

Okul Öncesi Dönemdeki Çocukların İletişim Becerilerinin Ölçülmesinde Analitik ve Bütüncül Dereceli Puanlama Anahtarlarının Karşılaştırılması

The Comparison of Analytic and Holistic Rubrics on Preschool Student's Communication Skill

Sema Sulak

Bartın Üniversitesi Eğitim Fakültesi Bartın

Bilge Akyüz

Dumlupınar Üniversitesi Eğitim Fakültesi Kütahya

Şule Ballı

Hacettepe Üniversitesi Eğitim Fakültesi Ankara

İlk Geliş Tarihi: 22.06.2016

Yayına Kabul Tarihi: 23.12.2016

Özet

Eldeki araştırmada, okul öncesi dönemdeki çocukların iletişim becerilerinin ortaya çıkarılmasında analitik ve bütüncül dereceli puanlama anahtarlarının etkililiği incelenmiştir. Bu amaçla, araştırmacı tarafından iletişim becerisini ortaya koyan ölçütler ve puanlama sistemi oluşturularak; analitik ve bütüncül olmak üzere iki farklı dereceli puanlama anahtarı geliştirilmiştir. Hazırlanan ölçme araçları, Bartın ili merkezde yer alan bir anaokulunda görev yapan üç okul öncesi öğretmeni tarafından kendi öğrencilerini değerlendirmek üzere uygulanmıştır. Analitik ve bütüncül dereceleme puanlarının farklılığını test etmek için Wilcoxon testine başvurulmuştur ve puan ortalamaları arasında anlamlı farklılık olmadığı görülmüştür. Analitik ve bütüncül dereceli puanlama anahtarlarının test-tekrar teste dayalı güvenilirlikleri hesaplanmıştır. Bu amaçla hazırlanan ölçme araçları iki hafta arayla aynı öğretmenler üzerinde uygulanmıştır. Hesaplanan korelasyon katsayılarının anlamlılığını test etmek için Fisher'in z testi kullanılmıştır. Analiz sonucunda analitik bütüncül puanlama anahtarının bütüncül dereceli puanlama anahtarından daha güvenilir olduğu bulunmuştur.

Anahtar Kelimeler: Analitik, bütüncül, dereceli puanlama anahtarı

Özet

In this study, the effectiveness of analytic and holistic rubrics were investigated in order to discover the communication skills of preschool students. In order to accomplish this task two different rubrics, analytic and holistic, were developed by the researchers. These instruments were given to preschool teachers working in Bartın to evaluate their students, and it was administered in two weeks intervals. To test the differences between analytic and holistic rubrics, Wilcoxon test was used, and it was found that there was no differences between mean scores. Additionally, Fisher's z was utilized to test the significance of correlation coefficients. The study

results revealed that analytic rubrics yielded more reliable results than holistic rubrics.

Keywords: *Analytic, holistic, rubric*

1. Giriş

Okul öncesi dönem temel becerilerin kazandırıldığı kritik bir dönemdir. Bu becerilerden birisi iletişim becerisidir ve iletişim becerilerinin geliştirilmesine katkı sağlaması bakımından doğru ölçülmesi önemlidir. Becerin ölçülmesi performans değerlendirme başlığı altında ele alınmaktadır; yani, üründen ziyade sürecin ölçülmesini içermektedir. Performansın ölçülmesinde dereceli puanlama anahtarı kullanılmaktadır. Dereceli puanlama anahtarı, performansa ait ölçütleri ve bu ölçütlere göre performansa ilişkin yargıya varmada kullanılan puanlama aracıdır (Haladyna, 1997; Moskal, 2000; Kan, 2007). Dereceli puanlama anahtarı, derecelendirme ölçütü, ölçütlerin tanımlanması ve puanlama stratejisi bölümlerinden oluşmaktadır (Popham, 1997). Derecelendirme ölçütünde, ölçülmek istenen performansı tanımlayan kriterler belirlenir, ölçütlerin tanımlanmasında tanımlanan kriterlere ilişkin açıklamalara yer verilir ve puanlama stratejisinde ise tanımlanan ölçütlerin eşleştirileceği puan tanımlamaları yapılır. Dereceli puanlama anahtarı puanlama stratejisine göre analitik ve bütüncül olarak ikiye ayrılmaktadır.

1) Analitik Dereceli Puanlama Anahtarı: Analitik dereceli puanlama anahtarı adından da anlaşılacağı gibi analitik seviyeleri değerlendirmek için uygundur. Öğrencilerin ürünlerindeki alt yetenekleri de göz önünde bulundurduğu için daha detaylı bir değerlendirme sağlar (Moskal, 2000). Bu puanlama anahtarında ölçülmek istenen özellik alt boyutlara parçalanarak; her bir boyut ayrı ayrı puanlanır. Sadler (2009)'in belirttiğine göre; analitik dereceli puanlama anahtarında öğrenci her puan seviyesinde kendinden ne beklediğini bilir.

2) Bütüncül Dereceli Puanlama Anahtarı: Bu puanlama anahtarında puanlayıcı bütün değerlendirme ölçütlerini bir arada göz önünde bulundurarak, bu ölçütleri tek bir ölçüt haline getirir (Popham, 1997). Bu puanlama anahtarında ölçülmek istenen özellik bir bütün olarak ele alınır (Carr, 2000). Bütüncül dereceli puanlama anahtarı, ortaya konan ürünlerin özelliklerini daha önceden belirlenmiş ölçütlere göre puanlar ve yüzeysel olarak performansı belirler (Gunning, 2006'den akt. Beyreli ve Arı, 2009). Bütüncül dereceli puanlama anahtarının eleştirilen yanı tüm ürüne genel olarak bir puan vermesidir. Hazırlanmasının kolay ve hızlı olmasının yanı sıra, detaylı açıklamalar olmadığı için kullanışlıdır (Moskal, 2003).

Moskal (2003) analitik dereceli puanlama anahtarının performansı bölümlere ayrılarak puanlandığını ve bütüncül dereceli puanlama anahtarının geniş bir ölçüt olarak performansının bütününe puanlandığını; kullanım alanına göre her ikisinin de işe koşulabilir ancak, hangisinin daha güvenilir olduğunun tartışmalı bir konu olduğunu belirtmiştir. Bunun yanı sıra, değerlendirme yaklaşımları arasındaki farklılıkların dik-

kate alınması ve ne zaman hangi yaklaşımın kullanılacağına karar verilmesi gerektiğini ifade etmiştir.

Birçok araştırma, dereceli puanlama anahtarlarının öğrenci ürünlerini zenginleştirdiğini ve güçlendirdiğini göstermektedir. Fakat bu puanlama araçlarının geçerliği ve güvenilirliği tartışılan bir konudur (Rezaei, Lovorn, 2010). Analitik ve bütüncül dereceli puanlama anahtarlarının karşılaştırılmasına ilişkin araştırmalara bakıldığında; Ghalib ve Al-Hattami (2015) İngilizcenin yabancı dil olarak öğretildiği sınıflarda İngilizce yazma dersinde öğrencilerin yazma seviyelerini belirlemek için analitik ve bütüncül dereceli anahtarlarını karşılaştırmıştır. Çalışmanın sonuçlarına göre; analitik puanlama anahtarının bütüncül puanlama anahtarından daha güvenilir olduğu sonucuna varılmıştır.

Şanlı (2010) bilimsel süreç becerilerinin ölçülmesinde bütünsel ve analitik puanlama anahtarlarının güvenilirliklerini karşılaştırmıştır. Analitik puanlama anahtarına ait tutarlılığın bütüncül puanlama anahtarına göre daha yüksek olduğu saptanmıştır.

Wiseman (2012), İngilizcenin ikinci dil olarak öğretildiği sınıflarda yazma performansının değerlendirilmesinde analitik ve bütüncül dereceli puanlama anahtarlarının karşılaştırmasını gerçekleştirmiştir. Araştırma sonunda, her iki dereceli puanlama anahtarının öğrencileri yeteneklerine göre ayırabildiği bulunmuştur. Fakat analitik dereceli puanlama anahtarının bütüncüle kıyasla daha detaylı ve hassas değerlendirmeye imkan sağladığı sonucuna varılmıştır.

Anıl ve Şahhüseyinoğlu (2008), analitik ve bütüncül dereceli puanlama anahtarını karşılaştırmalı olarak incelemiştir. Türkçe dersi okuma becerilerini ölçmeye yönelik hazırlanmış bütüncül ve analitik dereceli puanlama anahtarından elde edilen puanlar arasında anlamlı bir fark bulunamamıştır (akt:Şanlı,2010).

Carr (2000) İngilizce seviye tespit sınavını analitik ve bütüncül dereceli puanlama anahtarına göre puanlamıştır ve öğrencilerin yerleştikleri sınıf düzeyini yordama güçlerini karşılaştırmalı olarak incelemiştir. Öğrencinin toplam İngilizce yeterlik puanı üzerinde analitik veya bütüncül dereceli puanlama anahtarına göre puanlanmış olmasının etkili olmadığı sonucuna varılmıştır.

Klein ve diğerleri (1998), fen dersinde üç adet fen projesinin analitik ve bütüncül dereceli puanlama anahtarına göre değerlendirilmesini karşılaştırmıştır. Araştırma sonuçlarına göre, analitik dereceli puanlama anahtarının puanlanmasının daha uzun zaman aldığı ve daha tutarlı sonuçlar ortaya koyduğu bulunmuştur. Proje değerlendirme puanlarının analitik ve bütüncül dereceli puanlama anahtarına göre farklılık göstermediği görülmüştür.

Bugüne kadar yapılan çalışmalar dereceli puanlama anahtarı kullanmanın olumsuz sonuçlar ortaya koymadığını göstermiştir (Klein ve diğ, 1998, Şanlı, 2010; Wiseman, 2012;). Çalışmalarda analitik ve bütüncül dereceli puanlama anahtarının iç tutarlılığı

ve puanlayıcılar arası uyuma dayalı güvenilirlikleri incelenmiştir. Eldeki çalışmada ise farklı olarak test-tekrar teste dayalı puanlama güvenilirliği belirleme yöntemi incelenmiştir ve aşağıdaki sorulara cevap aranmıştır:

Araştırmanın problemi

Okul öncesi dönemdeki çocukların iletişim becerilerinin ölçülmesinde bütüncül ve analitik puanlama anahtarı ile elde edilen puanların ortalamaları ve güvenilirlik katsayıları arasında fark var mıdır?

Alt problemler

a) Okul öncesi dönemdeki çocukların iletişim becerilerinin ölçülmesinde bütüncül ve analitik puanlama anahtarı ile elde edilen puanların ortalamaları arasında fark var mıdır?

b) Okul öncesi dönemdeki çocukların iletişim becerilerinin ölçülmesinde bütüncül ve analitik puanlama anahtarlarının güvenilirlik katsayıları arasında fark var mıdır?

Araştırmanın Amacı

Eldeki çalışmada okul öncesi dönemdeki çocukların iletişim becerilerinin ölçülmesinde analitik ve bütüncül dereceli puanlama anahtarlarının karşılaştırmalı olarak incelenmesi amaçlanmıştır. Yapılan çalışmayla dereceli puanlama anahtarlarının beceriyi ölçmede bir farklılık ortaya koyup koymadığı belirlenmiştir. Ayrıca, güvenilirlik bakımından analitik ve bütüncül dereceli puanlama anahtarları arasında anlamlı bir fark olup olmadığı saptanmaya çalışılmıştır.

2. Yöntem

Araştırmanın verileri, Bartın ili merkezde yer alan bir anaokulunda görev yapan üç okul öncesi öğretmenini ve bu öğretmenlerin sınıfında bulunan 60-72 aylık toplam 56 öğrenciden elde edilmiştir.

Veri toplama aracı olarak; araştırmacı tarafından iletişim becerisini ölçmeye yönelik analitik ve bütüncül dereceli puanlama anahtarı geliştirilmiştir. Bu amaçla, 2015-2016 öğretim yılı Okul Öncesi Eğitim Programı incelenmiştir. Araştırmacı tarafından eğitim programındaki 60-72 aylık çocuklar için belirlenen “bilişsel gelişim, sosyal-duygusal gelişim ve dil gelişimi” alanlarında iletişim becerisini ölçmeye yönelik kazanımlar belirlenmiştir. Belirlenen kazanımların iletişim becerisini ölçmek için uygun olup olmadığını tespit etmek amacıyla uzman görüşü alınmıştır. Hazırlanan uzman değerlendirme formunda, araştırmacı tarafından seçilen kazanımlara ve “uygun, uygun değil” şeklinde iki kategorili puanlamaya yer verilmiştir. Bir psikolojik danışmanlık ve rehberlik alanından ve bir okul öncesi öğretmenliği alanından olmak üzere toplam iki akademisyenden belirlenen kazanımlardan gözlemlenebilir olan ve öğrencilerin iletişim becerisini yansıtan kazanımları belirlemeleri istenmiştir. Uzman

görüşleri doğrultusunda Temel İletişim İlke ve Teknikleri, Kendini İfade Etme ve Kişiler Arası İlişkiler alt başlıklarında toplam on kazanım belirlenmiştir.

Analitik puanlama anahtarında Temel İletişim İlke ve Teknikleri, Kendini İfade Etme ve Kişiler Arası İlişkiler ölçütleri belirlenmiş, bu ölçütlere ilişkin ayrı ayrı tanımlamalar yapılmış ve puanlama stratejisi olarak üçlü dereceli puanlama kullanılmıştır. Üçlü dereceli puanlama “genellikle, bazen ve hiç” olarak gösterilmiştir. Hazırlanan puanlama anahtarının bir ölçüte ait örneği Tablo 1’de sunulmuştur.

Tablo 1. Analitik Dereceli Puanlama Anahtarı

Temel İletişim İlke ve Teknikleri	Genellikle	Bazen	Hiç
Konuşma sırasında göz teması kurar			
Konuşurken jest ve mimikleri kullanır			
Konuşmayı başlatır, sürdürür ve sonlandırır			
Konuşmak için sırasını bekler			

Bütüncül puanlama anahtarında, Temel İletişim İlke ve Teknikleri, Kendini İfade Etme ve Kişiler Arası İlişkiler ölçütleri bir bütün olarak ele alınmıştır ve puanlama stratejisi olarak üçlü dereceli puanlama kullanılmıştır. Dereceli puanlama “iyi, orta ve kötü” olarak gösterilmiştir. Bütüncül dereceli puanlama anahtarının bir ölçütüne ilişkin örnek Tablo 2’de sunulmuştur.

Tablo 2. Bütüncül Dereceli Puanlama Anahtarı

Puanlama	Ölçüt
İyi (3)	<ul style="list-style-type: none"> - Temel iletişim ilke ve tekniklerine iyi düzeyde sahiptir. (göz teması kurar, jest-mimik kullanır, konuşmayı rahatlıkla başlatır, sürdürür) - Kendini ifade etmekte hiçbir zaman sıkıntı çekmez. (grup önünde rahat konuşur, duygu ve düşüncelerini rahatlıkla açıklar) -Kişiler arası ilişkiler kurmada başarılıdır. (başkalarıyla sorunlarını konuşarak çözer, kolaylıkla yeni arkadaş edinebilir).

Hazırlanan ölçme araçları Bartın ili merkezde yer alan bir anaokulunda 60-72 aylık öğrenci grubunda görev yapan öğretmenlere uygulanmıştır. Uygulamalardan önce öğretmenlere ölçme aracı ve puanlama ile ilgili bilgilendirme yapılmıştır. Öncelikli olarak bütüncül dereceli puanlama anahtarı uygulanmıştır ve öğretmenlerden öğrencilerinin iletişim becerilerini puanlaması istenmiştir. Daha sonra aynı öğretmenlerden analitik dereceli puanlama anahtarına göre değerlendirme yapması istenmiştir. Her iki ölçme aracı da üç hafta sonra aynı öğretmenlere tekrar uygulanmıştır.

Verinin analizinde, analitik ve bütüncül dereceli puanlama anahtarlarından elde edilen puanların farklılığını test etmek amacıyla Wilcoxon testi yapılmıştır. Her bir

ölçme aracı için iki hafta aralıkla yapılan uygulama sonuçlarına ilişkin Spearman-Brown Sıra Farkları korelasyon katsayısı hesaplanmıştır. Hesaplanan korelasyon katsayısının anlamlı farklılık gösterip göstermediğini belirlemek için Fisher'in z testi kullanılmıştır.

3. Bulgular ve Yorumlar

Analitik ve bütüncül dereceli puanlama anahtarından elde edilen puanların normallik testi yapılmıştır ve hesaplanan Kolmogorov-Smirnov testi sonucuna göre analitik dereceli puanlama anahtarından elde edilen puanların ($D=0,18$; $p<0,05$) ve bütüncül dereceli puanlama anahtarından elde edilen puanların ($D=0,46$; $p<0,05$) normal dağılmadığı görülmüştür. Analitik ve bütüncül dereceli puanlama anahtarından elde edilen puanların ortalamaları arasında anlamlı farklılık olup olmadığını belirlemek için Wilcoxon testi yapılmıştır ve elde edilen sonuçlar Tablo 3'de sunulmuştur:

Tablo 3. Analitik ve Bütüncül Dereceli Puanlama Anahtarı Puanlarının Farklılığını Test Eden Wilcoxon Testi Sonuçları

Analitik Puan-Bütüncül Puan	Sıra Ortalaması	Sıra Toplamı	z	p
Negatif Sıralar	8,64	121	-0,19	0,84
Pozitif Sıralar	15,71	110		

Tablo 3 incelendiğinde, analitik ve bütüncül dereceli puanlama anahtarından elde edilen puanların ortalamaları arasında anlamlı bir farklılık olmadığı görülmüştür [$z=-0,19$; $p>0,05$]. Bu durum, analitik ve bütüncül dereceli puanlama anahtarı sonunda öğrencilerin benzer puanlar aldığını ortaya koymuştur.

Analitik ve bütüncül dereceli puanlama anahtarlarının güvenilirliklerinin anlamlı farklılık gösterip göstermediğini belirlemek için öncelikle her bir puanlama anahtarı için iki hafta arayla elde edilen puanların korelasyon katsayıları hesaplanmıştır. Analitik dereceli puanlama anahtarına ait korelasyon katsayısı 0,84 ve bütüncül dereceli puanlama anahtarına ait korelasyon katsayısı 0,56 olarak hesaplanmıştır ve bu katsayıların anlamlı olduğu bulunmuştur [$p<0,05$]. Hem analitik hem de bütüncül puanlamaya ait test-tekrar teste dayalı güvenilirlik katsayılarının hangisinin daha anlamlı olduğunu belirlemek için Fisher'in z testi ile incelenmiştir. Bu amaçla, katsayılar, birbiriyle kıyaslanabilir hale getirmek için standartlaştırılmış z değerlerine dönüştürülmüştür. Standartlaştırılmış korelasyon katsayıları arasındaki fark aşağıdaki formülle hesaplanmıştır:

$$Z_{\text{fark}} = \frac{z_{r1} - z_{r2}}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}}$$

Yukarıdaki formülde, zr_1 , birinci grubun standartlaştırılmış korelasyon katsayısı değerini, zr_2 , ikinci grubun standartlaştırılmış korelasyon katsayısı değerini, N_1 ve N_2 ise birinci ve ikinci grubun kişi sayısını göstermektedir. Analitik ve bütüncül dereceli puanlama anahtarları için yapılan hesaplama sonuçları Tablo 4’de sunulmuştur:

Tablo 4. Analitik ve Bütüncül Dereceli Puanlama Anahtarının Güvenirliğinin Farklılığını Test Eden Fisher’in z testi Sonuçları

	Analitik	Bütüncül	Zr
Test- tekrar test korelasyon katsayısı	0,84	0,56	1,98*

Tablo 4 incelendiğinde; hesaplanan fark değerinin anlamlı farklılık gösterdiği görülmüştür. [$Zr=1,98 > 1,96$]. Buna göre, analitik dereceli puanlama anahtarının test-tekrar test güvenirliliği ile bütüncül dereceli puanlama anahtarının güvenirliliği arasında anlamlı bir farklılık olduğu sonucuna varılmıştır. Analitik dereceli puanlama anahtarının güvenirliliğinin daha yüksek olduğu bulunmuştur.

4. Sonuç ve Tartışma

Araştırma sonucunda analitik ve bütüncül dereceli puanlama anahtarından elde edilen puanların anlamlı farklılık ortaya koymadığı bulunmuştur. Bu durum, Anıl ve Şahhüseyinoğlu (2008)’nin sonuçları ile örtüşmektedir. Pek çok araştırmada (Şanlı, 2010; Wiseman, 2012; Ghalib ve Al-Hattami, 2015); analitik dereceli puanlama anahtarının daha tutarlı olduğu bulunmuştur, eldeki araştırmanın sonuçları da bu durumu desteklemektedir. Bu durumun, analitik dereceli puanlamada, ölçülecek performansın daha detaylı tanımlamasından kaynaklanmış olabileceği düşünülmektedir. Aynı zamanda ölçülen performansın özelliği de sonuç üzerinde etkili olabilir. Bu sonuçlar doğrultusunda, performans özelliklerinin dereceli puanlama anahtarları ile ilişkisinin araştırılması önerilmektedir. Ayrıca, bu çalışmada eksik bırakılan geçerliliğin karşılaştırmalı olarak incelenmesi de araştırma konusu yapılabilir.

5. Kaynakça

- Anıl, D., Şahhüseyinoğlu, D. (2008). Performansa Dayalı Değerlendirmede Puanlama Tutarlılığı: Holistik ve Analitik Dereceli Puanlama Anahtarının Karşılaştırılması 1. Ulusal Eğitimde ve Psikolojide Ölçme ve Değerlendirme Kongresi, AnkaraBeyreli, L., & Ari, G. (2009). The Use of Analytic Rubric in the Assessment of Writing Performance--Inter-Rater Concordance Study. *Educational Sciences: Theory and Practice*, 9(1), 105-125.
- Carr, N. T. (2000). A comparison of the effects of analytic and holistic rating scale types in the context of composition tests. *Issues in Applied Linguistics*, 11(2), 207-241.
- Ghalib, T. K., & Al-Hattami, A. A. (2015). Holistic versus Analytic Evaluation of EFL Writing: A Case Study. *English Language Teaching*, 8(7), 225-236. doi:10.5539/elt.v8n7p225
- Haladyna, T. M. (1997). *Writing Test Item to Evaluate Higher Order Thinking*. USA: Allyn & Bacon.

- Kan, A. (2007). Performans değerlendirme sürecine katkıları açısından yeni program anlayışı içerisinde kullanılabilecek bir değerlendirme yaklaşımı: Dereceli puanlama anahtarı puanlama yönergeleri. *Kuram ve Uygulamada Eğitim Bilimleri*, 7(1), 129-152.
- Klein, S. P., Stecher, B. M., Shavelson, R. J., McCaffrey, D., Ormseth, T., Bell, R. M., & Othman, A. R. (1998). Analytic versus holistic scoring of science performance tasks. *Applied Measurement in Education*, 11(2), 121-137.
- Moskal, B. M. (2003). Recommendations for developing classroom performance assessments and scoring rubrics. *Practical Assessment, Research & Evaluation*, 8(14), 1-8.
- Moskal, B. M. (2000). "Scoring rubrics: What, when and how?" *Practical Assessment, Research & Evaluation*, 7 (3) Available Online: <http://ericac.net/pare/getvn.asp?v=7&n=3>
- Popham, W. J. (1997). What's Wrong--and What's Right--with Rubrics. *Educational Leadership*, 55(2), 72-75.
- Rezaei, A. R., & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing*, 15(1), 18-39.
- Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34(2), 159-179.
- Şanlı, E.(2010). Bilimsel süreç becerilerinin ölçülmesinde bütünsel ve analitik puanlama anahtarlarının güvenilirliklerinin karşılaştırılması. *Yayınlanmamış Yüksek Lisans Tezi*, Ankara: *Ankara üniversitesi*
- Wiseman, C. S. (2012). A comparison of the performance of analytic vs. holistic scoring rubrics to assess L2 writing. *Iranian Journal of Language Testing*, 2(1), 59-92.