

## FEN ÖĞRETİMİNDE PERFORMANS DEĞERLENDİRMENİN BİLİMSSEL SÜREÇ BECERİLERİ, TUTUM VE KALICILIK AÇISINDAN YANSIMALARI

### REFLECTIONS OF PERFORMANCE ASSESSMENT ON SCIENCE PROCESS SKILLS, ATTITUDE AND RETENTION IN SCIENCE EDUCATION

İlke ÖNAL ÇALIŞKAN \*, Fitnat KAPTAN \*\*

**ÖZET:** Araştırmanın temel amacı, performans değerlendirme öğrencilerin bilimsel süreç becerileri ve tutum ve kalıcılık değişkenleri açısından yansımalarını incelemektir. Araştırma, Ankara İli Beytepe İlköğretim Okulu'ndan üç tane yedinci sınıfta (biri deney, ikisi kontrol grubu ) toplam 105 öğrenciyle gerçekleştirilmiştir. Ölçme araçları olarak, Önal (2005) tarafından geliştirilen Bilimsel Süreç Becerileri Testi (BSBT), literatürden alınan Fen Bilgisi Tutum Ölçeği (FBTÖ) kullanılmıştır. Deney grubunda süreç boyunca öğrencilere değerlendirme sürecinde performans dayanaklı etkinlikler verilmiş, bu etkinliklere ilişkin rubrikler geliştirilmiştir. Süreç sonunda, gruplara BSBT ve FTÖ yeniden uygulanarak tutum ve başarı açısından bir farklılık olup olmadığına bakılmıştır. Uygulamanın tamamlanmasından bir buçuk ay sonra, kalıcılık testi uygulanarak üç grupta bilimsel süreç becerilerindeki kalıcılığa bakılmıştır. Araştırma, deneysel araştırma yöntemine ait desenlerden ön-test son-test kontrol gruplu desene uygun olarak tasarlanmıştır. Araştırmada elde edilen veriler, t-testi ve tekrarlayan verilerde varyans analizi tekniği kullanılarak analiz edilmiştir. Uygulama sonrasında, gruplar arasında bilimsel süreç becerileri ve tutum açısından deney grubu lehine anlamlı bir farklılık olduğu saptanmıştır.

**Anahtar sözcükler:** Fen ve Teknoloji Eğitimi, Bilimsel Süreç Becerileri, Performans Dayanaklı Değerlendirme, Rubrik (Dereceli Puanlama Anahtarı)

**ABSTRACT:** The aim of this study was to identify the reflections of performance-based assessment on science process skills, attitudes and retention variables. Study was conducted in Ankara Beytepe Primary School with three seventh grade (one experimental and two control groups) totally 105 students. Science Process Skills Test (SPST) which was developed by Önal (2005) and Attitude towards Science Scale (ATSS) which was taken from literature was used as measurement tools. Performance activities were taken to students and rubrics were developed during the process. After this, SPST and ATSS were applied again to the students for identifying if there was a difference between the groups. Retention test was applied to groups to look for if science process skills were permanent. Research was pre-test post-test control group experimental design. Data were analyzed by t-test and repeated measures of ANOVA. There was a statistical significant difference of variables was found in favor of experimental group.

**Keywords:** Science and Technology Education, Science Process Skills, Performance Based Assessment, Rubrics (Graded Scoring Keys)

## 1. GİRİŞ

Bilim ve teknoloji çağı olarak nitelendirilen günümüzde adeta takip edilemez teknolojik gelişmelerin yanında, bu bilgi ve teknolojiyi insanlık ve çevre yararına kullanabilecek nitelikli insan gücüne büyük ihtiyaç duyulmaktadır. Yirmi birinci yüzyılın insan nitelikleri arasında düşünme, sorgulama, düşündüğünü doğru ve net bir biçimde ifade edebilme, bilime önem ve değer verme, bilimsel bilgileri günlük yaşantıya aktarabilme gibi özellikler bulunmaktadır (Önal, 2005). Bu özellikler bireye, kendi yaşantısı yoluyla ya da okullarda tasarlanan öğrenme-öğretme süreçleriyle kazandırılabilir. Eğitim, istedik doğrultuda davranış geliştiren ve girdi, süreç ve çıktılardan oluşan bir sistemdir (Baykul, 2000). Toplumların çağın gereklerine uygun bir şekilde gelişebilmeleri için eğitim sistemlerinin sürekli gelişim ve değişime açık olması, kendini yenileyen dinamik sistemler olması gereklidir (Ertürk, 1982; Bilen, 1999).

Bilgi çağının gerektirdiği niteliklerin başında olan üst düzey düşünme beceriyle bilimsel yöntem süreç becerilerinin kazandırıldığı derslerin başında Fen Bilgisi dersi gelir. Bu derste, çocukların içinde yaşadıkları çevreyi ve evreni bilimsel yönden ele alıp incelemeleri amaçlanır (Kaptan, 1999).

\* Öğr.Gör.Dr., Hacettepe Üniversitesi, ilkeonal@hacettepe.edu.tr

\*\* Prof.Dr., Hacettepe Üniversitesi, fitnat@hacettepe.edu.tr

Çağın ve eğitim sisteminin ihtiyaçları göz önünde bulundurularak 2005-2006 eğitim-öğretim yılında yürürlüğe giren yapılandırmacı öğrenme yaklaşımına göre oluşturulan “Fen ve Teknoloji Programı”nın vizyonu, bireysel farklılıkları ne olursa olsun bütün öğrencilerin fen ve teknoloji okur yazarı olarak yetişmesidir. Fen ve teknoloji okur yazarlığı, genel bir tanım olarak; bireylerin araştırma-sorgulama, eleştirel düşünme, problem çözme ve karar verme becerileri geliştirmeleri, yaşam boyu öğrenen bireyler olmaları, etraflarındaki dünya hakkındaki merak duygusunu sürdürmeleri için gerekli olan fenle ilgili beceri, tutum, değer, anlayış ve bilgilerin bir birleşimidir (MEB, 2004).

Fen eğitiminde kazandırılması öngörülen becerilerin başında bilimsel süreç becerileri gelmektedir. Bilim adamlarının doğayı incelemede kullandıkları becerilere ve düşünme süreçlerine bilimsel süreçler denir. Bu süreçler; temel ve deneysel süreçler olmak üzere ikiye ayrılmaktadır (TIMSS, 2003; OECD, 2002). Temel Süreçler, gözlem yapma, sınıflama, ölçme ve sayıları kullanma, uzay ve zaman ilişkilerini kurma, yordama, önceden kestirme; deneysel süreçler ise hipotez kurma ve yoklama, değişkenleri belirleme ve kontrol etme, yaparak tanımlama, model yaratma, deney düzenleme ve yapma gibi becerileri kapsamaktadır.

Lawson (1995), bilimsel süreç becerilerini günlük yaşamdan karşılaşılan problemler doğrultusunda bilimle ilgili sorular oluşturma, oluşturulan problemlerle ilgili değişkenleri tanımlama ve kontrol etme, hipotezler kurma ve bütünsel düşünme, tahminlerde bulunma, sonuçlar çıkarma, ilişkisel düşünme olarak tanımlar. Bilimsel süreç becerileriyle üst düzey bilişsel becerilerin öğrencilere kazandırılması, eğitim sisteminin her basamağında yapılacak yeniliklerle mümkün olacaktır. Bu doğrultuda süreç içerisinde yapılan eğitim öğretim etkinlikleri kadar ölçme ve değerlendirme etkinlikleri de önem kazanmaktadır.

Değerlendirme, ölçme sonuçlarını bir ölçüte vurarak, ölçülen nitelik hakkında bir değer yargısına varma sürecidir (Turgut, 1997). Lester’e (1995) göre öğrencilerin başarılarını ve becerilerini ölçmek için kullanılan yöntemlerin uygunluğu bireyin özelliklerine göre farklılık göstermektedir. Örneğin ölçüt dayanaklı testler, norm-referenced (standart dayanaklı) testleri gibi bir öğrencinin diğer öğrencilerle kıyaslanarak üstünlüğünü ölçmeyi belirli standartlara göre öğrencinin performansını ölçtüğünden, daha fazla tercih edilerek kullanılmaktadır.

Matematik ve fen disiplin alanlarında çoktan seçmeli testler gibi önem verilen kağıt-kalem testleri artık öğrencilerin yeteneklerini ölçmede yeterli olmamaktadır (Hamm ve Adams, 1991). Yeni değerlendirme anlayışı; güvenilir, performans temelli, işbirliğine dayalı, etkin ve gerçek yaşama ilişkin öğrenmeleri yansıtan, gerçekçi ve uygulanabilir özelliklere sahiptir (Spady ve Marshall, 1991).

Kutlu’ya (2002) göre dünyanın birçok ülkesinde olduğu gibi Türkiye’de de sosyoekonomik, kültürel, siyasal vb. sorunların temelinde eğitimden de gelen olumsuzlukların yer aldığı bilinmektedir. Olumsuzluğa yol açan eğitimle ilgili nedenlerden birisi de, okullarda ölçme ve değerlendirme uygulamalarının kullanım biçimidir. Öğretim sürecinde öğrencilerin belirli kaynaklara dayalı bilgileri ne kadar öğrendiğine (hatırlama, ezberleme, kopyalama vb.) dayalı bir ölçme-değerlendirme anlayışı sürdürülmektedir. Öğrenme ortamlarında öğrencilere bilimsel süreç becerileriyle üst düzey düşünme becerilerinin ölçme değerlendirme basamağında yoklanması ve geliştirilmesi için dönüt alınması konusunda klasik değerlendirme yaklaşımları olarak nitelendirilen yazılı kağıt- kalem testlerinin yetersiz kaldığı, bunun yanında ölçme ve değerlendirme sürecinde kullanılması önerilen yeni ölçme ve değerlendirme yaklaşımlarından performans dayanaklı değerlendirme ve dereceli puanlama anahtarı (rubrik) ile ilgili bilgilere bundan sonraki kısımda yer verilmiştir.

### 1.1. Performans Değerlendirme

Performans değerlendirmesi, öğrencinin çeşitli problem durumları karşısında gösterdiği tepkilerin teşhis edilerek eksiklik ve yeterliklerinin öğrenciyle birlikte belirlenmesini sağlayan bir değerlendirme biçimidir. Bu değerlendirme biçimi disiplinler arası bilgiyi birleştirme, analiz, yorumlama, yordama yapma gibi üst düzey düşünme becerilerini ölçer (Berk, 1994). Performans değerlendirmelerin güvenilirliğini arttırmak amacıyla rubrikler (dereceli puanlama anahtarı) kullanılır.

## 1.2. Rubrik (Dereceli Puanlama Anahtarı)

Harrington'a (1984) göre rubrik "Güvenilir bir kural, bir eseri incelerken ve değerlendirirken kullanılan ölçütler bütünüdür". Rubrik, öğrenci performansını tanımlayan ölçütleri ve farklı seviyelerdeki performansı bu ölçütlerle birlikte puanlama rehberidir. Çünkü rubrikler özel ölçütleri belirler, değerlendirme şeklinin bu ölçütlerle uyuşması için kesin ve net yönergeler tanımlarlar ve sık sık her seviyedeki performansa sayısal puanlar verir, öğretmenlere kapalı uçlu değerlendirmelerin sağladığı objektifliği, açık uçlu yöntemler olan performans ve portfolyo gibi değerlendirmeler için sağlamada da olanak tanırlar (Luft, 1999).

Taggart ve diğerlerine göre rubrikler iki alt başlık altında incelenebilir:

### 1.2.1. Analitik (Analytical) Rubrik

Analitik rubrik, bir değerlendirmenin farklı bölümlerindeki çok belirli cevapları puanlamada kullanılır. Değerlendirmenin her bölümünden bütün puanlar öğrencinin performansın seviyesini ve toplam puanını belirlemek için toplanır. Puanlama ölçütleri nasıl puan verilebileceği ya da verilemeyeceği konusunda son derece belirlidir ( Taggart ve diğerleri, 1998).

### 1.2.2. Holistik Rubrik (Bütünsel Puanlama Anahtarı)

Holistik rubrik, öğretmen öğrencinin cevabının toplam olarak değerlendirmek istediğinde kullanılır. Holistik rubrik, süreçten ziyade sonuçla ilgilidir. Analitik puanlamada yapıldığı gibi sonuca ulaşmak için aşılacak bireysel basamaklarla ilgilenmekten çok toplam performans ya da sonuçla ilgilenir (Taggart ve diğerleri, 1998). Rubriklerin güvenilirliğiyle ilgili çalışmalar, genellikle puanlayıcıların tutarlılığıyla ilgilidir. Rubriklerin güvenilirliğini arttırmak için iyi tanımlanmış performans birimleri oluşturulmalı, puanlama çizelgesinde seviye sayısı artırılmalı, rubrikler, alanında uzman birden fazla kişiye puanlatırılmalıdır (Haladyna, 1997).

Bu araştırma; fen öğretiminde, eğitimcilere günümüz bilgi çağının gerektirdiği insan gücüne ulaşmada bilimsel süreç becerilerinin geliştirilmesi ve yoklanmasına yönelik etkinliklerin ölçme değerlendirme basamağında kullanılmasıyla bu etkinliklerin objektif olarak puanlanabilmesinde kullanılan rubriklerin geliştirilmesi konusunda rehber olması beklenmektedir. Ayrıca, araştırma; ölçme ve değerlendirme bilim dallarının program geliştirme ve alan uzmanlarıyla yapacakları geçerlik ve güvenilirlik çalışmalarıyla şimdiye kadar uygulanan ölçme ve değerlendirme sistemlerine getireceği farklı vizyon ve çeşitlilik açısından da önem taşımaktadır (Önal, 2005).

## 2. YÖNTEM

Araştırma durum belirleme çalışması niteliğinde düzenlenmiş olup boylamsal araştırmanın bu çalışmaya konu olan nicel bölümü ön-test son-test kontrol gruplu yarı deneysel desene göre planlanmıştır. Araştırmanın problem ise, 'İlköğretim fen öğretiminde öğrenme sürecinde, performans değerlendirmenin bilimsel süreç becerileri, tutum ve kalıcılık açısından yansımaları nedir?' şeklinde ifade edilebilir.

Alt problemler;

1. İlköğretim fen öğretiminde performans değerlendirmenin uygulandığı grup ile uygulanmadığı grup arasında bilimsel süreç beceri puanları arasında anlamlı bir fark var mıdır?

2. İlköğretim fen öğretiminde performans değerlendirmenin uygulandığı grup ile uygulanmadığı grup arasında fenbilgisi dersine karşı tutum puanları arasında anlamlı bir fark var mıdır?

3. İlköğretim fen öğretiminde performans değerlendirmenin uygulandığı grup ile uygulanmadığı grup arasında kalıcılık puanları arasında anlamlı bir fark var mıdır? olmak üzere üç bölümde ele alınmıştır.

## 2.1. Çalışma Grubu

Bu araştırmada çalışma grubunu Ankara ili Çankaya ilçesi'nden Beytepe İlköğretim Okulu'ndan convenience (ulaşılabilir) örnekleme yoluyla belirlenen yedinci sınıf öğrencilerden oluşan biri deney diğer ikisi kontrol grubu olmak üzere üç öğrenci grubu oluşturmaktadır. Araştırma, deney grubunda 36, deney grubuyla aynı öğretmene ait kontrol grubunda 32, deney grubuyla farklı öğretmene ait kontrol grubunda 37 olmak üzere toplam 105 tane yedinci sınıf öğrencisiyle yürütülmüştür.

## 2.2. Deneysel İşlem Basamakları

1. Deney ve kontrol gruplarının denkliğinin belirlenmesi amacıyla her iki gruba da süreç öncesinde araştırmacılar tarafından geliştirilen Bilimsel Süreç Becerileri Testi ile öğrencilerin demografik özelliklerini, çalışma alışkanlıkları, fen'le ilgili görüşleri, geçen yılki fen dersi karne notları vs. özelliklerin bulunduğu 'Kişisel Bilgi Formu' ile Geban ve arkadaşları tarafından geliştirilen Fen Bilgisi Tutum Ölçeği uygulanarak tutum ve başarılarındaki denklige bakılmıştır.

2. Deney ve kontrol grupları düzenli olarak araştırmacı tarafından izlenmiştir. Araştırmacıya süreç içerisinde farklı bir araştırmacı gözlemci olarak eşlik etmiştir. Deney ve kontrol gruplarının denkliğini sağlamada bir başka değişken olarak deney ve kontrol gruplarındaki öğretmenlerin önceki yıllarda ders işleme tarzlarıyla ilgili ön görüşmeler yapılmış ve dersleri nasıl planladıkları tespit edilmiştir. Deney ve kontrol gruplarındaki iki farklı öğretmen süreç içerisinde içerik ve ders işleme yöntemleri açısından incelenmiş, dersler teybe kaydedilerek çözümlenmiş gözlem kayıtlarının analizler sonucunda iki farklı öğretmenin benzer yöntemler kullanıp aynı doğrultuda konuları işlediği saptanmıştır.

3. Deney grubunda 7.Sınıf I.Ünite olan 'Maddenin İç Yapısına Yolculuk' ünitesi süresince her konu sonunda öğrencilere değerlendirme etkinlikleri kapsamında konuyla ilgili oyun kurma, hikaye yazma, proje tasarlama, bilim insanının yaşamını araştırma, deney raporu oluşturma, drama gibi öğrencilerin bilimsel süreç becerilerini geliştirmeye dönük etkinlikler geliştirilmiştir. Proje tasarlama, bilim insanının yaşamını araştırma ve deney raporu oluşturma etkinliklerinde öğrencilerin gözlem, sınıflama, sıralama, uzay-zaman ilişkileri kurma gibi temel bilimsel süreç becerileriyle bilimsel araştırma yönteminin de aşamaları olan hipotez kurma, deney tasarlama, sonuçları yorumlama, rapor oluşturma gibi deneysel bilimsel süreç becerileri, drama etkinlikleriyle de ısınma ve oyun aşamalarındaki gözlem, sınıflama ve uzay-zaman ilişkileri kurma gibi bilimsel süreç becerilerinin yanında doğaçlama ve değerlendirme bölümlerinde yaratıcı, yansıtıcı, eleştirel düşünme ve günlük yaşam durumlarına uyarılma gibi üst düzey düşünme becerilerinin geliştirilmesi sağlanmıştır. Öncesindeki pilot çalışmalar altıncı sınıf "Uzayı Keşfediyoruz" ünitesinde 5 hafta süre ile gerçekleştirilmiştir. Toplamda 14 hafta süren deneysel süreçte 9 etkinlik ve bir tane de genel etkinlik değerlendirmesi olmak üzere öğrencilerden toplam 10 tane çalışma yapmaları istenmiş, her çalışma kağıdının sonuna değerlendirmede göz önünde bulundurmaları gereken nitelikleri listelenmeleri istenmiş, hafta başında öğrencilerden toplanan etkinlikler, deney grubu fen ve teknoloji öğretmeniyle bir araya gelinerek oluşturulan değerlendirme ölçütleriyle öğrencilerin yazmış oldukları değerlendirme ölçütleri bir araya getirilmiş, ortak olanlar bir grupta toplanarak her çalışma kağıdı için rubrikler oluşturulmuştur. Bu sırada kontrol gruplarına sürece dayalı performans değerlendirme etkinlikleri verilmemiş ve öğrencilerle birlikte rubrikler geliştirilmemiş, sadece öğretmenlerin ortak olarak hazırladıkları sınavlar uygulanmıştır. Kontrol grubunda da deney grubunda olduğu gibi öğretmen ve öğrenci merkezli öğretim yöntemleri birarada uygulanmış ancak değerlendirme boyutunda yalnızca sonuç odaklı yöntemler tercih edilmiştir.

4. Süreç sonunda deney grubu ve her iki kontrol grubuna BSBT ve Fen Bilgisi Tutum Ölçeği yeniden uygulanarak başarı ve tutuma ait son test puanları elde edilmiştir.

5. Hatırlamanın olmayacağı kadar uzun, yeni öğrenmelerin olmayacağı kadar kısa olan 1,5 aylık sürenin ardından BSBT'nin seçenek ve sorularının yerleri hatırlamayı önlemek amacıyla değiştirilerek deney ve kontrol gruplarına tekrar uygulanarak öğrenmedeki kalıcılığa bakılmıştır.

### 2.3. Veri Toplama Araçları

**2.3.1. Bilimsel Süreç Becerileri Testi (BSBT)-** 1999'da tekrarı yapılan TIMSS (Third International Mathematics and Science Study) ve OÖKSYS (Orta Öğretim Kurumları Seçme ve Yerleştirme Sınavı)'nin fen bilgisi sorularından bilimsel süreç becerilerini yoklayanları incelenmiş araştırmacı tarafından 50 sorudan oluşan öğrencilerin bilimsel yöntem süreç becerilerini yokladığı uzman görüşü alınıp redaksiyon çalışması yapılan 'Bilimsel Süreç Becerileri Testi' geliştirilmiştir. Kapsam geçerliğini sağlamak amacıyla ilköğretim 4'ten 8. sınıfa kadar fen üniteleri ve kazanımlar dikkate alınarak hazırlanmıştır. Geliştirilen testin pilot uygulaması Ankara ili Çankaya ilçesi Beytepe İlköğretim Okulu ve Yasemin Karakaya İlköğretim Okulu öğrencilerinden oluşan 136 kişilik öğrenci grubuna uygulanarak madde istatistikleri ITEMAN programında hesaplanarak madde gücü ve madde ayıricılıkları göz önünde bulundurularak başta 50 soru olarak belirlenen test, 36 soruya indirilmiştir. Testin güvenilirlik katsayısı 0.89 olarak tespit edilmiştir.

**2.3.2. Tutum Ölçeği:** Araştırmada literatürden alınan, geçerlik ve güvenilirlik çalışması yapılan 'Fen Bilgisi Tutum Ölçeği' (Geban ve diğerleri, 1994) kullanılmıştır. Fen Bilgisi Tutum Ölçeği'nde 5'li likert tipinde (tamamen katılıyorum, katılıyorum, kararsızım, katılmıyorum, hiç katılmıyorum) hazırlanmış 15 tane önerme bulunmaktadır. Ölçeğin güvenilirliği **0,83** olarak bulunmuştur.

**2.3.3. Performans Dayanaklı Etkinlikler ve Rubrikler:** Araştırmacı tarafından her etkinlik için süreç içerisinde fen bilgisi öğretmeni ve öğrencilerle birlikte dereceli puanlama anahtarı belirlenmiştir. Taggart, Phifer, Nixon ve Wood'un (1998) belirlediği süreçlerden yararlanılmıştır. Bu süreçler, uzman görüşü alma, performans ölçütlerini tanımlayarak listeleme, performans düzeylerini belirleyerek holistik dereceli puanlama anahtarı hazırlama şeklindedir. Uzman görüşünde, fen ve teknoloji ders öğretmeniyle araştırmacının puanları arasında Kendall w ilişkisine bakılmış ve korelasyon **0,784** çıkmıştır. Bu korelasyon **p=0,01** düzeyinde anlamlı bir değerdir. Bu değer öğrencilerin örnek bir etkinlik üzerinde verdikleri puanlarla araştırmacının puanı arasındaki ilişkiye bakıldığında **0,542** çıkmıştır. Bu değer düşük olması, öğrencilerin süreçle ilgili yeni uygulama yapıyor olmalarına bağlanabilir.

### 2.4. Verilerin Çözümlemesi

Araştırmaya ait alt problemlerin verilerinin çözümlemesinde ortalamalar arası fark istatistiği (t-testi) ve tekrarlayan verilerde varyans analizi (Repeated Measures of ANOVA) tekniğinden faydalanılmıştır. Verilerin analizinde SPSS 15.0 paket programı kullanılmıştır.

### 2.5. Araştırmanın Geçerliliği ve Güvenirliği

Ölçme araçları geçerlik ve güvenilirliklerinin yanı sıra, deneysel araştırmada göz önünde bulundurulması gereken iç geçerliği tehdit eden faktörler arasında sayılan araştırmacı yanlılığını ortadan kaldırmak amacıyla sürecin gözlemine farklı bir araştırmacı da dahil olmuş, sınıf ortamı ve öğretmenden kaynaklanabilecek yanlılıkları ortadan kaldırmak amacıyla her üç grupta da gözlem kayıtları tutularak oluşturulan kontrol listeleri kapsamında incelenmiş, veri toplama araçları her üç gruba eşit koşullarda uygulanmıştır (Fraenkel ve Wallen, 2003).

## 3. BULGULAR

Araştırmada, deney ve kontrol gruplarının denkliliğini sağlamak amacıyla; öğrencilerin fen ve teknoloji karne notu ortalamaları, anne-baba eğitim durumları, çalışma alışkanlıkları, fen ve teknoloji derslerindeki projeleri hazırlarken kullandıkları araç-gereçler, fen ve teknoloji derslerinde kavramlar

arası bağ kurma düzeyleri, öğrendikleri kavramları günlük yaşama transfer etmeleri, bilim ve teknolojiyle ilgili yayınları takip etme sıklıkları gibi demografik özellikleriyle fen ve teknoloji derslerindeki durumlarını belirlemeye yönelik sorulara benzer yanıtlar verdikleri yüzde ve frekans değerleriyle tespit edilmiştir. Bunun yanında ön bilimsel süreç becerileri puan ortalamalarının 36 üzerinden sırasıyla 36 kişilik deney grubunda 21,61; 32 kişilik birinci kontrol grubunda 21,47; 37 kişilik ikinci kontrol grubunda 22,14 olduğu saptanmıştır. Bu değerler  $p=.05$  düzeyinde anlamlı değildir. Bu durum, bilimsel süreç becerileri testi başarı puanları açısından deney ve kontrol gruplarının denk olduğunu göstermektedir. Ön tutum puanlarının ortalamalarının ise 75 üzerinden sırasıyla deney grubunda 55,28; birinci kontrol grubunda 54,41 ve ikinci kontrol grubunda 54,35 olduğu saptanmıştır. Bu değerler  $p=.05$  düzeyinde anlamlı değildir. Bu durum, fen dersine karşı tutum puanları açısından deney ve kontrol gruplarının denk olduğunu göstermektedir.

Araştırmanın üç probleminin analizinde tekrarlayan verilerde varyans analizi tekniği ve t-testi tekniğinden yararlanılmıştır. Deney ve kontrol gruplarının ön başarı, son başarı, kalıcılık başarısı, ön tutum ve son tutum değişkenleri açısından tekrarlanan verilerde varyans analizi sonucunda; değişkenlerin evrende normal dağıldığı, varyansların homojenliği (Sphericity testine göre  $p<0,05$ ) ve eş olasılıklı tesadüfi örnekleme sayıtları sağlanmıştır. Araştırmanın birinci alt problemine ilişkin değerler Tablo 3.1, 3.2 ve 3.3’de, ikinci alt probleme ilişkin değerler Tablo 3.4, 3.5 ve 3.6’da, üçüncü alt probleme ilişkin değerler Tablo 3.7, 3.8 ve 3.9’da belirtilmiştir.

**Tablo 3.1: 7/E Deney Grubu İçin Ön Başarı, Son Başarı Ve Kalıcılık Bilimsel Süreç Becerileri Değişkenleri İçin Tekrarlanan Verilerde Varyans Analizi**

Etki	Değer	F	Hipotez df	Hata df	Sig.	Kısmi		Observed Power(a)	
						Eta Squared	Noncent. Parameter		
7/E	Pillai's Trace	,772	57,650(b)	2,000	34,00	,000	,772	115,299	1,000
Deney	Wilks' Lambda	,228	57,650(b)	2,000	34,00	,000	,772	115,299	1,000
	Hotelling's Trace	3,391	57,650(b)	2,000	34,00	,000	,772	115,299	1,000
Grubu	Roy's Largest Root	3,391	57,650(b)	2,000	34,00	,000	,772	115,299	1,000

Tablo 3.1 e göre, çalışmanın başarı değişkeni açısından deney grubunda etkili olduğu da Wilk's Lambda'nın anlamlılık değerinin 0,05 den küçük olmasına dayanılarak söylenebilir. Wilk's Lambda=0,228;  $F(2,34)=57,65$ ;  $p<0,05$

**Tablo 3.2: 7/B Kontrol Grubu İçin Ön Başarı, Son Başarı Ve Kalıcılık Başarısı Değişkenleri İçin Tekrarlanan Verilerde Varyans Analizi**

Etki	Değer	F	Hipotez df	Hata df	Sig.	Kısmi		Observed Power(a)	
						Eta Squared	Noncent. Parameter		
7/B	Pillai's Trace	,082	1,565(b)	2,000	35,000	,223	,082	3,130	,309
Kontrol	Wilks' Lambda	,918	1,565(b)	2,000	35,000	,223	,082	3,130	,309
	Hotelling's Trace	,089	1,565(b)	2,000	35,000	,223	,082	3,130	,309
Grubu	Roy's Largest Root	,089	1,565(b)	2,000	35,000	,223	,082	3,130	,309

Tablo 3.2'ye göre, Ön başarı, son başarı ve kalıcılık başarısı açısından kontrol grupları arasında istatistiksel olarak anlamlı bir değişiklik görülmemektedir. Wilk's Lambda=0,228;  $F(2,34)=1,565$ ;  $p>0,05$

**Tablo 3.3: 7/C Kontrol Grubu İçin Ön Başarı, Son Başarı Ve Kalıcılık Başarısı Değişkenleri İçin Tekrarlanan Verilerde Varyans Analizi**

Etki	Değer	F	Hipotez df	Hata df	Sig.	Kısmi Eta Squared	Noncent. Parameter	Observed Power(a)	
7/C	Pillai's Trace	,025	,389(b)	2,000	30,000	,681	,025	,778	,107
Kontrol	Wilks' Lambda	,975	,389(b)	2,000	30,000	,681	,025	,778	,107
Grubu	Hotelling's Trace	,026	,389(b)	2,000	30,000	,681	,025	,778	,107
	Roy's Largest Root	,026	,389(b)	2,000	30,000	,681	,025	,778	,107

Tablo 3.3'e göre, ön başarı, son başarı ve kalıcılık başarısı açısından kontrol grupları arasında istatistiksel olarak anlamlı bir değişiklik görülmemektedir. Wilk's Lambda=0,228; F(2,34)=0,389; p>0,05

**Tablo 3.4: 7/E Deney Grubu İçin Ön Tutum Ve Son Tutum Değişkenleri İçin Tekrarlanan Verilerde Varyans Analizi**

Etki	Değer	F	Hipotez df	Hata df	Sig.	Kısmi Eta Squared	Noncent. Parameter	Observed Power(a)	
7/E	Pillai's Trace	,743	101,412 (b)	1,000	35,000	,000	,743	101,412	1,000
Deney	Wilks' Lambda	,257	101,412(b)	1,000	35,000	,000	,743	101,412	1,000
Grubu	Hotelling's Trace	2,897	101,412(b)	1,000	35,000	,000	,743	101,412	1,000
	Roy's Largest Root	2,897	101,412(b)	1,000	35,000	,000	,743	101,412	1,000

Tablo 3.4'e göre, çalışmanın tutum değişkeni açısından deney grubunda etkili olduğu da Wilk's Lambda'nın anlamlılık değerinin 0,05 den küçük olmasına dayanılarak söylenebilir. Wilk's Lambda=0,228; F(2,34)=101,412; p<0,05

**Tablo 3.5: 7/B Kontrol Grubu İçin Ön Tutum Ve Son Tutum Değişkenleri İçin Tekrarlanan Verilerde Varyans Analizi**

Etki	Değer	F	Hipotez df	Hata df	Sig.	Kısmi Eta Squared	Noncent. Parameter	Observed Power(a)	
7/B	Pillai's Trace	,001	,032(b)	1,000	36,000	,858	,001	,032	,054
Kontrol	Wilks' Lambda	,999	,032(b)	1,000	36,000	,858	,001	,032	,054
Grubu	Hotelling's Trace	,001	,032(b)	1,000	36,000	,858	,001	,032	,054
	Roy's Largest Root	,001	,032(b)	1,000	36,000	,858	,001	,032	,054

Tablo 3.5'e göre, ön tutum ve son tutum değerleri açısından kontrol grupları arasında istatistiksel olarak anlamlı bir değişiklik görülmemektedir. Wilk's Lambda=0,228; F(2,34)=0,32; p>0,05

**Tablo 3.6: 7/C Kontrol Grubu İçin Ön Tutum Ve Son Tutum Değişkenleri İçin Tekrarlanan Verilerde Varyans Analizi**

Etki	Değer	F	Hipotez df	Hata df	Sig.	Kısmi Eta Squared	Noncent. Parameter	Observed Power(a)	
7/C	Pillai's Trace	,004	,113(b)	1,000	31,000	,739	,004	,113	,062
Kontrol	Wilks' Lambda	,996	,113(b)	1,000	31,000	,739	,004	,113	,062
Grubu	Hotelling's Trace	,004	,113(b)	1,000	31,000	,739	,004	,113	,062
	Roy's Largest Root	,004	,113(b)	1,000	31,000	,739	,004	,113	,062

Tablo 3.6'ya göre, ön tutum ve son tutum değerleri açısından kontrol grupları arasında istatistiksel olarak anlamlı bir değişiklik görülmemektedir. Wilk's Lambda=0,228; F(2,34)=0,32; p>0,05

**Tablo 3.7: Deney Grubu Öğrencilerinin Son Başarı Ve Kalıcılık Başarı Ortalamaları Bağımlı Gruplarda T-Testi**

SINIF	BSB	ÖĞRENCİ SAYISI	SORU SAYISI	MIN	MAX.	$\bar{X}$	SS	R
Deney Grubu	SONTEST	36	36	28,08	5,116	0,956	0,452	0,654
	KALICILIK TEST	36	36	28,00	4,744			

Tablo 3.7'deki bağımlı gruplarda t-test sonuçlarına göre deney grubunun bilimsel süreç becerileri Son test ortalamaları ile kalıcılık test ortalamaları arasında istatistiksel olarak anlamlı bir fark bulunmamaktadır. Bu da deney sırasında kazanılan becerilerin kalıcı olduğunu göstermektedir.

**Tablo 3.8: Deney Grubuyla Aynı Öğretmenin Ders Verdiği Kontrol Grubu Öğrencilerinin Son Başarı ve Kalıcılık Başarı Ortalamaları Bağımlı Gruplarda T-Testi**

SINIF	BAŞARI	ÖĞRENCİ SAYISI	SORU SAYISI	$\bar{X}$	SS	R	T	P
I.Kontrol Grubu	SONTEST	32	36	21,53	3,951	0,977	-	0,403
	KALICILIK TESTİ	32	36	21,59	4,102			



**Tablo 3.9. Deney Grubuyla Farklı Öğretmenin Ders Verdiği Kontrol Grubu Öğrencilerinin Son Başarı Ve Kalıcılık Başarı Ortalamaları Bağımlı Gruplarda T-Testi (Paired Sample T-Test)**

SINIF	BAŞARI	ÖĞRENCİ SAYISI	SORU SAYISI	$\bar{X}$	SS	R	T	P	
II.Kontrol Grubu	SONTEST	37	36	22,14	4,744	0,980	-	1,550	0,130
	KALICILIK TESTİ	37	36	22,38	4,832				

Tablo 3.8 ve 3.9'daki bağımlı gruplarda t-test sonuçlarına göre her iki kontrol grubunda da Bilimsel Süreç Becerileri son test ve kalıcılık testleri ortalamalarında istatistiksel olarak anlamlı bir farklılık bulunmamaktadır. Bu durum süreç başından itibaren kontrol gruplarında bilimsel süreç becerileri anlamında herhangi bir gelişme bulunmadığını göstermektedir.

#### 4. YORUM/TARTIŞMA

Sonuç olarak; t-testi ve tekrarlanan verilerde tek yönlü varyans analizi sonuçlarına göre yapılan performans dayanaklı değerlendirme uygulamasının deney grubu lehine anlamlı bir etki yarattığı söylenebilir. Araştırmanın alt problemlerinden elde edilen bulgular; Geban (1991) ın kimya derslerini üç öğretim yöntemiyle ele aldığı araştırmasında, bilimsel süreç becerilerinin geliştirilmesine dönük performans dayanaklı etkinliklerin ve öğrencilerle birlikte belirlenen ölçütlerin işe koşulduğu sınıftaki başarının geleneksel laboratuvar öğretim yöntemi ve değerlendirme biçimleri kullanılan sınıftan daha yüksek olduğu bulgusuyla örtüşmektedir. Ayrıca Doğruöz (1998)'ün bilimsel süreç becerilerine dönük fen öğretim yönteminin öğrenci başarısı, fene karşı tutum ve bilimsel süreç becerileri değişkenlerine olan etkisini geleneksel öğretim yöntemiyle karşılaştırmalı incelediği çalışmasında, bilimsel süreç becerilerine yönelik uygulamalı, performans dayanaklı öğretim ve değerlendirmeye tabi tutulan öğrencilerin başarılarının geleneksel öğretimle öğrenen ve sonuç merkezli yaklaşımla değerlendirilen öğrencilerden daha yüksek olduğu da bu araştırmanın bulgularını doğrular niteliktedir. Tortop (2001)'un ise deneysel çalışmaların ilköğretim okullarındaki uygulama düzeyi, deneysel çalışmalardaki grup sayısı, deney raporunun hazırlanması gibi durumlar hakkında sonuçları ortaya koyduğu araştırmasında; deneysel ve uygulamalı çalışma ve değerlendirmelerin fen başarısını arttırdığı sonucuna ulaşması öğrenme ortamlarında performans değerlendirmenin önemini pekiştirmektedir. Toth, Suthers & Lesgold (2000)'ün bilimsel sorgulama sırasında temel akıl yürütme becerileri ve bunların öğretim ortamlarında kullanımı hakkında bilgi vermeyi hedefledikleri araştırmalarında; yansıtıcı sorgulama rubriklerinin öğrenme-öğretme sürecinde kullanımının, öğrencilerin bilimsel yöntem süreç becerilerini geliştirdiği sonucu araştırma bulgularını destekler niteliktedir. Gearhart ve diğerleri (2006) nin öğretmenlerin program ünitelerini desenleme, uygulama ve değerlendirmeye yönelik bir değerlendirme dosyası hazırlamalarına yönelik çalışmasında ölçüt belirlemenin öğretmenlerin derslere karşı özgüven ve tutumlarını olumlu etkilediği sonucu araştırmanın sonuç ve önerileriyle uyumludur. Andrade, Wang, Du ve Akawi (2009)'nin öz değerlendirmeyi de içeren kısa ve uzun süreli rubrik kullanımının cinsiyet ve öz yeterlik üzerine etkisini inceledikleri araştırmalarında rubrik kullanımının öğrencilerin özgüvenlerini yükselttiği ve Pilotti, Chodorow ve Petrov (2009) un araştırmasındaki özdeğerlendirmenin kavram öğretiminde etkili bir değerlendirme yaklaşımı olduğunu tespit etmeleri, bu çalışmada elde edilen bulgularla paralellik göstermektedir.

## 5. SONUÇLAR

İlköğretim öğrencilerinde bilimsel süreç becerilerinin geliştirilmesinde performans değerlendirmenin etkili olduğu gözlenmiştir. Deney grubundaki öğrenciler, araştırma sürecinde geliştirilen ve bilimsel süreç becerilerini ölçtüğü kabul edilen Bilimsel Süreç Becerileri Testi'nde (BSBT) süreç başındaki uygulamada kontrol gruplarıyla benzer başarı gösterirken, bilimsel süreç becerilerinin geliştirilmesine yönelik etkinliklerin uygulanarak, araştırma sırasında araştırmacı, fen bilgisi ders öğretmeni ve öğrenciler tarafından geliştirilen rubrikler aracılığıyla değerlendirme süreci sonrasında, kontrol gruplarındaki öğrencilerin ortalamasına göre anlamlı düzeyde bir artış göstermişlerdir.

İlköğretim öğrencilerinde fen bilgisi dersine karşı tutumlarda performans değerlendirmenin etkili olduğu gözlenmiştir. Deney grubundaki öğrenciler, araştırma sürecinde Fen Bilgisi Dersi Tutum Ölçeği'nde süreç başındaki uygulamada kontrol gruplarıyla benzer tutum puanı ortalaması sergilerken, bilimsel yöntem süreç becerilerinin geliştirilmesine yönelik etkinliklerin uygulanarak araştırma sırasında araştırmacı, fen ve teknoloji ders öğretmeni ve öğrenciler tarafından geliştirilen rubrikler aracılığıyla değerlendirme süreci sonrasında kontrol gruplarındaki öğrencilerin tutum puanı ortalamasına göre anlamlı düzeyde bir artış göstermişlerdir. Performans değerlendirme, öğrenci merkezli öğrenme-öğretme strateji, yöntem ve teknikleriyle paralellik gösteren ve süreç içerisinde öğrencilerin takip edilerek ölçütlerin öğrencilerle birlikte belirlendiği, öğrenme-öğretme sürecinden bağımsız olarak düşünülemeyen bir yaklaşımdır. Performans değerlendirme sürecinde öğrenciler grup etkinliklerine katılarak, model geliştirerek, deney tasarlayarak ya da özgün hikayelerini oluşturarak feni yaparak-yaşayarak deneyimleme fırsatı bulmuşlardır. Ayrıca değerlendirmede ölçütleri belirlerken sorumluluk aldıkları için fen derslerine ilişkin kaygıları azalarak tutumları olumlu yönde değişmiştir.

İlköğretim öğrencilerinde bilimsel süreç becerilerinin geliştirilmesi süresince kazanılan becerilerin kalıcılığında performans değerlendirmenin etkili olduğu gözlenmiştir. Deney ve kontrol grubu öğrencilerine, uygulamanın tamamlanmasından belirli bir süre sonra uygulanan kalıcılık testindeki ortalamayla son test ortalaması arasında anlamlı bir fark olmadığı, yani edinilen bilgi ve becerilerin kalıcı olduğu saptanmıştır.

## 6. ÖNERİLER

Araştırmanın sonuçları doğrultusunda eğitim alanında çalışan araştırmacılara ve fen ve teknoloji öğretmenlerine aşağıdaki öneriler sunulmuştur;

1. Performans değerlendirme gibi süreç odaklı değerlendirmeler öğrencilerin beceri gelişimine yönelik olduğundan uzun zaman ve emek ister. Bu nedenle öğretim programının kapsamı, bilgi yükü azaltılarak programda öğrencilerin bilimsel süreç becerilerinin geliştirilmesine ve değerlendirilmesine yönelik etkinliklere ağırlık verilebilir.

2. Öğrenciler performans değerlendirme sürecinde rubrik geliştirirken ölçüt belirlemede sorumluluk aldıklarından fen derslerine karşı kaygıları azalarak tutumları olumlu yönde etkilenmiştir. Bu sonuca göre, ilköğretimin ilk yıllarından itibaren öğrencilere süreç içerisinde ve sonunda yapılan sınavların analizini yapma alışkanlığı kazandırılabilir. Bu doğrultuda öğrencilere sınav sürecinde sınavda sorulan her sorunun kendisinden ne beklediğine ilişkin sorular yöneltilip sınav sonunda bu özellikleri listelemeleri istenebilir.

3. Öğrencilerin ölçüt belirlemede kendi öğrenme sorumluluklarını alarak fen öğrenmeye karşı isteklerinin artmasından hareketle öğrencilere verilen araştırma ödevi, proje, deney raporu gibi etkinlikleri hazırlama aşamasında nerelerde zorlandıkları ya da yaptıkları çalışmada hangi özelliklerin hangi önem derecesine sahip olduğunu hazırladıkları raporlarda belirlemelerinin istenebilir.

4. Performans değerlendirmeler, sonuç odaklı yazılı sınavlar ve çoktan seçmeli testlere göre daha subjektiftir. Performans değerlendirmelerin sınırlılığı kabul edilen subjektifliğini azaltmak, geçerlik ve güvenilirliğini arttırmak amacıyla aynı alandaki öğretmenler, öğrencilerle ilgili gözlemleri

doğrultusunda performans listeleri belirleyerek, aynı düzeydeki sınıflar için ortak rubrikler hazırlayabilirler.

5. Süreç odaklı ölçme ve değerlendirme yaklaşımlarının öğrenme ortamlarında kullanımı uzmanlık gerektirir. Milli Eğitim Bakanlığı aracılığıyla okullardaki ölçme değerlendirme uzmanlarıyla işbirliği yapılarak öğretmenlere, bilimsel süreç becerilerine yönelik çoktan seçmeli soruların hazırlanması, üst düzey düşünme becerilerini geliştiren etkinlikler tasarlanması, hazırlanan her etkinlik ya da ölçme aracının performans birimi ve değerlendirme yüzdesinin tanımlanması konularında hizmet içi eğitim seminerleri planlanıp uygulanabilir. Hizmet öncesinde ise, öğretmen adayları yeni ölçme-değerlendirme ve puanlama yollarıyla ilgili olarak detaylı bilgilendirilip bu konuda uygulamalar yaptırılabilir.

## KAYNAKLAR

- Andrade, H.L., Wang, X., Du, Y. ve Akawi, R.L.(2009). Rubric-referenced self-assessment and self-efficacy for writing. *The Journal of Educational Research*, 102 (4), 287-301.
- Baykul, Y. (2000). *Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması*. Ankara: ÖSYM Yayınları.
- Berk, R. A. (1994). Performance assessment of our own professional behavior. *Educational Measurement: Issues and Practice*, 13 (1), pp. 27-32, New York: Wiley Blackwell.
- Bilen, M. (1999). *Plandan Uygulamaya Öğretim*. Ankara: Anı Yayıncılık.
- Doğruöz, P. (1998). *Bilimsel işlem becerilerini kullanmaya yönelik yöntemin öğrencilerin akışkanların kaldırma kuvveti konusunu anlamalarına etkisi*. Yayımlanmamış Yüksek Lisans Tezi, Orta Doğu Teknik Üniversitesi, Ankara.
- Ertürk, S. (1982). *Eğitimde Program Geliştirme*. Ankara: Meteksan.
- Fraenkel, J.R. ve Wallen, N. (1993). *How to Design and Evaluate Research in Education*. New York: Mc. Graw Hill.
- Geban, Ö. (1990). *İki farklı öğretim yönteminin lise seviyesindeki öğrencilerin kimya başarılarına, bilimsel süreç becerilerine ve kimyaya karşı olan tutumlarına etkisi*. Yayımlanmamış Doktora Tezi, Orta Doğu Teknik Üniversitesi, Ankara.
- Geban, Ö., Ertepinar, H., Yılmaz, G., Altın, A. ve Şahbaz, F. (1994). Bilgisayar destekli eğitimin öğrencilerin fen bilgisi başarılarına ve fen bilgisi ilgilerine etkisi. *Birinci Ulusal Fen Bilimleri Eğitimi Sempozyumu Bildiri Özetleri Kitabı*. İzmir: Dokuz Eylül Üniversitesi.
- Gearhart, M., Nagashima, S., Clark, S., Schwab, C. ve Vendlinski, T., Osmundson, E., Herman, J. (2006). Developing expertise with classroom assessment in K-12 science: Learning to interpret student work. Interim findings from a two-year study. *Educational Assessment*, 11, 237-263.
- Haladyna, T. M. (1997). *Writing Test Items to Evaluate Higher Order Thinking*. USA: Allyn&Bacon A Viacom Company.
- Hamm, M. ve Adams, D. (1991). Portfolio: It's not just for artist anymore. *The Science Teacher*, 31 (8), 18-21.
- Harrington, T. F. (1984). Assessment of Abilities. *ERIC Digest, ERIC-RIEO, 19950101 Database: ERIC*, 1984. 12.04.2004 tarihinde [www.ericps.ed.uiuc.edu/eece/pubs/digests.html](http://www.ericps.ed.uiuc.edu/eece/pubs/digests.html) adresinden alınmıştır.
- Kaptan, F. (1999). *Fen Bilgisi Öğretimi*. Ankara: MEB Yayınları.
- Kutlu, Ö. (2004). Tek soruyla öğrenci performansının belirlenmesi. *Sabancı Üniversitesi Eğitimde Yeni Ufuklar Sempozyumu*. 08.09.2004 tarihinde <http://www.erg.sabanciuniv.edu/iok2004/bildiriler/Omer%20Kutlu.doc> adresinden alınmıştır.
- Lawson, A. E. (1995). *Science Teaching and The Development of Thinking*. USA: Wadsworth Inc.
- Lester, J. N. (1995). Assessing career development with portfolios. *ERIC Digest, ERIC-RIEO, 19950130, Database: ERIC*. 12.04.2004 tarihinde [www.ericps.ed.uiuc.edu/eece/pubs/digests.html](http://www.ericps.ed.uiuc.edu/eece/pubs/digests.html) adresinden alınmıştır.
- Luft, J. A. (1999). Rubrics. Design and use in science education. *Journal of Science Teacher Education*, 10(2): 107-121.
- MEB (2004). Talim ve Terbiye Kurulu Başkanlığı İlköğretim Okulu Fen ve Teknoloji Dersi Öğretim Programı. Ankara: Talim Terbiye Kurulu Yayınları.
- Önal, İ. (2005). *İlköğretim fen bilgisi öğretiminde performans dayanaklı durum belirleme uygulaması üzerine bir çalışma*. Yayımlanmamış Yüksek Lisans Tezi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- OECD (2002). *PISA 2000, Sample Tasks from the PISA 2000 Assessment*. New York: Mc Graw Hill.
- Pilotti, M., Chodorow, M. ve Petrov, R. (2009). The usefulness of retrieval practice and review-only practice for answering conceptually related test questions. *The Journal of General Psychology*, 136 (2), 179-204.
- Spady, W., ve Marshall, L. K. (1991). Beyond traditional outcomes-based education. *Educational Leadership*, 49, 67-72. Thousand Oaks, CA: Sage Publications.

- Taggart, G., Phifer, S.J., J. Nixon, A. & Wood, M. (1998). *Rubriks: A Handbook For Construction And Use*. Pennsylvania: Technomic Publishing Co, Inc.
- TIMSS (2003). *Assessment Frameworks and Specifications 2003*. USA: Boston College.
- Tortop, H. S. (2001). *İlköğretim ikinci kademesindeki öğrencilerin fen eğitiminde yapılan deneysel çalışmalar hakkındaki görüşlerinin değerlendirilmesi*. Yayımlanmamış Yüksek Lisans Tezi, Celal Bayar Üniversitesi, Manisa.
- Toth, E. E., Daniel, D. S. & Lesgold, A. M. (2000). *Mapping to know: The effects of representational guidance and reflective assessment on scientific inquiry*. *Journal Of Research In Science Teaching*, 86 (1), pp. 264-286, Wilmington, DE: Wiley Periodicals, Inc.
- Turgut, F. (1997). *Eğitimde Ölçme ve Değerlendirme Yöntemleri*. Ankara: Yargıcı Matbaası.

## Genişletilmiş Özet

Science and technology age requires people who have higher order thinking skills. These skills also cover the science process skills. Science classes are the ones which help to gain science process skills to the students. These classes aim to show the way to the children how to investigate their surrounding and universe in a scientific perspective (Kaptan, 1999).

Performance based assessment is a kind of assessment which provides to identify students' gaps and qualifying with students together. This assessment type measures the skills of integrating interdisciplinary knowledge, analyzing, synthesizing, interpreting and inferring. It contains trial reports that covers the students' written presentation, analyzing, explanation, measuring with summaries, experiments which provide to test students' abilities of understanding scientific concepts and following scientific procedures, presentations which deal with content of following fields and provides students to show their masteries about rules, group projects which provides students planning, research, group discussion and working on a complicated problem (Berk, 1994). Complementary measurement and assessment techniques like performance assessment and portfolio assessment can be more objective with rubrics.

Rubrics is a guidebook which defines different student levels with their related criteria. Rubrics identify special criteria, define accurate instructions for adapting assessment type with criteria, give points frequently to each level and provides effective and objective method for non-objective measurement types (Harrington, 1976).

This research study was based on quasi-experimental research and case study designs. This is a longitudinal research and quantitative part was handled out for this study. Research study group consisted of 105 seventh grade students with 3 groups (one experimental and two control groups) from Beytepe Elementary School in Ankara. These students were selected by convenient sampling. "Personal Information Form" which consists of questions to identify groups' demographical properties, study habits, views about science, previous year science grades, etc. was developed and applied to all groups for providing equality. Science process skills test, achievement test and attitude scale was applied to groups also for investigating equality.

Research question of this study is "Is there a significant mean difference between science process skill, attitude and retention scores of the students who were exposed to performance-based assessment and those who were exposed to traditional assessment?" and subquestions were identified in terms of this research question.

The instruments which were used in this study were summarized as follows;

1. Science Process Skills Test (SPST): This test was prepared by investigating TIMSS (Third International Mathematics and Science Study) and Entrance Examination of Secondary Schools (ÖKSYS). 50 questions which are skill-based and related to science concepts were prepared firstly and piloted in fourth through eighth grade elementary students. The number of questions of the test was reduced to 36 after doing item analysis. The reliability coefficient of the test was calculated as 0.89.

2. Attitude Scale towards Science: This scale was firstly constructed by Geban and others (1994) and consists of 15 items. The reliability of the scale was calculated as 0.89 during development and calculated for this study was 0.9.

Process based assessment procedures were piloted with one group in seventh grade students before the implementation process. After than, 14 week process was applied to experimental group during the unit of “A Trip to Inner Structure of Matter”. 36 seventh grade elementary students participated 10 activities during the implementation process and created their own rubricsc in experimental group. Although interactive process during assessment was conducted in experimental group, control group students had traditional assessment methodologies like written examinations or oral presentations without identifying any standards. After 14 weeks, researchers applied the tests as posttests to both experimental and control groups. Science Process Skills Test was applied to both experimental and control groups as retention tests 45 days after posttest. Both descriptive and inferential statistics were used in this study. Mean, frequencies and percentages were used for “Personal Information Form”; repeated measures of ANOVA and paired sample t-test were used for data analysis.

In terms of the first research question, repeated measures of ANOVA results showed that effect size was large (Partial Eta Square  $>0.14$ ). There was a significant difference between pre and post Science Process Skills Test results of experimental group. It could be said that treatment in experimental group was effective according to Wilk’s Lambda results. Repeated measures of ANOVA results of first and second control groups also showed that effect size was small (Partial Eta Square  $<0.14$ ). There was no significant mean difference between pre and post test results of first and second control groups.

Repeated measures of ANOVA results showed that effect size was large (Partial Eta Square  $>0.14$ ) related to second research question. There was a significant mean difference between Attitude towards Science Scale results in experimental group and it could be said that treatment in experimental group was effective for improving students’ attitudes towards science in terms of Wilk’s Lambda results. Repeated measures of ANOVA results also showed that effect size was small (Partial Eta Square  $<0.14$ ). There is no significant mean difference between pre and posttest results of both two control groups.

Paired t-test results showed that there is no significant mean difference between post and retention test results in experimental group and control groups in terms of third research question. It could be said that treatment in experimental group and control group was permanent for science process skills related to this question. Consequently, performance based assessment had a positive impact in experimental group in terms of the statistical analysis results.

Science process skills and attitudes are important variables for effective science learning. This research study is important for suggesting new ways for developing science process skills and attitudes toward science. Considering the load of science curricula, using performance-based activities and rubrics during measurement and evaluation process will provide permanent learnings in science. These kinds of activities and assessment help students to develop skills and attitudes. Asking students about their expectancies and constructing interactive environments about assessment processes will be useful for providing to develop self assessment capability of students in science. Science teachers could become together for preparing different criteria for different grade levels, interest areas...etc. In elementary schools. Preservice and inservice science teachers could be informed about performance based assessment by different training programs.

