu doğrultuda yapılan düzenlemeler puanlama anahtarlarının kapsam geçerliğine kanıt olarak kabul edilebilir.

Seçilen performans görevlerinin puanlanması amacıyla, araştırmacı tarafından kontrol listesi (Ek 9), dereceleme ölçeği (Ek 10) ve analitik rubrik (Ek 11) hazırlanmıştır. Hazırlanan araçlar yine daha önce çalışılan çalışma grubunda yer almayan altı sınıf öğretmeni tarafından incelenmiş ve son haline getirilmiştir. Araştırmanın verileri bu üç araç yardımıyla elde edilmiştir.

2.4. Verilerin Analizi

6 performans görevinin 45 puanlayıcı tarafından kontrol listesi, dereceleme ölçeği ve analitik rubrik yardımıyla puanlanmasından elde edilen veriler, 2, 3, 5, 10 puanlayıcının yaptığı puanlamalar dikkate alınarak genellenebilirlik kuramı kapsamında analiz edilmiş ve elde edilen bulgular incelenmiştir. Bu amaçla sırasıyla aşağıdaki istatistiksel çalışmalar yapılmıştır.

45 puanlayıcı arasından seçkisiz olarak 2, 3, 5 ve 10 puanlayıcıdan oluşan 100'er örneklem seçilmiştir. Örneğin, 2 puanlayıcı kombinasyonundan oluşan 100 farklı örneklem; 3 puanlayıcı kombinasyonundan oluşan 100 farklı örneklem vb. elde edilmiştir.

Seçilen 400 farklı örneklem için Genellenebilirlik Kuramı'nda analizler kontrol listesi, dereceleme ölçeği ve analitik rubrik için ayrı ayrı yapılmıştır. Araştırmada tümüyle çarpazlanmış desen $g \times pa \times p$ (g: görev, pa: puanlama anahtarı, p: puanlayıcı) kullanılmıştır. Yani her görev her bir puanlama anahtarı ile her bir puanlayıcı tarafından puanlanmıştır. Bu desen kullanılarak, 2, 3, 5 ve 10 puanlayıcı olması durumunda bu üç farklı ölçme aracına bağlı olarak yapılan puanlamaların güvenilirliği G katsayısı hesaplanarak incelenmiştir.

2.5. İşlem

Araştırmanın gerçekleştirilmesinde aşağıdaki aşamalar izlenmiştir.

1. Öğrencilerin hazırladıkları performans görevlerinin içerisinden seçilen 6 hikâye, araştırmaya gönüllü olarak katılan 45 sınıf öğretmenine 10-15 gün arayla verilmiştir. Kontrol listesi, dereceleme ölçeği ve analitik rubrik ve aynı zamanda 6 farklı hikâye (Ek 12) farklı sıralarda karışık olarak öğretmenlerin değerlendirmelerine sunulmuştur.
2. Elde edilen veriler bilgisayar ortamına aktarılmıştır.
3. 45 puanlayıcı içerisinden seçkisiz olarak 2, 3, 5 ve 10 puanlayıcılı 100'er örneklem çekmek için bir kod yazılmıştır.
4. Verilerin hesaplamaları için gerekli düzenlemelerin yapılmasını sağlamak amacıyla bir kod yazılmıştır (Ek 13).
5. Genellenebilirlik Kuramı'nda 2, 3, 5 ve 10 puanlayıcılı örneklemelerde yapılacak olan güvenilirlik kestirimi için birer kod yazılmıştır (Ek 14). 2, 3, 5 ve 10 puanlayıcılar için

ayrı ayrı olmak üzere yazılan kodlar birleştirilerek (Ek 15) 1200 örneklemin her biri için güvenilirlik katsayıları hesaplanmıştır.

6. Örneklemelerin seçilmesi, gerekli düzenlemelerin yapılması güvenilirliklerin hesaplanmasında R programından yararlanılmıştır.
7. Elde edilen güvenilirlik katsayılarının her durum için 100'er örneklemlerinin ortancaları ve standart hataları hesaplanmıştır.

BÖLÜM III

BULGULAR

Bu bölümde, araştırma sorularının her biri için yapılan analizler sonucu elde edilen bulgulara yer verilmiştir.

3.1. “Genellenebilirlik Kuramı’na göre, aynı performans görevlerinin; kontrol listesi, dereceleme ölçeği, analitik rubrik kullanılarak ve 2, 3, 5, 10 puanlayıcıdan oluşan 100’er örneklemin puanlama güvenilirlikleri nasıldır?” araştırma sorusuna ilişkin bulgular

Genellenebilirlik Kuramı’na göre, aynı performans görevlerinin; kontrol listesi, dereceleme ölçeği, analitik rubrik kullanılarak ve 2, 3, 5, 10 puanlayıcıdan oluşan 100’er örneklemin puanlama güvenilirlikleri aşağıdaki başlıklar altında verilmiştir.

i) 2 Puanlayıcı 100’er örneklem

Tablo 4, incelendiğinde, 2 puanlayıcı örneklemde kontrol listesi kullanılarak puanlama yapıldığında (2 kategorili) güvenilirlik kestirimleri 0,000 ile 0,800 arasında; analitik rubrik kullanılarak puanlama yapıldığında (3 kategorili) güvenilirlik kestirimlerinin 0,000 ile 0,880 arasında; dereceleme ölçeği kullanılarak puanlama yapıldığında (5 kategorili) güvenilirlik kestirimlerinin 0,000 ile 0,950 arasında değiştiği gözlenmektedir.

ii) 3 Puanlayıcı 100’er örneklem

Tablo 5, incelendiğinde, 3 puanlayıcı örneklemde kontrol listesi kullanılarak puanlama yapıldığında (2 kategorili) güvenilirlik kestirimleri 0,000 ile 0,869 arasında; analitik rubrik kullanılarak puanlama yapıldığında (3 kategorili) güvenilirlik kestirimlerinin 0,000 ile 0,787 arasında; dereceleme ölçeği kullanılarak puanlama yapıldığında (5 kategorili) güvenilirlik kestirimlerinin 0,000 ile 0,927 arasında değiştiği gözlenmektedir.

iii) 5 Puanlayıcı 100’er örneklem

Tablo 6, incelendiğinde, 5 puanlayıcı örneklemde kontrol listesi kullanılarak puanlama yapıldığında (2 kategorili) güvenilirlik kestirimleri 0,000 ile 0,960 arasında; analitik rubrik kullanılarak puanlama yapıldığında (3 kategorili) güvenilirlik kestirimlerinin 0,000 ile 0,804 arasında; dereceleme ölçeği kullanılarak puanlama yapıldığında (5 kategorili) güvenilirlik kestirimlerinin 0,000 ile 0,884 arasında değiştiği gözlenmektedir.

Tablo 4. Genellenebilirlik Kuramına göre, üç farklı puanlama anahtarı kullanılarak ve 2 puanlayıcıdan oluşan 100'er örneklemin puanlama güvenilirlikleri

Kategori Sayısı																			
No	2	3	5	No	2	3	5	No	2	3	5	No	2	3	5	No	2	3	5
1	0,000	0,000	0,000	21	0,000	0,217	0,095	41	0,209	0,360	0,355	61	0,385	0,487	0,537	81	0,506	0,667	0,759
2	0,000	0,000	0,000	22	0,003	0,240	0,108	42	0,219	0,364	0,388	62	0,385	0,491	0,543	82	0,516	0,675	0,766
3	0,000	0,000	0,000	23	0,006	0,247	0,129	43	0,238	0,378	0,389	63	0,391	0,498	0,546	83	0,529	0,675	0,772
4	0,000	0,000	0,000	24	0,013	0,266	0,139	44	0,252	0,407	0,406	64	0,391	0,506	0,553	84	0,529	0,677	0,773
5	0,000	0,000	0,000	25	0,015	0,266	0,144	45	0,270	0,423	0,410	65	0,394	0,508	0,569	85	0,540	0,682	0,787
6	0,000	0,000	0,000	26	0,054	0,288	0,147	46	0,276	0,423	0,417	66	0,401	0,517	0,577	86	0,556	0,683	0,788
7	0,000	0,000	0,000	27	0,056	0,298	0,148	47	0,300	0,430	0,430	67	0,403	0,536	0,585	87	0,599	0,690	0,788
8	0,000	0,000	0,000	28	0,056	0,300	0,163	48	0,300	0,431	0,461	68	0,406	0,536	0,599	88	0,599	0,701	0,790
9	0,000	0,000	0,000	29	0,069	0,309	0,191	49	0,302	0,431	0,468	69	0,429	0,541	0,600	89	0,630	0,704	0,809
10	0,000	0,000	0,000	30	0,084	0,311	0,232	50	0,302	0,438	0,468	70	0,431	0,553	0,605	90	0,643	0,719	0,815
11	0,000	0,000	0,000	31	0,092	0,324	0,237	51	0,315	0,440	0,479	71	0,433	0,554	0,608	91	0,675	0,728	0,827
12	0,000	0,000	0,000	32	0,122	0,334	0,238	52	0,315	0,442	0,479	72	0,435	0,567	0,635	92	0,684	0,737	0,829
13	0,000	0,000	0,000	33	0,153	0,343	0,259	53	0,330	0,446	0,483	73	0,435	0,577	0,641	93	0,696	0,748	0,841
14	0,000	0,000	0,000	34	0,154	0,343	0,268	54	0,339	0,446	0,484	74	0,454	0,578	0,659	94	0,717	0,753	0,883
15	0,000	0,000	0,000	35	0,162	0,345	0,268	55	0,347	0,446	0,490	75	0,456	0,588	0,680	95	0,727	0,772	0,884
16	0,000	0,032	0,014	36	0,170	0,347	0,275	56	0,352	0,473	0,499	76	0,460	0,600	0,685	96	0,740	0,772	0,890
17	0,000	0,063	0,022	37	0,170	0,348	0,280	57	0,364	0,473	0,511	77	0,464	0,612	0,690	97	0,751	0,780	0,899
18	0,000	0,134	0,022	38	0,202	0,355	0,311	58	0,367	0,476	0,514	78	0,473	0,621	0,696	98	0,788	0,794	0,930
19	0,000	0,207	0,078	39	0,203	0,355	0,311	59	0,382	0,480	0,531	79	0,474	0,634	0,698	99	0,796	0,797	0,950
20	0,000	0,216	0,078	40	0,203	0,356	0,327	60	0,382	0,487	0,533	80	0,490	0,665	0,757	100	0,800	0,880	0,950

Tablo 5. Genellenebilirlik Kuramına göre, üç farklı puanlama anahtarı kullanılarak ve 3 puanlayıcıdan oluşan 100'er örneklemin puanlama güvenilirlikleri

Kategori Sayısı																			
No	2	3	5	No	2	3	5	No	2	3	5	No	2	3	5	No	2	3	5
1	0,000	0,000	0,000	21	0,040	0,137	0,241	41	0,188	0,449	0,395	61	0,308	0,568	0,516	81	0,446	0,639	0,698
2	0,000	0,000	0,000	22	0,052	0,148	0,272	42	0,197	0,452	0,399	62	0,309	0,571	0,525	82	0,451	0,642	0,701
3	0,000	0,000	0,000	23	0,054	0,191	0,286	43	0,202	0,461	0,412	63	0,314	0,576	0,525	83	0,455	0,644	0,710
4	0,000	0,000	0,000	24	0,057	0,216	0,299	44	0,202	0,463	0,432	64	0,320	0,583	0,530	84	0,466	0,647	0,713
5	0,000	0,000	0,000	25	0,059	0,270	0,303	45	0,215	0,469	0,438	65	0,320	0,586	0,531	85	0,500	0,653	0,717
6	0,000	0,000	0,000	26	0,059	0,348	0,325	46	0,217	0,470	0,455	66	0,323	0,593	0,541	86	0,505	0,655	0,720
7	0,000	0,000	0,000	27	0,065	0,350	0,326	47	0,217	0,498	0,457	67	0,327	0,600	0,554	87	0,507	0,664	0,745
8	0,000	0,000	0,000	28	0,071	0,353	0,331	48	0,220	0,498	0,469	68	0,332	0,601	0,565	88	0,507	0,666	0,759
9	0,000	0,000	0,092	29	0,099	0,356	0,336	49	0,226	0,504	0,470	69	0,332	0,609	0,568	89	0,511	0,667	0,764
10	0,000	0,000	0,117	30	0,103	0,360	0,336	50	0,227	0,513	0,472	70	0,333	0,610	0,576	90	0,516	0,669	0,767
11	0,000	0,000	0,142	31	0,114	0,372	0,340	51	0,240	0,516	0,473	71	0,340	0,617	0,582	91	0,530	0,681	0,774
12	0,000	0,000	0,150	32	0,116	0,382	0,341	52	0,247	0,519	0,487	72	0,348	0,618	0,598	92	0,553	0,698	0,776
13	0,000	0,000	0,164	33	0,134	0,383	0,361	53	0,252	0,520	0,492	73	0,349	0,621	0,598	93	0,578	0,708	0,786
14	0,000	0,000	0,164	34	0,136	0,386	0,366	54	0,276	0,525	0,500	74	0,377	0,624	0,613	94	0,629	0,709	0,811
15	0,017	0,010	0,174	35	0,138	0,399	0,373	55	0,293	0,535	0,500	75	0,387	0,624	0,627	95	0,646	0,740	0,836
16	0,019	0,059	0,185	36	0,155	0,418	0,380	56	0,294	0,536	0,502	76	0,393	0,628	0,646	96	0,652	0,745	0,840
17	0,020	0,093	0,200	37	0,166	0,438	0,384	57	0,297	0,537	0,505	77	0,394	0,630	0,649	97	0,654	0,759	0,854
18	0,026	0,096	0,202	38	0,171	0,441	0,389	58	0,300	0,538	0,506	78	0,397	0,632	0,664	98	0,696	0,776	0,887
19	0,032	0,129	0,211	39	0,176	0,444	0,389	59	0,302	0,549	0,506	79	0,413	0,633	0,671	99	0,703	0,779	0,894
20	0,037	0,133	0,241	40	0,184	0,447	0,394	60	0,304	0,560	0,515	80	0,425	0,639	0,684	100	0,869	0,787	0,927

Tablo 6. Genellenebilirlik Kuramına göre, üç farklı puanlama anahtarı kullanılarak ve 5 puanlayıcıdan oluşan 100'er örneklemin puanlama güvenilirlikleri

Kategori Sayısı																			
No	2	3	5	No	2	3	5	No	2	3	5	No	2	3	5	No	2	3	5
1	0,000	0,000	0,000	21	0,638	0,373	0,371	41	0,791	0,557	0,537	61	0,865	0,623	0,646	81	0,902	0,685	0,719
2	0,000	0,076	0,000	22	0,639	0,386	0,398	42	0,800	0,557	0,546	62	0,868	0,627	0,650	82	0,902	0,685	0,720
3	0,084	0,118	0,024	23	0,650	0,396	0,401	43	0,811	0,560	0,546	63	0,871	0,629	0,656	83	0,903	0,685	0,720
4	0,197	0,154	0,095	24	0,667	0,404	0,424	44	0,811	0,562	0,552	64	0,872	0,629	0,657	84	0,907	0,691	0,727
5	0,343	0,159	0,159	25	0,703	0,408	0,427	45	0,817	0,569	0,556	65	0,872	0,637	0,659	85	0,907	0,693	0,740
6	0,387	0,178	0,179	26	0,720	0,427	0,436	46	0,825	0,571	0,562	66	0,874	0,637	0,660	86	0,908	0,700	0,745
7	0,434	0,192	0,211	27	0,723	0,441	0,458	47	0,825	0,577	0,568	67	0,874	0,639	0,671	87	0,908	0,705	0,753
8	0,446	0,199	0,215	28	0,727	0,446	0,469	48	0,826	0,580	0,581	68	0,878	0,640	0,673	88	0,908	0,709	0,756
9	0,451	0,201	0,217	29	0,731	0,465	0,470	49	0,826	0,582	0,592	69	0,878	0,644	0,677	89	0,911	0,715	0,761
10	0,491	0,202	0,220	30	0,735	0,470	0,483	50	0,828	0,585	0,603	70	0,888	0,645	0,678	90	0,913	0,720	0,764
11	0,513	0,213	0,236	31	0,738	0,483	0,492	51	0,832	0,592	0,607	71	0,888	0,655	0,680	91	0,917	0,731	0,773
12	0,520	0,218	0,247	32	0,740	0,504	0,505	52	0,833	0,596	0,608	72	0,888	0,660	0,681	92	0,918	0,744	0,773
13	0,559	0,227	0,257	33	0,747	0,509	0,515	53	0,834	0,606	0,617	73	0,891	0,660	0,681	93	0,919	0,745	0,780
14	0,563	0,241	0,267	34	0,747	0,512	0,520	54	0,838	0,609	0,626	74	0,892	0,661	0,685	94	0,919	0,756	0,782
15	0,577	0,308	0,269	35	0,752	0,523	0,524	55	0,853	0,611	0,638	75	0,894	0,662	0,693	95	0,925	0,765	0,782
16	0,580	0,310	0,290	36	0,752	0,524	0,525	56	0,855	0,614	0,639	76	0,895	0,669	0,712	96	0,930	0,766	0,801
17	0,606	0,314	0,332	37	0,760	0,526	0,527	57	0,858	0,618	0,639	77	0,895	0,671	0,714	97	0,932	0,768	0,805
18	0,606	0,353	0,333	38	0,761	0,532	0,531	58	0,860	0,619	0,643	78	0,897	0,672	0,715	98	0,938	0,793	0,820
19	0,609	0,363	0,359	39	0,773	0,537	0,533	59	0,863	0,622	0,643	79	0,899	0,679	0,716	99	0,940	0,793	0,860
20	0,612	0,372	0,366	40	0,780	0,546	0,536	60	0,864	0,623	0,645	80	0,901	0,681	0,719	100	0,960	0,804	0,884

Tablo 7. Genellenebilirlik Kuramına göre, üç farklı puanlama anahtarı kullanılarak ve 10 puanlayıcıdan oluşan 100'er örneklemin puanlama güvenilirlikleri

																			Kategori Sayısı			
No	2	3	5	No	2	3	5	No	2	3	5	No	2	3	5	No	2	3	5			
1	0,041	0,214	0,278	21	0,306	0,445	0,561	41	0,341	0,558	0,650	61	0,383	0,617	0,677	81	0,461	0,684	0,755			
2	0,118	0,261	0,346	22	0,307	0,445	0,570	42	0,342	0,562	0,651	62	0,385	0,622	0,677	82	0,462	0,685	0,757			
3	0,132	0,279	0,369	23	0,312	0,448	0,572	43	0,344	0,572	0,651	63	0,396	0,624	0,685	83	0,468	0,685	0,760			
4	0,161	0,282	0,412	24	0,313	0,451	0,575	44	0,349	0,577	0,651	64	0,399	0,625	0,698	84	0,469	0,686	0,761			
5	0,170	0,309	0,433	25	0,313	0,458	0,576	45	0,355	0,578	0,652	65	0,412	0,627	0,698	85	0,477	0,690	0,762			
6	0,223	0,327	0,444	26	0,313	0,460	0,588	46	0,356	0,580	0,653	66	0,415	0,627	0,704	86	0,478	0,690	0,763			
7	0,233	0,338	0,449	27	0,314	0,472	0,598	47	0,359	0,582	0,657	67	0,416	0,634	0,707	87	0,489	0,693	0,767			
8	0,235	0,357	0,465	28	0,317	0,489	0,602	48	0,359	0,587	0,658	68	0,418	0,635	0,708	88	0,491	0,697	0,768			
9	0,236	0,372	0,487	29	0,321	0,489	0,610	49	0,361	0,588	0,659	69	0,422	0,640	0,715	89	0,493	0,697	0,774			
10	0,245	0,375	0,490	30	0,326	0,502	0,621	50	0,361	0,592	0,660	70	0,427	0,642	0,723	90	0,499	0,697	0,779			
11	0,250	0,381	0,497	31	0,328	0,503	0,621	51	0,364	0,597	0,662	71	0,427	0,644	0,728	91	0,500	0,701	0,791			
12	0,254	0,398	0,505	32	0,329	0,528	0,627	52	0,365	0,600	0,663	72	0,436	0,648	0,729	92	0,506	0,705	0,792			
13	0,262	0,399	0,529	33	0,330	0,532	0,628	53	0,365	0,601	0,664	73	0,436	0,664	0,736	93	0,513	0,707	0,792			
14	0,263	0,411	0,534	34	0,331	0,540	0,639	54	0,367	0,602	0,665	74	0,442	0,664	0,737	94	0,530	0,714	0,793			
15	0,283	0,415	0,539	35	0,332	0,541	0,640	55	0,376	0,603	0,666	75	0,444	0,672	0,739	95	0,586	0,715	0,800			
16	0,285	0,415	0,540	36	0,332	0,542	0,640	56	0,376	0,604	0,667	76	0,447	0,676	0,743	96	0,600	0,716	0,831			
17	0,285	0,428	0,542	37	0,335	0,548	0,642	57	0,377	0,610	0,667	77	0,448	0,676	0,744	97	0,606	0,732	0,839			
18	0,298	0,428	0,545	38	0,336	0,550	0,642	58	0,377	0,610	0,672	78	0,449	0,679	0,744	98	0,610	0,733	0,851			
19	0,301	0,433	0,551	39	0,337	0,551	0,649	59	0,382	0,610	0,673	79	0,452	0,681	0,753	99	0,617	0,735	0,856			
20	0,305	0,439	0,561	40	0,338	0,556	0,650	60	0,382	0,612	0,676	80	0,455	0,682	0,755	100	0,630	0,756	0,865			

iv) 10 Puanlayıcı 100'er örneklem

Tablo 7, incelendiğinde, 10 puanlayıcı örneklemde kontrol listesi kullanılarak puanlama yapıldığında (2 kategorili) güvenilirlik kestirimleri 0,041 ile 0,630 arasında; analitik rubrik kullanılarak puanlama yapıldığında (3 kategorili) güvenilirlik kestirimlerinin 0,214 ile 0,756 arasında; dereceleme ölçeği kullanılarak puanlama yapıldığında (5 kategorili) güvenilirlik kestirimlerinin 0,278 ile 0,865 arasında değiştiği gözlenmektedir.

3.2. “Genellenebilirlik Kuramı’na göre, aynı performans görevlerinin; kontrol listesi, dereceleme ölçeği, analitik rubrik kullanılarak ve 2, 3, 5, 10 puanlayıcıdan oluşan 100'er örneklemde puanlama güvenilirliklerinin ortancaları ve standart hataları nasıldır?” araştırma sorusuna ilişkin bulgular

Yapılan çalışmalar sonucunda, Genellenebilirlik Kuramı’na göre, aynı performans görevlerinin; kontrol listesi, dereceleme ölçeği, analitik rubrik kullanılarak ve 2, 3, 5, 10 puanlayıcıdan oluşan 100'er örneklemde puanlama güvenilirliklerinin ortanca değerleri ve standart hataları elde edilmiştir. Bu değerler Tablo 8’de görülmektedir.

Tablo 8. Genellenebilirlik Kuramı’na ilişkin güvenilirliklerinin ortancaları ve standart hataları

Puanlayıcı sayısı		2 Kategorili	3 Kategorili	5 Kategorili
2	Ortanca	0,309	0,439	0,474
	SH	0,024	0,024	0,030
3	Ortanca	0,234	0,515	0,473
	SH	0,020	0,024	0,024
5	Ortanca	0,830	0,589	0,605
	SH	0,020	0,019	0,020
10	Ortanca	0,363	0,595	0,661
	SH	0,011	0,013	0,012

Tablo 8, 2 puanlayıcı örnekler açısından incelendiğinde, kontrol listesi kullanılarak puanlama yapıldığında (2 kategorili) Genellebilirlik Kuramı güvenilirlik kestirimlerinin ortancalarının 0,309 olduğu, standart hataların 0,024 olduğu; analitik rubrik kullanılarak puanlama yapıldığında (3 kategorili) Genellebilirlik Kuramı güvenilirlik kestirimlerinin ortancalarının 0,439 olduğu, standart hataların 0,024 olduğu; dereceleme ölçeği kullanılarak puanlama yapıldığında (5 kategorili) Genellebilirlik Kuramı güvenilirlik kestirimlerinin ortancalarının 0,474 olduğu, standart hataların 0,030 olduğu gözlenmektedir.

Tablo 8, 3 puanlayıcı örnekler açısından incelendiğinde, kontrol listesi kullanılarak puanlama yapıldığında (2 kategorili) Genellebilirlik Kuramı güvenilirlik kestirimlerinin ortancalarının 0,234 olduğu, standart hataların 0,020 olduğu; analitik rubrik kullanılarak puanlama yapıldığında (3 kategorili) Genellebilirlik Kuramı güvenilirlik kestirimlerinin ortancalarının 0,515 olduğu, standart hataların 0,024 olduğu; dereceleme ölçeği kullanılarak puanlama yapıldığında (5 kategorili) Genellebilirlik Kuramı güvenilirlik kestirimlerinin ortancalarının 0,473 olduğu, standart hataların 0,024 olduğu gözlenmektedir.

Tablo 8, 5 puanlayıcı örnekler açısından incelendiğinde, kontrol listesi kullanılarak puanlama yapıldığında (2 kategorili) Genellebilirlik Kuramı güvenilirlik kestirimlerinin ortancalarının 0,830 olduğu, standart hataların 0,020 olduğu; analitik rubrik kullanılarak puanlama yapıldığında (3 kategorili) Genellebilirlik Kuramı güvenilirlik kestirimlerinin ortancalarının 0,589 olduğu, standart hataların 0,019 olduğu; dereceleme ölçeği kullanılarak puanlama yapıldığında (5 kategorili) Genellebilirlik Kuramı güvenilirlik kestirimlerinin ortancalarının 0,605 olduğu, standart hataların 0,020 olduğu gözlenmektedir.

Tablo 8, 10 puanlayıcı örnekler açısından incelendiğinde, kontrol listesi kullanılarak puanlama yapıldığında (2 kategorili) Genellebilirlik Kuramı güvenilirlik kestirimlerinin ortancalarının 0,363 olduğu, standart hataların 0,011 olduğu; analitik rubrik kullanılarak puanlama yapıldığında (3 kategorili) Genellebilirlik Kuramı güvenilirlik kestirimlerinin ortancalarının 0,595 olduğu, standart hataların 0,013 olduğu; dereceleme ölçeği kullanılarak puanlama yapıldığında (5 kategorili) Genellebilirlik Kuramı güvenilirlik kestirimlerinin ortancalarının 0,661 olduğu, standart hataların 0,012 olduğu gözlenmektedir.

BÖLÜM IV

TARTIŞMA VE YORUM

Bu araştırmanın amacı, aynı performans görevinin farklı sayıda puanlayıcılar tarafından üç farklı teknikle puanlanmasından elde edilen puanların güvenilirliklerinin Genellenebilirlik Kuramı'na göre incelenmesidir.

Araştırmanın Genellenebilirlik Kuramı'na dayalı olarak gerçekleştirilen puanlayıcılar arası güvenilirlik çalışmaları birlikte değerlendirildiğinde, iki puanlayıcı ve 2 kategorili puanlama sözkonusu olduğunda elde edilen kestirimlerin 0,000 ile 0,800 arasında değiştiği; buna karşılık iki puanlayıcı ve 3 ve 5 kategorili puanlama ya da ikiden fazla puanlayıcı ve 2, 3 ve 5 kategorili puanlama sözkonusu olduğunda elde edilen kestirimlerin 0,000 ile 0,960 arasında değiştiği; en yüksek uyumun beş puanlayıcı ve iki kategorili puanlama durumunda ($G=0,960$) elde edildiği belirlenmiştir.

Araştırmanın Genellenebilirlik Kuramı'na dayalı olarak gerçekleştirilen puanlayıcılar arası güvenilirlik çalışmaları birlikte değerlendirildiğinde, güvenilirliklerin ortancalarının 0,234 ile 0,830 arasında değerler aldığı; güvenilirliklerin standart hata değerlerinin de 0,011 ile 0,030 arasında değiştiği belirlenmiştir.

Puanlayıcılar arası güvenilirliklerin ortancasının en yüksek olduğu durum Genellenebilirlik Kuramı'na göre beş puanlayıcı ve iki kategorili puanlamanın olduğu durum (0,830) olarak karşımıza çıkmaktadır.

Genellenebilirlik Kuramı'na göre elde edilen kestirimlerin iki kategorili puanlama dışında, puanlayıcı ve kategori sayısı arttıkça güvenirliliğin arttığı gözlenmiştir. Alanyazında, yüzeylerden biri ya da her ikisindeki koşul sayısının artırılması paydalardaki değerlerin azalmasına sebep olacağından güvenilirlik katsayıları artmasına sebep olacağı ifade edilmektedir (Brennan, 2001; Güler, Uyanık ve Teker, 2012; Büyükkıdık, 2012) Dolayısıyla araştırmada elde edilen bu bulguların alanyazın ile paralellik gösterdiği görülmektedir.

Güvenirliliklerin standart hataları incelendiğinde Genellenebilirlik Kuramı'na göre en düşük standart hata değerinin (0,011) on puanlayıcı ve bu sefer iki kategorili puanlama sözkonusu olduğunda elde edildiği gözlenmiştir.

Genellenebilirlik Kuramı'nda puanlayıcı sayısı arttıkça standart hata değerlerinin düştüğü gözlenmiştir. Özbek ve Keskin (2007), standart hata ne kadar küçük olursa popülasyona ait tahminlerin o kadar isabetli olacağını ve örneklem sayısı arttıkça standart hata değerlerinin azalacağını ifade etmişlerdir. Dolayısıyla elde edilen bulguların, alanyazındaki kaynaklarla paralellik gösterdiği söylenebilir.

Elde edilen bulgular, puanlayıcı sayısı açısından incelendiğinde, iki puanlayıcı için Genellenebilirlik Kuramı'nda kategori sayısı arttıkça güvenilirliklerin ortancasının yükseldiği ve beş kategorili puanlamadan elde edilen güvenilirliklerin ortancasının görece daha yüksek olduğu (0,474) belirlenmiştir.

Üç puanlayıcı için Genellenebilirlik Kuramı'nda yine üç kategorili puanlamadan elde edilen güvenilirliklerin ortancasının görece daha yüksek olduğu (0,515) belirlenmiştir.

Beş puanlayıcı için Genellenebilirlik Kuramı'nda iki kategorili puanlamadan elde edilen güvenilirliklerin ortancasının görece daha yüksek olduğu (0,830) ortaya konulmuştur.

On puanlayıcı için Genellenebilirlik Kuramı'nda kategori sayısı arttıkça güvenilirliklerin ortancasının yükseldiği ve beş kategorili puanlamadan elde edilen güvenilirliklerin ortancasının görece daha yüksek olduğu (0,515) gözlenmiştir.

Elde edilen bulgular, kategori sayısı (kontrol listesi, analitik rubrik, dereceleme ölçeği) açısından incelendiğinde, Genellenebilirlik Kuramı'nda beş puanlayıcı olması durumunda elde edilen güvenilirliklerin ortancasının görece daha yüksek olduğu (0,830) belirlenmiştir.

Üç kategorili puanlamalar için Genellenebilirlik Kuramı'nda puanlayıcı sayısı arttıkça güvenilirliklerin ortancasının yükseldiği ve yine on puanlayıcı olması durumunda elde edilen güvenilirliklerin ortancasının görece daha yüksek olduğu (0,595) belirlenmiştir.

Üç kategorili puanlamalar için Genellenebilirlik Kuramı'nda puanlayıcı sayısı arttıkça güvenilirliklerin ortancasının yükseldiği ve yine on puanlayıcı olması durumunda elde edilen güvenilirliklerin ortancasının görece daha yüksek olduğu (0,661) belirlenmiştir.

Puanlayıcı sayısı 5 ve kategori sayısı 2 olduğunda Genellenebilirlik Kuramı'nda güvenilirlik kestiriminin en yüksek değeri verdiği söylenebilir.

SONUÇ VE ÖNERİLER

Bu çalışmada, Genellenebilirlik Kuramı'na göre, aynı performans görevlerinin; kontrol listesi, dereceleme ölçeği, analitik rubrik kullanılarak ve 2, 3, 5, 10 puanlayıcıdan oluşan 100'er örneklemin puanlama güvenilirlikleri incelenmiş, elde edilen sonuçlar aşağıda özetlenmiştir.

- Genellenebilirlik Kuramı güvenilirlik kestirimlerinin ortanca değerleri incelendiğinde 5 puanlayıcının kontrol listesi kullanarak yaptıkları puanlamaların güvenilirliklerinin ortanca değeri hariç olmak üzere puanlayıcı sayısı ve aynı zamanda kullanılan ölçeğin kategori sayısı arttıkça ortanca değerlerinin de arttığı gözlenmiştir.
- Genellenebilirlik Kuramı'ndan elde edilen standart hataların puanlayıcı sayısı arttıkça azaldığı gözlenmiştir. En düşük standart hata değerlerinin 10 puanlayıcı olması durumunda elde edildiği saptanmıştır.
- Puanlayıcı sayısı 5 ve kategori sayısı 2 olduğunda Genellenebilirlik Kuramı'nda güvenilirlik kestiriminin en yüksek değeri verdiği söylenebilir.

Yapılan bu çalışmaya benzer çalışmalar yapacak olan araştırmacılar için aşağıda bazı önerilerde bulunulmuştur.

- Bu çalışmada elde edilen sonuçlara bağlı olarak araştırmacılar puanlayıcılar arası güvenilirlik çalışmalarında genellenebilirlik kuramına dayalı olarak iki kategorili ve beş puanlayıcının sözkonusu olduğu çalışmalar planlayabilirler.
- Bu araştırma, 2, 3, 5 ve 10 puanlayıcılı örneklemlerle gerçekleştirilmiştir. Benzer çalışmalar yapacak olan araştırmacılar, farklı değerlendirici sayıları kullanarak araştırmalarını düzenleyebilirler.

- Arařtırmada kullanılan farklı puanlama anahtarlarının madde sayıları sabittir fakat farklı madde sayıları kullanarak da gvenirlik katsayıları incelenebilir.
- Bu arařtırmada kullanılan ç farklı puanlama anahtarı dıřında farklı sayıda ve trde lme araları kullanılabilir.
- Bu arařtırma kapsamında gvenirliđin deđerlendirilmesinde, G kuramının ok deđiřkenli tmyle arpazlanmış $g \times pa \times p$ deseni ile analizi yapılmıřtır. Bu desen kullanılarak yapılan alıřmalarda i geerlik sorunu olabileceđi dřncesiyle benzer alıřmaların yuvalanmıř desen kullanılarak incelenmesi alana olduka yararlı bilgiler sađlayacaktır.

KAYNAKLAR

- Aiken, L. R. (2000). *Psychological testing and assessment (10th ed)*. USA: Allyn and Bacon.
- Alharby, Eid 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*. The Pennsylvania State University, Educational psychology, Doktora tezi.
- Anastasi, A. ve Urbina, S. (1997). *Psychological testing (7th ed.)*. USA: Macmillan Pub. Co. Inc.
- Arter, J. (2002). *Rubrics, Scoring Guides, and Performance Criteria*. Understanding Scoring Rubrics a Guide for Teachers (ed. Carol Boston). Washington: Office of Educational Research and Improvement.
- Atılğan, H. (2005). *Genellenebilirlik kuramı ve puanlayıcılar arası güvenilirlik için örnek bir uygulama*. Eğitim Bilimleri ve Uygulama, 4 (7), 95-108.
- Atılğan, H. (2004). *Genellenebilirlik kuramı ve çok değişkenlik kaynaklı Rasch modelinin karşılaştırılmasına ilişkin bir araştırma*. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı. Yayımlanmış Doktora Tezi.
- Atmaz, G. (2009). *Puanlama yönergesi (rubrik) kullanılması durumunda puanlayıcı güvenilirliğinin incelenmesi*. Mersin Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı. Yayımlanmış Yüksek Lisans Tezi.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik Test Teorisi ve Uygulaması*. Ankara: ÖSYM Yayınları.
- Brennan, R. L. (2001). *Generalizability theory*. USA: Springer-Verlag New York Inc.
- Büyükkıdık, S. (2012). *Problem çözme becerisinin değerlendirilmesinde puanlayıcılar arası güvenilirliğin klasik test kuramı ve genellenebilirlik kuramına göre karşılaştırılması*.

Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı. Yayınlanmış Yüksek Lisans Tezi.

Crocker, L. ve Algina, J. (1986). *Introduction to classical and modern test theory*. USA: CBS College Publishing.

Derkuş, E. (2009). *Puanlayıcılar arasındaki uzlaşmanın farklı tekniklerle incelenmesi*. Mersin Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı. Yayınlanmış Yüksek Lisans Tezi.

Deliceoğlu, G. (2009). *Futbol yetilerine ilişkin dereceleme ölçeğinin genellenebilirlik ve klasik test kuramına dayalı güvenilirliklerinin karşılaştırılması*. Ankara Üniversitesi Eğitim Bilimleri Enstitüsü Eğitimde Psikolojik Hizmetler Anabilim Dalı Ölçme ve Değerlendirme Bilim Dalı. Yayınlanmış Doktora Tezi.

Durand, M.; Loisel, P.; Poitras, S.; Mercier, R.; Stock, S. R.; Lemaire, J. (2004). *The interrater reliability of a functional capacity evaluation: The physical work performance evaluation. journal of occupational rehabilitation. Volume 14/2 119-129.*

Erkuş, A. (2006). *Sınıf öğretmenleri için ölçme ve değerlendirme kavramlar ve uygulamalar*. Ankara: Ekinoks Yayınları.

Erkuş, A. (2003). *Psikometri üzerine yazılar (1. Baskı)*. Ankara: Türk Psikologlar Derneği Yayınları.

Goodrich, A. H. (2005). Teaching with rubrics: the good, the bad and the ugly. *Current Issues in Education*, 4 (4).

Goodwin, L. D. (2001). Interrater Agreement And Reliability, *Measurement in Physical Education and Exercises Science*, 5(1), 13-14

Güler, N. (2008). *Klasik test kuramı genellenebilirlik kuramı ve Rash modeli üzerine bir araştırma*. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı. Yayınlanmış Doktora Tezi.

Güler, N. (2011). *Rasgele veriler üzerinde genellenebilirlik kuramı ve klasik test kuramı'na göre güvenirliliğin karşılaştırılması*. Eğitim ve Bilim, 36 (162).

- Güler, N. Uyanık, G. K. Taşdelen Teker, G. (2012). Genellenebilirlik kuramı. Ankara: Pegem Akademi.
- Hobart, C. ve Frankel, J. (1999) *A Practical Guide to Child Observation and Assessment*. Cheltenham: Stanley Thornes.
- Keeves, J. P. (1988). *Educational research, methodology, and measurement: an international handbook*. USA: Pergamon Press.
- Kutlu, Ö.; Doğan D.; ve Karakaya İ. (2009). *Öğrenci başarısının belirlenmesi performansa ve portfolyoya dayalı durum belirleme*, Ankara: Pegem Yayıncılık.
- Lane, S.; Sabers, D. (1989). Use of generalizability theory for estimating the dependability of a scoring system for sample essays. *Applied Measurement In Education*, 2(3),195-205.
- Lee, G ve Frisbie, D. A. (1999). Estimating reliability under a generalizability theory model for test scores composed of testlets. *Applied Measurement In Education*. 12(3), 237–255.
- Lee, Y. W. ve Kantor, R. (2007). Evaluating prototype tasks and alternative rating schemes for a new esl writing test through G theory. *International Journal of Testing*, 7(4), 353-385.
- Lord, M. F. ve Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*. New Jersey: Addison-Wesley Pub. Co. Inc.
- MEB (2009). Proje ve performans görevleri konulu genelge (No:2009/37). http://mevzuat.meb.gov.tr/html/7273_37.html web adresinden 05 Haziran 2013 tarihinde edinilmiştir.
- Moskal, B. M. ve Leydens, J. A. (2000). Scoring Rubric Development: Validity and Reliability. *Practical Assessment, Research & Evaluation*. <http://PAREonline.net/getvn.asp?v=7&n=10> web adresinden 24 Şubat 2012 tarihinde edinilmiştir.

- Nalbantođlu, F. (2009). *Performans ölçümlerinde genellenabilirlik kuramıyla farklı desenlerin karşılaştırılması*. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı. Yayınlanmış Yüksek Lisans Tezi.
- Ömür, S. (2009). *Dereceli puanlama anahtarıyla, genel izlenimle ve ikili karşılaştırmalar yöntemiyle yapılan değerlendirmelerin karşılaştırılması*. Mersin Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Ana Bilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı. Yayınlanmış Yüksek Lisans Tezi.
- Özbek, H. ve Keskin, S. (2007). *Standart sapma mı yoksa standart hata mı?* Van Tıp Dergisi: 14 (2):64-67
- Popham, J. W. (1997). *What's wrong and what's right with rubric*. Educational Leadership, 55, (2), 12.
- Tekindal, S (Editör). (2008). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem Akademi.
- Tobar, D. A., Stegner, A. J. ve Kane, M. T. (1999). The use of generalizability theory in examining the dependability of scores on the profil of mood state. *Measurement in Physical Education And Exercise Science*. 3 (3), 141-156.
- Turgut, M. F. (1997). *Eğitimde ölçme ve değerlendirme metotları (10. baskı)*. Ankara: Yargıcı Matbaası.
- Turgut, M. F. ve Baykul, Y. (2010). *Eğitimde ölçme ve değerlendirme (2. baskı)*. Ankara: Pegem Akademi.
- Yelbođa, A. (2007). *Klasik test ve genellenebilirlik kuramına göre güvenilirliđin bir iş performansı ölçeđi üzerinde incelenmesi*. Ankara Üniversitesi Eğitim Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı. Yayınlanmış Doktora Tezi.

EKLER

EK 1. 64 KAZANIMIN YER ALDIĞI KONTROL LİSTESİ

Sayın Öğretmenim;

Aşağıdaki sayfalarda İlköğretim 5. Sınıf Türkçe Dersine ait kazanımların bir kısmı yer almaktadır. Sizlerden istenen, bu kazanımların, öğrencilere bir performans görevi hazırlatmak için uygun olup olmadığını seçmenizdir. Eğer ilgili kazanımlarla alakalı bir performans görevi hazırlanabileceğini düşünüyorsanız “Uygun” kutucuğunu, düşünmüyorsanız “Uygun Değil” kutucuğunu işaretleyiniz. İlginiz ve desteğiniz için teşekkürler.

Sıra No	Kazanımlar	Uygun	Uygun Değil
1	Yazmak için hazırlık yapar.		
2	Yazma amacını belirler.		
3	Kelimeler, cümleler ve satırlar arasında uygun boşluklar bırakır.		
4	Anlamli ve kurallı cümleler yazar.		
5	Yazılarında noktalama işaretlerini doğru ve yerinde kullanır.		
6	Yazılarında imlâ kurallarını uygular.		
7	Sayfa düzenine ve temizliğine dikkat eder.		
8	Yazılarında sözlük ve imlâ kılavuzundan yararlanır.		
9	Sıralamalarda sayı, sembol ve işaretlerden yararlanır.		
10	Bütün yazılarını bitişik eğik yazı ile yazar.		
11	Konu dışına çıkmadan yazar.		
12	Yazılarında kelime tekrarıdan kaçınır.		
13	Yazma yöntemini belirler		
14	Kelimeleri yerinde ve anlamlarına uygun kullanır.		
15	Yeni öğrendiği kelimeleri yazılarında kullanır.		
16	Yazısına uygun başlık belirler.		
17	Yazısında alt başlıkları belirler.		
18	Yazısının giriş ve sonuç cümlelerini özenle seçer.		
19	Yazılarında söz varlığından yararlanır.		
20	Olayları, oluş sırasına göre yazar.		
21	Yazılarında mizahî öğelere yer verir.		
22	Yazılarında kelime türlerini işlevine uygun kullanır.		
23	Yazılarında farklı cümle yapılarına yer verir.		
24	Mantıksal bütünlük içinde yazar.		
25	Yazılarında genelden özele, özelden genele doğru yazar.		
26	Yazılarında ana fikre yer verir.		
27	Yazılarında yardımcı fikirlere yer verir.		
28	Yazılarında farklı düşünmeye yönlendiren ifadeleri kullanır.		
29	Yazılarında kendi yaşantısından ve günlük hayattan örnekler verir.		
30	Yazılarında uygun ifadeleri kullanarak destekleyici ve açıklayıcı örnekler verir.		
31	Yazısını anlam ve biçim açısından değerlendirir.		
32	Yazılarında karşılaştırmalar yapar.		
33	Yazılarında sebep-sonuç ilişkileri kurar.		

34	Yazılarında tanımlamalar yapar.		
35	İmza atar ve anlamını bilir.		
36	İzlenim ve deneyimlerine dayalı yazılar yazar.		
37	Duygu, düşünce ve hayallerini anlatan yazılar yazar.		
38	Kendisini, ailesini ve çevresini tanıtan yazılar yazar.		
39	Özet çıkarır.		
40	Kendisine ilginç ve çekici gelen karakter, olay, yer, an vb. ile ilgili görüş ve düşüncelerini yazar.		
41	Bir fikre katılıp katılmadığını nedenleriyle açıklayan yazılar yazar.		
42	Bir etkinliğin, işin vb. aşamalarını anlatan yönergeler hazırlar.		
43	Sorular yazar.		
44	Formları yönergelerine uygun doldurur.		
45	Dilek, istek ve şikâyetlerini yazılı olarak ifade eder.		
46	Eksik bırakılan metni yazarak tamamlar.		
47	Yazılarında önem belirten ifadeleri kullanır.		
48	Yazılarında ne, nerede, ne zaman, nasıl, niçin ve kim (5N 1K) öğelerini vurgular.		
49	Günlük, anı vb. yazar.		
50	Davetiye ve tebrik kartı yazar.		
51	Mektup yazar.		
52	Duyuru, afiş vb. yazar.		
53	Açıklayıcı ve bilgilendirici yazılar yazar.		
54	Hikâye yazar.		
55	Yazılarında betimlemeler yapar.		
56	Kısa oyunlar yazar.		
57	Şiir yazar.		
58	İkna edici yazılar yazar.		
59	Sorgulayıcı yazılar yazar.		
60	İş birliği yaparak yazar.		
61	Not alır.		
62	Vurgulamak istediği kelime ya da cümleleri farklı yazar.		
63	Planlı yazma yöntemine uygun yazar.		
64	İsteddiği bir konuda yazılar yazar.		

EK 2. PERFORMANS GÖREVİ ÖRNEKLERİ

- Öğrencilerden ‘elinizde sihirli bir değnek olsaydı dünyada neleri değiştirmek isterdiniz?’ konulu bir yazı yazmaları istenir.
- Öğrencilerden bir hikâye kitabı okumaları ve kitapta kendilerine en ilginç, en güzel gelen bölümü belirleyerek yazmaları istenir.
- Öğrencilere ‘gürültücü çocuk’ başlığı verilir ve bu başlığa uygun bir hikâye yazmaları istenir.
- Öğrencilerden konusu doğa ve çevre olan bir hikâye yazmaları istenir.
- Öğrencilere “üç arkadaşınızla ıssız bir adada yalnız kaldığınızı düşünün” ve bu durumu hikâyeleştirerek yazınız.
- Hayal gücünüzü kullanarak, deniz altında (bir deniz altının içinde) yaşadığınızı düşünerek bir hikâye yazınız.
- Bir resme bakar ve bu resmi temel alarak bir hikâye yazar.
- Arkadaşların ile ormanın derinliklerinde yaşadığını düşünerek bir hikâye yazınız.
- Saray, prenses, prens, palyaço, tavşan nesnelerinin içinde geçtiği bir hikâye yazınız.
- 2030 yılında olduğunu düşünerek sizin ailenizin ve yakınlarınızın içinde bulunduğu bir hikâye yazınız.

EK 3. PERFORMANS GÖREVİ

PERFORMANS GÖREVİ:HİKÂYE YAZIYORUZ

İçerik	Hikâye yazmak
Süre	2 Hafta
Beklenen Performans	<ul style="list-style-type: none"> • Araştırma yapma • Kısa bir hikâye yazma
Değerlendirme	<ul style="list-style-type: none"> • Kontrol Listesi • Dereceleme ölçeği • Analitik rubrik



Sevgili Öğrenciler;

Bu görevde sizlerden beklenmektedir. Bunlardan ilki, 5 adet kısa hikaye okuyup birinin özetini yazmanız, ikincisi ise size verilecek bir resmi dikkate alarak bir hikaye yazmanızdır. Bu görevi gerçekleştirebilmek için yapmanız gerekenler aşağıda açıklanmıştır.

Görevlerinizi tamamlamak için yapmanız gerekenler:

1. 5 tane kısa hikâye bulunuz.
2. Bulduğunuz hikâyeleri okurken olayların geçtiği yer, zaman, olayların kahramanları, anlatılış tarzlarına dikkat ederek okuyunuz.
3. Okuduğunuz 5 hikâyenin adını ve içlerinden bir tanesinin özetini görev süresi bittiğinde bir kâğıda yazarak öğretmeninize teslim ediniz. Bu görevi gerçekleştirmeniz için size verilen süre 15 gündür.
4. Bu göreviniz bittikten sonra sınıfınıza gelecek olan bir araştırmacı size bir tablo gösterecektir, hayal gücünüzü kullanarak bu tabloda gördüğünüz resmi anlatan bir hikâye yazınız.
5. Yazdığınız hikâyeyi belirtilen süre bitiminde araştırmacıya teslim ediniz.

EK 3. PERFORMANS GÖREVİ (Devamı)

EK 4. ÖĞRENCİLERİN HİKÂYE ÖZETİ ÖRNEKLERİ

Altın Tülek

Bir zamanlar bir aveli varmış. Bir gün dağa avlanmaya giderken, kenderle bir bac almış. Aveli bacı görünce, onu evine çağırmış. Sabah bacı avelinin üstüne atlamış ve aveliyi delik desik yapmış. Aveli bulmuşlar ama birinin yaraladığı belli değilmiş. Avelinin karısı tüleği duvara asmış. Avelinin ağız güçlü kuvvetli biri almış ve annesinden tüleği istemiş.

Madun istemeyerek:

- Öleceğimi bilsem sana tüleği vermem, demiş. Çünkü babam bunun için öldü demiş. Sana vermem demiş.

Günün birinde ağlan gıcice tüleği almış ve avela gitmiş!

Armana varınca, yine aynı bac almış:

- Babamın ben yardım ettim, sende yaralayacağım! demiş.

aveli:

Parrı bana yardım eder, demiş ve rızan alarak bacı evmiş.

aveli

aveli esi karısını almayan bacı evdiği için çok sevilmiş. Pastu götürüp Payzer'e göstermiş

Payzer:

- Bunu ne kadar satarsın, demiş.

dede = ne b

- Ne kadar ucrisen ucr satmam, demis.

Hayzer: ucrini cogurmus ve

Eger oglan size bu Postu vermeye altindan boller
mayacağı bir cera ucrin demis. Dede Postu ucr
memis ve ucris

Bir bag yap yedi gunde algunlasan, demis.

Dede bunan amannesine anlatmis annesi:

- Ben sana dememis miydin kirsay alur diye de
dede sehir dusunde geymeye baslamis bir ke:

- Ne oldu neden aglyarsun demis.

dede dardini anlatmis bir:

Uzulum, demis. Harba dogruca Hayzer'e gi
bagin nereden almasini istedigini sor. Ceresinde
hendekler kesdir. Sana bir tarlay al. Isine
bir dal bebek bag. Sana gasteriles bag
yerine git. Ucanur yat yedi gun sana olur, de

Bunları denemis ucrimler yetismis. Yine cera
vermisler Hayzer:

- Bana bildisinden saray yap demis.

Oglan bu emri alinca buren yanına gitmis.
Dede dardini anlatmis. Dede:

- Bunun guclugü yok ki, demis, ve anlatmis

Dayzer'e git bir gemi iste. İçinde yüz bin
şurup bir e badar ayran, en iyi tane mara
bulsun. Gemiyeye bin yala çık. İki dağ arasın
karınca suyu leşba tarala abut Anun yer
nehri yatığına şurup ve ayranı döküler
ve esbi il dilerini bulup vermisen maranga
da onu yapar demis.

Adacı bunları yapmış Dayzer hayretle d
balması.

Dayzer şöyle zar birşey dedi ki bu işi yapama.

Uzerin:

- Şimdi bamsu bayzerin kuru alda gel demis.

Adacı kuru yanına tekrar gider. Dura dardini

Dura:

Anca git Dayzerden büyükle bir gemi iste. Geminin
içinde en iyi tane dükkan yaptır. İler dük
anda ayır bir mal satılın güzel olsun, en iyi
ne dükkanı iste. Bunları satıcı olarak dükkanı be
barsına adam çibacak elinde şahinli şahin al
dükbandan leşeyler sat. Sara yine barsına adam
çibacak ada Pulu kalığını iste en son bir
adam daha gelecek onun güvercinini iste.

Kuşun tüyünü, balığın pulunu, Güvercinin tüyü
al demis ve sonra şehre gel prenses geldi
nde abama bodan ayala sarra Prenses P

Prensesin balundaki güvercin kabusuna haber vermeceğinin tıyını yak güvercin gitmesin. Prenses yere taş atınca gemi katacağın balığın Pulunu yak gemi katmas sonra bu hayat suyu ister güvercinin tıyını yak hayat suyu getirin, demi dedici yine bunları uygulandı ve 30 yaşın sınır almış dedici ile prenses eelenmişler mutlu mesut yaşamışlar.

≈ SON ≈

Masal adı

1- Altın Dolac

2- Sobacı Dolunbaga

3- Yıldız Pals

Üç Yararlı Öğüt

Esbi Cüm 'in küçük bir köyünde, yoksul bir adam varmış. Dilginin şakılla geçirdiği, her güzde de üzülürmüş. Düşünmüş taşınmış uzak ülkelere gidip para kazanmaya karar vermiş. Bir kararla bir sabah ailesiyle vedalaşmış, yola çıkmış. Gide gide büyük bir sahre varmış. Bir işe girerek çalışmaya başlamış. Dürüst ve çalışkan olduğu için sevilmiş, işinde yükselmiş ve çağa kalmadan yeterli kadar para kazanmış. Artık evine dönme zamanı geldiğini düşünmüş. Düşünmüş an parayı dönüştürme hangi tehlikeleri beklediğini bilmediği için korkmuş. O şehirde aklı ve bilgeliliği ile ün salmış biri varmış. Ora gitmiş ve altın almış. Bilge onun sakasını eliyle süzölüyor;

- Bak öğüt, demiş. Sana üç değerli öğüt vereceğim. Bunlardan şaşmadık yoluna devam edersen, evine zorlanmadan ulaşsın. Yok eğer birisi bile göz ardı edersen parandan da düşersün, canından da.

Ve sıralanmış üç öğütü:

- Birincisi, insanların baldırlık olarak gördüğü yerden uzak duracaksınız. İkincisi, büyük ağaç, küçük ağaçtan daha değerlidir, deniyacaksınız. Üçüncüsü de, "Keyle kalın zararla oturur." sözünü kimden almayacaksınız.

Adam bu üç öğütü sıradan görmüş ama gürede bilgeye tokluk bir edip ayırmış. Biriktirdiği parayı kaynaşa sokup yoluna devam etmiş. Bir süre gittikten sonra, bir uyarım benzerinde aşağıya bakınca orada insan süzünü görmüş. Aralarında bir şeyler konuşmuşlar. Belki baktıkları yerde garip bir şey almışmış. Meraka kapılarak gitmeyi düşünmüş ama birer yapı bilgenin birinci öğütünü hatırlayarak yola devam etmiş. Sana bir bakmış insanların küçüklüklerini içine düşünmüş. Adam çok şaşırmış ve yoluma devam etmiş. Bir süre gittikten sonra bakıldıkta beşarıncaşına yakın yağmaya başlamış ve adam büyük ağaçın dibine gitmeyi düşünmüş ama birer bilgenin ikinci öğütünü hatırlanmış ve küçük ağaçın dibine gitmiş.

Her taraf sülkeyoz gilei acılı veermiş ve yağmur fınmış. Sarra le
 simsek çarbmış ve a leiyük ağaç ortadan ikiye ayrılark yarmaya ke
 miş. Adam yine şaşırms ve yoluna devam etmiş. Bir süre sarra akine
 vemiş ve eee girmede öise karısını leaska leir adanla görmüş ve şir
 lenmiş. Sarra leirerin üçüncü eğidini hatırlamış ve eee salınca
 görmüş ve karısı leayına şarkmış. Meğensem a adam bendi öör leal
 şymış. Adam, baneına bir leirerin üç eğidini anlatmış ve leundan
 sarra malkı mesut yosmuşlar.

EK 5. HİKÂYELERİN PUANLANDIĞI 5 KATEGORİLİ PUANLANMA ANAHTARI

ÖRNEĞİ

	Çok iyi	İyi	Orta	Kötü	Çok kötü
1 No'lu Hikâye					
2 No'lu Hikâye					
3 No'lu Hikâye					
4 No'lu Hikâye					
5 No'lu Hikâye					
6 No'lu Hikâye					
7 No'lu Hikâye					
8 No'lu Hikâye					
9 No'lu Hikâye					
10 No'lu Hikâye					
•					
•					
•					
62 No'lu Hikâye					
63 No'lu Hikâye					
64 No'lu Hikâye					

EK 6. HİKÂYELERİN PUANLANDIĞI 3 KATEGORİLİ PUANLANMA ANAHTARI

ÖRNEĞİ

	İyi	Orta	Kötü
1 No'lu Hikâye			
2 No'lu Hikâye			
3 No'lu Hikâye			
4 No'lu Hikâye			
5 No'lu Hikâye			
6 No'lu Hikâye			
7 No'lu Hikâye			
8 No'lu Hikâye			
9 No'lu Hikâye			
10 No'lu Hikâye			
•			
•			
•			
62 No'lu Hikâye			
63 No'lu Hikâye			
64 No'lu Hikâye			

EK 7. HİKAYE YAZMA KAZANIMINI ÖLÇMEK İÇİN KULLANILACAK DENEMELİK ÖLÇÜTLER

1. Yazım kurallarına uygun mu?
2. Noktalama işaretleri yerinde ve doğru kullanılmış mı?
3. Cümle kuruluşu doğru mu?
4. Kullanılan materyali yorumlayabilmiş mi?
5. Ana fikir var mı?
6. Serim, düğüm ve çözüm bölümleri var mı, uygun mu?
7. Kenar boşlukları var mı, uygun mu?
8. Başlık var mı, uyumlu mu?
9. Kullandığı dil açısından zaman uyumu var mı?
10. Bitişik eğik yazı kullanmış mı?
11. Bütünlük var mı?
12. Kahramanlar oturmuş mu?
13. Olay belli bir yerde geçiyor mu, mekân uygun mu?
14. Mevsimi, zamanı aktarabilmiş mi?
15. Zaman sıralaması yapabilmiş mi?
16. Kelimeleri yerinde kullanabilmiş mi?
17. Akıcı bir dil kullanabilmiş mi?
18. Yazı okunaklı mı?
19. Konu anlaşılır mı?
20. Cümleler arası bağlantı sağlanmış mı?
21. Örneklendirme yapabilmiş mi?
22. Zaman, kişi, olay örgüsü var mı?
23. Kelimeleri doğru yazabilmiş mi?
24. Tamlama, deyim gibi kalıpları doğru kullanmış mı?
25. Serim, düğüm ve çözüm bölümleri kendi içinde tutarlı mı?
26. Gerekli yerlerde betimlemeler yapmış mı?

EK 8. ÖLÇÜTLERİN AĞIRLIKLANDIRILMASI İÇİN KULLANILAN TABLO

Sevgili Öğretmenim;

Aşağıda “Hikâye Yazmaya İlişkin Dereceli Puanlama Anahtarı” çalışması yer almaktadır. Sizden aşağıdaki tabloyu inceleyerek sorulan soruları cevaplamanız istenmektedir. Şimdiden verdiğiniz değerli destekleriniz için teşekkürler.

1. Ana ölçütlerin bu şekilde düzenlenmesi sizce uygun mu? () Uygun () Uygun Değil	2. Sizce yazılan hikâyeleri değerlendirmede biçim ve içerik ana başlıklarının ağırlıklandırılması nasıl yapılmalıdır? (%.....)Biçim (%.....)İçerik Lütfen her iki durum için de önerileriniz varsa belirtiniz.	Ana Ölçütler	Alt Ölçütler	Davranış Göstergeleri	Göstergelerin her birinin alt ölçütlere uygunluğunu değerlendiriniz.				
					4	3	2	1	0
I. BİÇİM	1. Alt ölçütlerin bu şekilde düzenlenmesi sizce uygun mu? () Uygun () Uygun Değil 2. Belirlenen alt ölçütlerin ağırlıklandırılması sizce nasıl olmalıdır? (%.....)Dilbilgisi ve imla (%.....)Sayfa düzeni Lütfen her iki durum için de önerileriniz varsa belirtiniz.....	A)DİL BİLGİSİ VE İMLA	B)SAYFA DÜZENİ	A1)Tamlama, deyim gibi kalıplar yerinde ve doğru kullanılmıştır.					
				A2)Bütün kelimeler doğru yazılmıştır.					
II. İÇERİK	1. Alt ölçütlerin bu şekilde düzenlenmesi sizce uygun mu? () Uygun () Uygun Değil 2. Belirlenen alt ölçütlerin ağırlıklandırılması sizce nasıl olmalıdır? (%.....)Anlatım (%.....)Düzenleme Lütfen her iki durum için de önerileriniz varsa belirtiniz.....	C)ANLATIM	D)DÜZENLEME	A3)Bütün kelimeler yerinde kullanılmıştır.					
				A4)Noktalama işaretlerinin kullanımı yerinde ve doğrudur.					
				B1)Sayfa kenarlarında boşluk bırakılmıştır.					
				B2)Bitişik eğik yazı kullanılmıştır ve yazımında yanlışlık yoktur.					
				B3)Yazı tamamen okunaklıdır.					
				C1)Cümle kuruluşları tamamen doğrudur.					
				C2)Hikâyeye başlık verilmiştir ve içerikle uyumludur.					
				C3)Kullanılan dil akıcıdır.					
				C4)Hikâyenin konusu açık ve anlaşılırdır.					
				C5)Cümleler arasındaki geçişler uygundur.					
C6)Verilen örnekler yerinde ve yeterlidir.									
C7)Yapılan betimlemeler yerinde ve uygundur.									
D1)Hikâye kullanılan materyale dayalı yazılmıştır.									
D2)Hikâyenin ana fikri açık ve anlaşılırdır.									
D3)Hikâyenin serim, düğüm ve çözüm bölümleri vardır ve uygundur.									
D4)Serim, düğüm ve çözüm bölümleri kendi içinde tutarlıdır.									
D5)Hikâyede anlam bakımından bir bütünlük sağlanmıştır.									
D6)Hikâyenin kahramanları belirgindir.									
D7)Hikâyenin geçtiği yer ve mekân belirgindir.									
D8)Hikâyenin geçtiği zaman ve mevsim belirgindir.									
D9)Hikâyede geçen olayların zaman sıralamasına dikkat edilmiştir.									
D10)Hikâyede zaman, kişi, olay örgüsü kurulmuştur.									

Varsa önerileriniz lütfen ekleyiniz.

EK 9. KONTROL LİSTESİ

Sayın Öğretmenim;

Bu form, İlköğretim 5. sınıf düzeyindeki öğrencilerin hikâye yazma becerilerini değerlendirebilmek amacıyla hazırlanmıştır. Sizden istenen her hikâyeyi okuduktan sonra, formdaki ana ölçüt, alt ölçüt ve alt davranış göstergelerini dikkate alarak hikâyeyi değerlendirmenizdir. Değerlendirmenizi yaparken her davranış göstergesini okuduktan sonra, ilgili hikâyede davranış göstergesi mevcutsa “Evet” mevcut değilse “Hayır” kutucuğunu (X) şeklinde işaretlemeniz gerekmektedir. Araştırmamıza sağladığınız katkı için teşekkür ederim.

Ana Ölçütler	Alt Ölçütler	Alt Ölçütlerin Davranış Göstergeleri	EVET	HAYIR
BİÇİM	DİL BİLGİ Sİ VE İMLA	Tamlama, deyim gibi kalıplar yerinde ve doğru kullanılmıştır.		
		Bütün kelimeler doğru yazılmıştır.		
		Bütün kelimeler yerinde kullanılmıştır.		
		Noktalama işaretlerinin kullanımı yerinde ve doğrudur.		
	SAYFA DÜZE Nİ	Sayfa kenarlarında boşluk bırakılmıştır.		
		Yazı tamamen okunaklıdır.		
İÇERİK	ANLATIM	Cümle kuruluşları tamamen doğrudur.		
		Hikâyeye başlık verilmiştir ve içerikle uyumludur.		
		Kullanılan dil akıcıdır.		
		Hikâyenin konusu açık ve anlaşılırdır.		
		Cümleler arasındaki geçişler uygundur.		
		Verilen örnekler yerinde ve yeterlidir.		
	DÜZENLEME	Yapılan betimlemeler yerinde ve uygundur.		
		Hikâye kullanılan materyale dayalı yazılmıştır.		
		Hikâyenin ana fikri açık ve anlaşılırdır.		
		Hikâyenin serim, düğüm ve çözüm bölümleri vardır ve uygundur.		
		Serim, düğüm ve çözüm bölümleri kendi içinde tutarlıdır.		
		Hikâyede anlam bakımından bir bütünlük sağlanmıştır.		
		Hikâyenin kahramanları belirgindir.		
		Hikâyenin geçtiği yer ve mekân belirgindir.		
		Hikâyenin geçtiği zaman ve mevsim belirgindir.		
		Hikâyede geçen olayların zaman sıralamasına dikkat edilmiştir.		
		Hikâyede zaman, kişi, olay örgüsü kurulmuştur.		

EK 10. DERECELEME ÖLÇEĞİ

Sayın Öğretmenim;

Bu form, İlköğretim 5. sınıf düzeyindeki öğrencilerin hikâye yazma becerilerini değerlendirebilmek amacıyla hazırlanmıştır. Sizden istenen her hikayeyi okuduktan sonra, formdaki ana ölçüt, alt ölçüt ve alt davranış göstergelerini dikkate alarak hikayeyi değerlendirmenizdir. Değerlendirmenizi yaparken her davranış göstergesini okuduktan sonra, her bir davranış göstergesine ilişkin uygun gördüğünüz puanı (0 - 4), ilgili davranış göstergesinin karşısında yer alan kutucuğu (X) şeklinde işaretleyerek vermeniz gerekmektedir. Araştırmamıza sağladığınız katkı için teşekkür ederim.

Ana Ölçütler	Alt Ölçütler	Alt Ölçütlerin Davranış Göstergeleri	Puan				
			4	3	2	1	0
BİÇİM	DİLBİLGİSİ VE İMLA	Tamlama, deyim gibi kalıplar yerinde ve doğru kullanılmıştır.					
		Bütün kelimeler doğru yazılmıştır.					
		Bütün kelimeler yerinde kullanılmıştır.					
		Noktalama işaretlerinin kullanımı yerinde ve doğrudur.					
	SAYFA DÜZENİ	Sayfa kenarlarında boşluk bırakılmıştır.					
		Yazı tamamen okunaklıdır.					
İÇERİK	ANLATIM	Cümle kuruluşları tamamen doğrudur.					
		Hikâyeye başlık verilmiştir ve içerikle uyumludur.					
		Kullanılan dil akıcıdır.					
		Hikâyenin konusu açık ve anlaşılırdır.					
		Cümleler arasındaki geçişler uygundur.					
		Verilen örnekler yerinde ve yeterlidir.					
		Yapılan betimlemeler yerinde ve uygundur.					
	DÜZENLEME	Hikâye kullanılan materyale dayalı yazılmıştır.					
		Hikâyenin ana fikri açık ve anlaşılırdır.					
		Hikâyenin serim, düğüm ve çözüm bölümleri vardır ve uygundur.					
		Serim, düğüm ve çözüm bölümleri kendi içinde tutarlıdır.					
		Hikâyede anlam bakımından bir bütünlük sağlanmıştır.					
		Hikâyenin kahramanları belirgindir.					
		Hikâyenin geçtiği yer ve mekân belirgindir.					
		Hikâyenin geçtiği zaman ve mevsim belirgindir.					
Hikâyede geçen olayların zaman sıralamasına dikkat edilmiştir.							
Hikâyede zaman, kişi, olay örgüsü kurulmuştur.							

EK 11. ANALİTİK RUBRİK

Sayın Öğretmenim;

Bu form, İlköğretim 5. sınıf düzeyindeki öğrencilerin hikâye yazma becerilerini değerlendirebilmek amacıyla hazırlanmıştır. Sizden istenen her hikayeyi okuduktan sonra, formdaki ana ölçüt, alt ölçüt ve alt davranış göstergelerini dikkate alarak hikayeyi değerlendirmenizdir. Değerlendirmenizi yaparken her davranış göstergesini okuduktan sonra, ilgili hikaye için her bir davranış göstergesine ilişkin uygun gördüğünüz puanı (İyi: 2 puan, Orta: 1 puan, Kötü: 0 puan) puan sütununa yazmanız gerekmektedir. Araştırmamıza sağladığımız katkı için teşekkür ederim.

Ana Ölçütler	Alt Ölçütler	Alt Ölçütlerin Davranış Göstergeleri			Puan
		İYİ (2 Puan)	ORTA (1 Puan)	KÖTÜ (0 Puan)	
BİÇİM	DİL BİLGİSİ VE İMLA	Tamlama, deyim gibi kalıplar yerinde ve doğru kullanılmıştır.	Tamlama, deyim gibi kalıplar kullanılmıştır ancak bazılarının kullanımı yanlıştır.	Tamlama, deyim gibi kalıplar yerinde ve doğru kullanılmamıştır.	
		Bütün kelimeler doğru yazılmıştır.	Bazı kelimeler yanlış yazılmıştır.	Kelimelerin birçoğu yanlış yazılmıştır.	
		Bütün kelimeler yerinde kullanılmıştır.	Bazı kelimeler yerinde kullanılmamıştır.	Kelimelerin birçoğunu yerinde kullanılmamıştır.	
	SAYFA DÜZENİ	Noktalama işaretlerinin kullanımı yerinde ve doğrudur.	Noktalama işaretleri kullanılmıştır fakat doğru ve yerinde değildir.	Noktalama işaretleri yerinde ve doğru kullanılmamıştır.	
		Sayfa kenarlarında boşluk bırakılmıştır.	Sayfa kenarlarında boşluk bırakılmıştır ancak uygun değildir.	Sayfa kenarlarında boşluk bırakılmamıştır.	
		Yazı tamamen okunaklıdır.	Yazı kısmen okunaklıdır.	Yazı okunaklı değildir.	
İÇERİK	ANLATIM	Cümle kuruluşları tamamen doğrudur.	Cümle kuruluşları yeterince doğru değildir.	Cümle kuruluşları doğru değildir.	
		Hikâyeye başlık verilmiştir ve içerikle uyumludur.	Hikâyeye başlık verilmiştir fakat içerikle yeterince uyumlu değildir.	Hikâyeye hiç başlık verilmemiştir ya da verilen başlık içerikle hiç uyumlu değildir.	
		Kullanılan dil akıcıdır.	Kullanılan dil yeterince akıcı değildir.	Kullanılan dil akıcı değildir.	
		Hikâyenin konusu açık ve anlaşılırdır.	Hikâyenin konusu kısmen açık ve anlaşılırdır.	Hikâyenin konusu açık ve anlaşılır değildir.	
		Cümleler arasındaki geçişler uygundur.	Cümleler arasındaki geçişler yeterince uygun değildir.	Cümleler arasındaki geçişler uygun değildir.	
		Verilen örnekler yerinde ve yeterlidir.	Verilen örnekler kısmen yerinde ve yeterlidir.	Verilen örnekler yerinde ve yeterli değildir.	
	DÜZENLEME	Yapılan betimlemeler yerinde ve uygundur.	Yapılan betimlemeler yeterince uygun değildir.	Betimleme yapılmamıştır ya da yapılan betimlemeler yerinde ve uygun değildir.	
		Hikâye kullanılan materyale dayalı yazılmıştır.	Hikâye kısmen kullanılan materyale dayalı yazılmıştır.	Hikâye kullanılan materyale dayalı yazılmamıştır.	
		Hikâyenin ana fikri açık ve anlaşılırdır.	Hikâyenin ana fikri vardır ancak yeterince açık ve anlaşılır değildir.	Hikâyenin bir ana fikri yoktur ya da anlaşılmamaktadır.	
		Hikâyenin serim, düğüm ve çözüm bölümleri vardır ve uygundur.	Hikâyenin serim, düğüm ve çözüm bölümleri vardır ancak yeterince uygun değildir.	Hikâyenin serim, düğüm ve çözüm bölümleri yoktur ya da hiç uygun değildir.	
		Serim, düğüm ve çözüm bölümleri kendi içinde tutarlıdır.	Serim, düğüm ve çözüm bölümleri kendi içinde yeterince tutarlı değildir.	Serim, düğüm ve çözüm bölümleri yoktur veya kendi içinde tutarlı değildir.	
		Hikâyede anlam bakımından bir bütünlük sağlanmıştır.	Hikâyede anlam bakımından yeterince bütünlük sağlanamamıştır.	Hikâyede anlam bakımından bütünlük sağlanamamıştır.	
		Hikâyenin kahramanları belirgindir.	Hikâyenin kahramanlarında kısmen belirsizlik vardır.	Hikâyenin kahramanları belirsizdir.	
		Hikâyenin geçtiği yer ve mekân belirgindir.	Hikâyenin geçtiği yer ve mekân yeterince belirgin değildir.	Hikâyenin geçtiği yer ve mekân belirgin değildir.	
		Hikâyenin geçtiği zaman ve mevsim belirgindir.	Hikâyenin geçtiği zaman ve mevsim yeterince belirgin değildir.	Hikâyenin geçtiği zaman ve mevsim belirgin değildir.	
DÜZENLEME	Hikâyede geçen olayların zaman sıralamasına dikkat edilmiştir.	Hikâyede geçen olayların zaman bakımından sıralamasında belirsizlik vardır.	Hikâyede geçen olayların zaman sıralamasına dikkat edilmemiştir.		
	Hikâyede zaman, kişi, olay örgüsü kurulmuştur.	Hikâyede zaman, kişi, olay örgüsü kurulmuştur fakat tutarlı ve açık değildir.	Hikâyede zaman, kişi, olay örgüsü yoktur veya tamamen tutarsızdır.		

EK 12. PUANLAYICILARIN PUANLADIĞI 6 HİKÂYE

Mehmet'in Ailesi

Bir durumu kâfi olan ve nefise kalabalık bir ailesi namına anlatıyor. O ailesi tek gün Kaynakları tavuklarında da yada yolda bir günün bir saatleri yemekte olarak ya da ciirece sekiyle sata lardı. O ailesi de parayı vardı. En büyükleri 5 sınıfa gittiğinde ise dersleri çok iyiydi adı Mehmet öğretmen Mehmet'i çok seviyor. Arasta gür duvarda. Türkçe öğrenmesi kâfi oldu. Bu durumda rasul fakihine merad ediyor. Mehmet 6. sınıfa geçince gelirken iyice kötülükte ve Mehmet'i okulları almak zorunda kalmışlardı. Diğer kâdılar İsmail, Bekir, Dursun, Emin kâfi çok üzüldü bir gün Allah'a dua edip duyarlardı. Allah'ın bir gün ailesine yardım edeceğini biliyorlardı. Bu arada Mehmet kâfide çalışmaya başlamış ananın oğlunu Mehmet'in ailesi Mehmet çok başarılı çok hızlı bir şekilde onu nice okula göndermişlerdir. Metin Bey gâfil çok basarılar kazanıyor. Her şeyden çok iyi bir durumda. İbrahim'i görseldi. Diğer müdürlerini İbrahim kâdiler. O fakih Mehmet'in tükür tükür çalıştığını görünce öyle bir elmanı kazanmışlar. Mehmet'in anasını arkanı ürete çekende ve böyle bir durumda aileleri rahatlamıştı.

Birlik aileden tavukları almıştı. Aile bir günün bir saatleri yemekte sen yetiyord. Mehmet Para birliği yeri bir yatak olacağına yatakta anımla sen bu yattığını yatakta bir yattığını dedi. Para tükür an kâdiler bir yatak aldı. sonra kâdiler iş buldular. anın ücret daha fazlaydı. Bir sâdiler güce bekçisi idi. Birlik durumları yavaş yavaş ilerliyordu. İsmail'e en iyi tanıdıkları daha çok tanıdıkları makinesi vardı. Emin tepesinde deligide anımla iş usta çözümler. Ustası anın arandı, ama Mehmet'in kâdilerin akıllı başlılar ve bilgisiyle istiyorlardı. Babaları şimdi daha kâdilerini dedi. ve kestirecekti. Ama Mehmet a bilgisiyle arabacı a kâdilerini hep iyiydi istiyorlardı.

Birlik durumları bir bilgisiyle alacak kadar iyiydi. O bilgisiyle aldılar, intemli kâdiler. Birlik günün bir saatleri yemekte anın hasta anını aldı. anın anına dayanarak intemli eder babaları kâdiler. Mehmet ve kâdilerini kâdiler.

İyilik Perisinin Büyük Yardımı

Çok uzak ülkelerde bir aile yaşamış. Bu aile birbirini çok sevmiş, saygısızlık yapmazlarmış. Bu aile çok fakir ama gururlu bir aileymiş. Elerinin bazı yerleri kırık, duvarları çatlak, camları kırılmış. Baba, karısını ve çocuklarını çok sevmiş. Her akşam yatmadan önce bir şemsiye alıp, çatlak tavandan akan küçük sudamlacıklarından korunmak için o şemsiyeyi düz tutar, öyle uyurlarmış. Hatta o kadar fakirlermiş ki evlerinde temizlik yaparmış, içeri bacaklar, horozlar giriyormuş.

Günlere bir gün en küçük çocuk olan Zezer'i bir bacak ısırması. Ancak hastaneye götünecek paraları yokmuş. Bu duruma bütün aile üyeleri çok üzülüyor, ancak buna rağmen bir şey yapamıyorlarmış.

Bir sabah tüm aile şaşkınlıkla uyanmış. O da ne? Zezer iyileşmiş. Bunu kimin yaptığını düşünmeye başlamışlar. Ama bulamamışlar. Bunu yapan o aileyi her gün camdan izleyen iyilik perisinin olduğunu asla tahmin edememişler.

Bu günden sonra iyilik perisi onlara istediğini yapmış. Bu aile iyilik perisinin yardımlarıyla mutlu mesut yaşamışlar.

Çocukların Yatağı

Bir zamanlar bir yatanmış. Bir tane fakir aile varmış. Burala eskiden bir evde yaşıyorlarmış. Evlerinde tek bir yatağı varmış. Bu aile altı çocukta oluşmaktadı. Bir de anne baba varmış. Evde sekiz kişilermiş. Bunlar yatağıta da, beş kişi bir tarafta, üç kişi diğer tarafta uyuyorlarmış. Çocuklar oku yorlarmış. Çocuklarına anne ve babası hep, çocukların okuyun dersin de benim gibi olmayın demişler. Çocuklar, rahatsızlıklarını kazarlarmışlar. İyi bir mesleğe sahibi olurlar. Ana baba, babalarını güd ve lemez ve eve yerleştirirler. Bu altı kardeş ise giden bir ev görmüşler. Gittikler babamışlar. Evia içinde bir - iki tane yatağı varmış. Bir yatağıta üç kişi, bir yatağıta dört kişi uyuyorlarmış. Eski zamanların hatırlanurlar. İse gittikler. Ama işde de o aileyi düşünmüşler. Sava o aileye yardım etmeye karar vermişler ve o aile eşyalar, kıyafetler, yiyecekler almışlar. Sava da işlerine dönmüşler.

Yardım etmek hem iyi bir iyilik hem de bir güzelliktir. Her insana iyilik yapması güzel bir şeydir. Her zaman iyilik yaparsak insanlar bizimle güzel dünya kalırsak her şeyin üstünde gelebiliriz.

YOKSULLUK

Bir zamanlar zengin aile vardı. Tam altı tane çocuğu vardı. Çocukları çok severlerdi. Büyük evi ve çocukları çok oyuncakları vardı. Babaları büyük şirkette müdür'dü. Bahçeleri ve hayvanı vardı. Babaları geldini çocukları onları karşıladı. Bir gün babaları yorgun ve üzgün bir şekilde geldi. Çocukları karşıladı. Babaları durun diyip mutfağa karısının yanına geldi. Çocuklarına gidin oyuncakları ile oynayın diyip gittiler. Her şey anlattı. Şirketin ortası kaçı. Masrafları üstüne yükü. Eğer ödemessek eve haciz gelebilir. Kadın başta artak almaz mı? dedi. Hiç kimse gelmiyor. Çocukların yanına gelerek çocuklar için aldığını anlatıyor. Çocuklar neyin var. Babası hiçbir şey yok.

Bir ay sonra
Kadın evde çocuklarıydı. Eve haciz geldi. Kadın hemen telefona sarıldı. Herkesi aradı. Adam'ın haberi vardı. İster gelmiş ve telefona bakmıyordu. Havale evde neredeyse hepsini aldılar. Adam evde ne olduğunu düşünüyordu. Adam eve gelerek uyuyarak ağladı ve uyudu kadın adamın yanına yattı ve sobayı yakıp çocukları için uyudular.

Murat'ın Hayatı

Günlerden bir gün Murat diye bir çocuk
 varmış. Murat'ın altı kardeşlemiş ve Murat onların
 son çocuğuymuş. Babası Emirhan Bey'in bir de sı-
 keti varmış. Bir gün işleri bitmiş ve iflas etmiş.
 İleri var, ne var, yoksa götürmüşler, artık su damlatan
 bir kulübece ekmiş, bir de yatağı bağlamış, som-
 suya ile yatıyormuş. Bir tarafta bir keti ve bir
 de kuzları varmış. Murat babası için çok üzülmüş-
 mü. Sonunda Murat'ın akırsa bir fikir gelmiş.
 Murat'ın annesi çok güzel köfte yaptığı için bir
 köfteci açmaya karar almış. Bir kaç köreca
 girip kulübeyi almışlar ve köfteci yapmışlar.
 İşleri çok ilerlemiş ve fabrika açmışlar. Bu arada
 Murat bir saat üzerinde çalışıyormuş. Bir kaç
 dereme ile icatını geliştirmiş. O şeye köfte
 yapmaya başlanmış. Ancak şehirden çok bir sevici
 şey çekmiş. İşleri, ciğerleri, büyük hay sandığı.
 Ar... Ancak sonunda malını kaybetmiş ve patlamış.
 Sonunda tekrar eski köfteci dönmüş ve fabrika
 kapatmış.

Son

EK 13. VERİ DÜZENLEMEK İÇİN KULLANILAN KOD ÖRNEĞİ

```
a<-read.table("d1.csv",sep=";",header=F)
```

```
p<-c(2)
```

```
sam<-a[sample(1:45,p,replace=F),]
```

```
m<-matrix(NA,6,23*p)
```

```
m[1,1:23]<-as.numeric(sam[1,5:27])
```

```
m[2,1:23]<-as.numeric(sam[1,82:104])
```

```
m[3,1:23]<-as.numeric(sam[1,159:181])
```

```
m[4,1:23]<-as.numeric(sam[1,236:258])
```

```
m[5,1:23]<-as.numeric(sam[1,313:335])
```

```
m[6,1:23]<-as.numeric(sam[1,390:412])
```

```
m[1,24:46]<-as.numeric(sam[2,5:27])
```

```
m[2,24:46]<-as.numeric(sam[2,82:104])
```

```
m[3,24:46]<-as.numeric(sam[2,159:181])
```

```
m[4,24:46]<-as.numeric(sam[2,236:258])
```

```
m[5,24:46]<-as.numeric(sam[2,313:335])
```

```
m[6,24:46]<-as.numeric(sam[2,390:412])
```


EK 14. GENELLENEBİLİRLİK KURAMI KOD ÖRNEKLERİ (2 PUANLAYICI)

```
a<-read.table("d1.csv",sep=";",header=F)
sam<-rbind(a[3,],a[41,])
p<-2
m<-matrix(NA,6,23*p)
```

```
m[1,1:23]<-as.numeric(sam[1,5:27])
m[2,1:23]<-as.numeric(sam[1,82:104])
m[3,1:23]<-as.numeric(sam[1,159:181])
m[4,1:23]<-as.numeric(sam[1,236:258])
m[5,1:23]<-as.numeric(sam[1,313:335])
m[6,1:23]<-as.numeric(sam[1,390:412])
```

```
m[1,24:46]<-as.numeric(sam[2,5:27])
m[2,24:46]<-as.numeric(sam[2,82:104])
m[3,24:46]<-as.numeric(sam[2,159:181])
m[4,24:46]<-as.numeric(sam[2,236:258])
m[5,24:46]<-as.numeric(sam[2,313:335])
m[6,24:46]<-as.numeric(sam[2,390:412])
```

```
m1<-m[,1:23]
m2<-m[,24:46]
xorthm<-(m1+m2)/2
```

```
xorthp1<-apply(m1,1,mean)
xorthp2<-apply(m2,1,mean)
xorth<-apply(m,1,mean)
```

```
xortms1<-apply(m1,2,mean)
xortms2<-apply(m2,2,mean)
xortm<-apply(xorthm,2,mean)
```

```
xortp1<-mean(xorthp1)
xortp2<-mean(xorthp2)
```

```
asa<-(xorthm)^2
bsa1<-(xorthp1)^2
bsa2<-(xorthp2)^2
csa<-m^2
sd<-c(5,22,1,110,5,22,110)
h<-6
m<-23
p<-2
```

```
pi<-c(m*p,h*p,h*m,p,m,h,1)
tort<-h*m*p*(mean(xorth)^2)
```

```

TTT<-
c(sum(xorth^2),sum(xortm^2),sum(xortp1^2+xortp2^2),sum(asa),sum(bsa1+bsa2),sum(xortm
s1^2+xortms2^2),sum(csa))
T<-pi*TTT
SS13<-T[1:3]-tort
SS4<-T[4]-T[1]-T[2]+tort
SS5<-T[5]-T[1]-T[3]+tort
SS6<-T[6]-T[2]-T[3]+tort
SS7<-T[7]-T[6]-T[5]-T[4]+T[3]+T[2]+T[1]-tort

SS<-c(SS13,SS4,SS5,SS6,SS7)
MS<-SS/sd

v1<-(MS[1]-MS[4]-MS[5]+MS[7])/(m*p)
if( !is.na( v1 ) && v1 < 0 )
{v1<-0}

v2<-(MS[2]-MS[4]-MS[6]+MS[7])/(h*p)
if( !is.na( v2 ) && v2 < 0 )
{v2<-0}

v3<-(MS[3]-MS[5]-MS[6]+MS[7])/(h*m)
if( !is.na( v3 ) && v3 < 0 )
{v3<-0}

v4<-(MS[4]-MS[7])/p
if( !is.na( v4 ) && v4 < 0 )
{v4<-0}

v5<-(MS[5]-MS[7])/m
if( !is.na( v5 ) && v5 < 0 )
{v5<-0}

v6<-(MS[6]-MS[7])/h
if( !is.na( v6 ) && v6 < 0 )
{v6<-0}

v7<-MS[7]
if( !is.na( v7 ) && v7 < 0 )
{v7<-0}
G<-(v1)/((v1)+(v4/m)+(v5/p)+(v7/h))
G

```

EK 14. (DEVAMI) GENELLENEBİLİRLİK KURAMI KOD ÖRNEKLERİ (10 PUANLAYICI)

```
a<-read.table("d1.csv",sep=";",header=F)
sam<-rbind(a[2,],a[8,],a[13,],a[19,],a[21,],a[32,],a[37,],a[39,],a[43,],a[45,])
```

```
p<-10
m<-matrix(NA,6,23*p)
```

```
m[1,1:23]<-as.numeric(sam[1,31:53])
m[2,1:23]<-as.numeric(sam[1,108:130])
m[3,1:23]<-as.numeric(sam[1,185:207])
m[4,1:23]<-as.numeric(sam[1,262:284])
m[5,1:23]<-as.numeric(sam[1,339:361])
m[6,1:23]<-as.numeric(sam[1,416:438])
```

```
m[1,24:46]<-as.numeric(sam[2,31:53])
m[2,24:46]<-as.numeric(sam[2,108:130])
m[3,24:46]<-as.numeric(sam[2,185:207])
m[4,24:46]<-as.numeric(sam[2,262:284])
m[5,24:46]<-as.numeric(sam[2,339:361])
m[6,24:46]<-as.numeric(sam[2,416:438])
```

```
m[1,47:69]<-as.numeric(sam[3,31:53])
m[2,47:69]<-as.numeric(sam[3,108:130])
m[3,47:69]<-as.numeric(sam[3,185:207])
m[4,47:69]<-as.numeric(sam[3,262:284])
m[5,47:69]<-as.numeric(sam[3,339:361])
m[6,47:69]<-as.numeric(sam[3,416:438])
```

```
m[1,70:92]<-as.numeric(sam[4,31:53])
m[2,70:92]<-as.numeric(sam[4,108:130])
m[3,70:92]<-as.numeric(sam[4,185:207])
m[4,70:92]<-as.numeric(sam[4,262:284])
m[5,70:92]<-as.numeric(sam[4,339:361])
m[6,70:92]<-as.numeric(sam[4,416:438])
```

```
m[1,93:115]<-as.numeric(sam[5,31:53])
m[2,93:115]<-as.numeric(sam[5,108:130])
m[3,93:115]<-as.numeric(sam[5,185:207])
m[4,93:115]<-as.numeric(sam[5,262:284])
m[5,93:115]<-as.numeric(sam[5,339:361])
m[6,93:115]<-as.numeric(sam[5,416:438])
```

```
m[1,116:138]<-as.numeric(sam[6,31:53])
m[2,116:138]<-as.numeric(sam[6,108:130])
m[3,116:138]<-as.numeric(sam[6,185:207])
m[4,116:138]<-as.numeric(sam[6,262:284])
```

```
m[5,116:138]<-as.numeric(sam[6,339:361])
m[6,116:138]<-as.numeric(sam[6,416:438])
```

```
m[1,139:161]<-as.numeric(sam[7,31:53])
m[2,139:161]<-as.numeric(sam[7,108:130])
m[3,139:161]<-as.numeric(sam[7,185:207])
m[4,139:161]<-as.numeric(sam[7,262:284])
m[5,139:161]<-as.numeric(sam[7,339:361])
m[6,139:161]<-as.numeric(sam[7,416:438])
```

```
m[1,162:184]<-as.numeric(sam[8,31:53])
m[2,162:184]<-as.numeric(sam[8,108:130])
m[3,162:184]<-as.numeric(sam[8,185:207])
m[4,162:184]<-as.numeric(sam[8,262:284])
m[5,162:184]<-as.numeric(sam[8,339:361])
m[6,162:184]<-as.numeric(sam[8,416:438])
```

```
m[1,185:207]<-as.numeric(sam[9,31:53])
m[2,185:207]<-as.numeric(sam[9,108:130])
m[3,185:207]<-as.numeric(sam[9,185:207])
m[4,185:207]<-as.numeric(sam[9,262:284])
m[5,185:207]<-as.numeric(sam[9,339:361])
m[6,185:207]<-as.numeric(sam[9,416:438])
```

```
m[1,208:230]<-as.numeric(sam[10,31:53])
m[2,208:230]<-as.numeric(sam[10,108:130])
m[3,208:230]<-as.numeric(sam[10,185:207])
m[4,208:230]<-as.numeric(sam[10,262:284])
m[5,208:230]<-as.numeric(sam[10,339:361])
m[6,208:230]<-as.numeric(sam[10,416:438])
```

```
m1<-m[,1:23]
m2<-m[,24:46]
m3<-m[,47:69]
m4<-m[,70:92]
m5<-m[,93:115]
m6<-m[,116:138]
m7<-m[,139:161]
m8<-m[,162:184]
m9<-m[,185:207]
m10<-m[,208:230]
xorthm<-(m1+m2+m3+m4+m5+m6+m7+m8+m9+m10)/10
```

```
xorthp1<-apply(m1,1,mean)
xorthp2<-apply(m2,1,mean)
xorthp3<-apply(m3,1,mean)
xorthp4<-apply(m4,1,mean)
xorthp5<-apply(m5,1,mean)
```

```
xorthp6<-apply(m6,1,mean)
xorthp7<-apply(m7,1,mean)
xorthp8<-apply(m8,1,mean)
xorthp9<-apply(m9,1,mean)
xorthp10<-apply(m10,1,mean)
xorth<-apply(m,1,mean)
```

```
xortms1<-apply(m1,2,mean)
xortms2<-apply(m2,2,mean)
xortms3<-apply(m3,2,mean)
xortms4<-apply(m4,2,mean)
xortms5<-apply(m5,2,mean)
xortms6<-apply(m6,2,mean)
xortms7<-apply(m7,2,mean)
xortms8<-apply(m8,2,mean)
xortms9<-apply(m9,2,mean)
xortms10<-apply(m10,2,mean)
xortm<-apply(xorthm,2,mean)
```

```
xortp1<-mean(xorthp1)
xortp2<-mean(xorthp2)
xortp3<-mean(xorthp3)
xortp4<-mean(xorthp4)
xortp5<-mean(xorthp5)
xortp6<-mean(xorthp6)
xortp7<-mean(xorthp7)
xortp8<-mean(xorthp8)
xortp9<-mean(xorthp9)
xortp10<-mean(xorthp10)
```

```
asa<-(xorthm)^2
bsa1<-(xorthp1)^2
bsa2<-(xorthp2)^2
bsa3<-(xorthp3)^2
bsa4<-(xorthp4)^2
bsa5<-(xorthp5)^2
bsa6<-(xorthp6)^2
bsa7<-(xorthp7)^2
bsa8<-(xorthp8)^2
bsa9<-(xorthp9)^2
bsa10<-(xorthp10)^2
```

```
csa<-m^2
sd<-c(5,22,9,110,45,198,990)
h<-6
m<-23
p<-10
```

```
pi<-c(m*p,h*p,h*m,p,m,h,1)
tort<-h*m*p*(mean(xorth)^2)
```

```
TTT<-
c(sum(xorth^2),sum(xortm^2),sum(xortp1^2+xortp2^2+xortp3^2+xortp4^2+xortp5^2+xortp6
^2+xortp7^2+xortp8^2+xortp9^2+xortp10^2),sum(asa),sum(bsa1+bsa2+bsa3+bsa4+bsa5+bs
a6+bsa7+bsa8+bsa9+bsa10),sum(xortms1^2+xortms2^2+xortms3^2+xortms4^2+xortms5^2+
xortms6^2+xortms7^2+xortms8^2+xortms9^2+xortms10^2),sum(csa))
```

```
T<-pi*TTT
SS13<-T[1:3]-tort
SS4<-T[4]-T[1]-T[2]+tort
SS5<-T[5]-T[1]-T[3]+tort
SS6<-T[6]-T[2]-T[3]+tort
SS7<-T[7]-T[6]-T[5]-T[4]+T[3]+T[2]+T[1]-tort
```

```
SS<-c(SS13,SS4,SS5,SS6,SS7)
MS<-SS/sd
```

```
v1<-(MS[1]-MS[4]-MS[5]+MS[7])/(m*p)
if( !is.na( v1 ) && v1 < 0 )
{v1<-0}
```

```
v2<-(MS[2]-MS[4]-MS[6]+MS[7])/(h*p)
if( !is.na( v2 ) && v2 < 0 )
{v2<-0}
```

```
v3<-(MS[3]-MS[5]-MS[6]+MS[7])/(h*m)
if( !is.na( v3 ) && v3 < 0 )
{v3<-0}
```

```
v4<-(MS[4]-MS[7])/p
if( !is.na( v4 ) && v4 < 0 )
{v4<-0}
```

```
v5<-(MS[5]-MS[7])/m
if( !is.na( v5 ) && v5 < 0 )
{v5<-0}
```

```
v6<-(MS[6]-MS[7])/h
if( !is.na( v6 ) && v6 < 0 )
{v6<-0}
```

```
v7<-MS[7]
if( !is.na( v7 ) && v7 < 0 )
{v7<-0}
```

```
G<-(v1)/((v1)+(v4/m)+(v5/p)+(v7/h))
G
```

EK 15. YAZILAN TÜM KODLARIN BİRLEŞTİRİLEREK HESAPLAMALARIN YAPILDIĞI KOD ÖRNEĞİ (2 PUANLAYICI)

```

a<- read.table( "d1.csv", sep=";", header = F )
kat<- matrix( NA, 100, 6 )

#####

s<- seq( 1:100 )
tt<- c(2)
for( t in 1:length( tt ) ) {
for ( i in 1:length( s ) ) {
sam <- a[ sample( 1:45, tt[t], replace=F ), ]
write.table( sam, file = paste( tt[ t ], "-", s[ i ], ".txt",sep="" ), sep=";", row.name=F,
col.name=F )

#####

p<- 2
m<- matrix( NA, 6, 23*p )

m[ 1, 1:23 ] <- as.numeric( sam[1, 57:79 ] )
m[ 2, 1:23 ] <- as.numeric( sam[1, 134:156 ] )
m[ 3, 1:23 ] <- as.numeric( sam[1, 211:233 ] )
m[ 4, 1:23 ] <- as.numeric( sam[1, 288:310 ] )
m[ 5, 1:23 ] <- as.numeric( sam[1, 365:387 ] )
m[ 6, 1:23 ] <- as.numeric( sam[1, 442:464 ] )

m[ 1, 24:46 ] <- as.numeric( sam[ 2, 57:79 ] )
m[ 2, 24:46 ] <- as.numeric( sam[ 2, 134:156 ] )
m[ 3, 24:46 ] <- as.numeric( sam[ 2, 211:233 ] )
m[ 4, 24:46 ] <- as.numeric( sam[ 2, 288:310 ] )
m[ 5, 24:46 ] <- as.numeric( sam[ 2, 365:387 ] )
m[ 6, 24:46 ] <- as.numeric( sam[ 2, 442:464 ] )

```

```
m1<- m[ ,1:23 ]
m2<- m[ ,24:46 ]
xorthm <-( m1 + m2 ) / 2

xorthp1 <- apply( m1, 1, mean )
xorthp2 <- apply( m2, 1, mean )
xorth <- apply( m, 1, mean )

xortms1 <- apply( m1, 2, mean )
xortms2 <- apply( m2, 2, mean )
xortm <- apply( xorthm, 2, mean )

xortp1<-mean(xorthp1)
xortp2<-mean(xorthp2)

asa<-(xorthm)^2
bsa1<-(xorthp1)^2
bsa2<-(xorthp2)^2
csa<-m^2
sd<-c(5,22,1,110,5,22,110)
h<-6
m<-23
p<-2

pi<-c(m*p,h*p,h*m,p,m,h,1)
tort<-h*m*p*(mean(xorth)^2)
```



```

TTT<-
c(sum(xorth^2),sum(xortm^2),sum(xortp1^2+xortp2^2),sum(asa),sum(bsa1+bsa2),sum(xortm
s1^2+xortms2^2),sum(csa))

T<-pi*TTT

SS13<-T[1:3]-tort
SS4<-T[4]-T[1]-T[2]+tort
SS5<-T[5]-T[1]-T[3]+tort
SS6<-T[6]-T[2]-T[3]+tort
SS7<-T[7]-T[6]-T[5]-T[4]+T[3]+T[2]+T[1]-tort

SS<-c(SS13,SS4,SS5,SS6,SS7)
MS<-SS/sd

v1<-(MS[1]-MS[4]-MS[5]+MS[7])/(m*p)
if( !is.na( v1 ) && v1 < 0 )
{v1<-0}
v2<-(MS[2]-MS[4]-MS[6]+MS[7])/(h*p)
if( !is.na( v2 ) && v2 < 0 )
{v2<-0}
v3<-(MS[3]-MS[5]-MS[6]+MS[7])/(h*m)
if( !is.na( v3 ) && v3 < 0 )
{v3<-0}
v4<-(MS[4]-MS[7])/p
if( !is.na( v4 ) && v4 < 0 )
{v4<-0}
v5<-(MS[5]-MS[7])/m
if( !is.na( v5 ) && v5 < 0 )
{v5<-0}
v6<-(MS[6]-MS[7])/h

```

```

if( !is.na( v6 ) && v6 < 0 )
{v6<-0}
v7<-MS[7]
if( !is.na( v7 ) && v7 < 0 )
{v7<-0}
kat[i,1]<-(v1)/((v1)+(v4/m)+(v5/p)+(v7/h))

#####

p<-2
m<-matrix(NA,6,23*p)

m[1,1:23]<-as.numeric(sam[1,31:53])
m[2,1:23]<-as.numeric(sam[1,108:130])
m[3,1:23]<-as.numeric(sam[1,185:207])
m[4,1:23]<-as.numeric(sam[1,262:284])
m[5,1:23]<-as.numeric(sam[1,339:361])
m[6,1:23]<-as.numeric(sam[1,416:438])

m[1,24:46]<-as.numeric(sam[2,31:53])
m[2,24:46]<-as.numeric(sam[2,108:130])
m[3,24:46]<-as.numeric(sam[2,185:207])
m[4,24:46]<-as.numeric(sam[2,262:284])
m[5,24:46]<-as.numeric(sam[2,339:361])
m[6,24:46]<-as.numeric(sam[2,416:438])

m1<-m[,1:23]
m2<-m[,24:46]

```

```
xorthm<-(m1+m2)/2
```

```
xorthp1<-apply(m1,1,mean)
```

```
xorthp2<-apply(m2,1,mean)
```

```
xorth<-apply(m,1,mean)
```

```
xortms1<-apply(m1,2,mean)
```

```
xortms2<-apply(m2,2,mean)
```

```
xortm<-apply(xorthm,2,mean)
```

```
xortp1<-mean(xorthp1)
```

```
xortp2<-mean(xorthp2)
```

```
asa<-(xorthm)^2
```

```
bsa1<-(xorthp1)^2
```

```
bsa2<-(xorthp2)^2
```

```
csa<-m^2
```

```
sd<-c(5,22,1,110,5,22,110)
```

```
h<-6
```

```
m<-23
```

```
p<-2
```

```
pi<-c(m*p,h*p,h*m,p,m,h,1)
```

```
tort<-h*m*p*(mean(xorth)^2)
```

```

TTT<-
c(sum(xorth^2),sum(xortm^2),sum(xortp1^2+xortp2^2),sum(asa),sum(bsa1+bsa2),sum(xortm
s1^2+xortms2^2),sum(csa))

T<-pi*TTT

SS13<-T[1:3]-tort
SS4<-T[4]-T[1]-T[2]+tort
SS5<-T[5]-T[1]-T[3]+tort
SS6<-T[6]-T[2]-T[3]+tort
SS7<-T[7]-T[6]-T[5]-T[4]+T[3]+T[2]+T[1]-tort

SS<-c(SS13,SS4,SS5,SS6,SS7)
MS<-SS/sd

v1<-(MS[1]-MS[4]-MS[5]+MS[7])/(m*p)
if( !is.na( v1 ) && v1 < 0 )
{v1<-0}
v2<-(MS[2]-MS[4]-MS[6]+MS[7])/(h*p)
if( !is.na( v2 ) && v2 < 0 )
{v2<-0}
v3<-(MS[3]-MS[5]-MS[6]+MS[7])/(h*m)
if( !is.na( v3 ) && v3 < 0 )
{v3<-0}
v4<-(MS[4]-MS[7])/p
if( !is.na( v4 ) && v4 < 0 )
{v4<-0}
v5<-(MS[5]-MS[7])/m
if( !is.na( v5 ) && v5 < 0 )
{v5<-0}
v6<-(MS[6]-MS[7])/h

```

```

if( !is.na( v6 ) && v6 < 0 )
{v6<-0}
v7<-MS[7]
if( !is.na( v7 ) && v7 < 0 )
{v7<-0}
kat[i,2]<-(v1)/((v1)+(v4/m)+(v5/p)+(v7/h))

#####

p<-2
m<-matrix(NA,6,23*p)

m[1,1:23]<-as.numeric(sam[1,5:27])
m[2,1:23]<-as.numeric(sam[1,82:104])
m[3,1:23]<-as.numeric(sam[1,159:181])
m[4,1:23]<-as.numeric(sam[1,236:258])
m[5,1:23]<-as.numeric(sam[1,313:335])
m[6,1:23]<-as.numeric(sam[1,390:412])

m[1,24:46]<-as.numeric(sam[2,5:27])
m[2,24:46]<-as.numeric(sam[2,82:104])
m[3,24:46]<-as.numeric(sam[2,159:181])
m[4,24:46]<-as.numeric(sam[2,236:258])
m[5,24:46]<-as.numeric(sam[2,313:335])
m[6,24:46]<-as.numeric(sam[2,390:412])

m1<-m[,1:23]
m2<-m[,24:46]

```

```
xorthm<-(m1+m2)/2
```

```
xorthp1<-apply(m1,1,mean)
```

```
xorthp2<-apply(m2,1,mean)
```

```
xorth<-apply(m,1,mean)
```

```
xortms1<-apply(m1,2,mean)
```

```
xortms2<-apply(m2,2,mean)
```

```
xortm<-apply(xorthm,2,mean)
```

```
xortp1<-mean(xorthp1)
```

```
xortp2<-mean(xorthp2)
```

```
asa<-(xorthm)^2
```

```
bsa1<-(xorthp1)^2
```

```
bsa2<-(xorthp2)^2
```

```
csa<-m^2
```

```
sd<-c(5,22,1,110,5,22,110)
```

```
h<-6
```

```
m<-23
```

```
p<-2
```

```
pi<-c(m*p,h*p,h*m,p,m,h,1)
```

```
tort<-h*m*p*(mean(xorth)^2)
```

```

TTT<-
c(sum(xorth^2),sum(xortm^2),sum(xortp1^2+xortp2^2),sum(asa),sum(bsa1+bsa2),sum(xortm
s1^2+xortms2^2),sum(csa))

T<-pi*TTT

SS13<-T[1:3]-tort
SS4<-T[4]-T[1]-T[2]+tort
SS5<-T[5]-T[1]-T[3]+tort
SS6<-T[6]-T[2]-T[3]+tort
SS7<-T[7]-T[6]-T[5]-T[4]+T[3]+T[2]+T[1]-tort

SS<-c(SS13,SS4,SS5,SS6,SS7)
MS<-SS/sd

v1<-(MS[1]-MS[4]-MS[5]+MS[7])/(m*p)
if( !is.na( v1 ) && v1 < 0 )
{v1<-0}
v2<-(MS[2]-MS[4]-MS[6]+MS[7])/(h*p)
if( !is.na( v2 ) && v2 < 0 )
{v2<-0}
v3<-(MS[3]-MS[5]-MS[6]+MS[7])/(h*m)
if( !is.na( v3 ) && v3 < 0 )
{v3<-0}
v4<-(MS[4]-MS[7])/p
if( !is.na( v4 ) && v4 < 0 )
{v4<-0}
v5<-(MS[5]-MS[7])/m
if( !is.na( v5 ) && v5 < 0 )
{v5<-0}
v6<-(MS[6]-MS[7])/h

```

```

if( !is.na( v6 ) && v6 < 0 )
{v6<-0}
v7<-MS[7]
if( !is.na( v7 ) && v7 < 0 )
{v7<-0}
kat[i,3]<-(v1)/((v1)+(v4/m)+(v5/p)+(v7/h))
#####KTK#####
p<-2

m<-matrix(NA,6,23*p)

m[1,1:23]<-as.numeric(sam[1,57:79])
m[2,1:23]<-as.numeric(sam[1,134:156])
m[3,1:23]<-as.numeric(sam[1,211:233])
m[4,1:23]<-as.numeric(sam[1,288:310])
m[5,1:23]<-as.numeric(sam[1,365:387])
m[6,1:23]<-as.numeric(sam[1,442:464])

m[1,24:46]<-as.numeric(sam[2,57:79])
m[2,24:46]<-as.numeric(sam[2,134:156])
m[3,24:46]<-as.numeric(sam[2,211:233])
m[4,24:46]<-as.numeric(sam[2,288:310])
m[5,24:46]<-as.numeric(sam[2,365:387])
m[6,24:46]<-as.numeric(sam[2,442:464])

p1<-m[,1:23]
p2<-m[,24:46]

```



```
pp1<-round(apply(p1,1,mean),0)
```

```
pp2<-round(apply(p2,1,mean),0)
```

```
kat[i,4]<-(( sum( pp1 & pp2 ) + sum( !pp1 & !pp2 ) )/( sum( pp1 & pp2 ) + sum( !pp1 & !pp2 ) + sum( !pp1 & pp2 )+sum( pp1 & !pp2 )))
```

```
#####
```

```
p<-2
```

```
m<-matrix(NA,6,23*p)
```

```
m[1,1:23]<-as.numeric(sam[1,31:53])
```

```
m[2,1:23]<-as.numeric(sam[1,108:130])
```

```
m[3,1:23]<-as.numeric(sam[1,185:207])
```

```
m[4,1:23]<-as.numeric(sam[1,262:284])
```

```
m[5,1:23]<-as.numeric(sam[1,339:361])
```

```
m[6,1:23]<-as.numeric(sam[1,416:438])
```

```
m[1,24:46]<-as.numeric(sam[2,31:53])
```

```
m[2,24:46]<-as.numeric(sam[2,108:130])
```

```
m[3,24:46]<-as.numeric(sam[2,185:207])
```

```
m[4,24:46]<-as.numeric(sam[2,262:284])
```

```
m[5,24:46]<-as.numeric(sam[2,339:361])
```

```
m[6,24:46]<-as.numeric(sam[2,416:438])
```

```
p1<-m[,1:23]
```

```
p2<-m[,24:46]
```

```
pp1<-round(apply(p1,1,mean),0)
```

```
pp2<-round(apply(p2,1,mean),0)
```

```

ppp<-cbind(pp1,pp2)
ppp
k<-3
kkk<-matrix(NA,6,k)
for(z in 1:6){

kkk[z,]<-c(
length(subset(ppp[z,],ppp[z,]==0)),
length(subset(ppp[z,],ppp[z,]==1)),
length(subset(ppp[z,],ppp[z,]==2)))
}
p<-2
h<-6

pj<-round(apply(kkk,2,sum)/(h*p),2)

pi<-round((apply(kkk^2,1,sum)-p)*(1/(p*(p-1))),2)
pitop<-sum(pi)
pe<-sum(pj^2)

P<-(1/(h*p*(p-1)))*(pitop*p*(p-1))

kat[i,5]<-(P-pe)/(1-pe)

#####

```

```
p<-c(2)
```

```
m<-matrix(NA,6,23*p)
```

```
m[1,1:23]<-as.numeric(sam[1,5:27])
```

```
m[2,1:23]<-as.numeric(sam[1,82:104])
```

```
m[3,1:23]<-as.numeric(sam[1,159:181])
```

```
m[4,1:23]<-as.numeric(sam[1,236:258])
```

```
m[5,1:23]<-as.numeric(sam[1,313:335])
```

```
m[6,1:23]<-as.numeric(sam[1,390:412])
```

```
m[1,24:46]<-as.numeric(sam[2,5:27])
```

```
m[2,24:46]<-as.numeric(sam[2,82:104])
```

```
m[3,24:46]<-as.numeric(sam[2,159:181])
```

```
m[4,24:46]<-as.numeric(sam[2,236:258])
```

```
m[5,24:46]<-as.numeric(sam[2,313:335])
```

```
m[6,24:46]<-as.numeric(sam[2,390:412])
```

```
p1<-m[,1:23]
```

```
p2<-m[,24:46]
```

```
pp1<-round(apply(p1,1,mean),0)
```

```
pp2<-round(apply(p2,1,mean),0)
```

```
ppp<-cbind(pp1,pp2)
```

```
ppp
```

```
k<-5
```

```
kkk<-matrix(NA,6,k)
```

```
for(z in 1:6){
```

```

kkk[z,]<-c(
length(subset(ppp[z,],ppp[z,]==0)),
length(subset(ppp[z,],ppp[z,]==1)),
length(subset(ppp[z,],ppp[z,]==2)),
length(subset(ppp[z,],ppp[z,]==3)),
length(subset(ppp[z,],ppp[z,]==4)))
}
p<-2
h<-6

pj<-round(apply(kkk,2,sum)/(h*p),2)

pi<-round((apply(kkk^2,1,sum)-p)*(1/(p*(p-1))),2)
pitop<-sum(pi)
pe<-sum(pj^2)

P<-(1/(h*p*(p-1)))*(pitop*p*(p-1))

kat[i,6]<-(P-pe)/(1-pe)

}

}

write.table(round(kat,3),"gler2.txt",sep=";",quote=F)

```

ÖZGEÇMİŞ

Mehtap AKTAŞ

13 Ocak 1982'de Kars'ta doğdu. İlk, orta ve lise öğrenimini Karsta tamamladı. 2004 yılında Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Sınıf Öğretmenliği Bölümünden Lisans derecesini aldı. 2005 yılında Kafkas Üniversitesinde araştırma görevlisi olarak göreve başladı. 2007 yılında Kafkas Üniversitesi Eğitim Fakültesi Eğitim Programları ve Öğretimi Bilim dalından Yüksek Lisans derecesini aldı. 2009 yılında Mersin Üniversitesi Eğitimde Ölçme ve Değerlendirme Ana Bilim Dalında lisansüstü eğitim almak için görevlendirildi. Halen Mersin Üniversitesinde araştırma görevlisi olarak çalışmaktadır.