

Sınıf Öğretmeni Adaylarının Ölçme ve Değerlendirmeye İlişkin İnançları

Preservice Elementary Teachers' Assessment Beliefs

Çavuş ŞAHİN* Pınar KARAMAN**

ÖZ: Bu çalışma, sınıf öğretmeni adaylarının ölçme-değerlendirmeye ilişkin inançlarını belirlemeyi amaçlamıştır. Bu amaçla, Genc (2005) tarafından geliştirilen Likert türü bir ölçek 264 sınıf öğretmeni adayına uygulanmıştır. Dört alt boyuttan (formatif, summatif, geleneksel ve alternatif değerlendirme) oluşan ölçek, toplam 34 tane madde içermektedir. Araştırmada elde edilen verilerin analizinde hem parametrik (t-testi) hem de non-parametrik (Kruskal Wallis ve Mann Whitney-U testi) testler kullanılmıştır. Öğretmen adaylarının ölçme-değerlendirmeye ilişkin inançlarına, cinsiyetlerine, mezun oldukları okul türlerine ve genel akademik başarı ortalamalarına göre karşılaştırılarak incelenmiştir. İnanç düzeylerinin, okul türlerine göre yapılan incelemesi istatistiksel olarak anlamlı bir farklılık ortaya koymazken, cinsiyetlere ve akademik başarı ortalamalarına göre yapılan analiz anlamlı bir farklılık olduğunu göstermiştir. Öğretmen adaylarının inanç düzeyleri ölçeğin alt boyutlarına göre karşılaştırıldığında, öğretmen adaylarının formatif değerlendirmeye olan inanç düzeylerinin en yüksek seviyede, alternatif değerlendirmeye olan inanç düzeylerinin ise en düşük seviyede olduğu bulunmuştur.

Anahtar sözcükler: Ölçme ve değerlendirme, ölçme- değerlendirmeyle ilişkin inançlar, sınıf öğretmeni adayları

ABSTRACT: The purpose of this study is to investigate preservice elementary teachers' assessment beliefs. To achieve that, a Likert-type scale developed by Genc (2005) was applied to 264 preservice elementary teachers. The scale including four subdimensions (formative, summative, traditional and alternative assessment) consists of a total of 34 items. The data were analyzed using both parametric (t-test) and non-parametric (Kruskal Wallis and Mann Whitney-U tests) statistical tests. Assessment beliefs of the preservice elementary teachers were investigated with respect to the following variables: gender, GPA (grade point average) and type of the secondary school graduated. As opposed to school type variable, the analysis of the data produced a statistically significant result on gender and GPA variables. The results revealed that preservice elementary teachers received the highest score on formative assessment subdimension and the lowest score on alternative assessment subdimension.

Keywords: Assessment, assessment beliefs, preservice elementary teachers

1. GİRİŞ

Ölçme ve değerlendirmenin eğitim alanındaki önemi son yıllarda daha çok vurgulanmaya başlanmıştır. Ölçme ve değerlendirme, öğrencilerin ne öğrendikleri hakkında bilgi ve kanıt toplama süreci olarak tanımlanmaktadır (Chen 2003; Wishon vd. 1998). Ölçme ve değerlendirmenin genel olarak amaç ve fonksiyonları şu şekilde sıralanabilir: (1) Öğrencilerin öğrenme yeteneğindeki güçlü ve zayıf yönlerini ortaya çıkarmayı sağlar, (2) Öğrencilerin öğrenme sürecindeki gelişiminin takip edilmesine yardım eder, (3) Sınıfta öğrenilenlerin bir değerlendirilmesinin yapılmasına olanak sağlar, (4) Öğrencilerin verilen kurumsal standartlara dayalı olarak çeşitli öğrenme gruplarına yerleştirilmesine katkı sunar (Heaton 1990; Popham 1995).

Öğrencilerin, öğrenme sürecinde kaydettiği ilerlemeyi veya gelişmeyi tespit ederken sadece bir değerlendirme yönteminden yararlanarak sağlıklı sonuçlara ulaşmak oldukça güçtür. Bu nedenle, farklı öğrenme yaklaşımlarına uygun olan farklı değerlendirme çeşitlerinin

* Doç. Dr., Çanakkale Onsekiz Mart Üniversitesi, Eğitim Fakültesi, Çanakkale-Türkiye, csahin25240@yahoo.com

** Doktora Öğrencisi, Çanakkale Onsekiz Mart Üniversitesi, Eğitim Fakültesi, Çanakkale-Türkiye, pkaraman1626@gmail.com

kullanılması daha verimli olacaktır (Brady 1997). Geleneksel değerlendirme, öğrencileri sadece cevapları ezberleyen (rote learning) pasif öğrencilere dönüştürdüğü için eleştiri konusu olmaktadır (Hart 1994). 1980'li yıllardan itibaren eğitimciler, öğrencilerin kendi cevaplarını açıklayabilecekleri, uygulayabilecekleri, kritik edebilecekleri ve savunacakları değerlendirme yöntemlerinin arayışı içerisine girmişlerdir (Janesick 2001). Eğitimde ölçme ve değerlendirme alanında son yıllarda ortaya çıkan en önemli reformların başında bilişsel (cognitive) kurama dayalı olan alternatif değerlendirme gelmektedir. Alternatif değerlendirme, öğrencilerin başarıları hakkında çok daha fazla ve gerçekçi bilgi elde etmeye olanak sağlayan, öğrencinin derse aktif katılımını teşvik eden ve akademik konuları daha derinlemesine anlamasına yardımcı olan sürece dayalı değerlendirmeleri ifade etmektedir (Stiggins 1994; Svinicki 2004). Davranışçı (behaviorist) kurama dayalı olan geleneksel değerlendirme ise, genel olarak daha çok öğrenme ürünlerini değerlendirmeyi hedefler. Bu hedefe ulaşmak için kullanılan en yaygın değerlendirme formatları ise çoktan seçmeli testler, eşleştirmeli testler, kısa cevaplı ve doğru-yanlış testlerdir. Geleneksel değerlendirmeden farklı olarak, alternatif değerlendirme, öğrenme süreci ve öğrenme ürünlerinin yanısıra öğretim işleyişine ve konularına da odaklanır. Alternatif değerlendirme yaklaşımında, öğrenme çok çeşitli yollarla gerçekleştiği için öğrenilenin değerlendirilmesi de birden fazla yolla gerçekleştirilir. Alternatif değerlendirme yaklaşımında çeşitli değerlendirme türleri kullanılarak anlamlı öğrenmenin ortaya çıkması sağlanır. En yaygın kullanılan alternatif değerlendirme türleri arasında akran değerlendirme, öz değerlendirme, portfolyo, gözlem, demonstrasyon, görüşme (interviews) sayılabilir (Genc 2005). Alternatif değerlendirme öğretmenler arasında genellikle, daha fazla zaman ve emek gerektiren bir yöntem olarak algılanır. Dolayısıyla, alternatif değerlendirmenin daha yaygın olarak uygulanması öğretmenlerin bu yolla yapılan değerlendirmenin daha etkili olacağına olan inançlarının güçlenmesine bağlıdır (Lee vd. 2012). Bu nedenle, öğretmenlerin değerlendirme yöntemlerine karşı geliştirdikleri inançlar onların öğretimlerinde yapacakları tercihler konusunda çok önemli bir rol oynarlar.

Geleneksel ve alternatif değerlendirme yaklaşımlarına ek olarak, öğretmenlerin derslerinde sıklıkla kullandıkları değerlendirme yöntemlerinin başlıcaları arasında formatif ve summatif değerlendirme yaklaşımları sayılabilir. Formatif değerlendirme, 1930'lu ve 1940'lı yıllarda bilişsel ve yapılandırmacı öğrenme teorileriyle birlikte ortaya çıkmıştır (Roos ve Hamilton 2004). Formatif değerlendirme, öğretim programının daha iyi planlanmasına ve öğrenme sürecinin geliştirilmesine odaklanırken, summatif değerlendirme ise yapılan öğretimin etkilerinin ve öğrenme çıktılarının ortaya konulmasıyla daha çok ilgilidir. Formatif değerlendirme, öğrenme sürecinde sadece öğrencilerin değil aynı zamanda öğretmenlerin de gelişimini hedefler. Özellikle, öğretmenlerin verdikleri derslerin bir değerlendirmesini yapmaları suretiyle kendilerini daha fazla geliştirmelerine yardımcı olur. Formatif değerlendirme yaklaşımında, öğrencilere düşündürücü sorular yönelme, tartışma ortamı oluşturma, akran ve öz değerlendirme yapma, gözlem yapma, görüşme yapma, portfolyo hazırlama ve çeşitli quizler uygulama gibi etkinlikler yer alabilir. Summatif değerlendirme, öğretmenin öğrencilerinin gösterdikleri başarı veya performansıyla ilgili bir karara varmasında veya yapılan öğretimin etkililiği konusunda bir sonuca ulaşması için yol gösterici bir rol oynar. Summatif değerlendirme yaklaşımında kullanılan etkinlikler arasında ünite ya da konu testleri, hazırlanan portfolyoların incelenmesi, dönem sınavları, final projeleri ve performans testleri sayılabilir. Formatif değerlendirme özellikle öğrencilerin performansları üzerinden dönütler sağlayarak öğrenimin hızlanmasına yardımcı olur.

Black ve William (1998), formatif değerlendirme yaklaşımının faydalarını çeşitli yollarla dönütü arttırmak (soru-cevap, sınıf tartışmaları, birebir görüşme gibi yöntemler), öğrencinin akran ve öz değerlendirmeye aktif katılımını sağlamak için fırsatlar oluşturmak, öğrencinin öğrenme işleyişi boyunca özsaygısını (self-esteem) ve motivasyonunu arttırmak olarak belirtmişlerdir. Bunlara ek olarak, formatif değerlendirme öğretmenlerin öğrencileri ile olan

ilişkilerinde yeni roller almasının ve sınıf uygulamalarında derin değişiklikler ortaya çıkmasının yolunu açar (Black 2003). Ancak, formatif değerlendirmenin sınıftaki uygulamalarında öğretmenler iki önemli konu ile baş etmek zorundadır. Bunlar; öğretmenin öğrenmeye olan doğal inancı ve öğretmenin olası tüm öğrencilerin öğrenimine olan inancıdır (Black ve William 1998). Dolayısıyla, ölçme-değerlendirmenin öğretmenler tarafından sınıflarında uygulanma yolları, onların öğrenme konusuna nasıl yaklaştıklarından doğrudan etkilenecektir. Formatif değerlendirmeye önem veren bir öğretmen gözlenen bireysel farklılıkların doğasına ek olarak öğrencilerin akıl yürütme ve problem çözmeleri için bilgi yapılarını nasıl geliştirmeleri gerektiği üzerine de odaklanmış olur (Genc 2005). Dikkat edildiği üzere, alternatif ve formatif değerlendirme daha çok yapılandırmacı kurama dayalı iken geleneksel ve summatif değerlendirmenin ise daha çok davranışçı kurama dayalı olduğu görülmektedir.

1.1. Öğretmenlerin İnançları ve Ölçme-Değerlendirme

İnançlar genel olarak bireysel duruşlardır. Pajares'e (1992) göre inançlar dağınık bir yapıya sahiptir. Bu dağınık yapının, tutumlar, değerler, yargılar, düşünceler, ideolojiler, algılamalar, kişisel kuramlar, zihinsel işleyişler, harekete geçme stratejileri, uygulama kuralları, perspektifler ve sosyal stratejiler gibi çok çeşitli kavramlarla ilişkili olduğu ifade edilmiştir. İnançlar, bireylerin yaşadıkları deneyimler, günlük hayatlarında karşılaştıkları olaylarla etkileşim ve katıldıkları olayların yorumları sonucu şekillenir (Al-Sharafi 1998; Hsieh 2002). Bireylerin sahip olduğu inançlar, tutumlara dönüşür, tutumlar niyetleri etkiler, niyetler ise eylemlere dönüşmeye öncülük edecek kararların temelini oluşturur (Bauch 1984). Nihayetinde, bireylerin karar verme süreçlerinde inançların önemli bir yeri vardır (Bandura 1997).

İnançlar, nitelikli öğretmen yetiştirmede anlaşılması gereken en önemli psikolojik yapılardan biridir. Araştırmalar incelendiğinde oldukça geniş bir yer tutan inanç sistemi kavramı, öğretmenlerin sınıflarındaki davranışlarını yöneten önemli bir faktördür. Öğretmenlerin inançlarının hem bakış açılarını hem de yargılarını etkilediği ve bunun da sınıflarındaki davranışlarına, eylemlerine ve uygulamalarına yansıdığı eğitim literatüründe birçok araştırmacı tarafından ifade edilmektedir (Abadiano ve Turner 2003; Ernest 1989; Nespor 1987; Rios 1996). Diğer taraftan ise, inançlar öğretmenlerin davranışlarına rehberlik eden ve şekillendiren yegane faktör değildir. Tutum, öz- yeterlik, öğretmenlik deneyimi, pedagojik formasyon bilgisi gibi faktörler veya psikolojik yapılar da öğretmenlerin sınıf içerisindeki eylemlerinde ve uygulamalarında etkilidir. Fakat, bireysel inançlar öğretmenlerin davranışlarını etkileyen faktörlerin en önemlisi olarak kabul edilmektedir (Genc 2005).

Araştırmacılar eğitim reformlarının başarıya ulaşmasında öğretmenlerin rollerinin önemli olduğunu ifade etmişlerdir. İnançlar, öğretmenlerin sınıf uygulamalarını doğrudan etkilediğinden dolayı eğitim reformlarının başarıya ulaşmasında merkezi rollere sahiptir. Öğretmenlerin sahip oldukları geleneksel inançlar eğitim alanında yapılacak yeniliklerin veya reformların hayata geçirilmesinin önünde bir engel oluşturmaktadır (Woolfolk 2001). Eğitim reformların başarılı bir şekilde uygulamaya geçirilebilmesi için değişikliğin ilk olarak öğretmenlerin inançlarında gerçekleşmesi gerekir (Seferoglu 2007). Bu durum, öğretmenlerin inançlarının öğretmen eğitimi aldıkları zamandan itibaren incelenmesinin önemini ortaya koymaktadır.

Öğretmenlerin profesyonel çalışma zamanlarının büyük bir kısmını ölçme ve değerlendirme etkinliklerini hazırlamak ve uygulamak için harcadıkları bilinen bir durumdur (Stiggins ve Conklin 1988, 1992). Ölçme ve değerlendirmenin sınıfta etkin bir şekilde kullanılabilmesi için öğretmenlerin bu alanda güçlü bilgi birikimine ve yeterli seviyede beceriye sahip olmalarının yanısıra öğretmenlerin ölçme ve değerlendirme konusunda olumlu bir tutum geliştirmeleri ve gelişmiş bir inanca sahip olmaları da gerekir. Fakat bütün bunların ötesinde, öğretmenlerin ölçme-değerlendirmeyi sınıflarında daha bilinçli ve etkin bir şekilde kullanabilmelerinin temelini oluşturan kavram ölçme-değerlendirme okuryazarlığıdır. Öğretmenlerin ne ölçtüğünü, neden ölçtüğünü, nasıl en iyi ölçebileceğini, ölçme-değerlendirme

işleyişinde hangi problemlerin oluşabileceğini ve bu problemlerin nasıl önlenebileceğini bilmesi gibi ve daha başka durumlar ölçme-değerlendirme okuryazarlığı (assessment literacy) ile ilgili konulardır (Stiggins 1995). Öğretmen adaylarının ölçme-değerlendirme okuryazarlığına ilk adımlarını atmalarını sağlamak maksadıyla öğretmen yetiştirme programlarında ölçme-değerlendirme dersi zorunlu olarak verilmektedir (Mertler 2004). Ancak, birçok araştırma öğretmen adaylarının ve hizmet içi öğretmenlerin ölçme-değerlendirme alanındaki bilgilerinin, becerilerinin ve tutumlarının yeterli derecede gelişmiş olmadığını ortaya koymaktadır (Alsarimi 2000; Daniel ve King 1998; Mertler 2003; Mertler ve Campbell 2005). Öğretmenlerin ölçme-değerlendirme okuryazarı bireyler olabilmesi için bu alana ilişkin bilgi ve becerilerinin yanında inançlarının ve tutumlarının da gelişmesi gerekir. Bundan dolayı, öğretmen adaylarının ilerideki meslek yaşamlarında yapacakları ölçme ve değerlendirme uygulamalarına rehberlik edecek olan inançlarının, bakış açılarının ve düşüncelerinin anlaşılması hayati bir öneme sahiptir.

1.2. Araştırmanın amacı

Bu çalışma ile sınıf öğretmeni adaylarının ölçme ve değerlendirmeye karşı geliştirdikleri inançların ortaya çıkarılması amaçlanmıştır. Araştırmanın alt amaçlarında ise aşağıdaki sorulara cevap aranacaktır:

1. Sınıf Öğretmeni adaylarının ölçme-değerlendirmeye ilişkin inanç düzeyleri ne seviyededir?
2. Sınıf Öğretmeni adaylarına uygulanan ölçeğin alt boyutları (formatif, summatif, geleneksel ve alternatif değerlendirme) arasında anlamlı bir ilişki var mıdır?
3. Sınıf Öğretmeni adaylarının ölçme-değerlendirmeye ilişkin inançlarında, cinsiyete, mezun olunan okul türüne ve akademik başarı ortalamasına göre anlamlı bir farklılık var mıdır?
4. Sınıf Öğretmeni adaylarına uygulanan ölçeğin alt boyutlarındaki (formatif, summatif, geleneksel ve alternatif değerlendirme) inanç düzeyleri arasında istatistiksel olarak anlamlı bir farklılık var mıdır?

2. YÖNTEM

Sınıf öğretmeni adaylarının ölçme-değerlendirmeye ilişkin inançlarının betimlenmesini amaçlayan bu çalışmada, tarama modeli kullanılmıştır. Karasar'a (1995) göre tarama modelleri, geçmişte ve halen var olan durumu olduğu şekliyle betimlemeyi amaçlayan araştırma yaklaşımıdır.

2.1. Evren ve Örneklem

Araştırmanın evrenini, Çanakkale Onsekiz Mart Üniversitesi Eğitim Fakültesi, İlköğretim Bölümü, Sınıf Öğretmenliği Anabilim Dalı'nda öğrenim gören öğretmen adayları oluşturmaktadır. Araştırmanın örneklemi ise; 2011-2012 öğretim yılında Çanakkale Onsekiz Mart Üniversitesi Eğitim Fakültesi İlköğretim Bölümü Sınıf Öğretmenliği Anabilim Dalı'nda 3. ve 4. sınıfta öğrenim gören öğretmen adayları oluşturmaktadır. Örneklem 3. ve 4. sınıflarda öğrenimine devam eden öğretmen adaylarının seçilmesinin sebebi, Sınıf Öğretmenliği programında ölçme ve değerlendirme dersinin 3. sınıftan itibaren verilmesinden dolayıdır. Dolayısıyla seçilen örneklem türü, amaçsal örneklem (purposeful sampling) yöntemlerinden biri olan ölçüt örneklemi (criterion sampling) temsil eder.

2.2. Çalışma Grubu

Çalışmaya katılan 3. ve 4. sınıf 264 tane sınıf öğretmeni adayının 183'ünü kızlar ve 81'ini ise erkekler oluşturmaktadır.

2.3. Ölçme Aracı

Bu çalışmada sınıf öğretmeni adaylarının ölçme-değerlendirme hakkında sahip oldukları inançları incelemek amacıyla Genc (2005) tarafından geliştirilen ölçek kullanılmıştır. Ölçme aracı Türkçe'ye çevrildikten sonra ifade yeterliliğini sağlamak amacıyla Türkçe alanında uzman kişinin görüşlerine başvurulmuştur. Geliştirilmiş ölçme aracının hipotez edilmiş modeli 34 maddeden oluşmuş olup formatif değerlendirme, summatif değerlendirme, geleneksel değerlendirme ve alternatif değerlendirme olmak üzere dört alt boyutu ölçtüğü ifade edilmiştir. 34 maddelik ölçme aracı öğretmen adaylarına uygulanarak ölçme ve değerlendirmeye ilişkin inançları ölçülmüştür. Öğretmen adayları ölçme aracında yer alan her bir maddeye ne derece katılıp katılmadıklarını 5 dereceli Likert türü ölçek üzerinde işaretlemişlerdir. Ölçekte yer alan seçenekler ve puan değerleri şu şekildedir: hiç katılmıyorum (1), katılmıyorum (2), kararsızım (3), katılıyorum (4) ve tamamen katılıyorum (5).

Araştırmanın örnekleminde yer alan 264 öğretmen adayının inançlarını ölçmek için Kaiser-Meyer-Olkin (KMO) testi ve Barlett'in küresellik derecesi (BKD) kullanılmıştır. KMO ölçümü, değişkenlerin korelasyonunu kullanarak örneklemdaki değişkenlerin korelasyon için yeterli olup olmadığını hesaplar. KMO değerinin 0.5'ten büyük çıkması, tatmin edici bir faktör analizinin yapılabileceğine işaret eder. Sınıf öğretmeni adaylarının ölçme-değerlendirme inançlarıyla ilgili toplanan veride KMO değeri 0.871'dir. Bu değer, faktör analizi için örneklem yeterliliğinin iyi derecede olduğunu göstermektedir. BKD testi ise, değişkenler arasında ilişki olup olmadığının bir ölçüsüdür. p değerinin 0.05'ten küçük çıkması halinde çalışmadaki değişkenler arasında yeterli korelasyonun var olduğu ve faktör analizi için yeterli güce sahip olduğu sonucuna varılır. Bu çalışmada, BKD testinin ($p < 0,01$) yüksek seviyede anlamlı bir değere sahip olduğu ve işletilen faktör analizinin güvenilir sonuçlar verdiğini göstermiştir. Maddelerin, ölçme aracının psikometrik özelliklerine olan etkisini ölçmek için uygulanan faktör analizinden elde edilen sonuçlar, gözlenen verinin faktör yük değerlerinin dört boyutlu örtük (gizli) yapıyı temsil ettiği görülmüştür.

Çalışmada yapılan faktör analizinde, faktör yük değerlerinin en yüksek 0.737 ve en düşük 0.376 olduğu görülmüştür. 0.5 ve daha yüksek çıkan faktör yük değerleri uygun değerler olarak kabul edilir (Costello ve Osborne 2005). Buna göre, yapılan faktör analizi sonucunda, faktör yük değerleri 0.5 değerinden düşük olan maddeler, teorik olarak belirlenen alt boyutlarla ilişki göstermeyen veya birden çok faktöre yüklenen maddeler oldukları düşünüldükçe gözardı edilmişlerdir. F1 (formatif değerlendirme) alt boyutunu temsil eden maddelerden 5. ve 8. madde, F2 (summatif değerlendirme) alt boyutunu temsil eden sorulardan 11. madde, F3 (geleneksel değerlendirme) alt boyutunu temsil eden maddelerden 21. ve 23. madde gözardı edilmiştir. F4 (alternatif değerlendirme) boyutunu temsil eden bütün maddelerin faktör yük değerlerinin istenilen değerlerde olduğu gözlenmiştir. Faktör analizinden elde edilen değerlere göre, 20 tane maddenin ölçekte yer almasının uygun olacağı sonucuna varılmıştır. Tablo 1'de ölçeğin döndürülmüş temel bileşenler analiz yöntemine göre faktör yük değerleri gösterilmiştir.

Tablo 1: Sınıf öğretmeni adaylarının ölçme-değerlendirme inançları hakkında 34 maddelik döndürülmüş temel bileşenler analizi yöntemine göre faktör yük değerleri

| Madde | F1 | Madde | F2 | Madde | F3 | Madde | F4 |
|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 | 0.583 | 11 | 0.376 | 18 | 0.501 | 26 | 0.561 |
| 2 | 0.615 | 12 | 0.684 | 19 | 0.464 | 27 | 0.641 |
| 3 | 0.636 | 13 | 0.569 | 21 | 0.441 | 28 | 0.500 |
| 4 | 0.652 | 14 | 0.589 | 23 | 0.425 | 29 | 0.652 |
| 5 | 0.363 | 16 | 0.662 | | | 30 | 0.728 |
| 6 | 0.594 | | | | | 31 | 0.659 |
| 7 | 0.519 | | | | | 32 | 0.691 |
| 8 | 0.468 | | | | | 34 | 0.502 |

Ölçme aracının güvenilirliğini belirlemek için Cronbach Alfa korelasyon katsayısı kullanılmıştır. George ve Mallery' e (1995) 0.5'ten büyük değerler korelasyon katsayısı için yeterli değerlerdir. Bu çalışmada kullanılan ölçme aracı için SPSS kullanılarak Croanbach Alfa korelasyon katsayısı 0.868 olarak hesaplanmıştır. Bu katsayıya göre ölçme aracının güvenilirlik düzeyinin oldukça iyi olduğu söylenebilir. İnanç ölçeğinin tüm alt boyutlarındaki Croanbach Alfa korelasyon katsayıları ölçüldüğünde ise, formatif değerlendirme alt boyutu için 0.765, summatif değerlendirme alt boyutu için 0.702, geleneksel değerlendirme alt boyutu için 0.663, alternatif değerlendirme alt boyutu için ise 0.822 değerleri bulunmuştur. Bu sonuçlara göre, uygulanan ölçeğin hem genel hem de alt boyutları için hesaplanan Cronbach Alpha güvenilirlik katsayısı değerlerinin oldukça iyi olduğu söylenebilir.

2.4. Verilerin Analizi

Araştırmada, parametrik istatistik testlerin uygulanabilmesi için gerekli olan varsayımların yerine gelip gelmediği test edilmiştir. Çalışmada analizi yapılacak değişkenlerin histogram grafiğinde normal dağılım gösterdikleri tespit edildiğinden dolayı parametrik test kullanılmasının uygun olduğuna karar verilmiştir. Araştırmanın amacına uygun olarak, parametrik istatistiksel testlerden bir tanesi olan t-testi kullanılmıştır. İlk olarak öğretmen adaylarının ölçeğin alt boyutları olan formatif, summatif, geleneksel ve alternatif değerlendirme hakkında sahip oldukları ortalama inanç düzeyleri hesaplanmıştır. Ayrıca, inanç ölçeğinin alt boyutları arasındaki korelasyon tespit edilmiştir. Öğretmen adaylarının cinsiyetleri arasında ölçme-değerlendirme inançlarına ilişkin istatistiksel olarak anlamlı bir farklılık olup olmadığını tespit etmek amacıyla bağımsız gruplar t-test analizi uygulanmıştır. Akademik başarı ortalamaları ve mezun olunan okul türü değişkenlerine göre inançlar arasındaki farklılıklar incelenirken, bazı gruplarda yeterli sayıda denek olmadığı tespit edilmiş ve non-parametrik istatistik uygulanmasına karar verilmiştir. Buna göre, bu değişkenlerin analizinde Kruskal Wallis ve Mann Whitney-U testleri kullanılmıştır.

3. BULGULAR

Tablo 2: Öğretmen Adaylarının Ölçeğin Alt Boyutlarına Göre İnanç Düzeyleri

| | N | \bar{X} | ss. |
|--------------------------|-----|-----------|------|
| Formatif Değerlendirme | 264 | 4.12 | 0.47 |
| Summatif Değerlendirme | 264 | 3.96 | 0.59 |
| Geleneksel Değerlendirme | 264 | 4.00 | 0.72 |
| Alternatif Değerlendirme | 264 | 3.88 | 0.55 |

Tablo 2'de, sınıf öğretmeni adaylarının her alt boyuttaki ölçme- değerlendirmeye ilişkin inanç düzeylerini belirlemek için hesaplanan ortalamalar en yüksek değerden en düşük değere doğru incelendiğinde şu sonuçlar elde edilmiştir: formatif değerlendirme (\bar{X})= 4.12), summatif değerlendirme (\bar{X})= 3.96), geleneksel değerlendirme (\bar{X})=4.0), alternatif değerlendirme (\bar{X})=3.88).

Tablo 3: İnanç Ölçeğinin Alt Boyutları Arasındaki Korelasyonlar

| Faktör | F1 | F2 | F3 | F4 |
|--------|---------|---------|---------|----|
| F1 | 1 | | | |
| F2 | 0.466** | 1 | | |
| F3 | 0.385** | 0.476** | 1 | |
| F4 | 0.494** | 0.329** | 0.405** | 1 |

*p<.001

Tablo 3'te yer alan F1 (formatif değerlendirme), F2 (summatif değerlendirme), F3 (geleneksel değerlendirme) ve F4 (alternatif değerlendirme) alt boyutları arasındaki korelasyon katsayılarının 0.329 ile 0.494 arasında pozitif değerler aldığı ve bu değerlerin istatistiksel olarak anlamlı olduğu görülebilir. Cohen'in (1988) de belirttiği gibi bu değerler pozitif yönde ve orta güçlükte korelasyon ilişkisi olduğunu gösterir. Tablo 3 dikkatlice incelendiğinde en yüksek anlamlı ilişkinin F1-F4($r=0.494$) alt boyutları arasında ve F2-F3($r=0.476$) alt boyutları arasında olduğu tespit edilmiştir.

Tablo 4: Cinsiyete Göre İnanç Seviyelerinin Bağımsız Gruplar t-test Sonuçları

| Grup | N | \bar{X} | ss. | t | p |
|-------|-----|-----------|------|------|-------|
| Kız | 183 | 4.04 | 0.39 | 3.67 | .000* |
| Erkek | 81 | 3.84 | 0.46 | | |

* $p<.05$

Sınıf öğretmeni adaylarının cinsiyetleri arasında ölçme-değerlendirme hakkında sahip oldukları inançlar arasında anlamlı bir farkın olup olmadığını incelemek amacıyla bağımsız gruplar için t-testi uygulanmıştır. Tablo 4'te özetlenen sonuçlara göre kız öğretmen adaylarının ölçme-değerlendirmeye karşı geliştirdikleri inanç düzeyleri (\bar{X}) =4.04) erkek adaylara (\bar{X}) =3.84) göre istatistiksel olarak anlamlı düzeyde yüksek çıkmıştır ($t = 3.67, p<0.05$).

Tablo 5: Cinsiyete Göre Alt Boyut İnanç Seviyelerinin Bağımsız Gruplar t-test Sonuçları

| | | N | \bar{X} | ss. | t | p |
|--------------------------|-------|-----|-----------|------|-------|--------|
| Formatif Değerlendirme | Kız | 183 | 4.15 | 0.43 | 2.025 | 0.04* |
| | Erkek | 81 | 4.03 | 0.54 | | |
| Summatif Değerlendirme | Kız | 183 | 4.03 | 0.57 | 2.93 | 0.004* |
| | Erkek | 81 | 3.8 | 0.6 | | |
| Geleneksel Değerlendirme | Kız | 183 | 4.11 | 0.63 | 3.7 | 0.000* |
| | Erkek | 81 | 3.76 | 0.85 | | |
| Alternatif Değerlendirme | Kız | 183 | 3.95 | 0.52 | 3.162 | 0.002* |
| | Erkek | 81 | 3.72 | 0.58 | | |

* $p<.05$

Ölçme-değerlendirmeye ilgili kız ve erkek öğretmen adaylarının inanç düzeyleri arasında istatistiksel olarak anlamlı bir farklılık tespit edildiğinden dolayı, ölçeğin alt boyutlarına göre inceleme yapılmıştır (Tablo 5). Tablo 2'deki bulgulara benzer nitelikte, hem kız hem de erkek öğrenciler en yüksek inanç puanlarını formatif değerlendirme alt boyutunda ve en düşük puanı ise alternatif değerlendirme alt boyutunda aldıkları görülmektedir. Bütün alt boyutlarda kızların aldıkları puanların ise erkeklerden daha yüksek olduğu tespit edilmiştir.

Tablo 6: Akademik Başarı Ortalamalarına Göre İnanç Puanlarının Kruskal Wallis Testi Sonuçları

| | Akademik Ortalamalar | N | Sıra Ortalaması | sd | χ^2 | p |
|-----------|----------------------|-----|-----------------|----|----------|--------|
| * $p<.05$ | 1.80-2.00 | 8 | 164.88 | 4 | 10.59 | *0.032 |
| | 2.01-2.50 | 66 | 146.08 | | | |
| | 2.51-3.00 | 166 | 129.72 | | | |
| | 3.01-3.50 | 21 | 103.00 | | | |
| mf | 3.51-4.00 | 2 | 29.75 | | | |

öğretmeni adaylarının akademik başarı ortalamalarına göre ölçme-değerlendirmeye karşı geliştirdikleri inanç düzeyleri arasında farklılık olup olmadığını tespit etmek için Kruskal Wallis analizi yapılmıştır. Bu test analizinin seçilmesinin temel nedeni ise, bazı gruplar içindeki denek sayılarının az olmasından dolayıdır. Öğretmen adaylarının akademik başarı ortalamaları ile ilgili veriler parametrik test varsayımlarına uygun olmadığı için non-parametrik Kruskal Wallis testi kullanılmıştır. Analiz sonuçlarına göre (Tablo 6), araştırmaya katılan öğretmen adaylarının akademik başarı ortalamalarına göre ölçme-değerlendirmeye ilişkin inançları, istatistiksel olarak anlamlı düzeyde farklılaşmaktadır ($\chi^2(4) = 10,59$, $p < .05$). Akademik ortalamaların hangi seviyelerde farklılık gösterdiğini belirlemek amacıyla Mann Whitney-U testi kullanılmıştır (Tablo 7).

Tablo 7: Akademik Başarı Ortalamalarına Göre İnanç Puanlarının Mann Whitney-U Testi Sonuçları

| Akademik Ortalamalar | N | Sıra Ortalaması | U | p |
|----------------------|----|-----------------|-----|--------|
| 2.01-2.50 | 66 | 47.53 | 460 | *0.021 |
| 3.01-3.50 | 21 | 62.90 | | |

*p<.05

Tablo 7’de görüldüğü gibi, akademik başarı ortalamaları “2.01-2.50” ve “3.01-3.50” arasındaki değerlere sahip olan sınıf öğretmeni adaylarının ölçme-değerlendirmeye ilişkin inançları arasında istatistiksel olarak anlamlı bir farklılık olduğu tespit edilmiştir (U=460, p<0.05). Sıra ortalamalarına bakıldığında, akademik başarı ortalaması “3.01-3.50” olan öğretmen adaylarının ölçme-değerlendirmeye olan inançlarının başarı ortalamaları “2.01-2.50” olan öğretmen adaylarına göre daha yüksek olduğu söylenebilir.

Tablo 8: Okul Türlerine Göre İnanç Puanlarının Kruskal Wallis Testi Sonuçları

| Mezun lise | N | Sıra Ortalaması | χ^2 | p |
|-----------------------|-----|-----------------|----------|-------|
| Genel liseler | 131 | 126.63 | 8.297 | 0.141 |
| Öğretmen liseleri | 10 | 153.65 | | |
| Çok programlı liseler | 13 | 125.88 | | |
| Meslek liseleri | 14 | 92.75 | | |
| Anadolu liseleri | 70 | 137.51 | | |
| Diğer | 24 | 156.96 | | |

Tablo 8’de gösterildiği gibi, sınıf öğretmeni adaylarının mezun oldukları ortaöğretim okul türlerine göre ölçme-değerlendirmeye ilgili inanç düzeyleri arasında istatistiksel olarak anlamlı bir farklılık olup olmadığı araştırılmıştır. Her okul türünde yeterli sayıda öğretmen adayı olmadığından dolayı Kruskal Wallis non-parametrik istatistik analizi kullanılmasının uygun olacağı düşünülmüştür. Kruskal Wallis testinden elde edilen bulgular, adayların ölçme-değerlendirme inanç puanlarının mezun oldukları ortaöğretim okul türlerine göre istatistiksel olarak anlamlı farklılık göstermediğini ortaya çıkarmıştır ($\chi^2=8.297$ ve $p>0.05$).

4. TARTIŞMA ve SONUÇ

Sınıf öğretmeni adaylarına uygulanan ölçme-değerlendirme inanç ölçeği formatif, summatif, geleneksel ve alternatif değerlendirme olmak üzere dört alt boyuttan oluşmuştur. Araştırmada sınıf öğretmenlerine uygulanan ölçeğin alt boyutları arasında anlamlı ilişki olup olmadığı araştırılmıştır. Elde edilen sonuçlar, alt boyutlar arasında istatistiksel olarak anlamlı, pozitif yönde ve orta derecede bir ilişkinin olduğunu göstermiştir. Ölçekte yer alan alt boyutlar arasında en güçlü ilişkinin F1 (formatif değerlendirme) ile F4 (alternatif değerlendirme) ($r=0.494$) alt boyutları arasında ve F2 (summatif değerlendirme) ile F3 (geleneksel değerlendirme) ($r=0.476$) alt boyutları arasında olduğu tespit edilmiştir. Summatif ve geleneksel değerlendirme yöntemlerinin davranışçı yaklaşıma daha yakın, formatif ve alternatif değerlendirme tekniklerinin ise yapılandırmacı yaklaşıma daha yakın olduğu düşünüldüğünde, elde edilen bulguların kuramsal yapıyı destekler nitelikte olduğu söylenebilir.

Bu araştırmadan elde edilen bulgular, sınıf öğretmeni adaylarının ölçme-değerlendirmeye ilgili inanç düzeylerinin, genel olarak yeterince gelişmiş bir seviyede olmadığını ortaya koymuştur. Sınıf öğretmeni adaylarının yeterince gelişmiş inanç düzeylerine sahip olmamalarında, ölçme- değerlendirmeyle ilgili tutumları ve bilgi seviyelerinin de etkili olduğu söylenebilir. Tittle'ın (1994) da belirttiği gibi, öğretmenlerin derslerinde kullandıkları ölçme-değerlendirme yaklaşımlarının oluşmasında öğretmenlerin inançları, tutumları ve bilgisi de önemli bir rol oynar. Bir başka ifadeyle, öğretmen adaylarının ölçme-değerlendirme konusundaki inanç düzeylerindeki göreceli düşüklük, bu alanda yeterince bilgi ve tecrübe birikimine sahip olmamak ile açıklanabilir. Nitekim Mertler (2003, 2004) tarafından yapılan bir çalışma ölçme ve değerlendirme dersini kısa bir süre önce tamamlamış öğretmen adaylarının ölçme-değerlendirme okuryazarlığı (assessment literacy) seviyelerinin, hizmet içi öğretmenlere kıyasla daha düşük bir değerde olduğunu ortaya koymuştur. Bundan çıkarılabilecek sonuç, kısıtlı saatlere sıkıştırılmış ve sadece teorik olarak verilen derslerin öğretmen adaylarının gelişimlerinde yeterince etkili olamadığı ve gelişkin bir anlayış geliştirmelerine olanak sağlayamadığı şeklinde olabilir. Dolayısıyla, öğretmen adaylarına verilen ölçme ve değerlendirme derslerinin daha çok uygulamaya yönelik etkinlikler içermesi, onların bu alanda daha gelişkin bir anlayışla meslek hayatlarına atılmalarına katkı sunacaktır. Balcı'nın (2003) da önerdiği gibi, öğretmen yetiştirme programlarında yer verilen ölçme ve değerlendirme derslerinin teorik bilginin yanısıra öğrencinin uygulama yapabileceği fırsatları da içermesi gerekir. Böylelikle, öğretmen adaylarının ölçme ve değerlendirme hakkında sahip olduğu bilgi ve deneyimlerinin gelişkin bir düzeye ulaşması, bakış açılarının ve inanç düzeylerinin de olumlu manada değişim göstermesine yol açacaktır.

Araştırmada, sınıf öğretmeni adaylarının ölçme-değerlendirmeye ilişkin inançlarının çeşitli demografik değişkenlere göre gösterdiği farklılıklar incelenmiştir. Ölçme-değerlendirme hakkındaki inançlara cinsiyet açısından bakıldığında, kız öğretmen adaylarının inanç düzeylerinin erkeklere oranla istatistiksel olarak anlamlı düzeyde farklılık gösterdiği tespit edilmiştir. Cinsiyet açısından ölçeğin tüm alt boyutlarına bakıldığında da kızların erkeklere oranla ölçme-değerlendirmeye ilgili inanç düzeylerinin daha gelişmiş olduğu görülmüştür. Birçok çalışmada, cinsiyete göre öğretmen adaylarının öğretmenlik mesleğine yönelik tutumları incelendiğinde kızların erkeklere oranla daha yüksek düzeyde puan aldığına rastlanmaktadır (Akpınar vd. 2006; Çapri ve Çelikkaleli 2008; Bozdoğan vd. 2007; Çetinkaya 2009; Baykara Pehlivan 2008; Baykara Pehlivan 2010). İnel ve diğerleri'nin (2010) yaptığı benzer bir çalışmada da sınıf öğretmeni adaylarının yapılandırmacı yaklaşıma ilişkin görüşleri ve tutumları incelenmiştir. Kız öğretmen adaylarının öğretmenlik mesleğinde geleneksel yöntemlere karşı yapılandırmacı yaklaşımı yansıtan alternatif değerlendirmeye karşı tutumlarının daha olumlu olduğu ortaya çıkmıştır. Ayrıca birçok çalışma kız öğretmen adaylarının öz-yeterlilik inançlarının daha güçlü olduğunu göstermektedir (Kalain ve Freeman 2004; Kiremit 2006; Başer vd. 2005; Özdemir 2008). Kız öğretmen adaylarının meslek ile ilgili tutumlarının, öz-

yeterliliklerinin daha gelişkin olmasına paralel olarak bu çalışmada da öğretmenlik meslek derslerinden bir tanesi olan ölçme-değerlendirmeye bakış açılarının ve inançlarının daha gelişkin olduğu ortaya çıkmıştır. Sonuçlar genel olarak, öğretmenlik mesleğinin kızlara daha uygun olduğunu gösterdiği şeklinde yorumlanabilir (Cheug 2008).

Sınıf öğretmeni adaylarının mezun oldukları ortaöğretim türlerine göre inanç düzeylerine bakıldığında ise istatistiksel olarak anlamlı bir farklılık ortaya çıkmamıştır. Özellikle, Anadolu Öğretmen Lisesi mezunu öğretmen adaylarının, ortaöğretim programlarında ölçme-değerlendirme dersinin de dahil olduğu öğretmenlik meslek bilgisi dersleri almalarına rağmen diğer öğretmen adaylarıyla aralarında istatistiksel olarak anlamlı bir farklılık ortaya çıkmadığı görülmüştür. Bu çalışmada sınıf öğretmeni adaylarının akademik başarı ortalamaları ile inanç düzeyleri arasında istatistiksel olarak anlamlı bir farklılık olduğu belirlenmiştir. Bu sebeple, sınıf öğretmeni adaylarının akademik başarı ortalamalarının ölçme-değerlendirmeyle ilgili inanç ya da bakış açılarında etkili olduğu söylenebilir.

Adayların ölçme-değerlendirme hakkındaki inançlarının, ölçekte yer alan alt boyutlarına (formatif, summatif, geleneksel ve alternatif değerlendirme) göre yapılan analizinde, formatif değerlendirmeye karşı sahip olunan inanç düzeyi, summatif değerlendirmeye kıyasla daha yüksek düzeydedir. Black vd. (2004), öğretmenlerin değerlendirme yaparken summatif yönünün daha ağır bastığını göstermektedir. Fakat araştırmalar formatif ve summatif değerlendirmenin birlikte kullanılmasıyla daha etkili sonuçlar elde edildiğini ortaya koymuştur (Black ve William, 1998). Öğretmen yetiştirme programlarında formatif değerlendirmenin çoğunlukla ihmal edildiği bilinen bir gerçektir (Dwyer 1994; Dwyer 1998). Uğurlu ve Akkoç'un (2011) çalışmaları da bu sorunu ortaya çıkarmıştır. Çalışmalarında öğretmen adaylarının formatif değerlendirme açısından yetiştirilmesi gerektiği üzerine vurgu yapmışlardır. Bu çalışmadan elde edilen sonuçlar ise öğretmen adaylarının formatif değerlendirmenin etkili olduğuna yüksek oranda inandığını ortaya koymaktadır. Buradan yapılacak çıkarım, öğretmen adaylarının formatif değerlendirme konusundaki inançlarının olumlu yönde bir gelişme gösterdiğine işaret eder. Buna karşın, öğretmen adaylarının alternatif değerlendirmeye olan inançlarının en düşük seviyede olduğu tespit edilmiştir. Öğretmen adaylarının bakış açılarının bu şekilde farklılık göstermesinin sebebi, Lee vd. (2012) de belirttiği gibi alternatif değerlendirmenin geleneksel değerlendirmeye kıyasla daha yeni bir yöntem olmasından kaynaklanıyor olabilir. Daha önceki birçok çalışma buna paralel olarak öğretmenlerin alternatif değerlendirme yöntemlerini yeterince benimseyemediklerini ve bu yöneme karşı daha olumsuz bir tutum geliştirdiklerini tespit etmiştir (Arslan vd. 2009; Karatepe vd. 2004; Gömleksiz ve Bulut 2007). Öğretmen adaylarının alternatif değerlendirmeyi yeterince etkili bulmamasının diğer bir sebebi ise, öğretmen adaylarının alternatif değerlendirme hakkında yeterli bilgi ve deneyime sahip olmamalarından kaynaklanabilir. Birgin ve Gürbüz'ün (2008) sınıf öğretmeni adaylarının ölçme ve değerlendirme alanındaki bilgi düzeylerini incelediği çalışması bunu destekler niteliktedir. Çalışmaları, sınıf öğretmeni adaylarının çoğunun alternatif değerlendirme metotları konusundaki bilgilerinin yeterli olmadığını göstermiştir. Araştırmalardan elde edilen sonuçlara göre, öğretmen adaylarının alternatif değerlendirme ile ilgili bilgi, tutum ve inanç düzeylerini geliştirmek için verilen ölçme-değerlendirme derslerinin alternatif değerlendirme yöntemi ile ilgili kapsamlı teorik bilgileri sağlaması yanında öğretmen adaylarına uygulama yapma fırsatı da sağlaması gerektiği söylenebilir. Öğretmen adayları bu kazanımları sadece ölçme-değerlendirme derslerinde değil aynı zamanda öğretmenlik uygulamalarında da sağlamalıdır. Şahin'in (2005) de belirttiği gibi öğretmen adayları meslekleri için gerekli bilgi ve davranışları ancak uygulama ile kazanır. Öğretmenlik uygulaması sırasında öğretmen adaylarına ölçme-değerlendirme alanında geleneksel değerlendirme tekniklerinin yanı sıra alternatif değerlendirme tekniklerini uygulama fırsatı da verilmelidir. Böylelikle, öğretmenlerin öğretmen eğitimi aldıkları andan itibaren farklı değerlendirme tekniklerine ve özellikle de alternatif değerlendirmeye karşı inanç ve tutumlarında olumlu bir gelişme sağlanabilir.

5. KAYNAKLAR

- Abadiano, H., & Turner, J. (2003). Thinking it through: re-examining our beliefs about assessment for diverse students. *The New England Reading Association Journal*, 39, 58-63.
- Akpınar, E., Yıldız, E. ve Ergin, Ö. (2006). Fen bilgisi öğretmen adaylarının öğretmenlik mesleğine yönelik tutumları. *Dokuz Eylül Üniversitesi Buca Eğitim Fakültesi Dergisi*, 19, 56-62.
- Al-sharafi, A. (1998). *An investigation of the beliefs and practice of foreign language teachers: A case study of five American high school foreign language teachers in Leon County*. Unpublished doctoral dissertation. College of Education of Florida State University.
- Alsarimi, A. M. (2000). *Classroom assessment and grading practices in the Sultanate of Oman*. Unpublished doctoral dissertation. University of Pittsburgh, Pennsylvania.
- Arslan, S., A., Kaymakçı, D., Y. ve Arslan, S. (2009). Alternatif ölçme-değerlendirme etkinliklerinde karşılaşılan problemler: Fen ve Teknoloji Öğretmenleri örneği. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi*, 28, 1-12.
- Balcı, M. (2003). Fen bilgisi öğretiminde ölçme-değerlendirme üzerine bir çalışma. *Milli Eğitim Dergisi*, 157, 23-29.
- Bandura, A. (1977). *Social Learning Theory*. New York: General Learning.
- Başer, N., Günhan, B. C. ve Yavuz, G. (2005). *İlköğretim öğretmen adaylarının ve öğretmenlerin öğretmen yeterlik algılarının karşılaştırılması üzerine bir araştırma*. XIV. Ulusal Eğitim Bilimleri Kongresi, 515-521. Denizli: Pamukkale Üniversitesi.
- Bauch, P. (1984). *The impact of teachers' instructional beliefs on their teaching: Implications for research and practice*. ERIC Digest. ED252954.
- Baykara, P. K. (2010). Öğretmen adaylarının öğrenme stilleri ve öğretmenlik mesleğine yönelik tutumları üzerine bir çalışma. *İlköğretim Online*, 9(2), 749-763.
- Birgin, O. ve Gürbüz, R. (2008). Sınıf Öğretmeni adaylarının ölçme ve değerlendirme konusundaki bilgi düzeylerinin incelenmesi. *Seçuk Üniversitesi Sosyal Bilimler Enstitüsü dergisi*, 20, 163-180.
- Black, P. (2003). *Formative and summative assessment: Can they serve learning together?* Paper presented at the annual convention of the American Educational Research Association, Chicago.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1), 9-21.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, 5, 7-74.
- Bozdoğan, A. E., Aydın, D. ve Yıldırım, K. (2007). Öğretmen adaylarının öğretmenlik mesleğine ilişkin tutumları. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, 8(2), 83-97.
- Brady, P. (1997). *Assessment and the curriculum*. In C. Cullingford (ed.), *Assessment versus evaluation* (pp. 8-23). London: Cassell.
- Chen, H. (2003). *A study of primary school English teachers' beliefs and practices in multiple assessments: A case study in Taipei City*. Unpublished master theses. Taipei: National Taipei Teachers College.
- Cheung, H. Y. (2008). Teacher efficacy: a comparative study of Hong Kong and Shanghai primary in-service teachers. *The Australian Educational Researcher*, 35(1), 103-123.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.) Hillsdale, NJ: Lawrence Erlbaum Associates.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, 10(7).
- Çapri, B. ve Çelikkaleli, Ö. (2008). Öğretmen adaylarının öğretmenliğe ilişkin tutum ve mesleki yeterlik inançlarının cinsiyet, program ve fakültelerine göre incelenmesi. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 9(15), 33-53.
- Çetinkaya, Z. (2009). Türkçe öğretmen adaylarının öğretmenlik mesleğine ilişkin tutumlarının belirlenmesi. *İlköğretim Online*, 8(2), 298-305.

- Daniel, L. G., & King, D. A. (1998). Knowledge and use of testing and measurement literacy of elementary and secondary teachers. *Journal of Educational Research*, 91, 331–343.
- Dwyer, C.A. (1994). Development of the Knowledge Base for the Praxis III: Classroom Performance Assessments Assessment Criteria. Princeton, NJ: Educational Testing Service.
- Dwyer, C. A. (1998). Assessment and classroom learning: Theory and practice. *Assessment in Education*, 5, 131 – 137.
- Ernest, P. (1989). The knowledge, beliefs and attitudes of the Mathematics Teacher: A model. *Journal of Education for Teaching*. 15, 113 – 133.
- Genc, E. (2005). *Development and validation of an instrument to evaluate science teachers' assessment beliefs and practices*. Unpublished doctoral dissertation. College of Education of Florida State University.
- George, D., & Mallery, P. (1995). *SPSS/PC: Step by step*. Belmont, CA: Wadsworth Press.
- Gömlüksiz, M. N. ve Bulut, İ. (2007). Yeni fen ve teknoloji dersi öğretim programının uygulamadaki etkililiğinin değerlendirilmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 32, 76-88.
- Hart, D. (1994). *Authentic assessment: A handbook for educators*. Menlo Park, CA: Addison-Wesley.
- Heaton, J. (1990). *Writing English language tests*. New York: Longman Inc.
- Hsieh, H. (2002). *Teachers' beliefs about English learning: A case study of elementary school English teachers in Taipei County*. Unpublished master thesis. Taipei: National Taipei Teachers' College.
- İnel, D., Türkmen, L. ve Evrekli, E. (2010). *Sınıf öğretmenleri adaylarının yapılandırmacı yaklaşıma ilişkin görüşlerinin ve tutumlarının incelenmesi: Uşak üniversitesi örneği*. Dokuz Eylül Üniversitesi Eğitim Fakültesi, Buca-İzmir, 9. Ulusal Fen ve Matematik Eğitimi Kongresi.
- Janesick, V. J. (2001). *The assessment debate: A reference handbook*. Santa Barbara, CA: ABC-CLIO.
- Karasar, N. (1995). *Araştırmalarda rapor hazırlama*. Ankara: Nobel Kitap Yayını.
- Karatepe, A., Yıldırım, H. İ., Şensoy, Ö. ve Yalçın, N. (2004). Fen bilgisi öğretimi amaçlarının gerçekleştirilmesinde yeni programın içerik boyutunda uygunluğu konusunda öğretmen görüşleri. *Kastamonu Eğitim Dergisi*, 12(2), 327-338.
- Kalaian, H. A., & Freeman, D. J. (1994). Gender differences in self-confidence and educational beliefs among secondary teacher candidates. *Teaching and Teacher Education*, 10(6), 647-658.
- Kiremit, H. Ö. (2006). *Fen bilgisi öğretmenliği öğrencilerinin biyoloji ile ilgili öz-yeterlik inançlarının karşılaştırılması*. Yayımlanmamış doktora tezi. İzmir: Dokuz Eylül Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Lee, S., A., Park, H., S., & Choi, Y. (2012). The relationship between communication climate and elementary teachers' beliefs about alternative assessment. *Journal of Pacific Rim Psychology*, 5(01), 11-18.
- Mertler, C.A. (2003). *Preservice versus inservice teachers' assessment literacy: Does classroom experience make a difference?* Paper presented at the meeting of the Mid-Western Educational Research Association, Columbus, OH.
- Mertler, C.A. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education*, 33, 49–64.
- Mertler, C. A., & Campbell, C. (2005). *Measuring teachers' knowledge and application of classroom assessment concepts: Development of the Assessment Literacy Inventory*. Paper presented at the meeting of the American Educational Research Association, Montreal, Canada.
- Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of Curriculum Studies*, 19, 317-328.
- Özdemir, S. M. (2008). Sınıf Öğretmenleri adaylarının öğretim sürecine ilişkin öz-yeterlilik inançlarının çeşitli değişkenler açısından incelenmesi. *Kuram ve Uygulamada Eğitim Yönetimi*, 54, 277-306.
- Pajares, M. F. (1992). Teacher beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62, 307-332.
- Popham, W. J. (1995). *Classroom assessment*. Needham Heights, MA: Allyn & Bacon.
- Rios, F. (1996). *Teacher thinking in cultural contexts*. New York: State University of New York Press.
- Seferoglu, S. S. (2007). Professional teaching standards: The case of Turkish teacher candidates. *World Applied Sciences Journal*, 2(4), s. 412-419.

- Stiggins, R. J. (1994). *Student-centered classroom assessment*. New York, NY: Macmillan.
- Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77, 238–245.
- Stiggins, R. J., & Conklin, N. F. (1988). *Teacher training in assessment*. Portland, OR: Northwest Regional Educational Laboratory.
- Stiggins, R. J., & Conklin, N.F. (1992). *In teachers' hands: Investigating the practices of classroom assessment*. Albany: State University of New York Press.
- Svinicki, M. D. (2004). Authentic assessment: Testing in reality. *New Directions for Teaching and Learning*, 100, 23–29.
- Şahin, Ç. (2005). Öğretmenlik uygulaması dersinde öğrenme-öğretme sürecinin cinsiyet açısından değerlendirilmesi. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 5(2).
- Tittle, C.K. (1994). Toward an educational psychology of assessment for teaching and learning: Theories, contexts, and validation arguments. *Educational Psychologist*, 29(3), 149–162.
- Uğurlu, R. ve Akkoç, H. (2011) Matematik öğretmen adaylarının ölçme-değerlendirme bilgilerinin gelişiminin tamamlayıcı-şekillendirici ölçme-değerlendirme bağlamında incelenmesi. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*. 30(2), 155-167.
- Wishon, P., Crabtree, K., & Jones, M. (1998). *Curriculum for the primary years: An integrative approach*. Upper Saddle River, NJ: Merrill, An Imprint of Prentice Hall.
- Woolfolk, A. (2001). *Educational psychology*. Needham Heights, MA: Allyn Bacon.

Extended Abstract

Assessment can be defined roughly as a process of gathering information about what students have learned (Chen, 2003; Wishon et al. 1998). In order to collect evidence about student progress or achievement, teachers need to use various methods, strategies or tools. That is because using only one assessment method is not sufficient for gaining a good understanding of the learning process. For that reason, using different assessment methods for different teaching approaches could be more efficient to determine the learning outcomes (Brady 1997). Traditional assessment approach has been criticized heavily by many researchers because of turning students into passive learners who only memorize answers (rote learners) (Hart 1994). Starting with 1980s, teachers began to look for better assessment tools that encourage students to explain, to apply, to critique and to justify their own responses (Janesick 2001). One of the most important developments occurred recently in educational assessment is alternative assessment approach built on the central premises of the cognitive theory. This assessment approach is considered superior by many researchers in providing the richer and more realistic information about student progress in the learning process. Furthermore, alternative assessment ensures the active participation of the students to the assessment process and helps them gain a deeper understanding of the academic content (Stiggins 1994; Svinicki 2004). While traditional assessment based on the behaviorist theory mostly targets the assessment of the learning outcomes, alternative assessment based on the cognitive theory focuses chiefly on the assessment of the learning process. Different assessment techniques used in alternative assessment approach creates a suitable environment for meaningful learning to flourish.

Many teachers consider that alternative assessment is generally more time consuming and labour intensive than traditional assessment. Therefore, the widespread use of alternative assessment by teachers depends on their beliefs in its effectiveness (Lee et al. 2012). For that reason, assessment beliefs held by classroom teachers play a crucial role in their teaching practices.

Teachers use both formative and summative assessment approaches in their classrooms frequently. Formative assessment has emerged in 1930s and 1940s with the major influence of the cognitive and constructivist theories (Roos & Hamilton 2004). Whereas formative assessment focuses primarily on the progress made by the students in the learning process, summative assessment aims to uncover the learning outcomes. Moreover, formative assessment targets not only students' learning but also teachers' improvement in the learning process.

Beliefs are personal dispositions. Beliefs held by individuals are shaped by personal experiences gained in a lifetime and interpretations of events encountered in daily routines (Al-Sharafi 1998; Hsieh, 2002). Beliefs owned by individuals are transformed into attitudes, these attitudes affect intentions, and the intentions affect decisions that lead to actions (Bauch 1984). The concept of the belief system discussed largely in education literature governs teachers' behaviors in their classrooms. Since beliefs influence teachers' classroom practices directly, reaching success in educational reforms requires important changes to occur first in teachers' beliefs. That is why investigating teachers' beliefs starting from their university education is a must.

Teachers spend a large amount of their professional time for classroom assessment activities (Stiggins & Conklin 1988, 1992). In addition to sophisticated beliefs and positive attitudes toward assessment, using assessment effectively in classroom requires teachers to have a sound knowledge and necessary skills, namely assessment literacy, in this area. Therefore, developing an understanding of preservice teachers' beliefs, perceptions and thoughts about classroom assessment is very crucial for improving the teaching and learning process in the schools.

This study aims to reveal the assessment beliefs of preservice elementary teachers. To achieve that, a Likert-type scale developed by Genc (2005) was applied to 264 preservice elementary teachers in Canakkale Onsekiz Mart University. The instrument comprised of four subdimensions (formative, summative, traditional and alternative assessment) used to determine preservice elementary teachers' beliefs about classroom assessment. Factor analysis was applied to the observed data to find out if factor loadings represented the hidden four subdimensions. The results indicated that the factor loading values represented the four subdimensions of the hidden construct. Some of the factor loadings lower than 0.5 and conflicted more than one subdimensions was excluded from the observed data. Cronbach Alpha Reliability Coefficient was calculated as 0.868 for overall scale. For individual subdimensions, reliability coefficient was found 0.765 for the formative assessment, 0.702 for the summative assessment, 0.663 for the traditional assessment, and 0.822 for the alternative assessment subdimension.

The data were analyzed using both parametric (t-test) and non-parametric (Kruskal Wallis ve Mann Whitney-U testi) statistical tests. The results of the study showed that preservice elementary teachers in general did not have sophisticated beliefs about classroom assessment. Based on the four subdimensions (formative, summative, traditional and alternative assessment) included in the scale, assessment beliefs of the preservice elementary teachers were investigated with respect to the following variables: gender, GPA (grade point average) and type of the secondary school graduated. As opposed to school type variable, the analysis of the data produced a statistically significant result on gender and GPA variables. The results revealed that elementary teacher candidates received the highest score on formative assessment subdimension and the lowest score on alternative assessment subdimension. Analysis of the correlations among subdimensions in the scale indicated that the highest correlations was found between formative-alternative assessment subdimensions ($r=0.494$) and summative-traditional assessment subdimensions ($r=0.476$). Finding the highest correlations between these pairs of subdimensions is consistent since summative and traditional assessment techniques are both closer to the assumptions of the behaviorist approach whereas formative and alternative techniques are both more inclined to the premises of the constructivist approach.

The results of the study suggested that the assessment beliefs held by preservice elementary teachers were limited in their sophistication these results found in the study were discussed further from different perspectives and some suggestions were made to improve preservice teachers' perceptions, thoughts, and beliefs about assessment.

Kaynakça Bilgisi:

Şahin, Ç. ve Karaman, P. (2013). Sınıf öğretmeni adaylarının ölçme ve değerlendirmeye ilişkin inançları. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi [Hacettepe University Journal of Education]*, 28(2), 394-407.

Citation Information:

Şahin, Ç., & Karaman, P. (2013). Preservice elementary teachers' assessment beliefs [in Turkish]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi [Hacettepe University Journal of Education]*, 28(2), 394-407.