

ÖĞRETİMDE SUNU BECERİLERİNİN DEĞERLENDİRİLMESİNDE DERECELİ PUANLAMA ANAHTARI (RUBRIC) KULLANILMASINA İLİŞKİN BİR ARAŞTIRMA *

Uzm. Ashhan ERMAN ASLANOĞLU

Ankara Özel Tefvik Fikret Okulları

Ölçme ve Değerlendirme Uzmanı

Yrd. Doç. Dr. Ömer KUTLU

Ankara Üniversitesi Eğitim Bilimleri Fakültesi

Ölçme ve Değerlendirme Anabilim Dalı Öğretim Üyesi

ÖZET

Bu çalışmanın amacı, dereceli puanlama anahtarı (rubric) kullanarak öğrencilerin üst düzey zihinsel becerilerini ifade etme yollarından biri olan sunu yapma davranışlarını değerlendirmektir.

Araştırma, bir özel okulun 24 öğrencisi üzerinde yürütülmüştür. Sunu yapma davranışı için; üç performans görevi, yol gösterici açıklamalar ve nasıl değerlendirileceğini gösteren dereceli puanlama anahtarı öğrencilere verilmiştir. Öğrencilerden bu üç görevden birini seçmesi ve buna sunu yapmak üzere hazırlanmaları istenmiştir. Sunuları, öğretmen ve sınıftaki öğrenciler değerlendirmiştir. Ayrıca her sunu kameraya kaydedilmiş, okuldaki Fen Bilgisi öğretmenlerine değerlendirmeleri için izletilmiştir. Çalışmanın sonunda araştırmaya katılan öğrenci ve öğretmenlere “dereceli puanlama anahtarı” ile değerlendirmenin eğitime katkısı bir anketle sorulmuştur.

Yapılan analizlerin sonucuna göre öğretmen ve izleyici öğrencilerin her üç performans görevine ait vermiş oldukları puanların ortalamalarıyla performans görevlerinin tümü için vermiş oldukları puanların ortalamaları arasında anlamlı bir fark olmadığı, öğretmenlerin sunulara verdikleri puanlar arasında yüksek bir uyum olduğu belirlenmiş ve bu puanların tutarlı olduğu sonucuna varılmıştır. Ayrıca sunu yapan öğrencilerin sınıf öğretmenlerinden aldıkları puanlarla, 1. dönem fen bilgisi dersinden aldıkları notlar arasında 0.72; izleyici öğrencilerin sunulara verdikleri puanlarla, sunu yapan öğrencilerin 1. dönem fen bilgisi notları arasında ise 0.49 düzeyinde anlamlı bir ilişki bulunmuştur. Anket sonuçları öğretmen ve öğrencilerin, dereceli puanlama anahtarı kullanmanın eğitime katkı getirdiğine inandıklarını ortaya koymaktadır.

Anahtar Sözcükler: *Dereceli puanlama anahtarı, durum belirleme, performans değerlendirme, üst zihinsel becerilerin değerlendirilmesi*

SUMMARY

The aim of this study is to evaluate, through the use of rubric, the student's presentational behaviours which is one of the ways of expression of their high level thinking.

The research has been conducted on 24 students of a private school. In order to exercise the presentational behaviour, the students were given 3 performance tasks, guiding explanations and the rubric showing how to make evaluation. The students were asked to chose one of these three

* Ankara Üniversitesi Eğitim Bilimleri Enstitüsünde Yüksek Lisans Tezi olarak hazırlanmıştır.

performance tasks and get prepared for it in terms of a presentation. The presentations were evaluated by the teacher and the students in class. Besides, each presentation was recorded in a camera and was displayed to the science teachers at school for their evaluations. At the end of the study, the contribution to education of evaluation by "rubric" was asked to the students and the teachers who took part in the research, via a questionnaire.

According to the results of the analysis made, it has been determined that there was no significant difference between the mean of the scores given by the viewer students and the teachers on each of the three performance tasks and the average of the scores they have given on the entire performance tasks; and that there was a high coherence in the scores given for the presentations by the teachers; and that these points were considered to be consistent. Furthermore, a relationship at 0.72 level was found between the marks which the students who made presentations were given by the class teachers and the marks they took from the first semester science lesson; as well as a relationship at 0.49 level between the points the spectator students gave for the presentations and the first semester science lesson marks of the students who made presentations. The results of survey stated that the teachers and the students believe that the use of rubric brings contribution to education.

Keywords: Rubric, assessment, performans assessment, evaluating higher order cognitive skills

GİRİŞ

Ölçme ve değerlendirme eğitim sürecinin ayrılmaz ve tamamlayıcı bir parçasıdır. Her alanla ilgili eğitimi planlama ve uygulama aşamaları incelendiğinde, ölçme ve değerlendirmenin eğitim sürecini tamamlayan bir uygulama olduğu görülecektir (Alison, 1999). Ölçme ve değerlendirme uygulamaları, öğretmenlere eğitim ve öğretim sürecine ilişkin pek çok konuda yardımcı olur. Öğretmen bu uygulamalar sayesinde programın ve öğretimin etkililiğini değerlendirebilir ve öğrencilerin öğrenme eksikliklerini ortaya koyarak onların başarılarını belirleyebilir (Stark, 1998).

Günümüzde öğrenci başarısını ölçmek ve değerlendirmek; birkaç sınavla öğrencilerin sınıflarını geçip-kaldıkları hakkındaki kararları veren geleneksel anlayış dışına taşmak durumundadır. Ölçme ve değerlendirme uygulamaları, okul programlarında yer alan davranışların öğrenilememesine ya da eksik öğrenilmesine neden olan, dolayısıyla öğrenci başarısının artmasına engel olan sistemdeki olumsuzlukları saptamada kullanılan bir araç ve bu olumsuzluklara dayanarak kararlar almaya yardımcı olan bir alan olarak görülmelidir. Bu nedenle, öğrencilerin okul hedeflerinde yer alan davranışları ne derece kazandıklarını ortaya koyan ölçme ve değerlendirme etkinlikleri yanında, öğrencilerin kazandıkları bu davranışları ne derece kullandıklarını gösterecek ölçme ve değerlendirme etkinliklerine gereksinim vardır. Her iki durumda da, öğrencilerin öğretim sürecindeki var olan durumlarının ortaya konması söz konusu

olmaktadır. Bu ise, eğitim süreci içinde durum belirleme (assessment) kavramının önemini ortaya çıkarmıştır (Kutlu 2002).

Durum belirleme, eğitimin iyileştirilmesini amaçlayan ve devamlılık gösteren bir işlemdir. Durum belirleme, beklentileri daha açık ve belirgin bir biçimde ortaya koymakta, eğitim kalitesi için uygun ölçütler ve yüksek standartlar önermektedir. Ayrıca gösterilen performansın bu beklenti ve standartlarla ne derece uyumlu olduğunu sistematik olarak analiz edip yorumlamakta ve elde edilen sonuçların performansı artırma, performansa ilişkin bir açıklamada bulunma ve belgeye dayandırmada kullanılmasını sağlamaktadır (Stiggins, 1994).

Günümüze kadar durum belirleme yaklaşımları daha çok kağıt-kalem testleriyle gerçekleştirilmekteydi. Ancak son yıllarda birçok öğretmen ve eğitim uzmanı öğrencilerin bazı klasik ölçme yöntemleriyle (çoktan seçmeli, doğru-yanlış, eşleştirmeli, tamamlamalı vb. sınavlar) değerlendirilmesinin daha az kullanılması konusunda aynı düşüncededirler. Çünkü bu tip değerlendirmeler öğrencilerde gözlenmesi istenen üst düzey zihinsel becerileri (yeni bir hikaye yazmak, hazırladığı bir çalışmayı sunmak gibi) ölçmede yetersiz kalmaktadır. Bu testler, öğrencinin sahip olduğu bilgileri ayrıntılı olarak nasıl kullandığı, karşılaştıkları sorun ya da problemleri nasıl çözdüğü ve hazırlayacakları ödevlerde bu bilgileri nasıl kullandığı konusunda, çok az bilgi vermektedirler. Eğitim ortamında görülen bu eksiklikler ölçme ve değerlendirme amacıyla kullanılması gereken yeni yaklaşımları gündeme getirmiştir. Bu yaklaşımlara "performansa

dayalı durum belirleme yaklaşımları” adı verilmektedir (ERIC, 1998).

Performansa dayalı durum belirleme, öğrencinin gerçekçi koşullarda, karmaşık ödevleri yaparken öğrendiği temel bilgileri ne kadar iyi kullandığını ölçmeye çalışır. Performans ölçümü böyle gerçekçi koşullar altında yapıldığında gerçekçi (authentic) değerlendirme olarak da adlandırılır. Performans ölçümü ile gerçek değerlendirmeyi kapsayan ve bunları geleneksel yazılı sınavlardan ayıran değerlendirme yöntemi “yeni yöntemler (yaklaşımlar)” olarak adlandırılır (Mehrens, 1992).

Yeni yaklaşımlar, farklı biçimlerde tanımlanmıştır. Stiggins (1991) yeni yaklaşımları, öğrenim görenin sahip olduğu bilgiyi belirlemek için uygulanan bir yöntem olarak tanımlamakta ve bu yöntem geleneksel durum belirlemeden farklılık göstermektedir. Bu yöntem örneği olarak; dereceli puanlama anahtarı (rubric)*, dosya oluşturma (portfolyo), konferans, günlük tutma, grup değerlendirmeleri vb. verilebilir (Brown ve Hudson, 1998). Son yıllarda “kavram haritaları” da yeni ölçme yaklaşımları içinde ele alınmaya başlamıştır (Şahin, 2003). Portfolyolar öğrencilerin herhangi bir derste başarıları, yetenekleri, çaba ve katkıları gösteren belgelerin bir dosyada tutulmasıdır. Konferanslar herhangi bir çalışmada ya da öğrenme süreci içerisinde öğrencilerin öğretmenleri ile yapmış olduğu toplantılardır. Grup değerlendirmeleri, öğrencilerin birbirlerinin çalışmalarını değerlendirmeleridir (O’Malley ve Pierce, 1996). Dereceli puanlama anahtarı ise, gözlemlerimize ait puanları tanımlanmış kategorilerden (ölçüt ya da ölçütler) uygun düşen boyuta kaydetmemizi sağlayan bir değerlendirme aracıdır (Haladyna, 1997).

Performans dayanaklı durum belirlemenin iki kritik bölümü vardır; bunlardan biri performans görevleri (performance task) ve dereceli puanlama anahtarıdır (Popham, 2000).

Performans görevi, öğrencinin listelenmiş yanıtlardan birini seçmesi yerine, kendi yanıtını oluşturmasını sağlayan bir durum belirleme etkinliğidir (Stark, 1998). Performans görevleri öğrencilerin yeni bilgiler inşa etmelerini gerektirir.

* Oxford İngilizce sözlük, “rubric” kavramını, 14 yy.’da bir kitabın farklı bölüm başlıklarını ifade eden bir kavram olarak kullandığını belirtmektedir. Bu çalışmada, bu kavrama Türkçe karşılık olarak, “dereceli puanlama anahtarı” önerilmiştir.

Öğrencilerin tamamlamaları istenen görevlerin çoğunun önceden belirlenmiş bir “doğru yanıtı” vardır. Öğrencilerin tamamlamalarını, yanıtları eşlemelerini veya çeşitli etkinliklerden oluşan bir listede doğru yanıt seçmelerini gerektiren görevlerin tümünde, önceden belirlenmiş tek yanıt formatı kullanılır. Bu görevlerde farklı yanıtlara yer yoktur. Belirli bir yanıt kesin doğrudur, diğerleri ise kesin yanlıştır. Hatta çoğu kompozisyon görevlerinin de oldukça belirli yanıtları vardır. Bundan dolayı bir soruya yanıt olarak bir veya iki paragraf yazmak öğrencilerin her zaman yeni bir bilgi yapılandırılmalarını gerektirmez. Etkili öğrenme için esas olan, bu bilgiyi yapılandırma yöntemidir ki bu da öğrencilere performans görevleri vererek gerçekleştirilebilir (Marzano, Pickering ve Mctighe, 1993).

Popham (1997) performans dayanaklı durum belirlemenin kritik bölümünü oluşturan dereceli puanlama anahtarını öğrencilerin, yazılı kompozisyonlarının, sözel sunumlarının veya bilimsel projelerinin değerlendirilmesinde bir çeşit puanlama ölçüsü olduğunu belirtmektedir. Popham, dereceli puanlama anahtarını, her bir çalışma için ölçütleri (ölçülecek boyutları) listeleyen ve çalışmada nelerin yapılacağını gösteren bir puanlama aracı olarak görmektedir. Popham’a göre dereceli puanlama anahtarı; değerlendirme ölçütleri, ölçüt tanımlamaları ve bir puanlama stratejisi olmak üzere üç bölümden oluşur.

1. *Değerlendirme ölçütleri:* Kabul edilebilir yanıtları kabul edilemez yanıtlardan ayırmak için kullanılır. Örneğin öğretmenler yazılı kompozisyonları değerlendirirken organizasyon, yapısal içerik, sözcük seçimi vb. gibi değerlendirilebilir ölçütler kullanırlar.

2. *Ölçüt tanımlamaları:* Öğrencilerin değerlendirilmek istenen yanıtlarındaki niteliksel farklılıkları tanımlama yolunu ifade eder. Örneğin bir kompozisyonda organizasyon değerlendirilecekse bu ölçütlerden en yüksek puanı alan öğrencinin kompozisyonu organizasyon açısından hiç hata içermemelidir.

3. *Puanlama stratejisi:* Puanlama bütünsel (holistic) ya da analitik (analytical) biçiminde olabilir. Dereceli puanlama anahtarlarından hangisinin kullanılacağı değerlendirmenin amacına bağlıdır (Popham 1997). Bazı durumlarda yapılan bir değerlendirmeyi bağımsız etkenlere (ölçüt) ayırtmak mümkün olamamakta, performansın farklı düzeylerinin ortaya çıkarılması için

belirlenmiş ölçütler arasında bir ayrışma bulunmamaktadır. Böyle durumlarda bütünsel puanlama anahtarı kullanılmalıdır (Brookhart, 1999). Analitik puanlama anahtarı ise, ölçülen bir yetenek boyutunun öğelere ayrıştırılabildiğinde ve daha ayrıntılı puanlama yapmak istendiğinde kullanılmaktadır (Haladyna, 1997).

Bu çalışma, üst düzey zihinsel becerileri ölçmede kullanılan dereceli puanlama anahtarı hakkında öğretmenlere ve araştırmacılara bilgi sunmak ve bu puanlama yoluyla sunum davranışlarının nasıl değerlendirileceğine bir örnek oluşturmak amacıyla gerçekleştirilmiştir. Eğitim ortamında, öğrencilerden çalışmalarını sunmaları sıkça istenen bir durumdur. Fakat yapılan sunuları değerlendirmede bir takım sorunlar yaşanmaktadır. Bunlardan ilki, ilgili becerinin ölçülmesinde kullanılacak sunum sorusunun hazırlanmasıdır. Öğrenciye araştırmaları için verilecek soruların öğrencilerin üst düzey zihinsel becerisini ortaya çıkarabilecek nitelikte sorular olması gerekmektedir. Bu çeşit soruları hazırlamak ise uzmanlık gerektirmektedir. Klasik ölçme uygulamaları kapsamında soru yazan öğretmenler için bu tür sorular yazmak zor olmaktadır. Sunum çalışmalarında karşılaşılan diğer bir güçlük ise, öğrencilere araştırma yapacakları konu hakkında yetersiz bilgiler verilmesidir. Bu durum, öğrencilerin akıllarında birçok soru işaretinin oluşmasına neden olmaktadır. Araştırmayı nasıl yapacaklarını ve kendilerinden neler istendiğini anlamayan öğrencilerin çalışmalarının da kalitesi düşmektedir. Ayrıca sunuların nasıl değerlendirileceğinin öğretmenler tarafından açık bir şekilde belirlenmemesi değerlendirmedeki yanlışlıkları da artırmaktadır. Tüm bu sorunlar göz önüne alındığında böyle bir çalışmanın öğretmenlere, öğrencilerinin sunularını değerlendirirken nasıl bir yol izleyeceği hakkında bilgi sağlayacağı düşünülmektedir. Ayrıca bu çalışmada, böyle bir puanlama anahtarı yardımıyla öğrencilerin değerlendirme sürecine katılmaları da öğrenci merkezli eğitim açısından önem taşımaktadır.

Araştırmanın Amacı

Bu çalışmanın amacı, dereceli puanlama anahtarı kullanarak öğrencilerin üst düzey zihinsel becerilerini ifade etme yollarından biri olan sunum yapma davranışlarını gözlemek ve elde edilen gözlem sonuçlarından çeşitli yorumlara ulaşmaktır.

Bu çalışmada aşağıdaki sorulara yanıt aranmıştır:

1. Öğrenci sunularına öğretmenlerin verdikleri puanlarla, izleyici öğrencilerin verdikleri puanlar arasında; 1., 2., 3. performans görevleri ve bu görevlerin tümü açısından fark var mıdır?

2. Öğretmenlerin sunulara verdikleri puanlar arasında bir uyum var mıdır?

3. İzleyici öğrencilerin sunulara verdikleri puanlar arasında bir uyum var mıdır?

4. Öğrencilerin 1. döneme ait fen bilgisi dersi notlarıyla, sınıf öğretmenin dereceli puanlama anahtarıyla sunulara verdiği puanlar arasında bir ilişki var mıdır?

5. İzleyici öğrencilerin dereceli puanlama anahtarıyla sunulara verdikleri puanlarla, sunu yapan öğrencilerin 1. döneme ait fen bilgisi dersi notları arasında bir ilişki var mıdır?

6. Öğretmenlerin dereceli puanlama anahtarının kullanılması konusundaki düşünceleri nelerdir?

7. Öğrencilerin sunularının değerlendirilmesinde, dereceli puanlama anahtarı kullanılmasıyla ilgili düşünceleri nelerdir?

Bu amaçlara ulaşmak için üç adet performans görevi oluşturulmuş, öğrencilerin sunu konusu bu görevlerden birini seçmeleri sağlanarak belirlenmiştir. Daha sonra bir grup öğretmen ve öğrenci sunum davranışlarını izlemiş, gözlemlerini dereceli puanlama anahtarına kaydederek puanlamıştır. Gözlem sonuçları çeşitli istatistiksel teknikler kullanılarak çözümlenmiştir.

YÖNTEM

Çalışma Grubu

Çalışma bir özel okulun 6. sınıfında okuyan 24 öğrenci ve aynı okulda öğretmenlik yapan 14 öğretmen üzerinde yürütülmüştür.

Performans Görevlerinin Hazırlanması

Performans görevleri hazırlanırken öncelikle öğrencilerin ilgisini çekecek, yüksek güdü düzeyiyle katılımlarını sağlayacak ve araştırarak hazırlanacakları konuların seçilmesine önem verilmiştir. Bunun için öğretmenlerden de görüş alınmış, performans görevlerinin fen bilgisi dersi kapsamında hazırlanmasına karar verilmiş ve üç farklı performans görevi hazırlanmıştır. Performans görevi, öğrencilerin üst düzey zihinsel becerisini

ortaya çıkaracak bir araştırma sorusu olarak tasarlanmıştır. Performans görevlerinin tanımlandığı formlar, bu çalışmanın ölçme aracı niteliğindedir (Örnek performans görevi Ek - 1). Performans görevlerinin son biçimi; Türkçe dil kurallarına uygunluk, bilimsel doğruluk ve ölçme-değerlendirme ilkelerine uygunluk bakımından denetimden geçirildikten sonra verilmiştir. Bunun için dil, konu ve ölçme-değerlendirme uzmanlarından gerekli destek sağlanmıştır. Ayrıca görevlerin ilgili yaş düzeyine uygunluğu ve birbirine eş güçlük düzeyinde olup olmadıkları hakkında da öğretmenlerden görüş alınmıştır.

Yönergelerin Oluşturulması

Öğrencinin performans görevlerini yerine getirirken dikkat etmesi gereken noktalarla, önüne çıkabilecek engelleri aşabilmesini sağlayacak yol gösterici açıklamalar bir yönerge biçiminde hazırlanmıştır (Ek-2). Ayrıca her performans görevine ait gerekli açıklamalar da görevin altında verilmiştir (Ek-1).

Dereceli Puanlama Anahtarının Hazırlanması

Dereceli puanlama anahtarı için Goodrich (2001)'in önerdiği basamaklardan yararlanılmıştır. Bu basamaklar şöyledir:

1. *Performansı belirlemede kullanılacak ölçütlerin listelenmesi:* Üç farklı performans görevine uygun sunuların değerlendirilmesinde kullanılacak dereceli puanlama anahtarı için beş ölçüt belirlenmiştir (EK-3). Bu ölçütler bu tür bir çalışmada öğrencinin yerine getirmesi istenen performanslar göz önüne alınarak belirlenmiştir.

2. *Kullanılacak dereceli puanlama anahtarına karar verilmesi:* Performans görevlerinin uygun olması ve sunuların ayrıntılı puanlanması istendiğinden araştırmada “analitik dereceli puanlama anahtarı” kullanılmıştır.

3. *Performans düzeylerinin belirlenmesi ve düzey tanımlamalarının yapılması:* Bu aşamada, en iyi ve en kötü performans düzeyleri (dereceleri) belirlenmiştir. Buna göre her ölçüt 0 ile 4 arasında derecelendirilmiştir. Gözlenmek istenen performansı gerçekleştiremeyen öğrencilerin alacağı puan 0, tam olarak gösteren öğrencilerin alacağı puan ise 4 olarak belirlenmiştir.

4. *Uzman görüşünün alınması:* Dereceli puanlama anahtarı hazırlandıktan sonra ilgili alan öğretmenlerinden, ölçme - değerlendirme uzmanla-

rından ve dil konusunda çalışan uzmanlardan görüşler alınmıştır.

Öğrencilerle ve Öğretmenlerle Ön Çalışmalar Yapılması

Öğrencinin sunu davranışı konusunda kendisinden beklenenleri tam olarak bilmesi için; seçtiği performans görevi, yönergeler ve dereceli puanlama anahtarı birlikte verilmiştir.

Öncelikle öğrencilerin üç performans görevinden birini ilgileri doğrultusunda seçmeleri istenmiştir. Bu seçimler sırasında öğrencilerin performans görevlerine eş sayıda dağılımlarına dikkat edilmiştir. Öğrencilere ve öğretmenlere dereceli puanlama anahtarının ne olduğu ve nasıl uygulanacağı anlatılmış, yönergelerde yazılanlar üzerinde konuşulmuştur. Öğrencilere, sunularını ders öğretmenin, sınıftaki öğrencilerin ve kendileri tarafından da değerlendirileceği söylenmiştir. Ayrıca, öğrencilere sunum yapılacak günler bildirilmiş ve onlardan kendilerine uygun olan bir günü seçmeleri istenmiş ve bu doğrultuda sunum tarihleri belirlenmiştir. Öğrencilere çalışmalarını sunmaları için uygun ortam sağlanmıştır. Buna göre, sunumunu bilgisayar yardımıyla yapacak öğrencilere okulun projeksiyon odası ayarlanmış, tepegöz gereksinimi olan öğrencilere tepegöz sağlanmıştır. Her sunum ders öğretmeni ve sınıftaki diğer öğrenciler tarafından puanlandırılmıştır. Ayrıca sunular, fen bilgisi dersini veren diğer on üç öğretmene izletilmek ve öğretmen değerlendirmeleri arasındaki tutarlılığa bakmak amacıyla kameraya kaydedilmiştir.

Verilerin Toplanması

Veriler üç temel kaynaktan elde edilmiştir:

1. kaynak, sunu davranışlarının değerlendirilmesinde kullanılan dereceli puanlama anahtarından elde edilen puanlardır. Bu puanlar hem öğrencilerden hem de fen bilgisi alanında görevli öğretmenlerden elde edilmiştir. Öğrenciler, sunuyu sınıfta izlemişler ve dereceli puanlama anahtarını öyle doldurmuşlardır. Öğretmenler ise, sunuyu video kayıtlarından bir sınıf ortamında izleyerek puanlamışlardır.

2. kaynak öğrencilerin okul notlarıdır. Bu notlar ilgili okulun bilgi işlem kayıtlarından sağlanmıştır.

3. kaynak ise öğretmen ve öğrencilere uygulanan ankete verilen yanıtlardır. Bu anketle öğrencilere ve öğretmenlere, eğitim ortamında

dereceli puanlama anahtarı kullanmanın yararları ve sakıncaları yanında, belirtmek istedikleri başka düşünceleri olup olmadığı sorulmuştur.

Verilerin Çözümlemesi

Veriler alt problemler temel alınarak çözümlenmiştir. Buna göre;

a) birinci alt probleme yanıt vermede ilişkisiz örneklem için t testi,

b) ikinci ve üçüncü alt problemlerde dile getirilen puanlar arasındaki uyumu bulmada Kendall'in w testi,

c) dördüncü ve beşinci alt problemlere yanıt aranırken Sperman Sıra Farkları Korelasyon Katsayısı,

d) altıncı ve yedinci problemleri yanıtlamada ise anketlere verilen yanıtlardan benzer olanlarının aynı grupta toplanması yolu kullanılmıştır.

BULGULAR VE YORUMLAR

Bu bölümde araştırma sorularına ilişkin bulgular ve yorumlar verilmiştir. Bulguların sunulduğunda, araştırma soruları sırasıyla ele alınıp yorumlanmıştır.

1. problem: Öğrenci sunularına öğretmenlerin verdikleri puanlarla, izleyici öğrencilerin verdikleri puanlar arasında; 1., 2., 3. performans görevleri ve bu görevlerin tümü açısından fark var mıdır?

Bu probleme yanıt aramak için ilişkisiz örneklem t testi kullanılmıştır. 1. performans görevine ait t-testi sonuçları Çizelge 1'de verilmiştir.

Çizelge 1. Öğretmen ve öğrencilerin 1. performans görevine verdikleri puanlara ilişkin t-testi sonuçları

	N	\bar{X}	S_x	sd	t
Öğrenci	8	14,5	2,6	14	0,67
Öğretmen	8	13,4	3,9		

Çizelge 1'e göre öğretmen ve izleyici öğrencilerin 1. performans görevine ait vermiş oldukları puanların ortalamaları arasında anlamlı bir fark olmadığı görülmektedir ($t_{(8)}=0,67$, $p>0,01$).

2. performans görevine ait t-testi sonuçları Çizelge 2'de verilmiştir.

Çizelge 2. Öğretmen ve öğrencilerin 2. performans görevine verdikleri puanlara ilişkin t-testi sonuçları

	N	\bar{X}	S_x	sd	t
Öğrenci	8	15,75	2,82	14	0,51
Öğretmen	8	14,83	4,05		

Çizelge 2, öğretmen ve izleyici öğrencilerin 2. performans görevi için vermiş oldukları puanların ortalamaları arasında anlamlı bir fark olmadığını göstermektedir ($t_{(8)}=0,51$, $p>0,01$). 3. performans görevine ait t-testi sonuçları Çizelge 3'te verilmiştir.

Çizelge 3. Öğretmen ve öğrencilerin 3. performans görevine verdikleri puanlara ilişkin t-testi sonuçları

	N	\bar{X}	S_x	sd	t
Öğrenci	8	15,63	2,07	14	1,38
Öğretmen	8	14,25	1,91		

Çizelge 3'e göre öğretmen ve izleyici öğrencilerin 3. performans görevine verdikleri puanların ortalamaları arasında anlamlı bir fark görülmemektedir ($t_{(8)}=1,38$, $p>0,01$). Her üç performans görevine ait t-testi sonuçları Çizelge 4'te verilmiştir.

Çizelge 4. Öğretmen ve öğrencilerin her üç performans görevine verdikleri puanlara ilişkin t-testi sonuçları

	N	\bar{X}	S_x	sd	t
Öğrenci	24	15,29	2,48	46	1,32
Öğretmen	24	14,17	3,37		

Çizelge 4, öğretmen ve izleyici öğrencilerin her üç performans görevine vermiş oldukları puanların ortalamaları arasında anlamlı bir fark olmadığını göstermektedir ($t_{(24)}=1,32$, $p>0,01$).

Bu bulgular, öğrencilerin, kendilerine değerlendirmede yardımcı olacak dereceli puanlama anahtarı gibi bir yanıt anahtarı verildiğinde öğretmenler gibi değerlendirme yapabildiklerini ve dereceli puanlama anahtarında verilen ölçütleri öğretmenleri gibi algıladıklarını göstermektedir. Bu durum öğrenci merkezli eğitim yapısı içinde dereceli puanlama anahtarlarının kullanılabilmesiyle ilgili ipuçları sunmaktadır.

2. problem: Öğretmenlerin sunulara verdikleri puanlar arasında bir uyum var mıdır?

Araştırmanın bu alt probleminde öğretmenlerin puanlama uyumluluğuna yanıt bulmak için Kendall'ın w testi kullanılmış, sonuç Kendall $w = 0,83$ bulunmuştur. Bu değer öğretmenlerin sunulara verdikleri puanlar arasında yüksek bir tutarlılık bulunduğunu göstermektedir. Başka bir deyişle sunulara öğretmenlerin verdikleri puanlar birbirleriyle uyumludur. Bu da dereceli puanlama anahtarının nesnellüğünün bir göstergesi olabilir.

3. problem: İzleyici öğrencilerin sunulara verdikleri puanlar arasında bir uyum var mıdır?

Bu probleme yanıt bulabilmek için Kendall'ın w testi uygulanmış, sonuç Kendall $w = 0,42$ bulunmuştur. Bu değer, izleyici öğrencilerin sunulara verdikleri puanlar arasında yüksek bir uyum olmadığını göstermektedir. Uyumun düşük çıkmasının nedeni öğrencilerin bu tür bir değerlendirmeyi ilk kez yapıyor olmaları ve arkadaşlarını değerlendirirken yanlı davranmaları olabilir. Öğrencilerin bu tür uygulamalara katıldıkça puan vermedeki uyumlarının artacağı düşünülebilir.

4. problem: Öğrencilerin 1. döneme ait fen bilgisi dersi notlarıyla, sınıf öğretmenin dereceli puanlama anahtarıyla sunulara verdiği puanlar arasında bir ilişki var mıdır?

Bu araştırma sorusuna yanıt bulmak için Sperman Sıra Farkları Korelasyon katsayısı hesaplanmıştır. Buna göre öğrencilerin 1. dönem fen bilgisi dersi notlarıyla dereceli puanlama anahtarından aldıkları puanlar arasında 0,72 düzeyinde bir ilişki bulunmuştur. Bu değer, öğrencilerin 1. dönem fen bilgisi dersi notlarıyla dereceli puanlama anahtarından aldıkları puanlar arasında yüksek bir ilişki olduğunu yani fen bilgisi

notları yüksek olan öğrencilerin sunu becerilerinden de yüksek, fen bilgisi notları düşük olan öğrencilerin ise sunu becerilerinden düşük puan aldığını göstermektedir.

5. problem: İzleyici öğrencilerin dereceli puanlama anahtarıyla sunulara verdikleri puanlarla, sunu yapan öğrencilerin 1. döneme ait fen bilgisi dersi notları arasında bir ilişki var mıdır?

Bu araştırma sorusu için Sperman Sıra Farkları Korelasyon Katsayısı hesaplanmıştır. Buna göre izleyici öğrencilerin dereceli puanlama anahtarıyla verdikleri puanlarla birinci döneme ait fen bilgisi notları arasında 0,49 düzeyinde anlamlı bir ilişki bulunmuştur. Buna göre öğrencilerin dereceli puanlama anahtarıyla verdikleri puanların, birinci döneme ait fen bilgisi notlarıyla orta düzeyde bir ilişki gösterdiği belirlenmiştir. Korelasyon katsayısının yüksek çıkmamasının nedeni öğrencilerin bu tür bir değerlendirmeyi ilk kez yapıyor olmaları ve arkadaşlarını değerlendirirken yanlı davranmaları olabilir.

6. problem: Öğretmenlerin dereceli puanlama anahtarının kullanılması konusundaki düşünceleri nelerdir?

Araştırmanın bu sorusuna yanıt bulmak için, öğretmenlerin anketlere verdiği yanıtlardan yararlanılmış ve birbiriyle ilgili olan yanıtlar gruplandırılmıştır. Buna göre, öğretmenlerin dereceli puanlama anahtarı kullanılmasının yararlarına ilişkin görüşleri şöyledir:

- Verilen performans görevleri öğrencinin düzeyine uygundu ve çok ilgi çekiciydi, bu da öğrencilerin araştırmaya motive olmalarını sağladı.
- Dereceli puanlama anahtarı nesnel değerlendirme ortamı sağladı.
- Öğrenci, çalışmasının hangi ölçütlere göre değerlendirileceğini bildiğinden daha özenli hazırlık yaptı.
- Puanlama sonucunda öğrenci hangi alanlarda eksik olduğuna dair bir geri bildirim almış oldu.
- Öğrencilerin daha araştırmacı bir davranış içine girmelerini sağladı.
- Ölçütlerin iyi belirlenmiş olması puanlama yapmada çok işimize yaradı.

- Öğrencilerin çalışmada aktif rol üstlenmesi öğrenmelerin kalıcı olmasını sağladı.

Dereceli puanlama anahtarının sakıncasına ilişkin öğretmenler her ölçüt hakkında ayrıntılı düşünmek zorunda kaldıklarını ve puanlamanın çok zaman aldığını dile getirmişlerdir. Dereceli puanlama anahtarı konusunda çalışmaları olan Popham (2000) ve Haladyna (1997) analitik dereceli puanlama anahtarının bu olumsuzluğuna dikkat çekmişlerdir.

7. *problem:* Öğrencilerin sunularının değerlendirilmesinde, dereceli puanlama anahtarının kullanılmasıyla ilgili düşünceleri nelerdir?

Bu araştırma sorusu için öğrenci anketleri okunmuş ve elde edilen sonuçlar gruplandırılmıştır. Öğrencilerin dereceli puanlama anahtarının yararlarına ilişkin düşünceleri şöyledir:

- Verilen performans görevleri çok ilginçti ve böylelikle bizim çalışmaya motive olmamızı sağladı.
- Verilen performans görevlerinin yanıtı bir kaynaktan bulunabilecek tarzda değildi dolayısıyla araştırma yapma becerimiz gelişti.
- Hayal gücümüzü çalışmanın içine katabilmemizi sağladı.
- Dereceli puanlama anahtarı sayesinde bizden ne istendiğini anladık, bu da kafamızdaki soru işaretlerini baştan yok etti.
- Bilgiler daha kalıcı ve anlaşılır oldu.
- Bir konu hakkında değişik düşüncelerin olduğunu görmemizi sağladı.
- Puanlamaya bizim de katılmamız sunuları daha iyi anlamamıza neden oldu.
- Bizi fazla sayıda kaynak kullanmaya teşvik etti.
- Sunu sonrasında eksiklerimi görmeme ve bir sonraki çalışmada bu hataları yapmamam için ders almama neden oldu.

Sakıncalarına ilişkin düşünceleri ise, sınıftaki bazı öğrencilerin sunuları değerlendirirken nesnel olma konusunda güçlük yaşadıklarını, bazen sevdikleri arkadaşlarına yüksek, sevmediklerine düşük puan verme yönünde bir eğilim gösterdiklerini söylemişlerdir.

Sonuç

Araştırmaların bulguları ışığında dereceli puanlama anahtarının okul öğrenmelerinin ölçülmesinde ve değerlendirilmesinde önemli görev üstleneceği ve okul öğrenmelerine katkı getireceği düşünülmektedir. Öğrencilerin mevcut bilgi ve becerilerini, belirlenen ölçütler açısından değerlendirebilmesi yanında, bizzat öğrencinin kendisini değerlendirme sürecine katması açısından da önemlidir.

Olumlu yanlarına rağmen dereceli puanlama anahtarına dayalı değerlendirmelerin okullarımızdaki uygulamalarında dikkatli davranmak gerekmektedir. Dereceli puanlama anahtarının kullanılması ve elde edilen verilerin yorumlanması hakkında çok az bilgisi olan öğretmenler bu tür puanlama anahtarını yanlış kullanabilecekler ve elde ettikleri sonuçlardan yanlış kararlara ulaşabileceklerdir. Dereceli puanlama anahtarı, son zamanlarda sıkça konuşulan ve üzerinde araştırma yapılan konulardan birisidir. Dolayısıyla internet ortamında birçok dereceli puanlama anahtarı örneği bulunmaktadır. Bu örnekler bir dereceli puanlama anahtarının yapısının ne olduğu ve nasıl kullanıldığı hakkında öğretmenleri bilgilendirmesi açısından önemlidir. Fakat bunların olduğu gibi kullanılması da doğru değildir. Nitekim bazı öğretmenler bu tür puanlama anahtarlarını aynen kullanarak öğrencilerine not vermektedirler. Oysa bu tür araçların geliştirilmesi kadar kullanılması da uzmanlık istemektedir. Bunun yanında bazı öğretmenler dereceli puanlama anahtarını kullanmayı kendilerine bir yük olarak görmekte ve puanlama işlemini gelişigüzel yapabilmektedirler. Öğretmen yetiştiren programlarda ölçme ve değerlendirme dersinin kaldırılmış olması “yeni ölçme ve değerlendirme yaklaşımları”yla ilgili bilgilerin uygulamaya konmasını ve yaygınlaştırılmasını oldukça zorlaştırmaktadır. Bu olumsuz etkenlere rağmen yeni ölçme ve değerlendirme yaklaşımlarının öneminin vurgulanması ve öğretmenlerin bilgilendirilerek kullanımının yaygınlaştırılması okul eğitimine katkı getirecektir.

KAYNAKÇA

- Alison, A. T. (1999). **Authentic Assessment**. England: Friends & Friends Puplicing.
- Brookhart, S. M. (1999). *The Art and Science of Classroom Assessment: The Missing Part of*

- Pedagogy. ASHE-ERIC **Higher Education Report** (Vol. 27, No.1). Washington, DC: The George Washington University, Graduate School of Education and Human Development.
- Brown, J. D., Hudson T. (1998). **The alternatives in language assessment**. *Tesol Quarterly*, 32, 657.
- ERIC (1998, August). Implementing Performance Assessment in The Classroom.
- Web:
www.ed.gov/databases/ERIC_Digests/ed423312.html
- Goodrich, A. H. (2001). The effects of instructional rubrics on learning to write. **Current Issues in Education**, 4, (4).
- Haladyna, T. M. (1997). **Writing Test Item to Evaluate Higher Order Thinking**. USA: Allyn & Bacon.
- Kutlu, Ö. (2002). **Öğretmen Yetiştirme Programlarının Yeni Ölçme ve Değerlendirme Yaklaşımlarını İçerme Düzeyi**. Uluslararası Katılımlı 2000'li Yıllarda I. Öğrenme ve Öğretme Sempozyumu. 29-31 Mayıs, Marmara Üniversitesi, Atatürk Eğitim Fakültesi.
- Marzano, R. J.; Pickering, D. & Mctighe, J. (1993). **Assessing Student Outcomes**. Virginia: ASCD Publishing.
- Mehrens, W. A. (1992). Using Performance Assessment for Accountability Purposes. **Educational Measurement**. 11, (4) 3-9.
- O'Malley, J. M., Pierce, V. D. (1996). **Authentic Assessment**. Boston: Addison-Wesley Publishing Company, Inc.
- Popham, J. W. (1997). What's wrong and what's right with rubric. **Educational Leadership**. 55, (2), 12
- Popham, J. W. (2000). **Modern Educational Measurement**. Needham: Allyn & Bacon.
- Stark, J. F. (1998). **Measurement and Evaluation in Education**. Ohio: Great & Great Publishers.
- Stiggins, R. J. (1991). Facing The Challenges of New Era of Educational Assessment. **Applied Measurement in Education**. 4, 263-272.
- Stiggins, R. J. (1994). **Student-Centered Classroom Assessment**. New York: Macmillan Publishing Company.
- Şahin, B. (2003). Matematik Dersinde Kavram Haritası Yöntemini Kullanarak Öğrenci Başarısının Değerlendirilmesine İlişkin Bir Araştırma. Yayınlanmamış Doktora Tezi. Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü.

Ek - 1 : 2. PERFORMANS GÖREVİ

İçerik Düzeyi	Sınıf Düzeyi	Beklenen Performans	Puanlama Yöntemi
Fen Bilgisi	İlköğretim 6. sınıf	Öneri Geliştirme ve Sorun Çözme Becerisi	Dereceli Puanlama Anahtarı (Rubric)

Uzayda yaşanabilecek yeni bir gezegen bulunduğunu ve bu gezegende yeni bir yaşam kurma görevinin de size verildiğini düşünün. Amacınız insanlık için az sorunlu ve mutlu bir düzen kurmaktır.

Buna göre, dünyamızda insanların yaşadığı temel sorunları düşünerek, bu yeni gezegen de sorunsuz bir düzen için ne tür önlemler alırdınız?

Sorunsuz bir düzeni kurmak için neler yapardınız?

DİKKAT :

Bu sorunun yanıtı için öncelikle, dünyamızda yaşanan sorunların neler olduğunu belirleyiniz.

Bu sorulara neden olan ortak etkenleri düşününüz.

Daha sonra ne tür önlemler alacağınızı belirleyiniz ve alacağınız önlemleri gerekçelere dayandırınız.

Ek - 2: PERFORMANS GÖREVLERİNE İLİŞKİN YOL GÖSTERİCİ AÇIKLAMALAR

Öğrencilerin;

1. Verilen konulardan ilgi duydukları birini seçerek bilgi toplamaları,
2. Topladıkları bilgileri, grup önünde tartışılacak biçime getirmeleri (Gerekirse bunun için sunu yazılı duruma getirilebilir),
3. Sunuyu sınıfta yazılı bir metni okuyarak değil, konuşarak yapmaları,
4. Sunuyu 10 dakikalık bir sürede gerçekleştirmeleri,
5. Görüşleri, çelişkili açıklamalarla değil, tutarlı biçimde dile getirmeleri,
6. Gerekli durumlarda tablo, grafik ve resimlerden yararlanmaları (Bunun için poster, harita, tepegöz gibi araçlar kullanılabilir),
7. Görüşleri destekleyecek farklı kaynaklardan yararlanmaları,
8. Görüşleri, dikkatle ve ilgiyle dinletecek biçimde sergilemeleri,
9. Sunu sırasında kendine güvenen ve inandırıcı bir görünüm içinde olmaları,
10. Önerilerde bulunurken kendine özgü düşünceler geliştirerek, yeni görüşlere ulaşmaları,
11. Sununun bir örneğini yazılı olarak teslim etmeleri

beklenmektedir.

Ek - 3: DERECELİ PUANLAMA ANAHTARI (RUBRİK)

Öğrencinin

Adı Soyadı :
 Numarası :
 Sınıfı :

Puan	İÇERİK
4	Konu ile ilgili öne sürülen düşünceler mantıklı gerekçelerle desteklenmişti. Konuyla ilgili çelişkili açıklamalar yapılmadı.
3	Konu ile ilgili öne sürülen düşünceler mantıklı gerekçelerle desteklenmişti ama yeterli değildi. Konuyla ilgili çelişkili açıklamalar yapılmadı.
2	Konuyla ilgili öne sürülen düşünceler mantıklı gerekçelerle desteklenmişti ama yeterli değildi, anlatımda ardışık olmayan çelişkili açıklamalar yapıldı.
1	Konuyla ilgili öne sürülen mantıklı düşünceler oldukça azdı. Anlatımda ardışık olmayan çelişkili açıklamalar yapıldı.
0	Konu ile ilgili öne sürülen düşünceler mantıklı gerekçelere dayandırılmamıştı. Konuyla ilgili çelişkili açıklamalar yapıldı.

Puan	ANLATIM
4	Konu, dinleyicinin anlayabileceği, ilgisini çekebileceği şekilde sunuldu.
3	Konu, dinleyicinin anlayabileceği şekilde sunuldu fakat çok ilginç değildi.
2	Konu, dinleyici tarafından zor anlaşıldı, fakat sunum ilginçti.
1	Konu dinleyici tarafından zor anlaşıldı, sunum ilginç değildi.
0	Dinleyici konuyu hiç anlamadı, sunum ilginç değildi.

Puan	MATERYAL KULLANIMI
4	Sunum hedefe yönelik dört farklı materyalle desteklendi.
3	Sunum hedefe yönelik üç farklı materyalle desteklendi.
2	Sunum hedefe yönelik iki farklı materyalle desteklendi.
1	Sunum hedefe yönelik bir farklı materyalle desteklendi.
0	Sunumda materyal kullanılmadı.

Puan	İLETİŞİM
4	Akıcı bir dil kullanıldı. Dinleyiciyle göz teması kuruldu. Ses tonu ve beden dili doğru biçimde kullanıldı.
3	Belirtilen özelliklerden üçü yerine getirildi.
2	Belirtilen özelliklerden ikisi yerine yerine getirildi.
1	Belirtilen özelliklerden biri yerine getirildi.
0	Belirtilen özelliklerin hiç biri yerine getirilmedi.

Puan	ZAMANI KULLANMA
4	Verilen süre içinde sunuyu tamamladı.
3	Verilen süreye +/- 2 dakika uymadı.
2	Verilen süreye +/- 3 dakika uymadı.
1	Verilen süreye +/- 4 dakika uymadı.
0	Verilen sürenin sürenin oldukça altında veya üstünde sunumu tamamladı