



## Yazılı Anlatım Becerilerinin Değerlendirilmesi İçin Dereceli Puanlama Anahtarı Geliştirme Çalışması

### Development of an Analytical Rubric for Assessing Writing Skills

Safiye BİLİCAN DEMİR\* Özen YILDIRIM\*\*

• Geliş Tarihi: 08.07.2019 • Kabul Tarihi: 01.09.2019 • Yayın Tarihi: 12.09.2019

#### Öz

Bu araştırmanın amacı, ilköğretim 5. sınıf düzeyindeki öğrencilerin yazılı anlatım becerilerini değerlendirmek üzere analitik dereceli puanlama anahtarı (DPA) geliştirmek ve bu DPA'nın geçerlik ve güvenilirlik kanıtlarını sunmaktır. Araştırma verileri, Ankara ve Denizli'deki görev yapan 22 öğretmenin 10 öğrencinin yanıtlarını değerlendirmesiyle elde edilmiştir. Geliştirilen analitik DPA'nın geçerlik kanıtı için alan öğretmeni ve ölçme ve değerlendirme uzmanlarının her bir ölçüt ve tanımlamalara ilişkin görüşleri arasındaki uyum, basit uyum yüzdesi ile değerlendirilmiştir. Puanlayıcılar arası güvenilirlik için Kendall's W katsayısı ve puanlayıcı içi tutarlılık için uyum yüzdesi hesaplanmıştır. Araştırma bulguları DPA'da yer alan ölçütler ve bunlara ilişkin tanımların uygunluğu ve yeterliği konusunda uzmanlar arasında görüş birliği olduğunu; aynı puanlayıcı ve farklı puanlayıcılar tarafından yapılan değerlendirmeler arasında yüksek düzeyde tutarlılık olduğunu göstermiştir. Araştırma bulguları ilgili literatür çerçevesinde tartışılmış ve önerilerde bulunulmuştur.

**Anahtar sözcükler:** yazılı anlatım becerisi, analitik dereceli puanlama anahtarı, puanlayıcılar arası uyum.

#### Atıf:

Bilican Demir, S. ve Yıldırım, Ö. (2019). Yazılı anlatım becerilerinin değerlendirilmesi için dereceli puanlama anahtarı geliştirme çalışması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 47,457-473. doi: 10.9779/pauefd.588565.

\* Dr. Öğr. Üyesi, Kocaeli Üniversitesi Eğitim Fakültesi Ölçme ve Değerlendirme ABD, ORCID: <https://orcid.org/0000-0001-9564-9029>, safiyebilican@gmail.com

\*\* Sorumlu yazar: Dr. Öğr. Üyesi, Pamukkale Üniversitesi Eğitim Fakültesi Ölçme ve Değerlendirme ABD, ORCID: <https://orcid.org/0000-0003-2098-285X>, ozenyildirim@pau.edu.tr

**Abstract**

The aim of this study is to develop an analytical rubric to evaluate the writing skills of 5th graders and to provide evidence of validity and reliability of the developed rubrics. Data were obtained from 22 teachers in Ankara and Denizli provinces in Turkey. For the validity study, the fit between experts about each criteria and definitions was evaluated with the percentage of simple agreement. The Kendall's W coefficient was calculated for inter-rater reliability and the percent agreement for intra-rater consistency was calculated. The findings of the study indicate that there is a high level of agreement among the experts on the appropriateness and adequacy of the criteria and the definitions in the DPA; there is a high level of consistency between the evaluations made by the same rater and different raters. The findings of the research have been discussed and suggestions have been made within the framework of the relevant literature.

**Keywords:** writing skills, analytical rubric, inter-rater reliability.

**Cited:**

Bilican Demir, S. & Yıldırım, Ö. (2019). Development of an analytical rubric for assessing writing skills. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 47,457-473. doi: 10.9779/pauefd.588565.

## Giriş

Günümüzde hızla gelişen teknoloji ve dijital dünya, kişiler arası iletişimin yoğunlaşmasına neden olmuş ve buna bağlı olarak sözlü ya da yazılı anlatım yollarının kullanım sıklığı artış göstermiştir. Bu bağlamda hızlı değişim ve gelişime ayak uydurmak üzere anlatım becerileri daha iyi gelişmiş bireylerin yetiştirilmesi önemli hale gelmiştir. Bu gereksinime uygun olarak ilk ve ortaokullardaki Türkçe derslerinin amaçlarından biri öğrencilerin gördüklerini, okuduklarını, dinlediklerini/izlediklerini, deneyimlerini doğru, etkili ve etkin bir biçimde anlatabilmelerini sağlamaktır (Şengül, 2011). Anlatım becerilerinden biri de yazılı anlatım becerisidir.

Yazma, düşünceleri ifade edebilme amacıyla gerekli sembol ve işaretleri kurallara uygun kullanma ve düşünce üretebilme becerisidir (Akyol, 2006). Kişilerin yazma ile ilgili becerileri, dinleme, konuşma ve okuma becerileriyle yakından ilişkilidir (Coşkun, 2007). Yazılı anlatım becerileri, bireylerin bildiklerini, duyduklarını ve öğrendiklerini ifade etmelerine olanak sağladığından, bireylerin akademik, iş ya da sosyal yaşantısındaki başarısında önemli bir yer tutar (Norman ve Spencer, 2005). İlgili araştırmalar, bireylerin farklı alanlardaki akademik başarılarının bireylerin yazma becerisiyle yakından ilişkili olduğunu göstermiştir (Lan, Hung ve Hsu, 2011; Ungan, 2007). Yazma becerisi, doğasına uygun olarak birçok beceriyi içerecek biçimde ele alınmalıdır. Çünkü giriş gelişme ve sonuç gibi bölümleri içeren bir yazı oluşturma süreci çok yönlü ve karmaşık süreçleri içermektedir (Evans, 2001). Bu süreç, yazılan konunun içeriğini bilme, içeriği işlemeye yarayan süreç bilgisi, yazı türünü belirleyen plan, sözdizimsel formları ve noktalama işaretlerinin kullanımını kapsayan yazı üretme bilgisini uygulama (Hillocks, 1987) aşamalarından oluşur. Bunun yanında yazma, dil bilgisi, analiz, söz sanatlarını kullanma vb. çalışmaları da içerir (Zamel, 1982).

Öğrencilerin yazılı anlatım becerilerinin geliştirmek üzere yapılacak etkinliklerin amaca uygun bir biçimde yürütülmesinde ölçme-değerlendirme çalışmalarının önemi büyüktür. Öğrencilerin yazılı anlatım ürünlerini değerlendirmek birçok değişkenin göz önünde bulundurulması gereken karmaşık bir süreçtir. Yazma becerilerine yönelik ölçme ve değerlendirme çalışmalarının, hem yazılı anlatım ürününün içeriğini hem de yazma sürecinin işleyişini çeşitli boyutlarda ele alan ve öğrencilerin yazılı anlatım ürünlerini bu boyutlarda tek tek değerlendirmeye olanak sağlayan araçlarla yapılması önemli hale gelmektedir. Ancak araştırma bulguları öğretmenlerin, öğrenci ürünlerini ölçme ve değerlendirmeye yönelik uygun ve nitelikli (geçerli-güvenilir) araçları geliştirme ve kullanma konusunda önemli eksiklerinin olduğunu göstermektedir (Epeçan ve Erzen, 2008; TED,2009). Bunun yanında öğrencilerin yazılı anlatım çalışmalarının değerlendirilmesinde hangi ölçütlerin dikkate alınacağı da önemli bir konudur. Okurer'e (1967) göre, bir yazılı ürün hem teknik-form hem de içerik-düşünce tarafı ile birlikte bir değer ifade eder. Biçimsel unsurlara ait yanlış ile konunun ele alınmasına ilişkin yapılan yanlışları ayrı ayrı değerlendirmek gerekir. Göçer (2011) tarafından yapılan araştırmanın bulguları, öğretmenlerin öğrencilerin yazılı çalışmalarında aradığı temel özelliklerin plan, yazının düzgünlüğü, kâğıdın tertip-düzeni, yazım ve noktalama gibi biçime yönelik olduğunu; konuyu ele alma, düşüncede özgünlük, anlatım gücü, düşüncelerin akıcılığı, uygun anlatım biçimlerini seçme, paragraflar arası uygun geçişler yapma ve anlam bütünlüğü gibi içeriğe yönelik unsurlara fazla önem vermediklerini göstermiştir. Oysaki bir metnin uyumlu olabilmesi için kendini oluşturan çeşitli bölümlerin bir dilsel bütünlük sağlayacak

biçimde birbirine bağlanması (bağdaşlık) (Onursal, 2003), metnin, kavramların ve düşüncelerin yapılanması ve birbiriyle ilintisi yoluyla okurda bütünlük duygusu oluşturması (tutarlılık) (Hyland, 2002) gibi yazının metinsellik boyutunu ilgilendiren ölçütlerin de dikkate alınması önemlidir. Bu yanında, yazma amacın metne hangi türde, hangi işleve göre yansıtılacağı, okuyucunun zihninde amaçlanan izlemi oluşturmak için hangi anlatım yollarına başvurulacağı, art alan bilgisinin metne nasıl yansıtılacağı, duruma uygun olma şartlarının neler olduğu, metinlerarası ilişkilerin nasıl kurulacağı gibi daha teknik becerilerin de öğrencilere süreç içerisinde öğretilmesi gerekmektedir (Çeçen, 2011). Ancak araştırma bulguları, öğretmenlerin değerlendirme aşamasında standart ölçütlerden ziyade kendi oluşturdukları ölçütlerini kullanmayı tercih ettiklerini göstermektedir (Elma ve Bütün, 2015; Karakoç Öztürk, 2014).

Öğretmenler, öğrencilerin yazılı anlatımlarını sadece not vermek için değil, yazma becerilerindeki eksikleri/yanlışları tespit etmek ve bunların düzeltilmesine yardımcı olmak için değerlendirmelidir. Yazılı anlatım etkinliği sonrası öğretmenlerin geri bildirimde bulunmaları, öğrencilerin yaptıkları hataların tekrarlanma olasılığını azaltacak ve öğrencilerin başarılı oldukları yönleri görmesiyle anlatım becerilerinin daha da gelişmesini sağlayacaktır (Coşkun, 2007). Yazma becerisinin geliştirilmesinde öğrencilere yazma fırsatı sağlamak, özellikle planlama ve gözden geçirme aşamaları olmak üzere yazma süreçlerine odaklanmak, başarılı yazmak için ölçütleri açıklamak gibi uygulamalar oldukça etkilidir (Karadağ ve Kayabaş, 2011). Ancak Wiener (2003), öğretmenlerin öğrencilerine birçok yazma ödevi verdiklerini ancak bu ödevlerin nasıl yapılacağına ilişkin bilgi vermedikleri ve rehberlik edemediklerini ifade etmektedir. Benzer şekilde Ungan (2007) tarafından yapılan araştırma ülkemizde yazma çalışmalarının, öğretmenlerin zaman öldürmek için kullandıkları, öğrencilerin birinci ders yazdıkları metinleri ikinci ders okudukları, performansları hakkında dönüt almadıkları uygulamalar biçiminde olduğunu göstermiş ve okullarda, tamamen çoktan seçmeli maddelerden oluşan sınavlar ile öğrencilerin yazılı anlatım becerilerinin gelişmediği vurgusu yapılmıştır. Ancak bu tür madde formatıyla yapılan bir ölçümede öğrencilerden, seçenekler içerisindeki doğru yanıtı bulmaları beklendiği için bu yolla öğrencilerin iyi bir yazılı metin üretip üretmediklerine ilişkin bir değerlendirme yapmak doğru bir yaklaşım olmayacaktır. Çünkü öğrenciler, başkaları tarafından oluşturulmuş tümceler ya da paragraflarla ilgili sorularla karşılaşmakta ve sahip oldukları bilgileri kullanarak bu soruları yanıtlamaktadır. Bu durumda öğrenciler bir tümce ya da paragraf üretmedikleri için kendi becerilerini sergilemek yerine ancak üretilmiş olan tümce ya da paragraflar temelinde noktalama, yazım, tutarlılık vb. alanlara ilişkin sahip oldukları bilgileri işe koşabilmektedirler (Ülper, 2009a). Bunun için yazılı anlatım uygulamalarında öğrencilere gerekli zaman tanınmalı, dönüt ve düzeltmelere gereken önem verilmeli, öğretmenlerimiz bilgilendirilmelidir.

Türkiye’de yazma becerilerini izlemeye ve ölçmeye yönelik belirgin bir yazma yaklaşımı olmadığından bu konuda yapılan uygulamalar ve bu uygulamaların ne kadar işe yaradığı konusundaki kaynaklar sınırlıdır (Karatay, 2013). Oysa başarılı bir yazma ürünü ortaya çıkarmak için yazma öğretiminin; planlama, belli aralıklarla bunu düzenleme, gözden geçirme, düzeltme ve tekrar yazma gibi, bazen ileriye bazen geriye doğru atılan adımlarla ilerlemesi gerekmektedir (Karatay, 2011). Flower ve Hayes (1981) yazma becerisinin bilişsel sürecini dört temel süreç içerisinde değerlendirmektedir. Araştırmacılara göre yazma süreci yazarın düzenlediği farklı düşünme süreçlerinin bir kümesidir, bu süreç hiyerarşik bir düzene sahiptir ve bir süreç diğer süreçleri kapsayabilir. Yazma süreci amaç odaklı bir süreçtir ve bu süreci yazar

yüksek düzey ve bunları destekleyen alt amaçlarına göre yönlendirir. Yazma becerisinin karmaşık yapısı dikkate alındığında öğrencilerin yazılı anlatım becerilerini değerlendirirken nasıl bir ölçme-değerlendirme yaklaşımı benimseneceği ve kullanılacak ölçme araçları önemli hale gelmektedir. Yazılı anlatım becerilerinin birden çok beceri içermesi ve bu becerinin değerlendirilmesi ile ilgili yukarıda yapılan tartışmalar dikkate alındığında, ilgili becerinin ölçülmesinde bireylerin performanslarıyla ilgili değerlendirme ölçütlerini ve ilgili ölçütlere ilişkin tanımlamaları belirgin bir biçimde içermesi bakımından dereceli puanlama anahtarlarını (DPA, rubrik) kullanmak nesnel ve ölçüt odaklı bir yaklaşım olacaktır. Bu araçları kullanarak yazılı bir metne ait hem biçimsel hem de konunun ele alınmasına ilişkin boyutlar ayrıntılı olarak değerlendirilebilir. Ülper (2009b) bilişsel süreç modeline göre, öğrencilerin yazma becerilerinin geliştirilmesi amacıyla yaptıkları çalışmalarında analitik türde DPA'larla öğrenci başarısını değerlendirmişler ve bu araçların öğrencilerin yazma becerileri hakkında daha kullanışlı bilgi sağladığını, sınıf uygulamalar için daha uygun olduğunu ve güvenilir olduğunu vurgulamışlardır.

DPA'lar öğrencilerin farklı alanlardaki performanslarının değerlendirilmesinde ölçülecek özelliğin boyutlarını gösteren bir çeşit puanlama ölçüsüdür ve değerlendirme ölçütleri, ölçüt tanımlamaları ve bir puanlama stratejisinden oluşmaktadır (Popham, 1997). DPA'lar, öğrenciden beklenen performansı ve değerlendirme sürecinde hangi davranışların önemli olduğunu açık ve belirgin hale getirmekte ve böylece değerlendirme sürecine ilişkin öğrenci ve öğretmen için ortak bir algı oluşmaktadır (Arter ve McTighe, 2001). Her bir performans düzeyinde öğrenciden ne beklenildiğinin tanımlanması, DPA'nın değerlendirme aracı yanında bir öğretim aracı olarak da kullanılmasına da olanak sağlar (Moskal ve Leydens, 2000). DPA'da açık tanımlanan ölçütler sayesinde öğrenciler kendi çalışmaları ile ilgili güçlü ve zayıf yanlarını görebilmekte öğretmenler ise hem velilere hem de öğrencilere bu konuda etkili geri bildirim verebilmektedir. Bu bağlamda DPA'lar sürece ve sonuca dayalı değerlendirmelerde kullanılabilen ölçme araçlarıdır. Süreç içerisinde öğrenciler öğrenme hedefini yerine getirirken nasıl bir yol izlemeleri gerektiğini DPA'ları kullanarak belirleyebilir. Diğer taraftan süreçte öğretmenin öğrenciyi izlemesi için de bir rehberdir. Sonuca dayalı değerlendirmede öğrenci ve öğretmen, öğrenci ürünü belirlenen ölçütlere göre değerlendirmenin yanı sıra öğrencilerin, akranlarının çalışmalarını değerlendirmesine de imkân sunar. Bunlara ek olarak DPA'ların en önemli avantajlarından biri de öğrenci çalışmalarına ilişkin geçerli ve güvenilir değerlendirme yapmaya olanak sağlamasıdır (Mertler, 2001;Popham, 1997). Çünkü amaca uygun yapılandırılmış bir DPA kullanılarak yapılacak puanlama, ölçme işlemini değerlendirmeyi yapan kişiden ve puanlamanın yapıldığı zamandan bağımsız hale getirebilmektedir (Moskal ve Leydens, 2000). Bilindiği gibi ölçülmesi amaçlanan özellik dışında puanlayıcıdan kaynaklanan ve ölçme sonuçlarının doğruluğunu ve tekrarlanabilirliğini olumsuz etkileyebilen durumlar söz konusudur. Bu yüzden öğrenci performansının değerlendirilmesinde dereceli puanlama anahtarı kullanmak puanlayıcıdan kaynaklan ölçme hatalarını en aza indirecek ve yapılan ölçme işleminin güvenilirliği artacaktır (Dunbar, Brooks ve Miller, 2006). Yazma becerisinin çok yönlü karmaşık yapısı düşünüldüğünde DPA'ların öğretmenlerin yazma öğretimini yeniden düzenlemede işlerine en çok yarayacak puanlama anahtarları olduğu söylenebilir (Cristie, Enz ve Vukelich, 2003). Ancak araştırma bulguları öğretmenlerin öğrencilere ait yazılı metinleri değerlendirirken dereceli puanlama anahtarları kullanım sıklığının düşük olduğunu göstermektedir (Göçer, 2005; Karakoç Öztürk, 2014; 2016).

İlgili alan yazın incelendiğinde öğrencilerin dil becerilerinin ölçülmesi kapsamında okuma, dinleme ve konuşma becerilerinin ölçülmesinde dereceli puanlama anahtarının sıklıkla kullanıldığı (Bozkurt ve Arıca-Akkök, 2019; Tok ve Erdoğan, 2017; Yıldız, 2018) görülmüştür. Ayrıca alan yazın incelendiğinde ilgili araştırmalarda (Ağan Haykır, 2014; Aktaş ve Alıcı, 2018; Topuzkanamış, 2014) yazılı anlatım becerilerinin kompozisyon ya da hikâyeler üzerinden değerlendirildiği görülmektedir. Bu bakımdan yazılı anlatım becerilerini farklı zihinsel düzeylere ve farklı metin türlerine göre değerlendiren bir çalışmaya rastlanılmamıştır. Dereceli puanlama anahtarlarının alanyazında sözü edilen olumlu yönleri ve ilgili becerinin ölçülmesinde ortaya çıkan eksiklikler dikkate alınarak, ilköğretim 5. sınıf düzeyindeki öğrencilerin yazılı anlatım becerilerini değerlendirmek ve öğretmenin çalışmaları konusunda öğrencilere destekleyici geri bildirim vermesinde yardımcı olacak bir dereceli puanlama anahtarı (rubrik) geliştirilmesi amaçlanmıştır.

## **Yöntem**

### **Çalışma Grubu**

Araştırma verileri, seçkisiz olmayan örnekleme içerisinde yer alan uygun örnekleme ve ulaşılabilirlik dikkate alınarak Ankara ve Denizli ilinde bir devlet okulunda görev yapan öğretmenlerden ve ilköğretim 5. Sınıf öğrencilerinden elde edilmiştir. İlgili sınıflardaki 27 öğrenciye uygulama yapılmış bu öğrenciler içerisinde rastgele örnekleme seçilen 10 öğrencinin yanıtları değerlendirilmiştir. Değerlendirilen yazılı anlatımların fazlalığı nedeniyle (iki metin, 9 soru ve her soru için altı ölçüt) öğrenci sayısı 10 ile sınırlı tutulmuştur. Öğrencilerin 9'u kız, 1'i erkek öğrencidir. Değerlendiriciler (puanlayıcılar), ilköğretim Türkçe öğretmenliği alanında lisans eğitimi almış öğretmenlerdir. Birinci metni 10 öğretmen, ikinci metni 12 öğretmen puanlamıştır. Öğretmenlerin 4'ü erkek 18'i kadındır.

### **Veri Toplama Aracı**

Araştırmada öğrencilerin yazılı anlatım becerilerinin ölçülmesi amacıyla iki metin seçilmiştir. Bu metinlerden biri hikâye diğeri ise bilgi verici tarzda yazılmıştır. Metinlerin seçiminde, öğrencilerin gelişim, sınıf ve yaş düzeyi göz önünde tutulmuştur.

Seçilen iki metnin okunabilirlik düzeyleri ve eğitim düzeyine uygunluğu için farklı okunabilirlik formülleri incelenmiştir (Gobbledgook formülü, Gunning-Fog İndeks formülü, ARI formülü, Flesh reading Ease Score, Ateşman formülü, Çetinkaya ve Uzun formülü vb.). Araştırma dilinin Türkçe olması ve diğer formüllerin yabancı bir dil için geliştirilmesinden kaynaklı Flesch'in okunabilirlik formülünü Türkçeye uyarlayan Ateşman (1997) 'ın ve Çetinkaya (2010)'un geliştirdikleri hesaplamasından yararlanılarak metinlerin okunabilirlikleri incelenmiştir. Ateşman (1997)'nin hesaplamasına göre iki metinde orta güçlükte, Çetinkaya ve Uzun'un hesaplamasına göre de 5-7. sınıfta bağımsız okumaya uygun metinler olarak belirlenmiştir. Hesaplamalar sonucunda ilgili iki metnin zorluk düzeylerinin ve eğitim düzeyinin 5. sınıf düzeyine uygun olduğu söylenebilir. Bununla birlikte Türkçe eğitimi alanında üç uzmandan ve iki Türkçe öğretmeninden ilgili metinlerin bilişsel gelişim, sınıf ve yaş düzeyi bakımından uygunluğu hakkında görüş alınmıştır.

Metinlerin ilgili sınıf düzeyine uygunluğu konusunda görüş alındıktan sonra öğrencilerin yazılı anlatım becerilerini ölçmek amacıyla açık uçlu ve uzun yanıtı olmak üzere bir metin için dört ve diğeri için beş soru geliştirilmiştir. Sorular, Uluslararası Okuma

Becerilerinde Gelişim Projesi (Progress in International Reading Literacy Study (PIRLS))’de kullanılan dört farklı aşamalı bilişsel düzey ve ilköğretim 5. Sınıf öğretim programı dikkate alınarak hazırlanmıştır (Mullis ve Martin, 2015; MEB,2019). PIRLS’ün uluslararası düzeyde okuma becerilerinin ölçülmesinde yaygın kullanılan geniş ölçekli uygulama olması ve uygulanan testlerin geçerliği ve güvenilirliğinin önceden test edilmesinden dolayı bu sınıflama dikkate alınmıştır. Bu sınıflama, basit zihinsel düşüncelerden karmaşık zihinsel düşüncelere doğru şöyle tanımlanmıştır (Mullis ve Martin, 2015):

1. Metinde açıkça anlatılan düşünceleri bulma (doğrudan çıkarım yapma)
2. Metinde açıkça anlatılmayan düşünceleri bulma (yorumlama)
3. Metinde geçen olayları kişisel bilgi ve deneyimlerle ilişkilendirme
4. Metnin içeriğini ve dilini inceleme ve değerlendirme

Bunun yanında MEB 2019 5. sınıf Türkçe dersi öğretim programı da incelenmiş ve soruların öğretim programındaki yazma becerisi ile ilgili kazanımları yoklamasına önem verilmiştir. Bu amaçla, araştırmanın amacın uygun yedi kazanım belirlenmiştir. Bu kazanımlar aşağıda verilmiştir:

1. Bilgilendirici metin yazar.
2. Yazma stratejilerini uygular.
3. Büyük harfleri ve noktalama işaretlerini uygun yerde kullanır.
4. Yazdıklarını düzenler
5. Görüş ve düşüncelerini yazar.
6. Yazdıklarında ses olaylarına uğrayan kelimeleri doğru kullanır
7. Yazılanlara uygun geçiş ve bağlantı ifadeleri kullanır.

Soruların hazırlanmasından sonra, ilköğretim 5. sınıf düzeyi için zorluk, anlaşılabilirlik ve ölçülebilirlik bakımından soruların uygun olup olmadığını belirlemek amacıyla yukarıda belirtilen alan uzmanlarıyla birlikte üç ölçme ve değerlendirme uzmanından görüş alınmıştır

### ***Yazma etkinliğinin uygulaması***

Metinlerin uzun ve soruların zorluk düzeylerinin farklı olmasından dolayı araştırmacılar tarafından uygulamanın iki güne yayılması tercih edilmiştir. Öğrencilerin metni okuma ve soruları yanıtlama sürecine daha fazla güdülenmeleri/odaklanmaları amacıyla uygulama, Türkçe dersinde, dersin öğretmeni tarafından yapılmıştır. Öğrencilere yazma etkinliği için bir ders saati süre verilmiştir.

### ***DPA'nın geliştirilmesi***

Öğrencilerin yazılı anlatım becerilerini değerlendirmek üzere bir DPA hazırlanmıştır. DPA hazırlanırken Goodrich (2001) tarafından önerilen adımlar izlenmiştir. Bu adımlara ilişkin açıklamalar aşağıda verilmiştir.

a) Yazılı anlatım becerisinin değerlendirilmesinde kullanılacak ölçütlerin belirlenmesi:

DPA'nın geliştirilmesi sırasında ilgili alanyazın taranmış ve öğrencinin yazılı anlatım becerisini hem biçimsel hem de içerik açısından değerlendirmeye olanak sağlayan Raimes (1983)'in ölçütleri dikkate alınmıştır. Bu ölçütlere ilişkin kısa açıklamalar aşağıda verilmiştir:

1. İçerik: Sözcük, cümle, deyim vb.'nin konuyla ilişkili olması
2. Söylem Yönetimi: Konuşmayı zenginleştirecek bağlaç, deyim, açıklama bulunması
3. Anlaşılabilirlik: Sözcük ve düşünce tekrarları, anlatım bozuklukları olmaması
4. Yazım Kuralları: Yazım kurallara uyulması, sözcüklerin doğru yazılması ve ayrılması
5. Bağlama Uygun Dil Kullanma: Bağlamlara dayalı verilen göreve uygun dil kullanma
6. Dil bilgisi: Çekim ekleri, özne yüklem uyumu vs. yönden doğru olması

b) DPA türünün belirlenmesi:

DPA'lar bütünsel veya analitik türde olabilirler. Analitik DPA'lar, öğrenci performansı hakkında daha detaylı geri bildirim vermesi (Mertler, 2001) bakımından bütüncül DPA'lara göre daha sıklıkla tercih edilmektedir. Ayrıca analitik DPA'ların en önemli avantajlarından biri değerlendirme sürecinde hem puanlayıcı içi hem de puanlayıcılar arası güvenilirliği sağlama konusunda bütüncül DPA'lara göre daha iyi olmasıdır (Knoch, 2009). Bu bakımdan bu çalışmada analitik DPA tercih edilmiştir.

c) Davranış düzeylerine ilişkin tanımların yapılması:

Belirlenen her bir ölçüt yaş düzeyi de dikkate alınarak en az puandan (1-başlangıç düzeyinde), en fazla puana (4-tam olarak başarılı) doğru derecelendirilmiştir. 4 puan alan bir öğrenci ilgili ölçütte beklenen performansı tam olarak göstermiş; 1 puan alan ise gösterememiştir. Güvenilir bir puanlama için bu yaş grubunda dördüncü derecelendirme önerilmektedir (Haladyna, 1997). Ölçütler ve derecelendirme dikkate alınarak ayrıntılı tanımlamalar yapılmıştır.

d) Uzman görüşü alınması:

İlgili DPA, üç alan uzmanı, iki Türkçe öğretmeni ve üç ölçme değerlendirme uzmanına gönderilmiş ve uzmanlardan DPA'yı kapsam geçerliği (içerik, yapı ve ölçütler) bakımından "yeterli, kısmen ve yeterli değil" kategorilerinde incelemeleri istenmiştir. Uzmanlardan gelen öneriler doğrultusunda DPA tekrar gözden geçirilerek düzenlenmiştir.

### ***Yazılı anlatım becerilerinin DPA'ya göre değerlendirilmesi***

Öğrencilerin yazılı anlatım becerilerinin puanlanmasında puanlayıcıdan kaynaklı karışabilecek kişisel yanlılık, ilk izlenim ve mantık hatasını (Popham, 2000) en aza indirmek için öğretmenlerin öğrencileri tanımamasına dikkat edilmiş, öğrenciler farklı isimlerle kodlanmıştır. Öğrenciye ait yanıtlar ve DPA araştırma kapsamında yer alan öğretmenlere verilmiştir. Değerlendirme öncesi öğretmenlere araştırmanın amacı açıklanmış ve puanlamada kullanacakları DPA tanıtılmıştır. Öğretmenlerin kodlamada bir sorun yaşamaması ve veri girişinde herhangi bir hata olmaması için, her iki metin için ölçütlerin ve öğrencilere ait tanıma kodlarının yer aldığı bir veri kodlama formu öğretmenlere verilmiştir.



## Verilerin Analizi ve Yorumlanması

İlgili uzmanların geliştirilen analitik DPA'da yer alan ölçütlerin ve ölçütlere ilişkin tanımların uygunluğu konusunda yaptıkları değerlendirmeler arasındaki uyum, uyuşum yüzdesi kullanılarak hesaplanmıştır. Bu amaçla Miles-Huberman güvenilirlik formülü ve Tavşancıl ve Aslan (2001) tarafından da önerilen aşağıdaki eşitlik kullanılmıştır:

$$\text{Güvenirlik} = \frac{\text{Uzlaşma Sayısı}}{\text{Uzlaşma Sayısı} + \text{Uzlaşmama Sayısı}}$$

Bu hesaplama sonucunda elde edilen uyuşum yüzdesinin %70'in üzerinde çıkması, araştırma için güvenilir kabul edilmektedir (Miles ve Huberman;1994, Tavşancıl ve Aslan, 2001).

DPA'dan elde edilen puanların güvenilirliğini test etmek için ayrıca puanlayıcılar arasındaki tutarlılığa bakılmıştır. Ölçütlere dayalı çoklu puanlamanın yapıldığı bu tür ölçme araçlarından elde edilen puanların güvenilirliklerin belirlenmesinde, puanlayıcılar arası güvenilirlik katsayısının belirlenmesi önerilmektedir (Moskal ve Leydens, 2000). Her bir soru ve öğrencilerin testten aldıkları toplam puan için puanlayıcıların yaptıkları değerlendirmeler arasındaki tutarlılık (inter-rater) Kendall's W analiziyle test edilmiştir. Kendal's W istatistiği parametrik olmayan bir tekniktir ve ikiden fazla puanlayıcılar arasındaki tutarlılığı ortaya koymaktadır. Kendall's W katsayısı 0 ve 1 arasında değerler almaktadır. Hesaplanan değerler sıfıra yakın olması puanlayıcılar arasında tutarsızlığı, bire yakın olması ise puanlayıcılar arası tutarlılığı işaret etmektedir (Howell, 2002). Ancak puanlayıcılar arası uyuşumun en az 0.80 düzeyinde olması beklenir (Şencan, 2005).

Ayrıca aynı puanlayıcının iki farklı zamanda yaptığı puanlamalar arasında fark olup olmadığını (zamana karşı güvenilirlik) belirlemek üzere puanlayıcı içi (intra-rater) uzlaşma katsayısı hesaplanmıştır. Bu amaçla öğrenci yanıtları arasından tesadüfen seçilen bir öğrenciye ait yanıtlar, yine tesadüfen seçilen bir puanlayıcı tarafından üç hafta arayla tekrar kodlanmıştır. Bu hesaplama için de yukarıda belirtilen uyuşum yüzdesi formülü kullanılmıştır.

## Bulgular

### Geçerlik Bulguları

Geliştirilen analitik DPA'nın geçerliğini değerlendirmek üzere uzman görüşüne dayalı kanıtlar toplanmıştır. Geçerliği sağlamak üzere DPA'ya hazırlanırken sistematik geliştirme adımları izlenmiş ve sürecin ilgili aşamalarında (metin seçimi, metine dayalı soruların yazılması, DPA'da yer alan ölçütler ve bunlara ilişkin tanımlar) uzmanlardan amaca ve kapsama uygunluk bakımından görüşler alınmıştır. DPA'da yer alan ölçütlerin uygunluğu ve ölçütlere ilişkin tanımlamaların yeterliği konusunda uzmanların görüşleri arasındaki uyum yüzdeslerinin %87 ile %100 arasında değiştiği belirlenmiştir. Hesaplanan bu uyum yüzdesleri, geliştirilen DPA'nın kapsam geçerliğine kanıt olarak değerlendirilmiştir.

### Güvenirlik Bulguları

DPA'da yer alan ölçütler dikkate alınarak her bir soru için puanlayıcılar arası uyumu gösteren Kendall's W katsayıları hesaplanmış ve elde edilen değerler Tablo1'de verilmiştir.

**Tablo1. Puanlayıcılar arası Kendall's W katsayısı**

Uyuşum Katsayısı	Sorular				
	1	2	3	4	5
1. metin Kendall's W	0.866	0.917	0.919	0.811	-
2. metin Kendall's W	0.805	0.677	0.762	0.794	0.723

Tablo 1 incelendiğinde, birinci metine ait dört sorunun DPA kullanılarak puanlanmasından elde edilen uzmanlar arası uyum katsayılarının yüksek olup 0.81 ile 0.92 arasında değiştiği görülmektedir. İkinci metindeki sorular için değerlendiriciler arası uyum katsayısı ikinci soru için orta düzeyde (0.68), geriye kalan dört soru için yüksek düzeyde (0.72-0.80) bulunmuştur. Ancak birinci metinde aynı bilişsel özelliği ölçen ikinci sorunun puanlanmasına ilişkin uyum katsayısının ikinci metindeki soruya göre daha yüksek olması bu tür soruların değerlendirilmesinde DPA'nın güvenilirliğine ilişkin destekleyici bilgi vermektedir.

Her iki metinde yer alan bilişsel özellikler dikkate alındığında DPA'nın "metinde açıkça anlatılan düşünceleri bulma (doğrudan çıkarım yapma)", "metinde açıkça anlatılmayan düşünceleri bulma (yorumlama)", "metinde geçen olayları kişisel bilgi ve deneyimlerle ilişkilendirme" ve "metnin içeriğini ve dilini inceleme ve değerlendirme" becerilerini ölçen soruların puanlanmasında güvenilir sonuçlar verdiği belirlenmiştir.

Değerlendiriciler tarafından öğrencilere verilen toplam puanlar arasındaki uyum katsayıları da hesaplanmış ve ilgili bulgular Tablo2'de verilmiştir.

**Tablo 2. Toplam puanlar için hesaplanan puanlayıcılar arası Kendall's W katsayısı**

Uyuşum Katsayısı	Kensdal's W
1. metin	0.931
2. metin	0.706

Tablo 2 incelendiğinde, ikinci metin için hesaplanan puanlayıcılar arası uyum katsayısının (0.71) birinci metin için hesaplanan uyum katsayısından (0.93) daha düşük olduğu görülmektedir. Geliştirilen DPA'nın güvenilirliğine kanıt olarak ayrıca aynı puanlayıcının iki farklı zamanda yaptığı puanlamalar arasındaki tutarlılık incelenmiş ve ilgili puanlayıcının aynı öğrenciye ait yanıtları ilk ve son incelemesi arasındaki uyum yüzdesi 1. metin için %91 ve 2. metin için % 85 olarak hesaplanmıştır. Bu durum puanlamayı yapan öğretmenin kendi içinde de tutarlı olduğunun bir göstergesidir.

## Tartışma ve Sonuç

Bu araştırmada ilköğretim 5. Sınıf öğrencilerin yazılı anlatım becerilerinin değerlendirilmesi amacıyla analitik bir dereceli puanlama anahtarı geliştirilmiş ve puanlama anahtarının geçerlik ve güvenilirlik kanıtları ortaya konulmuştur (Bakınız Ek1).

Geliştirilen DPA, öğrencilerin farklı bilişsel düzeylerde yazılı anlatım becerilerini ölçmek üzere altı ölçüt içermektedir. Uzman görüşüne dayalı bulgular ilgili DPA'da yer alan ölçütlerin ve ölçütlere ilişkin açıklamaların öğrencilerin yazılı anlatım becerilerini değerlendirmek üzere yeterli ve uygun olduğunu göstermiştir. Uzmanların görüşleri arasındaki tutarlılık ilgili DPA'nın kapsam geçerlik kanıtı olmuştur. Puanlayıcılar arası güvenilirlik katsayısı Kendall's W katsayısı yoluyla hesaplanmış ve elde edilen değerler iki farklı metin türü için ve

farklı bilişsel düzeyleri ölçen soruların puanlanmasında ilgili DPA'nın ölçütlerinde puanlayıcılar arası güvenilirliğinin yüksek olduğunu göstermiştir. Ayrıca metinler için hesaplanan aynı puanlayıcının farklı zamanlarda yaptığı puanlamalar arasındaki uyum düzeyi de yüksek çıkmıştır. Bu bulgulara dayalı olarak, analitik puanlama anahtarı kullanımının puanlayıcılar arasındaki uyumsuzlukları tamamen ortadan kaldırmada yeterli olmadığı; ancak puanlamalar arasında tutarlılığı arttırarak objektiflik düzeyini arttırdığı yorumu yapılabilir. Bu konuda yapılan çalışmalar (Bıkmaz Bilgen ve Doğan, 2017; Büyükkıdık ve Anıl, 2015; Knoch, 2009) analitik dereceli puanlama anahtarları kullanmanın puanlayıcılar arasındaki güvenilirliği arttırdığını göstermektedir.

Ancak genel olarak 1. metin için hesaplanan güvenilirlik katsayılarının 2. Metin'e göre daha yüksek olduğu görülmüştür. Her ne kadar ölçütlere dayalı tanımlamaların uygunluğu uzman görüşlerine dayalı olarak değerlendirilse de ilgili ölçütlere öğrenciden beklenen performansın çerçevesi puanlamayı yapan öğretmenin konuyla ilgili bilgi düzeyine, konuyu algılayışına ve yorumlayışına göre değişmesi kaçınılmaz olacaktır. Bilindiği gibi puanlayıcı etkisi bu tutarsızlığın ortaya çıkmasında en önemli etkenlerden biridir (Cronbach, 1990). İlgili araştırma bulguları, puanlayıcı katılığı/ cömertliği veya halo etkisi gibi puanlayıcı etkisini ortadan kaldırma konusunda dereceli puanlama anahtarlarının yetersiz kaldığını göstermiştir (Güler ve Gelbal, 2011; İlhan, 2015). Puanlayıcı etkisi yanında bu durumun ortaya çıkmasında metin türlerindeki farklılık da önemli bir etken olabilir. Örneğin 1. metin'in didaktik ve bilgi verici türde bir metin olması, ilgili ölçütlerin öğretmenler tarafından daha nesnel anlaşılmasını sağlamış olabilir. 2. metin ise hikâye türünde yazılmış olması ve ilgili soruların öğrenci kestirimlerini-değerlendirmelerini gerektirdiği bir yapıda olması puanlayıcıların ölçüt tanımlarını farklı algılamasına neden olabilir. Bıkmaz Bilgen ve Doğan (2017) tarafından yapılan çalışmada da analitik puanlama anahtarlarının güvenilirliği incelenmiş ve puanlayıcı-ölçüt arasındaki etkileşimin puanlayıcılar arasındaki tutarsızlığa neden olduğu vurgusu yapılmıştır. Bu konuda yapılan başka çalışmalarda da (Köse, Usta ve Yandı, 2016; Yüzüak, Yüzüak ve Kaptan, 2015) puanlayıcılardan kaynaklı tutarsızlıklar vurgulanarak puanlayıcıların bazen nesnel bazen yanlı olabileceği belirtilmiştir.

Bu araştırma bulgularına dayalı olarak uygulayıcılara ve diğer araştırmacılara önerilerde bulunulmuştur. İlk olarak geliştirilen DPA, kompozisyon ve hikâye yazma gibi daha uzun metinlere dayalı yazılı anlatım türlerinde denenmediği için kullanıcıların DPA'yı bu tür yazılı anlatımlarda tekrar test ederek daha dikkatli kullanması önerilmektedir. Öte yandan özellikle 2. metin için ortaya çıkan puanlayıcılar arasındaki uyumsuzluğun nedenlerini (ölçüt-puanlayıcı etkileşimi) daha iyi değerlendirebilmek için uyuşum katsayılarına göre daha detaylı bilgi veren log linear analiz veya çok düzeyli Rasch modellerin kullanılması önerilebilir. Bu çalışmada ilgili DPA'nın güvenilirliği puanlayıcılar arası ve puanlayıcı içi uyuşum katsayıları yoluyla hesaplanmıştır. Ayrıca Genellenebilirlik Kuramı ve değişmezlik (invariance) çalışmalarıyla bu DPA'nın geçerlik ve güvenilirliğine ilişkin başka kanıtlar da elde edilebilir.

## Kaynakça

- Ağın Haykır, H. (2012). *İlköğretim 6, 7 ve 8. sınıf öğrencilerinin okuduğunu anlama becerisi ile yazılı anlatım becerisi arasındaki ilişki*, Yayınlanmamış yüksek lisans tezi, Ahi Ervan Üniversitesi, Sosyal Bilimler Enstitüsü, Kırşehir.
- Aktaş, M ve Alıcı, D. (2018). Yazılan hikâyeyi değerlendirmeye yönelik analitik rubrik geliştirme: geçerlik ve güvenilirlik çalışması. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 14(2), 597-610. DOI: 10.17860/mersinefd.424198.
- Akyol, H. (2006). *Yeni programa uygun Türkçe öğretim yöntemleri*. Ankara: Kök Yayıncılık.
- Arter, J. A.ve McTighe, J. (2000). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance* (Experts in Assessment Series). California: Corwin Press.
- Atesman, E. (1997). Türkçede okunabilirliğin ölçülmesi. *Dil Dergisi*, 58, 71-74.
- Bıkmaz Bilgen, Ö. ve Doğan, N. (2017). Puanlayıcılar arası güvenilirlik belirleme tekniklerinin karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 8(1), 63-78. DOI: 10.21031/epod.294847.
- Bozkurt, Ü. B.ve Arıca-Akkök, E. (2019). Anadili Türkçe olan yetişkin konuşucular için konuşma becerisi derecelendirme ölçeğinin geliştirilmesi. *İlköğretim Online*, 18(1), 416-436. DOI: 10.17051/ilkonline.2019.527649.
- Büyükkıdık, S ve Anıl, D. (2015). Performansa dayalı durum belirlemede güvenirliliğin genellenebilirlik kuramında farklı desenlerde incelenmesi. *Eğitim ve Bilim*, 40 (177), 285-296. DOI: 10.15390/EB.2015.2454.
- Coşkun, E. (2007). Yazma becerisi. A. Kırkkılıç ve H. Akyol (Ed.), *İlköğretimde Türkçe öğretimi* (Sayfa: 49-91). Ankara: Pegem A Yayıncılık
- Christie, J., Enz, B. ve Vukelich, C. (2003). *Teaching language and literacy*. (Second edition). Boston: Pearson Education Inc.
- Cronbach, L.I. (1990). *Essentials of psychological testing*. New York: Harper and Row.
- Çeçen, A. (2011). Yazma Eğitimi. M. Özbay (Ed.), *Yazma eğitimi açısından metin bilgisi*. Ankara: PegemA Yayıncılık.
- Çetinkaya, G. (2010). *Türkçe metinlerin okunabilirlik düzeylerinin tanımlanması ve sınıflandırılması*, Ankara Üniversitesi Sosyal Bilimler Enstitüsü, Doktora Tezi, Ankara.
- Dunbar, N.E., Brooks, C. F ve Miller, K. T.( 2006). Oral communication skills in higher education: Using a performance-based evaluation rubric to assess communication skills. *Innovative Higher Education*, 31(2), 115–28. DOI: /10.1007/s10755-006-9012-x
- Elma, C. ve Bütün, E. (2015). İlkokul ve ortaokul öğrencilerinin yazılı anlatım becerilerine ilişkin öğretmen görüşleri. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 15 (2),104-131. DOI: 10.17240/aibuefd.2015.15.2-5000161315
- Epçaçan, C. ve Erzen, M. (2008). İlköğretim Türkçe dersi öğretim programının değerlendirilmesi, *Uluslararası Sosyal Araştırmalar Dergisi*, 1 (4),182-202.
- Evans, J. (2001). Introduction: Learning and teaching the complexities of writing. Evans, Janet (Ed.), *Writing in the Elementary Classroom: A Reconsideration*, Portsmouth, NH: Heinemann.
- Flower, L. Ve Hayes, J.R. (1981) A cognitive process theory of writing. *College Composition and Communication*, 32, 365-387. <http://dx.doi.org/10.2307/356600>
- Goodrich, H. (2001). The effects of instructional rubrics on learning to write. *Current Issues in Education*, 4(4), 1-21. İnternet'ten 24 Mayıs 2019'da [http://scholarsarchive.library.albany.edu/etap\\_fac\\_scholar/6](http://scholarsarchive.library.albany.edu/etap_fac_scholar/6) adresinden alınmıştır.
- Göçer, A. (2005). *İlköğretim ikinci kademe Türkçe öğretiminde ölçme ve değerlendirme*. Yayınlanmamış Doktora Tezi, Atatürk Üniversitesi, Sosyal Bilimler Enstitüsü, Erzurum.

- Göçer, A. (2011). Öğrencilerin yazılı anlatım çalışmalarının Türkçe öğretmenlerince değerlendirilmesi üzerine. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi*, 30(2), 71-97. İnternet'ten 10 Mayıs 2019'da <https://dergipark.org.tr/download/article-file/188013> adresinden alınmıştır.
- Güler, N. ve Gelbal, S. (2010). Klasik test kuramı ve çok değişkenlik kaynaklı Rasch modeli üzerine bir çalışma. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 38, 108-125. İnternet'ten 21 Mart 2019'da [http://www.aniyayincilik.com.tr/main/pdfler/38/7\\_guler\\_nese.pdf](http://www.aniyayincilik.com.tr/main/pdfler/38/7_guler_nese.pdf) adresinden alınmıştır.
- Haladyna, T. M. (1997). *Writing Test Item to Evaluate Higher Order Thinking*. USA: Allyn & Bacon.
- Hillocks, G. J. (1987). Synthesis of research on teaching writing. *Educational Leadership*, 44 (8), 71-82.
- Howell, D.C. (2002). *Statistical methods for psychology* (5th ed.). Pacific Grove CA: Duxbury.
- Hyland, K. (2002). Authority and invisibility: authorial identity in academic writing. *Journal of Pragmatics*, 34(8), 1091-1112. Doi: 10.1016/S0378-2166(02)00035-8
- İlhan, M. (2015). *Standart ve SOLO taksonomisine dayalı rubrikler ile puanlanan açık uçlu matematik sorularında puanlayıcı etkilerinin çok yüzeyli rasch modeli ile incelenmesi*. Yayımlanmamış doktora tezi, Gaziantep üniversitesi, Eğitim Bilimleri Enstitüsü, Gaziantep.
- Karadağ, R., ve Kayabaşı, B. (2011). Yazılı anlatım becerilerinin geliştirilmesinde okuma metinleri olarak gazete köşe yazılarının kullanılması. *Turkish Studies-International Periodical for the Languages, Literature and History of Turkish or Turkic*, 6(3), 989-1010. İnternet'ten 15 Mayıs 2019'da [https://www.researchgate.net/profile/Ruhan\\_Karadag/publication/271022310\\_Using\\_Newspapers\\_to\\_Improving\\_Writing\\_Skills/links/563894fd08ac7f7eb1859a6b.pdf](https://www.researchgate.net/profile/Ruhan_Karadag/publication/271022310_Using_Newspapers_to_Improving_Writing_Skills/links/563894fd08ac7f7eb1859a6b.pdf) adresinden alınmıştır.
- Karakoç Öztürk, B. (2014). Türkçe öğretmenlerinin yazma eğitimine ve öğrencilerin öykü yazma becerilerine ilişkin görüşleri. *International Journal of Language Academy*, 2(4), 170-194.
- Karakoç Öztürk, B. (2016). Yazma becerisine ilişkin süregelen uygulamaların türkçe dersi öğretim programı (2006) çerçevesinde değerlendirilmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 16(4), 1921-1945. İnternet'ten 15 Mayıs 2019'da <https://dergipark.org.tr/download/article-file/291948> adresinden alınmıştır.
- Karatay, H. (2011). 4+1 Planlı yazma ve değerlendirme modelinin öğretmen adaylarının yazılı anlatım tutumlarını ve yazma becerilerini geliştirmeye etkisi. *International Periodical For the Languages, Literature and History of Turkish or Turkic*, 6(3), 1029-1047. İnternet'ten 24 Mart 2019'da <http://www.turkishstudies.net/DergiTamDetay.aspx?ID=2622> adresinden alınmıştır.
- Karatay, H. (2013). Süreç temelli yazma modelleri: 4+1 Planlı yazma ve değerlendirme modeli. M. Özbay (Ed.) *Yazma eğitimi* (s. 21-40). Ankara: Pegem.
- Knoch, U. (2009). Diagnostic assessment of writing: A comparison of two rating scales. *Language Testing*, 26(20), 275-304. DOI: 10.1177/0265532208101008
- Köse, İ. A., Usta, H. G. ve Yandı A. (2016). Sunum yapma becerilerinin çok yüzeyli rasch analizi ile değerlendirilmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 16(4), 1853-1864. İnternet'ten 12 Mart 2019'da <https://dergipark.org.tr/download/article-file/291940> adresinden alınmıştır.
- Lan, Y., Hung, C. ve Hsu, H. (2011). Effects of guided writing strategies on students' writing attitudes based on media richness theory. *TOJET: The Turkish Online Journal of Educational Technology*, 10(4), 148-164. İnternet'ten 19 Mart 2019'da <https://files.eric.ed.gov/fulltext/EJ946620.pdf> adresinden alınmıştır.
- MEB (2019). Türkçe dersi öğretim programı (1-8. Sınıflar). İnternet'ten 12 Mart 2019'da <http://mufredat.meb.gov.tr/ProgramDetay.aspx?PID=663> adresinden alınmıştır.
- Mertler, C. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25). İnternet'ten 14 Nisan 2019'da <http://PAREonline.net/getvn.asp?v=7&n=25> adresinden alınmıştır.

- Miles, M. B. ve Huberman, A. M. (1994). *Qualitative data analysis*. (2. baskı). Thousand.
- Moskal, B. M. ve Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation*, 7(10).
- Mullis, I. V. S. ve Martin, M. O. (Eds.). (2015). PIRLS 2016 Assessment Framework (2nd ed.). İnternet'ten 21 Mart 2019'da <http://timssandpirls.bc.edu/pirls2016/framework.html> adresinden alınmıştır.
- Norman, K. A. ve Spencer, B. H. (2005). Our lives as writers: examining preservice teachers' experiences and beliefs about the nature of writing and writing instruction, *Teacher education quarterly*, 32 (1), 25-40.
- Okurer, C. (1967). *Kompozisyon öğretimi bakımından İngiltere maarifi hakkında bir inceleme (Neticeler ve Bazı Teklifler)*. İstanbul: Milli Eğitim Basımevi.
- Onursal, D. (2003). Türkçe metinlerde bağdasıklık ve tutarlılık. A. Kıran, E. Korkut, S. Ağıldere (Yayına Hazırlayanlar), *Günümüz dilbilim çalışmaları* (Sayfa: 121-132). İstanbul: Multilingual Yayınları.
- Popham, W. J. (1997). What's wrong-and what's right-with rubrics. *Educational Leadership*, 55(2), 72-75.
- Popham, W.J. (2000). *Modern educational measurement: Practical guidelines for educational leaders* (3rd Ed.). Boston: Allyn and Bacon.
- Raimes, A. (1983). *Techniques in teaching writing*. Oxford: Oxford University Press.
- Sencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik* (1. Baskı). Ankara: Seçkin Yayıncılık.
- Şengül, M.(2011). *İlköğretim II. kademe Türkçe öğretiminde yazma becerilerine yönelik öğretim ve ölçme-değerlendirme yaklaşımlarının uygulamadaki etkililiğinin değerlendirilmesi*. Yayınlanmamış doktora tezi, Fırat Üniversitesi, Eğitim Bilimleri Enstitüsü, Elazığ.
- Tavşancıl, E. & Aslan E. (2001). *İçerik analizi ve uygulama örnekleri*. İstanbul: Epsilon Yayınları.
- TED (2009). Öğretmen yeterlikleri - özet rapor. Ankara: Türk Eğitim Derneği Yayınları.
- Tok, R. ve Erdoğan, Ö. (2017). İlkokul 2. 3. ve 4. sınıf öğrencilerinin yazma becerilerinin incelenmesi. *YYÜ Eğitim Fakültesi Dergisi (YYU Journal of Education Faculty)*, 14(1), 1003-1024.
- Topuzkanamış, E. (2014). *Yazma stratejileri öğretiminin Türkçe öğretmenliği birinci sınıf öğrencilerinin yazılı anlatım başarısı ve yazma kaygısına etkisi*. Yayınlanmamış doktora tezi, Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Ungan, S. (2007). Yazma becerisinin gelişimi ve önemi. *Sosyal Bilimler Enstitüsü Dergisi*, 23(1), 461-472.
- Ülper, H. (2009a). *Yazma becerilerini ölçmede (notlandırmada) kullanılan farklı ölçme (notlandırma) yaklaşımlarının geçerlikleri üzerine kompozisyon yazmanın amaçları bağlamında karşılaştırmalı bir inceleme*. I. Ulusal Eğitimde ve Psikolojide Ölçme ve Değerlendirme Kongresi (14-16 Mayıs, 2008), Ankara, Türkiye.
- Ülper, H. (2009b). *Bilişsel süreç modeline göre hazırlanan yazma öğretimi izlencesinin öğrenci başarısına etkisi*. Yayınlanmamış doktora lisans tezi. Ankara Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara
- Wiener, H. S. (2003). *Any child can write*, New York: Oxford University Press.
- Yıldız, S. (2018). Akademik dinleme becerisi yeterliliği ölçęi: Geçerlik ve güvenilirlik çalışması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18 (2), 1210-1230.
- Yüzüak, A. V., Yüzüak, B. ve Kaptan, F. (2015). Performans görevinin akran gruplar ve öğretmen yaklaşımları doğrultusunda çok-yüzeyle rasch ölçme modeli ile analizi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(1), 1-11. DOI: doi.org/10.21031/epod.57425
- Zamel, V. (1982). Writing: the process of discovering meaning. *TESOL QUARTERLY*, 16 (2), 195-209. DOI: doi.org/10.2307/3586792

## Extended Abstract

### Introduction

Writing skills are related to different life skills and have an important role in an individual's success. In order to develop this skills, the place of measurement and evaluation methods is important. Studies indicated that teachers have significant deficiencies in the development and the use of appropriate and qualified tools for measuring and evaluating student products. In writing-based assessments, teachers prefer to use criteria that evaluate the visual aspect of their writing rather than standard criteria. In addition, teachers give many writing assignments to their students, but they do not give enough information about how to write assignments and do not provide enough guidance for them. Also, teachers do not give sufficient feedback during the assessment of the writing skills. It is correct to use rubrics in the assessment of writing skills in order to include evaluation criteria related to individuals' performances and also definitions related to these criteria. Through clearly defined criteria in rubrics, students can see the strengths and weaknesses of their products, and teachers can give feedback to them and their families. In the literature, it was determined that there were rubrics for reading, listening and speaking skills within the scope of measuring language skills of students, but there was no Turkish study found to evaluate their written skills according to different cognitive levels and different text types. In the study, it is aimed to develop a rubric that will help the students at the 5th grade to evaluate their writing skills and to provide supportive feedback to the students.

### Methodology

The data were obtained from 22 Turkish teachers working in a public school in Ankara and Denizli and 10 students at the 5th grade level. Two texts were selected in order to measure the writing skills of the students. The opinions of three Turkish language experts and two Turkish teachers about the appropriateness of the texts for the students' cognitive development, grade and age level were obtained. The experts were asked whether the questions were appropriate in terms of difficulty, clarity and measurability. During the development process of rubric, the steps of determining criteria, type, definitions of behavior levels and obtaining expert opinions were followed respectively. The coefficients of agreement between the experts were calculated and the scale was found to be reliable for the criteria and definitions related to the rubric. In addition, in order to test the reliability of the scores obtained from the rubric, the correlation coefficient between the raters was calculated by Kendall's W statistic. The difference between the scores made by the same rater at two different times was examined in order to support the reliability of scale, in other words intra-rater agreement was calculated.

### Findings

Evidence based on expert opinion was collected to assess the validity of the rubric. During the the rubric development stage, a systematic path was followed. The experts stated that definitions related to the criteria was revealed to be adequate and the correlation coefficient between experts was high. When Kendall's W coefficients were analyzed for each question, the agreement between the raters was found to be high for the first and second texts, but the coefficient of agreement was lower for the second question in the second text. As another proof

of the reliability of the developed rubric, the consistency between the scores made by a rater at two different times was found to be high.

### **Discussions**

The developed rubric includes six criteria to measure writing skills of students at different cognitive levels. The findings based on expert opinions showed that the criteria and explanations related to the criteria of the relevant rubric were sufficient and appropriate to evaluate the students' writing skills. The level of consistency between the raters' scores and between scores made by the same rater at different times were found to be enough high. Based on these findings, it was found that analytical rubric was not sufficient to completely eliminate discrepancies between raters; however, it can be said that it increases the level of objectivity by increasing the consistency between the raters.

It is inevitable that the performance level expected from the student will change according to the knowledge level, perception and interpretation of the teachers. Rubrics have some shortcomings in eliminating the rater effects, such as rater rigidity / generosity and halo effects. Besides the rater effect, the difference in text types can be an important factor in the emergence of this situation. The fact that text 1 is a didactic and informative text may have contributed to a more objective understanding of the criteria by teachers.



## Ek1 Yazılı Anlatım Analitik Dereceli Puanlama Anahtarı

Ölçütler	Başlangıç Düzeyinde (1)	Geliştirilmesi Gerekir (2)	Kabul Edilebilir (3)	Tam Olarak Başarılı (4)	Başarı Puanı
<b>İçerik</b>	Yazılanlar (seçilen sözcükler, cümleler, deyimler, örnekler vb.) verilen konuyla ya da olayla ilişkili değildir ve birbirini desteklememektedir. Belirlenen bağlamlarda yazılanların neredeyse tümü, hatalı ve iletisini aktarmada yetersizdir.	Yazılanlar (seçilen sözcükler, cümleler, deyimler, örnekler vb.) verilen konuyla ya da olayla çok az ilişkili ve yeteri kadar destekleyici değildir. Belirlenen bağlamlarda yazılanların çok azı, doğru ve iletisini aktarmada yeterlidir.	Yazılanlar (seçilen sözcükler, cümleler, deyimler, örnekler vb.) verilen konuyla ya da olayla oldukça ilişkili ve birbirini yeterince destekleyicidir. Belirlenen bağlamlarda yazılanların çoğu doğru ve iletisini aktarmada yeterlidir.	Yazılanlar (seçilen sözcükler, cümleler, deyimler, örnekler vb.) verilen konuyla ya da olayla tam olarak ilişkili ve birbirini destekleyicidir. Belirlenen bağlamlarda yazılanların tümü doğru ve iletişimi aktarmada yeterlidir.	
<b>Söylem Yönetimi</b>	Belirlenen bağlamların neredeyse tümünde yazmayı zenginleştirecek bağlayıcılar, dilsel ifadeler, açıklamalar bulunmamaktadır..	Belirlenen bağlamların çok azında yazmayı zenginleştirecek bağlayıcılar, dilsel ifadeler, açıklamalar bulunmaktadır.	Belirlenen bağlamların çoğunda yazmayı zenginleştirecek bağlayıcılar, dilsel ifadeler, açıklamalar bulunmaktadır.	Belirlenen bağlamların çoğunda yazmayı zenginleştirecek bağlayıcılar, dilsel ifadeler, açıklamalar bulunmaktadır.	
<b>Anlaşılabilirlik</b>	Yazının her bölümünde; sözcük ve düşünce tekrarlarına, anlatım bozukluklarına rastlanmaktadır.	Yazının bazı bölümlerinde sözcük ve düşünce tekrarlarına, anlatım bozukluklarına rastlanmaktadır.	Yazının bir bölümünde sözcük ve düşünce tekrarlarına, anlatım bozukluklarına rastlanmaktadır.	Yazının bölümlerinde sözcük ve düşünce tekrarlarına, anlatım bozukluklarına rastlanmamaktadır.	
<b>Bağlama Uygun Dil Kullanma</b>	Belirlenen bağlamların neredeyse tümünde verilen göreve uygun dil kullanımını sergileyemiyor.	Belirlenen bağlamların çok azında verilen göreve uygun dil kullanımını sergileyebiliyor.	Belirlenen bağlamların çoğunda verilen göreve uygun dil kullanımını sergileyebiliyor.	Belirlenen bağlamların tümünde verilen göreve uygun dil kullanımını sergileyebiliyor.	
<b>Yazım Kuralları</b>	Yazım ve noktalama kurallarına, sözcük yazımlarına, hiç dikkat etmemiştir.	Yazım ve noktalama kurallarına genellikle uyulmamış, sözcük yazımlarına yeterince dikkat etmemiştir.	Yazım ve noktalama kurallarına genellikle uyulmuş, sözcükleri doğru yazmış ve ayırmıştır.	Yazım ve noktalama kurallarına uyulmuş, sözcükleri doğru yazmış ve ayırmıştır.	
<b>Dil Bilgisi</b>	Düzeyine uygun kurduğu tümceler neredeyse tümü, dil bilgisel yönden hatalıdır.	Düzeyine uygun kurduğu tümceler çok azı dil bilgisel yönden doğrudur.	Düzeyine uygun kurduğu tümceler çoğu dil bilgisel yönden doğrudur.	Düzeyine uygun kurduğu tümceler tümü dil bilgisel yönden doğrudur.	