

Öğrenilen Bilgilerin Hatırlanmasında Sınavların Etkisi

Elif ATABEK-YİĞİT¹, Fatime BALKAN-KIYICI²

Geliş Tarihi: 02.12.2015

Kabul Ediliş Tarihi: 21.03.2016

ÖZ

Bu çalışmada sınavların öğrenilen bilgilerin hatırlanması üzerindeki etkisi incelenmiştir. Çalışma grubunu bir devlet üniversitesinin eğitim fakültesinde öğrenim gören 72 fen bilgisi öğretmen adayını oluşturmaktadır. Öğretmen adaylarına “Enerji kaynakları ve çevre üzerindeki etkileri” ile ilgili bilgilerinin sorgulandığı bir soru formu ön test şeklinde uygulanmıştır. Forma verilen cevapların puanlanması sonucunda; üç deney, bir kontrol olmak üzere dört grup oluşturulmuştur. Ön test uygulamasından bir hafta sonra, deney grubunu oluşturan gruplardan birinci gruba çoktan seçmeli test, ikinci gruba doğru-yanlış testi, üçüncü gruba ise boşluk doldurma testi deneysel işlemler olarak uygulanmış kontrol grubuna ise herhangi bir işlem uygulanmamıştır. Deneysel sürecin bir hafta sonrasında ise aynı soru formu son test olarak tekrar uygulanmıştır. Uygulanan ön test ve son test sonucunda elde edilen puanlar istatistiksel olarak analiz edilmiştir. Bu çalışmanın sonucunda öğrenilen bilgileri hatırlanmasında sınavların etkili oldukları bulunmuştur. Ayrıca doğru-yanlış testlerinin çoktan seçmeli testlere göre daha fazla test etkisi oluşturduğu tespit edilmiştir.

Anahtar kelimeler: test etkisi, hatırlama, ölçme ve değerlendirme

Effects of Tests to Recall the Learned Information

ABSTRACT

In this study, effect of tests to recall the learned information was examined. Seventy-two pre-service science teachers studying at an education faculty of a university were the participants of this study. Three experimental and 1 control group were formed according to scores from a question form which was used as pre test. Experimental groups has received different types of tests e.g., multiple choice test (MCT), true-false test (TFT) and filling the blanks test (FBT) while control group has received no test. After completing the experimental period, same form was applied as post-test. Findings of this study reveals that there was a significant difference between post test scores of experimental groups and control group. Therefore tests have an effect to help students to recall the learned information. Also, there was a significant difference between post test scores of FBT group and MCT group. Therefore different types of tests can cause different testing effects, FBT form more testing effect than MCT.

Keywords: testing effect, recalling, measuring and evaluation

¹ Yrd.Doç.Dr., Sakarya Üniversitesi, eatabek@sakarya.edu.tr

² Doç.Dr., Sakarya Üniversitesi, fbalkan@sakarya.edu.tr

GİRİŞ

Eğitim-öğretim içerisinde birçok aşamayı barındıran bir süreçtir. Bu süreçteki her aşama tartışılmaz bir öneme sahip olup, bu aşamalardan birisi de ölçme ve değerlendirmedir. Eğitim reform hareketleri incelendiğinde; birçok gelişmiş ülkenin çağın gerekliliği olarak daha nitelikli bireyler yetiştirebilme sürecinde bu çabaların bir parçası olarak ölçme ve değerlendirmeye de büyük önem verdikleri görülmektedir (Creswell 2005). Ölçme ve değerlendirme çoğu birey tarafından akla ilk gelen şekliyle ne biliniyor bilinmediğinin ortaya çıkarılması olarak ele alınmaktadır (Naylor ve Keogh 1999). Ancak bu algılayış ölçme ve değerlendirmenin işlevi için yetersiz kalmaktadır. Eğitim ve öğretim sürecinin ayrılmaz bir parçası olan ölçme ve değerlendirme, belirlenen eğitim hedeflerine ulaşıp ulaşılmadığına karar vermede kullanılmaktadır. Ancak bunun yanında değerlendirmenin bir amacı da, öğrencinin öğrenmesini geliştirmek ve desteklemektir. (Turgut ve Baykul 2015).

Ölçme ve değerlendirmede yaygın olarak kullanılan birçok teknik ve araçlar yer almakta ve öğrenciler sıkça yapılaş şekillerine ve özelliklerine göre farklı şekillerde gerçekleştirilen sınavlarla karşı karşıya kalmaktadır. Bu sınavların yapılması sürecinde çoktan seçmeli, doğru-yanlış, boşluk doldurma testleri gibi testler çok önemli bir yer tutmaktadır.

Öğretim çalışmalarının sonuçlarını değerlendirmek maksadıyla en fazla tercih edilen sınav şekillerinden birisi çoktan seçmeli testlerdir. Bunun sebebi, çoktan seçmeli testlerin daha kısa sürede değerlendirilebilmesi ve büyük sınıflara kolaylıkla uygulanabilmesidir. Ayrıca, çoktan seçmeli testler objektif ve geçerli bir değerlendirme için notların elde edilebilmesine olanak sağlarlar. Bununla birlikte, öğrencilerin cevapları seçenekler arasından seçiyor olması derinlemesine bir öğrenmenin olup olmadığını değerlendirmede eksik kalır. Öğrencinin cevabı bilmiyor olmasına rağmen seçeneklerden birini işaretleyerek doğruyu bulma şansı bu değerlendirme aracının güvenilirliğini azaltır ve öğrencinin cevabı seçenekler arasından seçiyor olması kendi cevabını ifade edebilme ve düşüncelerini organize edebilme yeteneğinin gelişmesine mani olur ya da bunun ölçülmesine fırsat vermez (Chang, Yeh ve Barufaldi 2010).

Eğitim sürecindeki temel amaç öğrenilen bilgilerin ezberden uzak, kalıcı olmasını sağlamak olmalıdır. Bu anlamda öğrenilen bilgilerin aradan uzun zaman geçtikten sonra da hatırlanabiliyor olması önemlidir. Test ve sınavlar öğrencilerin ne bildiğini ölçmede kullanılacak araçlar olmakla beraber, sonuçları çok iyi olmasa hatta geri dönüt verilmesede dahi konunun daha uzun süre sonunda hatırlanması doğrultusunda büyük bir olumlu etkiye sahiptir. Bu şekilde *test etkisi* denilen test ile performansın artırılması deneysel psikoloji alanında pek çok araştırmaya konu olmuştur ve sınavların hatırlama üzerinde etkileri olduğu uluslararası literatürde yer alan pek çok çalışma ile ortaya konulmuştur (Agarwal ve ark. 2008; Chang, Yeh ve Barufaldi 2010; McDaniel ve ark. 2007a; McDaniel ve ark. 2007b; Roediger ve Karpicke 2006a; Roediger ve Karpicke

2006b; Wheeler, Ewers ve Buonanno 2003). Ancak eğitim alanında yeterli çalışmalar mevcut değildir (Agarwal ve ark. 2008; Roediger ve Karpicke 2006b). Deneysel psikoloji alanındaki çalışmalarda çalışma materyali olarak çoğunlukla kelime listeleri kullanılmıştır (Karpicke ve Zaromb 2010). Bununla birlikte eğitim alanındaki çalışmalarda ise daha çok düz metinler çalışma materyali olarak tercih edilmiştir.

Roediger ve Karpicke (2006a) bilimsel konular ile ilgili düz metinleri çalışma materyali olarak kullandıkları çalışmalarında testlerin ve tekrar çalışmanın hatırlamaya olan etkisini incelemişlerdir. Çalışmaya katılan öğrenciler üç gruba ayrılmış ve bir grup yalnızca tekrar çalışma gerçekleştirirken, diğer iki grup tekrar çalışmaya ilaveten hatırlama testleri (bir grup bir hatırlama testi ve diğer grup ise üç hatırlama testi) yapmışlardır. 5 dakika 2 gün ve 1 hafta sonunda yapılan son testlerden elde edilen sonuçlara göre tekrar çalışma yapan grup 5 dakika sonra gerçekleştirilen son teste en fazla hatırlayabilirken 1 hafta sonra yapılan son teste en az hatırlayabilmiştir. Üç test gerçekleştiren grup ise tek test gerçekleştiren gruba göre tüm sürelerin sonunda daha fazla hatırlayabilmiştir. Benzer şekilde Nungester ve Duchastel (1982 akt: Roediger ve Karpicke 2006b) çoktan seçmeli testlerin ve kısa cevaplı testlerin etkisini araştırdıkları çalışmalarında öğrencileri üç gruba ayırmışlar, birinci gruptan metni yalnızca bir kez çalışmalarını, ikinci gruptan metni iki kez çalışmalarını ve üçüncü gruptan ise metni bir kez çalıştıktan sonra çoktan seçmeli ve kısa cevaplı sorulardan oluşan testi cevaplamalarını istemişlerdir. İki hafta sonra yapılan hatırlama testinde (hatırlama testi de çoktan seçmeli ve kısa cevaplı sorulardan oluşmaktadır) test yapan grubun en fazla hatırladığı, bunu iki kez çalışan grubun izlediği ve bir kez çalışan grubun ise en az hatırlayabildiği sonucu ortaya çıkmıştır. Araştırmacılar daha uzun sürede etkinin ne şekilde olacağını incelemek amacıyla beş ay sonra yeniden hatırlama testi gerçekleştirmişler ve yine aynı sonuç ile karşılaşmışlardır. Buradan hareketle testlerin hatırlama üzerinde etkili oldukları, tekrar çalışmanın kısa süre sonunda daha fazla hatırlamaya yardımcı olduğu ancak uzun zaman sonra hatırlama koşulunda en az etkili olduğu ve testlerin uzun süre sonunda hatırlamada daha etkili olduğu söylenebilir. Kısa süre sonunda etkisi daha fazla olan tekrar çalışma pek çok öğrenci tarafından tek çalışma yöntemi olarak uygulanmakta ve uzun süre sonundaki etkisizliğinin farkına varılmamaktadır.

Hatırlama testleri genel olarak serbest geri çağırma (free recall), ipucu ile geri çağırma (cued recall) ve tanıma (recognition) şeklinde uygulanmaktadır. Bunlar arasında geri çağırma testlerinin tanıma testlerinden daha etkili olduğu yapılan pek çok çalışma ile gösterilmiştir (Carpenter ve DeLosh 2006). Bu tekniklerin kullanıldığı deneysel psikoloji alanındaki çalışmalara ek olarak eğitim alanında değerlendirme aracı olarak sıklıkla kullanılan kısa cevaplı testler ve çoktan seçmeli testlerin de test etkisini ortaya koymaya çıkarmaya çalışan araştırmalarında (Butler ve Roediger 2007; McDaniel ve ark. 2007a) kullanıldığı olmuştur. Kullanılan testlerin çeşidine bağlı olarak test etkisinin değişip değişmediği de araştırmacılar tarafından üzerinde çalışılan bir konu olmuştur.

Duchastel 1981’de yaptığı çalışmada (akt: Roediger ve Karpicke 2006b) kısa cevaplı ve çoktan seçmeli testler kullanarak hangi testin daha fazla hatırlamaya yardımcı olduğunu incelemek amacıyla öğrencileri üç gruba ayırmıştır. Bir grubun metni çalışırken, diğer grup metni çalıştıktan sonra kısa cevaplı sorulardan oluşan testi cevaplamış bir diğer grup ise metni çalıştıktan sonra çoktan seçmeli sorulardan oluşan testi cevaplamıştır. İki hafta sonra kısa cevaplı sorulardan oluşan son testin sonuçlarına göre test yapan gruplar yalnızca metni çalışan gruplara göre daha fazla hatırlayabilmişler ve kısa cevaplı test gerçekleştiren grup çoktan seçmeli test gerçekleştiren gruptan daha fazla hatırlayabilmiştir. Bir başka çalışmada McDaniel ve ark. (2007b) kısa cevaplı sınavların çoktan seçmeli sınavlara göre daha fazla hatırlamaya yardımcı olduğunu ortaya koymuşlardır.

Literatürdeki çalışmalara bakılarak öğrencinin kendi cevaplarını bizzat kendisinin ürettiği türdeki sınavların (kısa cevaplı test, kompozisyon yazma) cevabın seçenekler arasından seçildiği sınavlara (çoktan seçmeli test) göre daha fazla hatırlamaya sebep olduğu söylenebilir.

Test etkisini araştırmak için yapılan çalışmaların çoğu deneysel psikoloji alanında ve laboratuvar ortamında gerçekleştirilmektedir. Bununla birlikte sınıf ortamında da test etkisinin incelendiği çalışmalar mevcuttur (Campbell ve Mayer 2009; Leeming 2002; McDaniel ve ark. 2007a; McDaniel ve ark. 2007b). Laboratuvar ortamı ve sınıf ortamı çok çeşitli şekillerde birbirinden farklıdır. Öğrencinin öğrenmekle sorumlu olduğu bilgiler sınıf ortamında laboratuvar ortamına göre daha fazladır. Öğrenilecek materyaller sınıf ortamında çok farklı şekillerde (kitaplarda, dersin anlatımında, sınıf içi tartışmalarda vb.) sunulur. Ayrıca öğrenciler sınav öncesi çalışma miktarları, çalışmaya başlama süreleri, ders materyaline olan ilgileri ve öğrenme motivasyonları açısından da farklılıklar gösterirler. Tüm bunlardan dolayı test etkisi sınıf ortamında daha farklı bir şekilde işler (Roediger ve Karpicke 2006b).

Test etkisi, Heisenberg belirsizlik prensibinin psikolojideki küçük bir versiyonu gibi düşünülebilir. Bu prensibe göre nasıl ki bir elektronun yalnızca yerini ölçmek yerinin değişimine sebep olursa, benzer şekilde hafızadan bir bilgiyi geri çağırarak da hatırlamayı güçlendirici bir etki oluşturur ve test edilen materyalin uzun süreli kalıcılığına sebep olur (Roediger ve Karpicke 2006b).

Test etkisinin hatırlamaya nasıl etkisi olduğu ile ilgili çeşitli açıklamalar mevcuttur. Bu konuda yapılmış literatürü inceleyen ilk çalışmalardan birisi olan Dempster’e (akt: Roediger ve Karpicke 2006b) göre test etkisi iki farklı şekilde izah edilmektedir. İlki, test etkisinin öğrencinin sınav ile birlikte öğrenilen konuya yeniden maruz kalması dolayısıyla oluştuğu yönündedir. Diğer açıklamaya göre ise sınavlar zihindeki bilginin pek çok zihinsel süreç ile geri çağırılması sonucu etkili olurlar ve test etkisi bu şekilde ortaya çıkar. Ancak, ilk açıklama sınavların uzun süre sonunda hatırlamaya olan pozitif etkisini açıklamada yetersiz kalır.

Anlamalı öğrenmenin temelini üç bilişsel süreç oluşturur. Bunlardan birincisi öğrenenlerin bilgi için dikkatlerini yönlendirmesi süreci yani “seçim”dir. İkincisi, öğrenenlerin bu bilgiyi birleştirmesi, diğer bir deyişle “organize etme” sürecidir. Sonuncusu ise öğrenilenlerin var olan bilgiye eklenmesi süreci yani “entegrasyon”dur (Mayer 2008). Sınavlar bu süreçlerin her birine yardımcı olabilirler. Örneğin, sınav uygulanacağını bilerek öğrenciler dikkatlerini bilgiye yönlendirebilirler. Böylece seçim sürecine katkı sağlanmış olur. Benzer şekilde, pek çok konudan sorulara cevap verirken organize etme ve entegrasyon süreçleri işler. Son olarak geri dönüt alındığında ise öğrenciler kendi cevaplarını yeniden düşünür, organize eder ve eski bilgilerle entegrasyonu gerçekleştirirler (Campbell ve Mayer 2009). Böylece sınavların anlamlı öğrenme süreçlerinde etkili olduklarını ifade edebilmek mümkündür.

Daha ziyade öğrencilerin ne bildiğini veya ne kadar bildiğini ölçmeye yönelik kullanılan sınavların öğrenilen bilgilerin hatırlanmasında da etkisi olduğu uluslararası literatürde pek çok çalışma ile tespit edilmiş durumdadır (Butler ve Roediger 2007; Chang, Yeh ve Barufaldi 2010; Leeming 2002; McDaniel ve ark. 2007a). Ancak ülkemizde bu anlamda yapılan yeterli çalışma bulunmamaktadır. Yukarıdaki açıklamalar doğrultusunda bu çalışmada öğrenilen bilgilerin hatırlanmasında sınavların etkisi incelenmiştir. Sınavların hatırlama üzerindeki etkisi ile ilgili çalışmalar henüz okunan veya öğrenilen kelime listeleri veya düz metinler kullanılarak yapılmasına rağmen bu çalışmada öğretiminin üzerinden bir yıldan fazla zaman geçen bir konunun tekrar hatırlanmasında sınavların etkisi ele alınmıştır. Bu yönüyle ele alındığında da bu çalışma test etkisinin incelendiği diğer çalışmalardan farklıdır. Bu çalışmada aşağıdaki sorulara cevap aranmıştır.

1. Sınavlar, daha önceden öğrenilen bilgilerin hatırlanması üzerinde etkili midir?
2. Değişik türdeki sınavların oluşturacakları (eğer varsa) test etkisi farklı mıdır?

YÖNTEM

Bu çalışmada nicel araştırma yöntemlerinden öntest-sontest kontrol gruplu yarı deneysel desen kullanılmıştır. Yarı deneysel yöntemde katılımcıların deney ve kontrol gruplarına atanmasında rastgele bir yaklaşım kullanılmaz (Creswell 2005). Bu desende hazır gruplardan ikisi belli değişkenler üzerinden eşleştirilmeye çalışılır (Büyüköztürk, Çakmak, Akgün ve Karadeniz 2008). Bu çalışmada da deney ve kontrol gruplarını oluşturan öğretmen adayları gruplara puan ortalamaları eşit olacak şekilde atanarak gruplar oluşturulmuştur.

Çalışma Grubu

Bu çalışmada, çalışma grubunu Sakarya Üniversitesi Eğitim Fakültesi İlköğretim Bolumu Fen Bilgisi Öğretmenliği 4.sınıflarında öğrenim gören 72 öğretmen adayı oluşturmaktadır. Bu öğretmen adayları öğrenim hayatlarının üçüncü senesinde “Çevre Eğitimi” isimli bir ders almışlar ve bu derste çevre sorunları ve

enerji kaynakları konuları da ele alınmıştır. Katılımcılar araştırmaya gönüllü olarak katılmışlardır.

Çalışma Konusu

Özellikle son zamanlarda etkileri daha fazla hissedilen küresel sorunlar sebebiyle çevre kirlenmesi ve kirliliğin küresel etkileri konusunun önemi daha fazla anlaşılmakta ve pek çok alandan araştırmacının çalışmalarına konu olmaktadır. Geleceğimizi şekillendirecek çocuklarımızı yetiştirecek olan öğretmen adaylarının da bu konu hakkında bilgi sahibi olması gerekliliği açıktır. Bu bağlamda öğretmen adaylarına lisans eğitimleri süresince verilmekte olan çevre derslerinin de ne kadar önem taşıdığı bir kez daha ifade edilebilir. Araştırmada uygulama yapmak üzere “Enerji kaynakları ve çevre üzerindeki etkileri” konusu bu sebeple seçilmiştir.

Veri Toplama Süreci

Veri Toplama Aracı

Bu çalışmanın verileri öğretmen adaylarına çalışmanın başında ve sonunda “Enerji kaynakları ve çevre üzerindeki etkileri” ile ilgili bilgilerinin sorgulandığı sorulardan oluşan bir soru formuna verdikleri cevapların puanlanmasından elde edilmiştir.

Öğretmen adayları, ön testten altıkları puanlar göz önünde bulundurularak ve her bir grubun ortalaması yaklaşık olarak eşit olacak şekilde gruplara seçkisiz olarak atanmıştır. Üç deney ve bir kontrol olmak üzere toplam dört grup oluşturulmuş ve kontrol grubuna arada herhangi bir uygulama yapılmazken, deney grubunu oluşturan 3 ayrı grubun her birine ise şu deneysel işlemler uygulanmıştır:

1. Deney grubuna: Çoktan seçmeli test (ÇST)
2. Deney grubuna: Doğru-yanlış testi (DYT)
3. Deney grubuna: Boşluk doldurma testi (BDT)

Uygulama Süreci

Çalışmada, izlenen yol şu şekilde kısaca özetlenebilir:

- Öncelikle, öğretmen adaylarının “Enerji kaynakları ve çevre üzerindeki etkileri” ile ilgili ön bilgilerinin yoklanması ve çalışma sonundaki değişimin görülebilmesi amacıyla bu konuyla ilgili olarak bir soru formu hazırlanmıştır. Bu formda yer alan sorular, yenilenebilir ve yenilenemeyen enerji kaynaklarının neler olduğunu ve bunların çevre sorunları üzerinde nasıl bir etkiye sahip olduğunu sorgular niteliktedir. Bu forma araştırmacılar dışında 2 ayrı alan uzmanının görüşleri doğrultusunda son hali verilmiştir. Bu form ön test ve son test olarak kullanılmıştır.
- “Enerji kaynakları ve çevre üzerindeki etkileri” konusu ile ilgili olarak deneysel işlem sürecinde kullanılmak üzere çeşitli testler oluşturulmuştur. Testlerin her biri aynı bilgiyi ölçmeyi amaçlayan sorular içermektedir. Bu testler;
 - **Çoktan Seçmeli Test (ÇST):** “Enerji kaynakları ve çevre üzerindeki etkileri” konusunda iki ayrı alan uzmanı (bir

biyoloji alanında doktor unvanına sahip kişi ve bir fen eğitimi alanında doktor unvanına sahip kişi) tarafından 38 adet soru hazırlanmış, testin güvenilirlik çalışmaları sonucunda ve uzman görüşleri neticesinde 13 soru elenerek 25 sorudan oluşan testin son haline ulaşılmıştır. Testte yer alan her soru 4 seçenektir ve her bir soru için doğru cevap 4 puan değerindedir. Bu testte yer alan sorulara örnek verecek olursak,

▪ *Petrol arayacak olsanız, hangi seçenekteki yeri arardınız?*

- *Eskiden çöl olan bir yer*
- *Eskiden deniz olan bir yer*
- *Bir ova*
- *Kayalık bir bölge*

- ***Doğru-Yanlış Testi (DYT):*** Çoktan seçmeli testte yer alan sorulara paralel olarak doğru ve yanlış pek çok ifade yazılmış ve uzman görüşleri (bir biyoloji alanında doktor unvanına sahip kişi ve bir fen eğitimi alanında doktor unvanına sahip kişi) doğrultusunda bu testin 25 maddeden oluşan son hali oluşturulmuştur. Testte 13 doğru, 12 yanlış ifade yer almaktadır. Bu testte yer alan sorulara örnek verecek olursak,

() *Petrol aranacak bir yerin eski jeolojik çağlarda deniz/okyanus olması gerekir.*

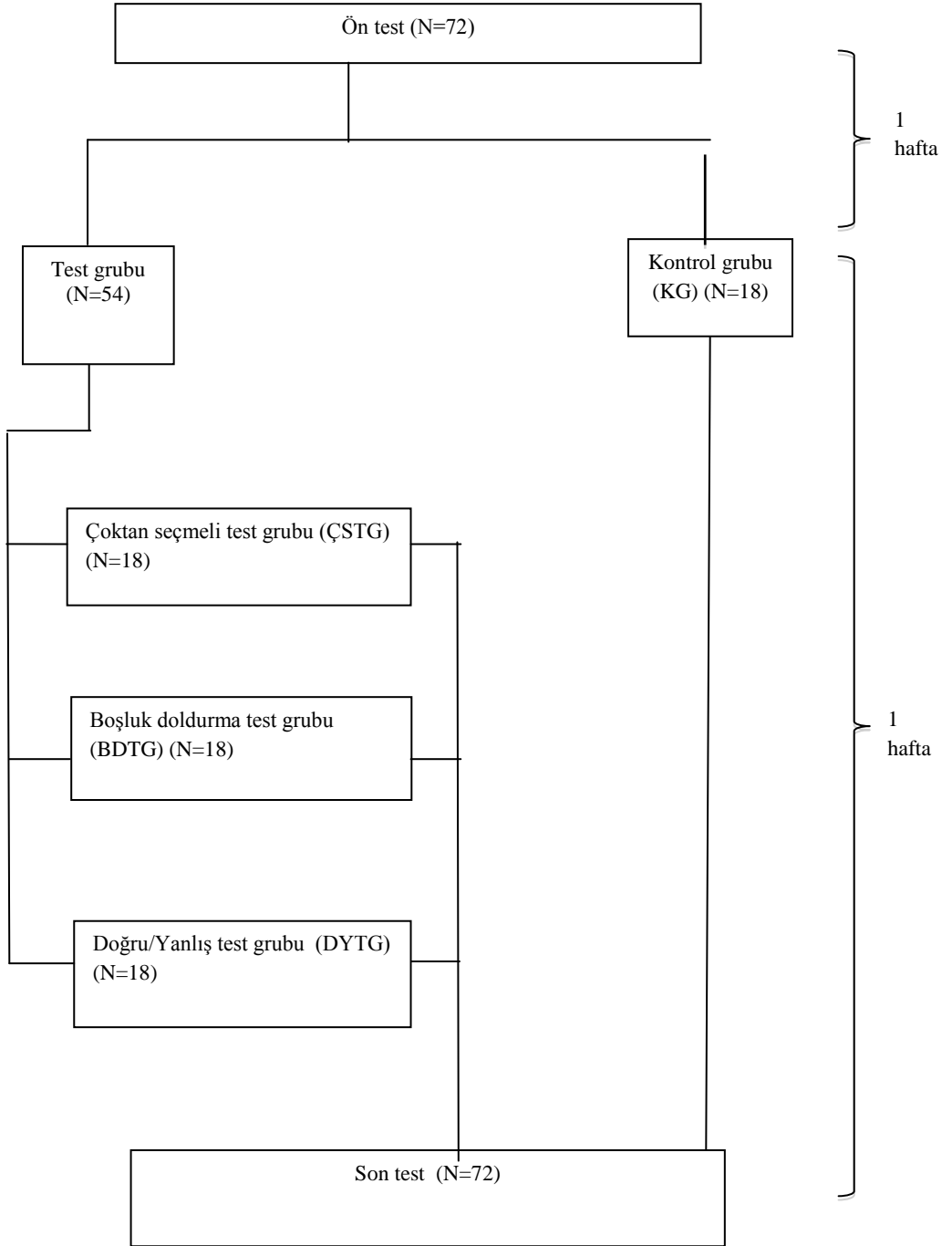
- ***Boşluk Doldurma Testi (BDT):*** Çoktan seçmeli testte yer alan sorulara paralel olarak ifadeler yazılmış ve bu ifadelerin bir kısmı boşluk şeklinde bırakılmıştır. Boşluklara gelebilecek ifadeler ise testin sonunda bir bölüm halinde verilmiştir. Test için uzman görüşleri (bir biyoloji alanında doktor unvanına sahip kişi ve bir fen eğitimi alanında doktor unvanına sahip kişi) alınarak son hali oluşturulmuştur. 25 maddeden oluşan bu testte yer alan sorulara örnek verecek olursak,

▪ *Petrol aranacak olan bir yerin eski jeolojik çağlarda olması gerekir.*

- Öğretmen adaylarına ön test olarak uygulanan formlar değerlendirilmiş ve aldıkları puanlara göre ortalamaları yaklaşık olarak aynı olacak şekilde gruplar eşleştirilerek dört grup oluşturulmuştur.
- Oluşturulan gruplar; çoktan seçmeli test grubu (ÇSTG), doğru-yanlış test grubu (DYTG), boşluk doldurma test grubu (BDTG) ve kontrol grubu (KG) olarak adlandırılmıştır.
- Ön testin uygulanmasından 1 hafta sonra deneysel işlem aşamasına geçilmiştir. Üç farklı deney grubunda bulunan öğretmen adaylarından birinci deney grubu olan ÇSTG'nu oluşturan öğretmen adaylarının çoktan seçmeli testi, ikinci deney grubu olan DYTG'ndeki öğretmen adaylarının doğru-yanlış testini ve üçüncü deney grubu olan BDTG'ndeki öğretmen

adaylarının ise boşluk doldurma testini çözmeleri istenmiştir. Kontrol grubunda bulunan öğretmen adaylarına ise herhangi bir uygulama yapılmamıştır.

- Deneysel işlemin 1 hafta sonrasında ön test olarak uygulanan soru formu bu kez son test olarak tekrar uygulanmış ve öğretmen adaylarının cevapları değerlendirilerek çalışmanın verileri elde edilmiştir.
- Çalışma planı, şekil 1’de verildiği gibidir.



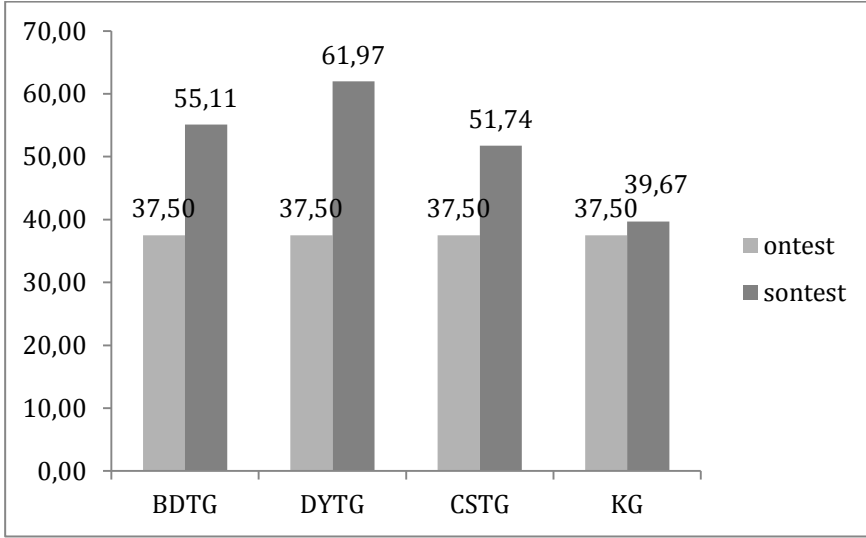
Şekil 1. Araştırma Çalışma Planı

Verilerin Analizi

Çalışma sonucunda elde edilen verilerin analizinde MS Excel ve SPSS programları kullanılmıştır. Öğretmen adaylarının ön testten aldıkları puanlara göre ortalamaları aynı olacak şekilde grupların oluşturulması sürecinde MS Excel programından faydalanılmıştır. Bu şekilde her bir grubun ortalaması yaklaşık olarak 37,5 olacak şekilde dört grup oluşturulmuştur. Sınavların hatırlama üzerinde etkisi olup olmadığını anlayabilmek için ön test ve son test sonuçlarının karşılaştırılması ve hangi tür sınavın daha fazla etkiye sahip olduğunun belirlenmesi için son test puanlarının kıyaslanması için SPSS programından faydalanılmıştır. Bu amaçla tek yönlü ANOVA testi ile analiz gerçekleştirilmiştir.

BULGULAR

Verilerin analizi sonucunda aşağıdaki bulgulara ulaşılmıştır. Her bir grubun ön test ve son test puan ortalamaları Şekil 2’de gösterilmiştir.



Şekil 2. Grupların Ön-test ve Son-test Puan Ortalamaları

Şekil 2 incelendiğinde ortalamaları eşit olacak şekilde oluşturulan dört gruptan DYTG’nun en yüksek son test puan ortalamasına sahip olduğu (61.97), bunu sırasıyla BDTG (55.11) ve ÇSTG (51.74)’nin izlediği görülmektedir. Deneysel işlemler süreci boyunca diğer gruptaki öğretmen adaylarından farklı olarak herhangi bir uygulama gerçekleştirilmeyen KG’nun son test ortalaması ise 39.67 ile en düşük ortalamadır.

Sınavların öğretmen adaylarının daha önceden öğrenmiş oldukları bilgileri hatırlamasında nasıl bir etkiye sahip olduğunu incelemek amacıyla, grupların son test puanları bağımlı değişken olarak kullanılarak tek yönlü ANOVA testi gerçekleştirilmiş ve sonuçları Tablo 1’de verilmiştir.

Tablo 1. Grupların Son Test Puanlarının Karşılaştırılmasına İlişkin ANOVA Testi Sonuçları

	KT	SD	KO	F	p
Gruplar arası	4701.96	3	1567.065	12.029	.000
Gruplar içi	8858.363	68	130.270		
Toplam	13559.559	71			

Tablo 1'e göre grupların son test puan ortalamaları arasında anlamlı bir farklılık ($F(3,68)=12.029, p<.05$) vardır. Dolayısıyla, bu sonuç sınavların daha önceden öğrenilen bilgileri hatırlama üzerinde etkisi olduğu şeklinde yorumlanabilir. Sınav türleri arasında hatırlama açısından bir farklılık olup olmadığının görülmesi için ise çoklu karşılaştırma analizi (Tukey testi) gerçekleştirilmiş ve sonuçları Tablo 2'de verilmiştir.

Tablo 2. Grupların Son Test Puanlarına Göre Çoklu Karşılaştırmasına İlişkin Analiz Sonuçları

	Test Grubu	Test Grubu	Ort.fark	p
Tukey HSD	BDTG	DYTG	-6.86	.281
		CSTG	-3.37	.812
		KG	15.45*	.001*
	DYTG	BDTG	6.86	.281
		CSTG	10.23*	.044*
		KG	22.30*	.000*
	CSTG	BDTG	3.37	.812
		DYTG	-10.23*	.044*
		KG	12.07*	.012*
	KG	BDTG	-15.45*	.001*
		DYTG	-22.30*	.000*
		CSTG	-12.07*	.012*

Çoklu karşılaştırma analizinin sonuçlarından tüm test gruplarının (DYTG, CSTG, BDTG) son test puanları ile KG'nun son test puanları arasında anlamlı düzeyde bir farklılık ($p<.05$) olduğu anlaşılmaktadır. Ayrıca DYTG ile CSTG'nun son test puan ortalamaları arasında DYTG'nun lehine anlamlı düzeyde bir farklılık ($p<.05$) vardır.

TARTIŞMA ve SONUÇ

Sınavların önceden öğrenilen bilgilerin hatırlanmasında etkili olup olmadıklarının incelendiği bu araştırmanın sonuçları incelendiğinde; tüm grupların son test puanlarında ön test puanlarına göre bir artış olduğu tespit edilmiştir. Ön test puan ortalamaları eşitlenerek oluşturulan gruplardan DYTG

en yüksek son-test puan ortalamasına (61.97) sahiptir (en fazla hatırlama gerçekleşmiştir). Bu test grubunda yer alan öğretmen adaylarının son test puan ortalamaları ön test puan ortalamalarına göre %65.2'lik bir artış göstermiştir. Bunu sırasıyla BDTG (55.11) ve ÇSTG (51.74) izlemektedir. Bu test gruplarındaki öğretmen adaylarının puanları sırasıyla %46.9 ve %37.9'luk artış göstermiştir. Bu sonuçlardan literatürdeki pek çok çalışmada (Agarwal ve ark. 2008; Chang, Yeh ve Barufaldi 2010; McDaniel ve ark. 2007a; McDaniel ve ark. 2007b; Roediger ve Karpicke 2006a; Roediger ve Karpicke 2006b; Wheeler, Ewers ve Buonanno 2003) olduğu gibi sınavların öğrenilen bilgileri hatırlamaya yardımcı olduğu yorumu yapılabilir. Çalışmada kontrol grubu, deneysel işleme maruz kalmayan öğretmen adaylarının puanlarında bir değişim olup olmayacağını incelemek ve deney gruplarının sonuçlarını anlamlandırmak için oluşturulmuştur. Kontrol grubunun son test puanındaki %5.7'lik artış öğretmen adaylarının aynı sınava tekrar maruz kalmalarının bir sonucu olarak yorumlanabilir. Tüm deney gruplarının son test puanlarındaki artışın bu değerden fazla olması ise, test etkisinin aynı materyale maruz kalma sonucunda olduğu fikrinin yanlış olduğunu destekler şekildedir (Carpenter ve De Losh 2006). Öğrenilen bilgilerin hatırlanmasında sınavların nasıl bir etkiye sahip olduğu ile ilgili çeşitli açıklamalar söz konusu iken, sınav uygulanması ile zihinsel pek çok sürecin (seçme, ilişki kurma, organize etme, entegrasyon ve geri çağırma) aynı anda gerçekleştiği ve hatırlamanın bunun neticesinde ortaya çıktığı görüşü kabul edilen bir açıklamadır (Roediger ve Karpicke 2006b). Böylece, sınavı cevaplandıran öğretmen adaylarının hatırlamalarında anlamlı düzeyde oluşan artışın sebebinin testte yer alan sorulardaki ipuçları ile öğrendikleri bilgiler arasında ilişki kurmaları ve bu bilgileri geri çağırmaları sonucu olduğu ifade edilebilir.

Sınav türünün test etkisini oluşturmada etkili olup olmadığını araştırmak amacıyla yapılan son test puanlarının çoklu karşılaştırması sonucunda DYTG'nun son test puanları ile CSTG'nun son test puanları arasında anlamlı bir farklılık ($p < .05$) tespit edilmiştir. Bu durumda test etkisinin yalnızca konuya yeniden maruz kalma sonucunda oluştuğuna yönelik olan açıklamanın yetersiz oluşu bir kez daha ortaya çıkmış olmaktadır. Tüm gruplardaki öğretmen adayları testler için eşit zaman harcamış olmalarına karşın DYTG'ndeki öğretmen adaylarının diğer gruplardakilere göre neden daha fazla hatırlayabildiği (test etkisinin bu durumda neden daha farklı olduğu), test etkisinin yalnızca konuya yeniden maruz kalınması sonucunda oluştuğu açıklaması ile izah edilemez. Test içerisindeki ipuçlarının (ya da öğrenilen bilgiler ile kurulabilecek ilişkilerin) sayısı arttıkça test etkisi fazlaşmakta ve daha fazla hatırlama sağlanabilmektedir. Çoktan seçmeli testte ve boşluk doldurma testinde doğru cevap seçenekler arasında ya da testin içerisinde bulunmaktadır. Bununla birlikte DYT'inde öğretmen adaylarının konu ile ilgili ifadelerin doğru ya da yanlış olduğuna karar vermeleri beklenmektedir. Böylelikle DYT'inde öğretmen adaylarının soruları doğru cevaplayabilmek için CST veya BDT'ine göre daha fazla zihinsel süreç becerilerini kullanmaları gerekmektedir. Bunun sonucunda da DYT'inin CST'e göre daha fazla test etkisi oluşturduğu ifade edilebilir.

Kıscacası daha fazla zihinsel süreç becerilerinin kullanımı gerektiren DYT gibi testler daha fazla test etkisine neden olmakta ve daha fazla hatırlama gerçekleştirebilmektedir. Çoktan seçmeli test türündeki sınavlarda öğrencinin soruya doğru cevap verebilmesi için, doğru cevabı seçenekler arasından bulması beklenir. Bu durum deneysel psikoloji alanında gerçekleştirilen tanıma testlerine benzemektedir. Kısa cevaplı test (örneğin bu çalışmada uygulanan doğru-yanlış testi veya boşluk doldurma testi gibi) türündeki sınavlarda ise öğrencinin soruya doğru cevap verebilmesi için, konuyla ilgili olarak daha önceki bilgilerini analiz etmesi ve cevabını üretmesi beklenir. Bu durum da deneysel psikoloji alanındaki geri çağırma testlerine benzer. Geri çağırma türündeki sınavlar, tanıma türündeki sınavlara kıyasla daha fazla zihinsel süreç kullanımı gerektirir ve daha fazla hatırlamaya yardımcı olabilir (Carpenter ve DeLosh 2006; Kang ve ark. 2007; Karpicke ve Zaromb 2010; McDaniel ve ark. 2007b). Benzer şekilde McDaniel ve arkadaşları (2007a), Butler ve Roediger (2007) ve Kang ve arkadaşları (2007) de yaptıkları çalışmalar sonucunda kısa cevaplı testlerin (doğru yanlış testleri gibi) çoktan seçmeli testlere göre daha fazla test etkisi oluşturduklarını ifade etmişlerdir.

Test etkisini araştıran çalışmaların çoğu deneysel psikoloji alanında ve laboratuvar ortamında gerçekleştirilmektedir. Eğitim alanında ve sınıf ortamında gerçekleştirilen çalışmalar daha sınırlı kalmaktadır. Her iki durumda da genel olarak kelime listeleri veya düz metinler çalışma materyali olarak seçilmekte ve uygulama hemen gerçekleştirilmektedir. Bu çalışmanın diğer çalışmalardan önemli bir farkı çalışma materyali olarak öğretiminin üzerinden bir yıldan fazla zaman geçen bir konunun seçilmiş olmasıdır.

Çalışmanın sonuçlarına göre test etkisi bilgilerin öğrenilmesinin üzerinden uzun zaman geçtiğinde de oluşmakta ve hatırlama gerçekleşmektedir. Uygulamanın hemen gerçekleştirildiği ve test etkisinin belirli periyotlarla araştırıldığı çalışmalarda da testlerin uzun süreli kalıcılığa sebep olduğu ve tekrar çalışmaya göre daha etkili oldukları sonucu ortaya çıkmıştır (Agarwall ve ark. 2008; Kang ve ark. 2007).

Bu çalışmanın sonuçlarından hareketle, eğitim-öğretim faaliyetlerinde testlerin yalnızca ölçme ve değerlendirme araçları olarak değil aynı zamanda öğretimi destekleyici olarak kullanılmasının etkili ve kalıcı öğrenmeyi artıracığı ifade edilebilir. Testler öğrenilen konuların üzerinden belli bir zaman geçtikten sonrasında bile bilgilerin hatırlanmasını sağlıyorsa, öğretim sürecinin içerisinde destekleyici olarak yer almalarının daha çok katlı sağlayacağı söylenilebilir. Bu durumda test etkisi ile ilgili daha fazla çalışmanın yapılması ve testlere eğitim-öğretim programlarında yer verilirken öğretimi destekleyici özelliklerinin göz önünde bulundurulması daha fazla yer verilmesi önerilebilir.

KAYNAKLAR

- Agarwall, P.K. Karpicke, J.D. Kang, S.H.K. Roediger, H.L. & McDermott, K.B., (2008). Examining the testing effect with open- and closed-book tests. *Applied Cognitive Psychology*, 22, 861-876.
- Butler, A.C. & Roediger, H.L. (2007). Testing improves long term retention in a simulated classroom setting. *European Journal of the Cognitive Psychology*, 19(4-5), 514-527.
- Büyüköztürk, Ş., Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. ve Demirel, F. (2008). *Bilimsel Araştırma Yöntemleri*. 2. Baskı. Ankara: Pegem A Yayıncılık
- Campbell, J. & Mayer, R.E. (2009). Questioning as an instructional method: does it affect learning from lectures?. *Applied Cognitive Psychology*, 23, 747-759.
- Carpenter S.K. & DeLosh, E.L. (2006). Impoverished cue support enhances subsequent retention: support for the elaborative retrieval explanation of the testing effect. *Memory&Cognition*, 34(2), 268-276.
- Chang, C.Y. Yeh, T.K. & Barufaldi, J.P. (2010) The positive and negative effects of science concept tests on student conceptual understanding. *International Journal of Science Education*, 32(2), 265-282.
- Creswell, J.W. (2005). *Research design: Qualitative, quantitative and mixed methods approaches*. London: Sage Publications, Inc.
- Kang, S.H.K. McDermott, K.B. & Roediger, H.L. (2007). Test format and corrective feedback modify the effect of testing on long term retention. *European Journal of the Cognitive Psychology*, 19(4-n5), 528-558.
- Karpicke, J.D., & Zaromb, F.M. (2010). Retrieval Mode Distinguishes the Testing Effect from the Generation Effect. *Journal of Memory and Language*. 62,227-239.
- Leeming, F.C.(2002). The exam-a-day procedure improves performance in psychology classes. *Teaching of Psychology*, 29(3), 210-212.
- Mayer, R. E. (2008). *Learning and instruction* (2nd ed.). Upper Saddle River, NJ: Merrill Prentice Hall Pearson.
- McDaniel, M.A. Anderson, J.L. Derbish, M.H. & Morrisette, N. (2007a). Testing the testing effect in the classroom. *European Journal of the Cognitive Psychology*, 19(4-5), 494-513.
- McDaniel M.A. Roediger, H.L. & McDermott, K.B. (2007b). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review*, 14(2), 200-206.
- Naylor, S., & Keogh, B. (1999). Constructivism in Classroom: Theory into Practice. *Journal of Science Teacher Education*. 10(2), 93-106.
- Roediger, H.L. & Karpicke, J.D. (2006a). Test-enhanced learning taking memory tests improves long term retention. *Psychological Science*, 17(3), 249-255.
- Roediger, H.L. & Karpicke, J.D. (2006b). The power of testing memory basic resarch and implications for educational practice. *Perspectives on Psychological Science*, 1(3), 181-210.
- Turgut, M.F. ve Baykul, Y. (2015). *Eğitimde Ölçme ve Değerlendirme*, Pegem Akademi, Ankara.
- Wheeler, M.A. Ewers, M. & Buonanno, J.F. (2003). Different rates of forgetting following study versus test trials. *Memory*, 11(6), 571-580.

SUMMARY

Evaluation, one of the stages of education processes, plays a vital role to determine whether or not to reach to the educational goals. However it is also used to improve and to support the development of the students. Improvement of students' performances by tests is known as testing effect and is a subject to many studies in the field of clinical psychology. However there is a few studies about testing effect in the field of education. Therefore this study aims to determine,

1. whether or not the tests have an effect to recall the learned information?
2. if the testing effect (if there is any) is different for different types of tests?

Pre-test post test control group model was used in this study. Study group composed of 72 science teacher candidates and the study topic was chosen as "energy sources and their effects on the environment" which can be viewed as an important topic for everyone. Data were obtained through a question form which examines the study topic. According to the scores from the pre-test, four groups (3 experimental and 1 control) with the same average score were formed (by using MS Excel). For the experimental process, one week later the pre-test, experimental groups were received different types of tests i.e., first group received multiple choice test (MCT), second group received true-false test (TFT) and the third group received fill-in-the-blanks test (FBT). Control group has received no test. One week later, the question form which was applied as pre-test was given to the pre-service science teachers as post-test and the scores from pre-test and post-test were analyzed through SPSS.

According to the post test scores, true-false test group (TFTG) has the highest score (61.97) and then fill-in-the-blanks test group (FBTG) follows with an average of 55.11 and then multiple-choice test group (MCTG) has an average of 51.74. Control group has the lowest average score (39.67) from post-test.

Post test scores of the groups were used as dependent variable and one-way ANOVA analysis was performed in order to examine the effects of tests on recalling the learned information. According to the results of this analysis there is a meaningful difference ($F(3,68)= 12.029, p<.05$) between post test scores of the groups. Therefore it can be said that tests have an effect to recall the learned information. Tukey analysis was performed in order to examine if there were any meaningful difference according to the types of tests. Results of this analysis reveals that there is meaningful difference ($p<.05$) between all the test groups (TRFT, MCTG, FBTG) post test scores and control group post test score. Additionally there is a meaningful difference ($p<.05$) between post test scores of FBTG and MCTG in favour of TFTG.

It was found with this study that there is an increase in all group's test scores. TFTG's post test scores has risen up 65.2% and reached to 61.97, the highest scores of all four groups. FBTG has risen up 46.9% and reached to 55.11 and MCTG has risen up 37.9% and reached to 51.74. There was also an increase in

post test score of the control group but it was comparatively very small (%5.7). All experimental groups' post test scores were higher than control groups' score. Therefore testing effect can not be merely explained as an extra exposure to the same material. From these results it can be said that tests can help to recall the learned information. There are many studies in the literature (Agarwal et al. 2008; Chang, Yeh and Barufaldi 2010; McDaniel et al. 2007a; McDaniel et al. 2007b; Roediger and Karpicke 2006a; Roediger and Karpicke 2006b; Wheeler, Ewers and Buonanno 2003) which also claims that tests improve the recalling. Multiple comparison analysis results showed that there is a meaningful difference ($p < .05$) between TFTG scores and MCTG scores in favour of TFTG. This result reveals that the testing effect does not only occur because of exposure to the material i.e. since all the pre-service science teachers (in all the groups) expose to the testing material for the same amount of time. However it can be said that as the number of clues (relations between the learned information and the test material) in the testing material increases testing effect increases. The answer is given in between the choices for MCTs or in test for FBTs. However pre-service science teachers were asked to decide whether the statements were true or false for TFT. Therefore pre-service science teachers in the TFTG needed to use more mental processes which brought bigger testing effect. In summary, the findings of this study reveals that tests can help to recall the learned information and between different types of tests TFTs can cause better results since it needs more mental processes.