

## **Eđitim Programlarını Deęerlendirme Ölçeęi: Geęerlik ve Güvenirlik Çalışması**

**Gökhan BAŞ<sup>1</sup>**

<sup>1</sup> Yrd. Doç. Dr., Nięde Üniversitesi Eđitim Fakóltesi, Nięde, Türkiye.  
gokhanbas51@gmail.com

### **Özet**

Bu çalışmanın temel amacı, eđitim programlarını deęerlendirebilmek için kullanılabilen olan geęerli ve güvenilir bir ölçme aracı geliřtirmektir. Arařtırmada geliřtirilen ölçek, ilköęretim okulları ve liselerde görev yapmakta olan öęretmenler ( $n = 200$ ) üzerinde uygulanmıştır. Ölçeęin geęerlięini belirlemek üzere, aımlayıcı faktör analizi ile güvenilirlik analizi yapılmıştır. Aımlayıcı faktör analizi sonuçlarına göre geliřtirilen ölçek toplam dört boyuttan ve 35 maddeden oluşmuştur. Geliřtirilen ölçek; hedefler, içerik, öęretme-öęrenme süreci ve ölçme-deęerlendirme boyutlarından meydana gelmiştir. Arařtırmada ölçeęin Cronbach Alpha güvenilirlik katsayısı 0.92, Spearman-Brown iki yarı testi sonucu ise 0.90 olarak hesaplanmıştır.

**Anahtar Kelimeler:** Eđitim Programı, Program Deęerlendirme, Öęretmenler, Ölçek Geliřtirme Çalışması.

### **Curriculum Evaluation Scale: Validity and Reliability Study**

#### **Abstract**

The purpose of the study was to develop a valid and reliable scale in order to evaluate curriculum. The research was carried out on elementary and high school teachers ( $n = 200$ ). In order to determine the validity of the

scale the exploratory and confirmatory factor analyses were conducted. According to the findings of the exploratory factor analysis, the scale consisted of 35 items and four sub-dimensions. The scale was consisted of aims, content, teaching-learning process, and measurement-evaluation sub-dimensions. The Cronbach's Alpha value of the total scale was found as 0.92 and the Spearman-Brown split half correlation of the scale was calculated as 0.90 in the study.

**Keywords:** Curriculum, Curriculum Evaluation, Teachers, Scale Development Study.

## 1. Giriş

Eğitim programı kavramı üzerindeki düşünceler yüzyıllar öncesinden başlamıştır (Ornstein ve Hunkins, 1993). Eğitim programı kavramı, 1820'den beri kullanılmasına karşın, Franklin Bobbitt'in 1918 yılında yazmış olduğu "Eğitim Programı" (The Curriculum) isimli kitabı, eğitim programı alanının başlangıcı olarak kabul edilmektedir (Korkmaz, 2007). Diğer bir başka görüşe göre, eğitim programı kavramının kullanılmasının M.Ö. I. yüzyıla kadar uzandığı belirtilmektedir. Julius Caesar ve askerlerinin Roma'da yarış arabalarının üzerinde yarıştığı oval biçimdeki koşu pisti olarak kullanılan bu kavram, koşu pisti olarak bilinen somut bir kavramdan, bugün eğitim programı anlamında kullanılan soyut bir kavrama doğru geçişi sağlamıştır. Bu süreçte, eğitim programı "izlenen yol" anlamında eğitimde de kullanılmaya başlanmıştır (Demirel, 2005). Bu açıdan, kimi yazarlar eğitim programını izlemekten yola çıkarak "izlençe" olarak, kimi yazarlar da yetiştirmek kavramından yola çıkarak "yetişek" olarak nitelendirmektedirler (Ertürk, 1972; Sönmez, 2007). Ancak, günümüze kadar "eğitim programı" kavramı daha yaygın bir kullanım olarak devam ede gelmiştir. Eğitim programı kavramı kapsamlı ve çok boyutlu

olduğu için tanımında güçlükler ve çeşitlilikler ile karşılaşmaktadır. Bu nedenle ilgili alanyazında eğitim programının birbirinden farklı tanımlarına rastlanılmaktadır. Bu tanımların çoğunlukla yazarların eğitim anlayışlarına, felsefelerine ve planlı eğitimin hangi boyutlarda olabileceği görüşlerine bağlı olarak değişmekte oldukları görülmektedir (Erden, 2000).

Okulun ve öğretmenin yol göstericiliği altında, okulun içi ve dışındaki bütün öğrenme ve öğretim faaliyetlerini içeren (Varış, 1978) eğitim programı kavramını, Saylor, Alexander ve Lewis (1981) eğitilecek olan bireylere öğrenme yaşantıları kazandırma planı olarak tanımlarken, Taba (1962) ise onu bir öğrenme planı olarak ele alarak tanımlamıştır. Ertürk (1972), eğitim programını yetişek olarak tanımlamakta ve yetişegi düzenli öğrenme yaşantıları bütünü olarak tanımlamaktadır. Varış'a (1978) göre eğitim programı, bir eğitim kurumunun, çocuklar, gençler ve yetişkinler için sağladığı, Milli Eğitimin ve kurumun amaçlarının gerçekleşmesine dönük tüm faaliyetlerdir. Günümüz eğitim programcılarında Demirel'e (2005) göre ise eğitim programı, öğrenene okulda ve okul dışında planlanmış etkinlikler yoluyla sağlanan öğrenme yaşantıları düzeneğidir. Görüldüğü gibi, eğitim programı kapsayıcı ve okulda yürütülen tüm eğitsel etkinliklerin tamamıdır. Eğitim programı bu haliyle, bir eğitim basamağında bir dersin öğretimiyle ilgili olan öğretim programını ve ders programını da kapsamaktadır. Öğretim programı, bir dersin okulda ve okul dışında öğretimi ile ilgili tüm etkinliklerin düzenlendiği yapı olarak, ders programı da bir ders süresi içinde planlanan hedeflerin bireye nasıl kazandırılacağını gösteren tüm etkinliklerin yer aldığı bir plan olarak tanımlanabilmektedir (Demirel, 2005). Benzer bir şekilde öğretim programı, belli bir öğretim basamağında çeşitli sınıf ve derslerde okutulacak konuları, bunların amaçlarını, her dersin sınıflara göre haftada kaç saat okutulacağını ve öğretim metotlarını, tekniklerini gösteren kılavuzdur (Büyükkaragöz, 1997).

Eğitim programları, ulusal ya da uluslararası düzeyde kaliteli bir eğitim sistemi kurma, ülkenin kalkınmasını ve gelişmesini sağlayacak nitelikli insan gücü yetiştirme, toplumsal ve kültürel değerlerin korunmasını ve geliştirilmesini destekleme gibi amaçlara yönelik olarak geliştirilirler (Özdemir, 2009). Yani, eğitim sistemleri öğretim programları ile işlerlik kazanmaktadırlar (Erden, 2000). Varış (1978), eğitim programını operasyonel bir kavram olarak yorumlamakta ve programların sürekli olarak geliştirilmesi gerektiğini dile getirmektedir. Bu bağlamda, program geliştirme ve değerlendirme faaliyetlerinin odak noktasını program tasarımının teşkil ettiği söylenebilir. Program tasarımı ilkelerine uygun olarak hazırlanan programlar, uygulamada işlerlik kazanmaktadırlar. Eğitim programı tasarımı hazırlanırken, birinci aşamada alınacak kararlara temel teşkil edecek veriler konu alanı, toplum ve bireyin ihtiyaçlarının analizi ile elde edilmektedir. Bu veri toplama kaynaklarından hangisinin öncelik taşıdığı, tasarımın dayandığı eğitim felsefesi ve modeline göre değişiklik göstermektedir (Erden, 2000).

Eğitim programı tasarısı, uygulamaya esas olan bir kitap ya da kılavuz kaynaktır. Program tasarısı ilkelerine uygun olarak hazırlanan programlar uygulamada işlerlik kazanırlar. Program geliştirme ve değerlendirme faaliyetinin odak noktası tasarı olmakla birlikte uygulama göz önünde bulundurulmaksızın gerçekçi bir değerlendirme yapmak mümkün değildir (Demirel, 2005). Diğer yandan, programların uygulayıcıları okul yöneticileri ve öğretmenlerdir. Uygulama sırasında çeşitli nedenlerle tasarımın olduğu gibi uygulanması mümkün olmayabilir ya da tasarımın hazırlanması sırasında göz önünde bulundurulmayan bazı faktörler tasarımın olduğu gibi uygulanmasını engelleyebilir. Bu nedenlerden ötürü programın etkililiği hakkında yargıda bulunabilmek için programın uygulanması sürecini de gözlemek ve uygulama sırasında da bilgi toplamak gerekir

(Özçelik, 2010). Hazırlanan ve uygulamaya konulan programların değerlendirilmesi ise, gözlem ve çeşitli ölçme araçları ile eğitim programlarının etkililiği hakkında veri toplama, elde edilen verileri programın etkililiğinin işaretçileri olan ölçütlerle karşılaştırıp yorumlama ve programın etkililiği hakkında karar verme süreci ile mümkün olabilmektedir. Bu bağlamda, uygulanan programların yetersizlikleri giderildikçe, yani programlar geliştirildikçe eğitimin niteliğinin artması beklenmektedir. Dolayısıyla, eğitim programları gelişigüzel bir biçimde asla geliştirilemezler. Zira, programları daha etkili hale getirecek doğru kararların alınabilmesi, bu kararların bilimsel çalışmalarla araştırılmasına ve uygulamaların değerlendirilmesine bağlıdır (Erden, 2000).

Program değerlendirme, kısaca, programın sağlamlığına karar verme işi olarak tanımlanmaktadır (Baykul, 2000). Geliştirilen bir eğitim programının sağlamlığını, etkililiğini, öğrenme düzeyini ve öğrenci başarısını görebilmek için çeşitli değerlendirmeler yapılmaktadır (Yüksel ve Sağlam, 2012). Program değerlendirmede, ölçülmek istenen özelliğe ilişkin kararlar almak için bilgi toplanmakta ve bu bilgiler kullanılarak programın etkililiği hakkında bir sonuca ulaşılmaktadır (Worthen, Sanders ve Fitzpatrick, 1997). Eğitim programının etkililiği hakkında bilgi edinebilmek için öğretim programlarının mutlaka değerlendirmeye tabi tutulması gerekmektedir (Sönmez ve Alacapınar, 2015). Bu nedenle programın hedeflerine ulaşılmasını engelleyen ya da yetersiz kalan öğelerin saptanması ve gerekli düzeltmelerin yapılabilmesi için değerlendirme çalışmasının yapılması gerekmektedir (Uşun, 2012).

İlgili alanyazında eğitim programlarının değerlendirildiği pek çok çalışmanın yapıldığı görülmekle birlikte, bu programların değerlendirilmesinde kullanılacak olan ölçüklerin sayısının oldukça

sınırlı olduğu tespit edilmiştir (örn., Acar, Bıçak, Tutkun ve Pektaş, 2009; Cusimano ve Garg, 1996; Fraser, 1977; Göle ve Temel, 2015). Bu çalışmalardan bazılarının doğrudan tıp eğitiminde kullanılabilecek bir şekilde geliştirildiği (örn., Cusimano ve Garg, 1996), bazılarının eski tarihli olduğu (Fraser, 1977), bazılarının ise kongre veya sempozyum bildirisi (örn., Acar, Bıçak, Tutkun ve Pektaş, 2009; Göle ve Temel, 2015) olarak sunulup, bu ölçeklerin henüz yayımlanarak araştırmacılarla paylaşılmadıkları anlaşılmıştır. Aynı zamanda, ilgili alanyazında ders öğretim programlarının değerlendirilmesine dönük olarak geliştirilen ölçeklerin de olduğu görülmekle birlikte (örn., Adıgüzel ve Özüdoğru, 2014; Gömleksiz ve Bulut, 2006a, 2006b; Öksüz, 2015), bu ölçeklerin eğitim programlarını genel olarak değerlendirme noktasından uzak oldukları, dolayısıyla belli bir alana özgü olarak geliştirildikleri görülmüştür. Yapılan çalışmada ise eğitim programlarının genel anlamda değerlendirilmesine yönelik bir ölçek geliştirilmiş, bu ölçek ile tüm eğitim programlarının belli bir ders alanına bağlı kalmaksızın değerlendirilmesi mümkün kılınmıştır. Bunun yanı sıra, geliştirilen ölçeğin bir eğitim programının tüm boyutlarına odaklanarak, programın tüm boyutlarının değerlendirilmesine imkân tanımakta; bu da, programın boyutları bağlamında bir sonuca varılmasını kolaylaştırmaktadır. Sonuçta, gerek alanyazında eğitim programlarının değerlendirilmesine yönelik olarak kullanılabilecek ölçeklerin oldukça az sayıda olması, gerekse alanyazındaki çoğu ölçeğin ders öğretim programı bağlamında geliştirilmiş olmasının böylesi bir çalışmaya olan ihtiyacı artırdığı düşünülmektedir.

### **1.1. Araştırmanın Amacı**

Yapılan bu araştırmanın amacının, ülkemizde uygulanan eğitim programlarını değerlendirmede kullanılabilecek olan geçerli ve güvenilir bir ölçme aracı geliştirmek olduğu söylenebilir. Eğitim programlarının değerlendirilmesine hizmet etmesi beklenen bu çalışmanın, eğitim

politikacılarına, öğretmenlere, okul yöneticilerine ve program geliştirme uzmanlarına kullanabilecekleri geçerli ve güvenilir bir ölçek sunması ve bu konuda yapılacak olan diğer araştırmalara da ışık tutması beklenmektedir.

## 2. Yöntem

Yapılan bu araştırma, bir ölçek geliştirme çalışmasıdır. Yapılan bu çalışma, öğretmenlerin eğitim programlarına ilişkin görüşlerini tespit edebilmek amacıyla geliştirilmiş olan “Eğitim Programlarını Değerlendirme Ölçeği”nin geçerlik ve güvenilirliğini test etmek amacıyla yapılmıştır.

### 2.1. Çalışma Grubu

Araştırmanın çalışma grubunu, ilköğretim ve lise öğretmenleri oluşturmaktadır. Araştırmada geliştirilen ölçek, Niğde ili merkez ilçesinde devlete bağlı ilköğretim okulları ve liselerde görev yapmakta öğretmenler ( $n = 200$ ) üzerinde uygulanmıştır. Araştırmaya katılan öğretmenlerin demografik nitelikleri incelendiğinde, bu öğretmenlerin; %53'ünün ( $n = 106$ ) kadın, %47'sinin ( $n = 94$ ) ise erkek, %11'inin ( $n = 22$ ) 1-5 yıllık, %16'sının ( $n = 32$ ) 6-10 yıllık, %26.5'inin ( $n = 53$ ) 11-15 yıllık, %31'inin ( $n = 62$ ) 16-20 yıllık ve %15.5'inin ( $n = 31$ ) ise 20 yıl ve üzeri mesleki kıdeme sahip olduğu, aynı zamanda %12.5'inin ( $n = 25$ ) yüksekokul/eğitim enstitüsü, %80.5'inin ( $n = 161$ ) lisans ve %7'sinin ( $n = 14$ ) ise lisansüstü eğitim mezunu olduğu görülmüştür.

### 2.2. Ölçeğin Geliştirilmesi

Öncelikle, ilgili alanyazın (Tyler, 1950; Taba, 1962; Ertürk, 1972; Varış, 1978; Saylor, Alexander ve Lewis, 1981; Doll, 1986; Olivia, 1988; Ornstein ve Hunkins, 1993; Wiles ve Bondi, 1993; Tanner ve Tanner, 1995; Büyükkaragöz, 1997; Erden, 2000; Demirel, 2005; Henson, 2006 McNeil, 2006; Marsh ve Willis, 2007; Korkmaz, 2007; Özdemir, 2007; Çubukçu,

2008) taranarak, alanyazındaki eğitim programı değerlendirme yaklaşımlarına dayalı olarak deneme amaçlı tutum maddeleri yazılmıştır. Yazılan maddelerden oluşan toplam 35 maddelik bir havuz oluşturulmuştur. Daha sonra, oluşturulan ölçek maddelerinin kapsam geçerliğinin sağlanabilmesi için alan uzmanları tarafından incelenmiştir. Ölçeğin kapsam geçerliliği bağlamında öncelikle uzmanların görüşlerine başvurulmuş ölçekte yer alan maddelerin uygunluluk düzeyleri tespit edilmeye çalışılmıştır. Uzman görüşleri ışığında, havuzdaki 35 maddenin tümünün eğitim programlarını değerlendirmeye ilişkin oldukları belirlenmiş ve geri bildirimler ışığında yanlış anlamaları giderecek şekilde ölçek maddeleri üzerinde gerekli düzeltmeler yapılmıştır. Bu haliyle ölçeğin ön uygulama formu 35 maddeden oluşmuştur. Ölçekteki maddeler; “kesinlikle katılmıyorum”, “katılıyorum”, “kararsızım”, “katılıyorum” ve “kesinlikle katılıyorum” şeklinde belirtilen beşli dereceleme ölçeğinde düzenlenmiştir. Ölçekteki maddeler için en olumsuz 1, en olumlu 5 olmak üzere düzenlenmiştir. Ölçeğin yapı geçerliğini sağlayabilmek amacıyla temel bileşenler analizine dayalı olarak açımlayıcı faktör analizi (Tabaschinck ve Fidell, 2001; Çokluk, Şekercioğlu ve Büyüköztürk, 2010) uygulanmıştır.

Ölçeğin geçerlik ve güvenirlik çalışmalarının ilk aşamasında açıklayıcı faktör analizi çalışmasına yer verilmiştir. Açımlayıcı faktör analizinin değerlendirilmesinde Kaiser-Meyer-Olkin (KMO) katsayısı ve Bartlett küresellik testi ile incelenmiştir. İlgili alanyazında, KMO değerinin 0.60'dan yüksek olması dağılımın faktör analizi için yeterli olduğu ve Bartlett küresellik testinin anlamlı çıkmasının verilerden anlamlı faktör çıkarılabileceğinin bir göstergesi olduğu belirtilmektedir (Reuterberg ve Gustafsson, 1992; Kline, 1994; Tabaschinck ve Fidell, 2001). KMO değerinin 0.80-0.90 arasında çıkması çok iyi olarak değerlendirilirken, 0.90 ve üzerinde çıkması mükemmel olarak kabul edilmektedir. Ayrıca, faktör



analizinde özdeğeri 1 ve 1'den daha büyük olan faktörler önemli faktörler olarak kabul edilmektedir (Murphy ve Davidshofer, 1991; Thompson, 2004). İlgili alanyazında, faktör örüntüsünün oluşturulmasında 0.30 ile 0.40 arasında değişen faktör yüklerinin alt kesme noktası olarak alınabileceği belirtilmektedir (Şencan, 2005; Çokluk, Şekercioğlu ve Büyüköztürk, 2010). Alanyazında, alt faktör yük değeri sınırının 0.30 ve 0.40 arasında olması gerektiğine ilişkin görüşler bulunmasına karşın, Tabaschinck ve Fidell (2001), bu değer en asgari 0.32 olması gerektiğine dikkat çekmektedir. Ancak Ferguson ve Takane (1989), faktör örüntüsünün oluşturulabilmesi için 0.40'ın alt kesme noktası olarak alınması gerektiğini belirtmektedir. Bu sebeple, bu çalışmada alt kesme noktası olarak 0.40 kabul edilmiştir. Diğer taraftan, bu çalışmada faktör sayısına her hangi bir sınırlama getirilmemiş ve özdeğeri 1.00'dan büyük olan faktörler ölçeğe alınmıştır. Alanyazında, faktör analizinde özdeğeri 1 ya da 1'den daha büyük olan faktörler önemli faktörler olarak kabul edilmektedir (Büyüköztürk, 2007). Bu çalışmada özdeğer 1.00 olarak alınmıştır. Aynı zamanda, çalışmada ölçeğin iç tutarlılığını test edebilmek amacıyla Cronbach Alpha güvenilirlik katsayısı ile Spearman-Brown iki yarı testi hesaplanarak, ölçeğin geçerlik ve güvenilirlik çalışmaları bitirilmiştir.

### **2.3. Verilerin Çözümlemesi**

Ölçeğin geçerlilik ve güvenilirlik analizleri kapsamında öncelikle açımlayıcı faktör analizi ve madde analizi çalışmalarına yer verilmiştir. Araştırmanın açımlayıcı faktör analizi ve madde analizi çalışmaları SPSS 20.0 paket programıyla gerçekleştirilmiştir. Araştırma kapsamında kullanılan tüm istatistiksel işlemlerde anlamlılık düzeyi 0.05 olarak kabul edilmiştir.

### 3. Bulgular

Araştırmanın bu kısmında, elde edilen bulgular; açımlayıcı faktör analizine ilişkin bulgular ve güvenilirliğe ilişkin bulgular başlıkları altında incelenecektir.

#### 3.1. Açımlayıcı Faktör Analizine İlişkin Bulgular

Oluşturulan deneme ölçeği seçilen öğretmen grubuna uygulandıktan sonra, eğitim programlarını değerlendirme ölçeğinin yapı geçerliğinin belirlenmesi için açımlayıcı faktör analizi kullanılmıştır. Yapılan açımlayıcı faktör analizi sonunda, madde toplam korelasyon değeri 0.40 ve üzerindeki maddeler ölçekte bırakılmıştır. Yapılan analizde, ölçekteki 35 maddenin de faktör yük değerlerinin 0.40'ın üzerinde olduğu tespit edilmiştir. Ölçeğin, 0.40 ve üzerinde olan tüm maddelerine ikinci bir faktör analizi daha uygulanmış ve bu kez ölçeğin KMO değeri ve Bartlett küresellik testi sonucu bulunmuştur. Yapılan analiz sonucunda, ölçeğin KMO değerinin 0.940 olduğu saptanmıştır. KMO değeri, dağılımın faktör analizi için yeterli olup olmadığını test etmek amacıyla kullanılmakta olup, 0.80 ve 0.90 aralığı çok iyi, 0.90 ve üzeri değerler ise mükemmel olarak nitelendirilmektedir (Kline, 1994; Büyüköztürk, 2007). Ayrıca, KMO değerinin 1'e yakın bir değer alması halinde çalışma grubunun sayısının yeterli olduğu kanısına varılmaktadır (Murphy ve Davidshofer, 1991; Kline, 1994). Bunun yanında, hesaplanan Bartlett küresellik testi sonucuna göre ölçeğin bu değeri ( $\Delta\chi^2 = 3710.726/sd = 595$ ) olarak bulunmuştur. Bartlett küresellik testinin anlamlı çıkması ölçüm yapılan değişkenin evren parametresinde çok değişkenli olduğunu göstermektedir (Thompson, 2004). Elde edilen Bartlett küresellik testi sonucunun anlamlı çıkması, faktör analizine geçilebileceğini göstermektedir (Tatlıdil, 2002). Sonuç olarak, bu çalışmada elde edilen KMO değeri mükemmel anlama gelen 0.940 şeklinde, Bartlett küresellik

testi sonucu da anlamlı olarak bulunmuştur. Çalışmada gerçekleştirilen bu analizlerin bir neticesi olarak açılımlı faktör analizinin yapılabileceğine karar verilmiştir. Faktör analizinde özdeğeri 1 ya da 1'den daha büyük olan faktörler önemli faktörler olarak kabul edilmektedir (Büyüköztürk, 2007). Bu araştırmada özdeğer 1.00 olarak alınmış ve bu değerden büyük dört boyut belirlenmiştir. Bu amaçla, ölçeğe uygulanan döndürülmüş bileşenler analizi sonuçlarına ilişkin olarak elde edilen boyutlar ve bu boyutlarda yer alan maddelere ilişkin faktör yükleri Tablo 1'de verilmiştir.

**Tablo 1. Eğitim Programlarını Değerlendirme Ölçeğindeki Maddelerin Faktör Yükleri**

Maddeler	Faktörler			
	Hedefler	İçerik	Öğrenme- Öğretme	Ölçme- Değerlendirme
M6.	.747			
M3.	.691			
M5.	.665			
M27.	.616			
M16.	.583			
M7.	.575			
M2.	.570			
M1.	.570			
M10.	.550			
M9.	.531			
M12.		.683		
M28.		.661		
M18.		.644		
M17.		.609		
M30.		.581		
M23.		.578		
M15.		.542		
M8.		.538		
M13.		.516		
M14.		.447		
M25.			.637	
M31.			.608	
M34.			.597	
M22.			.596	
M21.			.593	
M19.			.570	

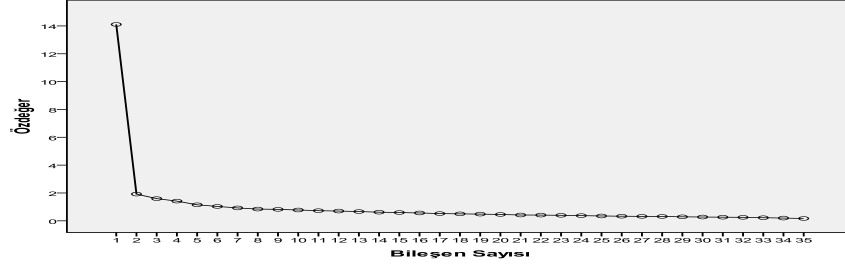
---

M24.	.554	
M33.	.484	
M26.	.457	
M35.	.414	
M11.		.787
M29.		.636
M4.		.525
M20.		.518
M32.		.468

---

*Not.* Açıklanan Varyans; Hedefler = %16.511; İçerik = %16.093; Öğrenme-Öğretme Süreci = %13.135; Ölçme-Değerlendirme = %8.651; Toplam = %54.390

Tablo 1'deki bulgular doğrultusunda ölçeğe ilişkin birinci faktörde (hedefler) yer alan maddelerin yüklerinin 0.747 ile 0.531 arasında değiştiği, ikinci faktördeki (içerik) maddelerin faktör yük değerlerinin 0.683 ile 0.447 arasında değiştiği, üçüncü faktördeki (öğrenme-öğretme süreci) maddelerin faktör yük değerlerinin 0.637 ile 0.414 arasında değiştiği ve dördüncü faktördeki (ölçme-değerlendirme) maddelerin faktör yük değerlerinin ise 0.787 ile 0.468 arasında değiştiği görülmektedir. Diğer taraftan, birinci faktörün açıkladığı varyans oranı %16.511, ikinci faktörün açıkladığı varyans oranı %16.093, üçüncü faktörün açıkladığı varyans oranı %13.135 ve dördüncü faktörün açıkladığı varyans oranı ise %8.651 olarak bulunmuştur. Ayrıca, ölçeğin toplam açıkladığı varyans oranı ise %54.390 olarak saptanmıştır. Faktör analizinde %40 ile %60 arasında değişen varyans oranları ideal olarak kabul edilmektedir (Kline, 1994). Eğitim programlarını değerlendirme ölçeğinin madde analiz çözümlemesi varimax faktör analiz tekniği ile yapılan döndürme işlemi sonunda, ölçeğin dört boyutlu olduğuna karar verilmiştir. Ancak, ortaya çıkan bu durumu daha net görmek amacıyla Cattel'in scree sınaması (Kline, 1994) yapılarak maksimum anlamlı faktör sayısı ile ilgili olarak Grafik 1 elde edilmiştir.



**Şekil 1. Scree Sınaması Grafiği**

Grafikte dikey eksen özdeğer miktarlarını, yatay eksen ise faktörleri göstermektedir. Grafik faktörlerin özdeğerleriyle eşleştirilmesi sonucunda bulunan noktaların birleştirilmesi ile elde edilir. Grafikte yüksek ivmeli, hızlı düşüşlerin yaşandığı faktör önemli faktör sayısını vermektedir. Yatay çizgiler faktörlerin getirdikleri ek varyansların katkılarının birbirine yakın olduğunu göstermektedir (Büyüköztürk, 2007; Çokluk, Şekercioğlu ve Büyüköztürk, 2010). Scree sınama grafiğinde (bkz. Şekil 1), grafik eğrisinin hızlı bir düşüş gösterdiği nokta dördüncü faktörün olduğu yerdir. Beşinci noktadan sonra eğitim bir plato yapmaktadır. Zira, bu noktadan sonraki faktörlerin varyansa yaptıkları katkı hem küçük, hem de yaklaşık olarak aynıdır (Fabrigar, Wegener, MacCallum ve Strahan, 1999). Bu sebeple geliştirilen ölçeğin dört faktörlü olduğu söylenebilir. Ayrıca belirlenen dört faktöre ilişkin korelasyon değerleri Tablo 2’de verilmiştir.

**Tablo 2. Eğitim Programlarını Değerlendirme Ölçeğinin Faktörleri Korelasyon Matrisi**

Faktörler	X	SS	1	2	3	4
Hedefler	4.17	7.77	-	.736**	.721**	.622**
İçerik	3.83	8.25	.736**	-	.753**	.673**
Öğretme-Öğrenme Süreci	3.94	8.21	.721**	.753**	-	.784**
Ölçme-Değerlendirme	3.71	4.62	.622**	.673**	.684**	-

Not. \*\*  $p < 0.01$

Tablo 2’de verilen eğitim programlarını değerlendirme ölçeğinin faktörleri arasındaki ilişkiye bakıldığında, ölçeğin faktörleri arasında yüksek düzeyde pozitif ve doğrusal bir ilişkinin olduğu görülmektedir. Dolayısıyla, ölçeğin faktörleri arasında pozitif düzeyde anlamda bir tutarlılığın olduğu anlaşılmaktadır.

### 3.2. Güvenirlğe İlişkin Bulgular

Ölçeğin iç geçerliliğinin tespiti için bağımsız gruplar *t* testi uygulanmıştır. Öncelikle ölçek puanları küçükten büyüğe doğru sıralanmış ve grubun %27’lik alt ve üst kısmı hesaplanmıştır. %27’lik alt ve üst dilimde kalan katılımcıların aldıkları toplam puanlar Tablo 3’de karşılaştırılmıştır.

**Tablo 3. Eğitim Programlarını Değerlendirme Ölçeğinin İç Geçerliliği**

Gruplar	n	X	SS	sd	t	p
Üst <sub>%27</sub>	54	4.66	5.42	109	-	0.000*
Alt <sub>%27</sub>	54	3.01	22.13		19.283	

Not. \*\*  $p < 0.01$

Tablo 3 incelendiğinde, madde toplam puanlarında üst grubun ortalamasının ( $X = 4.66$ ,  $SS = 5.42$ ), alt grubun ortalamasından ( $X = 3.01$ ,  $SS = 22.13$ ) daha yüksek olduğu görülmektedir. Bu anlamda, ortaya çıkan sonuç itibarıyla üst grup lehine anlamlı bir fark vardır ( $p < 0.05$ ) denilebilir. Bu durum, soruların kendi içinde ayırt ediciliğinin yüksek olduğunu ve iç geçerliliğe sahip olduğunu göstermektedir. Diğer taraftan, hazırlanan eğitim programlarını değerlendirme ölçeğinin güvenirlği iki yolla hesaplanmıştır. Bunlar; Cronbach Alpha katsayısı ve Spearman Brown iki yarı testidir. Bunlardan birincisi olan Cronbach Alpha iç tutarlılık katsayısı hesaplanmış ve ölçeğin güvenirlilik katsayısı 0.92 olarak bulunmuştur. İlgili alanyazında birden fazla derecelendirmeye sahip ölçeklerine ilişkin güvenirlilik

çalışmalarında 0.60 ile 0.70 düzeyindeki katsayıların yeterli olduğu ileri sürülmektedir (Cronbach, 1990). Bu açıdan, ölçeğin elde edilen güvenirlik katsayısı mükemmel olarak değerlendirilebilir. Bu nedenle ölçeğin güvenilir bir yapıya sahip olduğu belirlenmiştir. İkincisi ise testi yarılama yöntemiyle ölçeğin güvenirliği hesaplanmış, testin Spearman Brown iki yarı testi korelasyonu 0.90 olarak bulunmuştur. Bu katsayı ilgili alanyazında iyi kabul edilen değerler içerisinde. Alanyazında 0.80'in üzerindeki değerler iyi olarak nitelenmektedir (Büyüköztürk, 2002). Testi yarılama güvenirliği, bir testi iki kere uygulamanın veya bir testin iki eşdeğer formunun hazırlanmasının güç olduğu ve testin tek bir değişkeni ölçtüğü durumlarda başvurulması gereken bir yöntemdir (Özen, Gülaçtı ve Kandemir, 2006). Testi yarılama yöntemi ile elde edilen güvenirlik iç tutarlılık olarak da bilinmektedir. Ayrıca, geliştirilen eğitim programlarını değerlendirme ölçeğinin dört boyutuna ilişkin güvenirlik katsayıları ise Tablo 4'de sunulmaktadır.

**Tablo 4. Eğitim Programlarını Değerlendirme Ölçeği Faktörlerinin Güvenirlik Katsayıları**

Faktörler	Maddeler	Güvenirlik
Hedefler	1 2 3 5 6 7 9 10 16 27	0.89
İçerik	8 12 13 14 15 17 18 23 28 30	0.88
Öğretme-Öğrenme Süreci	19 21 22 24 25 26 31 33 34 35	0.87
Ölçme-Değerlendirme	4 11 20 29 32	0.78

Tablo 5'deki bulgulara bakıldığında, ölçeğe ait boyutların Cronbach Alpha güvenirlik katsayılarının 0.89 ile 0.78 arasında değiştiği görülmektedir. Güvenirlik analizinde Alpha değerinin en az 0.70 olması gerektiği (Cronbach, 1990; Kline, 1994) göz önünde bulundurulduğunda, ölçeğin tümünün yanı sıra her bir boyutunun da oldukça güvenilir değerlere sahip oldukları söylenebilir.

#### 4. Tartışma

Bu çalışmanın amacını, ülkemizde uygulanmakta olan eğitim programlarını değerlendirmede kullanılacak geçerli ve güvenilir bir ölçme aracı geliştirmek teşkil etmiştir. Toplam 35 maddeye sahip olan eğitim programlarını değerlendirme ölçeği ilköğretim ve lise öğretmenlerine uygulanmış ve yapılan açımlayıcı faktör analizi sonucunda ölçeğin dört boyutlu bir yapıya sahip olduğu saptanmıştır.

Ölçekteki maddelerden faktör yükü 0.40 ve bu değer üstünde olan maddeler işler durumda kabul edilerek analiz için seçilmiştir. Yapılan faktör analizi sonucunda, ölçeğin toplam 35 maddeden oluştuğu saptanmıştır. Bu araştırmada özdeğer 1.00 olarak alınmış ve özdeğeri bu değerden büyük olan toplam dört faktör belirlenmiştir. Yapılan çözümlenmeye göre, deneme ölçeğindeki maddelerin onu birinci faktörde, onu ikinci faktörde, onu üçüncü faktörde ve beşi de dördüncü faktörde toplanmıştır. Döndürülmüş temel bileşenler analizi sonuçlarına göre ölçeğin dört boyutlu bir yapıya sahip olduğu belirlenmiştir. Geliştirilen ölçeğin dört boyuttan meydana geldiği belirlenmiş olup, bu boyutlar; hedefler, içerik, öğretme-öğrenme süreci ve ölçme-değerlendirme biçiminde şekillenmiştir. Ölçeğin bütün olarak Cronbach Alpha güvenilirlik katsayısı 0.92 olarak hesaplanmıştır. Ölçeğin birinci boyutu olan “hedefler”in güvenilirlik katsayısı 0.89, ikinci boyutu olan “içerik”in güvenilirlik katsayısı 0.88, üçüncü boyutu olan “öğretme-öğrenme süreci”nin güvenilirlik katsayısı 0.87 ve dördüncü boyutu olan “ölçme-değerlendirme”nin güvenilirlik katsayısı ise 0.78 olarak hesaplanmıştır. Güvenirlik analizinde Alpha değerinin en az 0.70 olması gerektiği (Anderson, 1988; Tavşancıl, 2005) göz önünde bulundurulduğunda, ölçeğin tümünün yanı sıra, her bir alt boyutunun da güvenilir değerlere sahip olduğu ifade edilebilir. Bunun yanında, öğretmenlerin ölçeklerden aldıkları puanlar ile ölçeğin boyutları arasında pozitif ve doğrusal bir ilişkinin olduğu tespit



edilmiştir. Buradan hareketle, boyutlar arasında bir tutarlılığın olduğu söylenebilir. Yapılan analizler sonucunda elde edilen değerler itibari ile geliştirilen ölçeğin geçerli, güvenilir ve kuramsal temeli sağlam bir ölçek olduğu belirtilebilir.

İlgili alanyazın gözden geçirildiğinde, eğitim programlarının değerlendirilmesinde kullanılabilecek olan ölçeklerin var olduğu, ancak bu ölçeklerin sayısının oldukça sınırlı olduğu tespit edilmiştir (örn., Acar, Bıçak, Tutkun ve Pektaş, 2009; Cusimano ve Garg, 1996; Fraser, 1977; Göle ve Temel, 2015). İfade edilen ölçek geliştirme çalışmaları değerlendirildiğinde, bu çalışmalardan bazılarının doğrudan tıp eğitiminde kullanılabilecek bir şekilde geliştirildiği (örn., Cusimano ve Garg, 1996), bazılarının eski tarihli olduğu (Fraser, 1977), bazılarının ise kongre veya sempozyum bildirisi (örn., Acar, Bıçak, Tutkun ve Pektaş, 2009; Göle ve Temel, 2015) olarak sunulup, bu ölçeklerin henüz yayımlanarak araştırmacılarla paylaşılmadıkları anlaşılmıştır. Aynı zamanda, ilgili alanyazında rastlanılan ölçek geliştirme çalışmalarının çoğunluğunun Türkiye dışında geliştirildiği, bu yüzden bu ölçeklerin kültürel özelliklerinin dikkatlice incelenerek kullanılmasının uygun olacağı düşünülmekte olup, Türkiye bağlamında geliştirilen eğitim programlarını değerlendirme ölçeklerinin ise hem sayıda az, hem de bu çalışmaların kongre veya sempozyumlar dışında yayımlanarak ilgililerle henüz paylaşılmadıkları tespit edilmiştir. Dolayısıyla, Türkiye bağlamında geliştirilerek yayımlanarak, ülkemizde eğitim programlarını değerlendirmek için kullanılabilecek olan geçerliği ve güvenilirliği sağlanmış olan bir ölçeğin varlığından şu an için söz etmek pek olası gözükmemektedir. Sonuçta, yapılan bu ölçek geliştirme çalışmasının Türkiye’de program geliştirme alanındaki önemli bir boşluğu doldurduğu ve bu konuda yapılacak gelecek çalışmalara ışık tutabileceği düşünülmektedir. Her ne kadar eğitim programlarının değerlendirilmesinde

kullanılabilecek olan ölçeklerin var olduğu, ancak bu ölçeklerin sayısının oldukça sınırlı olduğu tespit edilse de, bu çalışmaların hali hazırdaki ölçek geliştirme ile karşılaştırılması oldukça anlamlıdır. Örneğin, Cusimano ve Garg (1996) tarafından geliştirilen ölçek sağlıkla ilgili mesleklerdeki eğitim programlarının değerlendirilmesinde kullanılmakta olup, ilgili ölçek hedeflerin hazırlanmasında yardımcı olan eğitim felsefeleri, öğrenme deneyimlerinin seçimi ve organizasyonu ve programın değerlendirilmesi süreci olan toplam üç boyuttan meydana gelmektedir. Acar, Bıçak, Tutkun ve Pektaş (2009) tarafından hazırlanan ölçek ise eğitim programlarını değerlendirebilmek için geliştirilmiştir. Üniversitede eğitim fakültelerinde öğrenim görmekte olan ilköğretim bölümü öğretmen adayları örnekleminde gerçekleştirilen eğitim programlarını değerlendirme ölçeği, bu araştırmadaki benzer bir biçimde, toplam dört boyuta sahip olmakta olup, bu boyutlar; kazanımlar, öğrenme-öğrenme süreçleri, içerik ve değerlendirme boyutlarından meydana gelmektedir. Toplam 27 maddeye sahip olan ölçeğin Cronbach Alpha güvenirlik katsayısı 0.92 olarak hesaplanmıştır. Acar, Bıçak, Tutkun ve Pektaş (2009) tarafından geliştirilen ölçek pek çok açıdan hali hazırdaki ölçek geliştirme çalışmasına benzemekte olup, her iki ölçekte de programın öğelerine dönük değerlendirmenin esas alındığı sonucuna ulaşılmaktadır. Her ne kadar her iki ölçek de pek çok açıdan birbirine benzese de, özellikle ölçeklerin geliştirilmesinde dikkate alınan çalışma grupları birbirinden tamamıyla olmasa da farklılık göstermektedir. Acar, Bıçak, Tutkun ve Pektaş (2009) tarafından geliştirilen ölçek ilköğretim bölümü öğretmen adayları çalışma grubu dâhilinde geliştiriliş olmasına rağmen, hali hazırda geliştirilen ölçek ise doğrudan görev yapmakta olan öğretmenler üzerinde gerçekleştirilmiştir. Bunun yanı sıra, hali hazırdaki ölçek gerek ilköğretim, gerekse lise düzeyinde görev yapmakta olan öğretmenler üzerinde gerçekleştirilmiş olup, eğitim programlarının değerlendirilmesinde kullanılabilecek olan bu ölçeğin doğrudan eğitim

programları ile sürekli ilgili olan ve onu uygulamada kullanan öğretmenler üzerinde uygulanmış olması oldukça önemli ve anlamlıdır. Zira, eğitim programları ile sürekli ilgili olup, onları uygulamaya yansıtan öğretmenler ile eğitim fakültelerinde öğrenim görmekte olan öğretmen adaylarının eğitim programının değerlendirilmesine yönelik sahip oldukları bilgi, tutum ve algı birbirinden farklılık gösterebilir. Dolayısıyla, eğitim programlarının onlara şu an için uzak bir grup olan öğretmen adaylarınca değil, onları bizzat uygulamada kullanan öğretmenler tarafından değerlendirilmesi oldukça manidardır. Aynı zamanda, öğretmen adayları üzerinde geliştirilen eğitim programlarını değerlendirme ölçeğinin daha sonra görev yapmakta olan öğretmenler üzerinde kullanılmasının geçerlik anlamında bir sorun teşkil edebileceği düşünülebilir. Sonuçta, bu ölçeğin görev yapmakta olan öğretmenler üzerinde ileri geçerlik ve güvenilirlik çalışmalarının yapılarak kullanılması ifade edilen bu sorunun aşılmasında önemli bir husus olarak değerlendirilebilir. Göle ve Temel (2015) tarafından geliştirilen “program niteliği değerlendirme ölçeği” ise doğrudan bir eğitim programı değerlendirme ölçeği olmasa da, bir eğitim programının niteliğine ilişkin kriterleri içerisinde barındırması açısından oldukça önemli görülmektedir. Göle ve Temel (2015) tarafından hazırlanan bu ölçek içerisinde görev yapmakta olan okul öncesi eğitimi öğretmenlerini barındıran çalışma grubu kullanılarak geliştirilmiştir. Göle ve Temel (2015) tarafından geliştirilen bu ölçek toplam 61 maddeden ve “öğretmen-çocuk arasındaki ilişkiler”, “program müfredatı”, “eğitim süreci”, “aile katılımı” ve “değerlendirme” olmak üzere beş boyuttan meydana gelmektedir. Göle ve Temel (2015) tarafından geliştirilen bu ölçek içerisinde programın niteliğine ilişkin farklı türdeki kriterleri içermekte olup, doğrudan programın öğelerinin değerlendirildiği bir yapıya sahip bulunmamaktadır. Bu açıdan, hali hazırda geliştirilen ölçek ile benzerliği olmasa da, bir eğitim programının niteliğinin belirlenmesinde dikkate alınması gereken kriterleri içermesi açısından

oldukça anlamlı görülmektedir. Sonuç olarak, ilgili alanyazında farklı kültürel bağlamlarda geliştirilmiş olan eğitim programı değerlendirme ölçekleri ile karşılaşılmış olup, bu ölçeklerin hali hazırdaki çalışmada geliştirilen ölçek ile bazı benzerlik ve farklılıklarının olduğu sonucuna ulaşılmıştır. Bu çalışmanın; Türkiye bağlamında şu anda eğitim programlarının değerlendirilmesinde kullanılabilir olan yayımlanmış olan bir ölçek geliştirme çalışmasının bulunmaması, bu araştırmanın çalışma grubunda eğitim programlarını uygulamada kullanmakta olan hem ilköğretim hem de lise öğretmenlerini dikkate alması, bir eğitim programının bütün öğelerini içerisinde barındırması gibi hususlar açısından oldukça önemli bir yere sahip olduğu ve sahip olduğu psikometrik özellikler ile eğitim programlarının değerlendirilmesinde rahatlıkla kullanılabilir olduğu düşünülmektedir.

Eğitim programlarını değerlendirme ölçeğinin geçerlik ve güvenilirlik çalışmasına ilişkin bulgular değerlendirildiğinde, geliştirilen bu ölçeğin öğretmenlerin eğitim programlarını değerlendirmede kullanılması önerilmektedir. Nitekim, programın etkili bir biçimde uygulanması, değerlendirilmesi ve düzeltilmesi faktörleri de, kendilerine bu konuda büyük sorumluluklar düşen öğretmenlerin program değerlendirmeye yönelik görüşlerinin incelenmesi gereksinimini gündeme getirmektedir (Demir ve Fer, 2009). Her ne kadar uzman gruplarla hazırlanmış olsa da programların ikinci veya üçüncü gözlerle izlenmesi bir zorunluluk olarak değerlendirilmektedir (Acar, Bıçak, Tutkun ve Pektaş, 2009). Eğitim programları ile içli-dışlı olan ve onları uygulamaya yansıtan öğretmenlerce programların sürekli olarak değerlendirilmesi oldukça önemli görülmektedir. Hazırlanarak okullara uygulanmak üzere gönderilen eğitim programlarının onların birer uygulayıcısı olan öğretmenler tarafından değerlendirilmemesi programların niteliğinin, eksiklik veya yanlışlıklarının, olumsuz işleyen yön

ve boyutlarının görülememesine yol açabilir. Bu da sonuçta, sabit, günün veya zamanın koşullarına uymayan, uygulanamayan, işlevsel olmayan, öğrenciyi geliştirmeyen, vb. eğitim programlarının hazırlanmasına yol açabilir. Böylesi özelliklere sahip olan bir eğitim programının ise entropiden kaçması ve sağlıklı bir toplum ve eğitim sistemi oluşturması mümkün gözükmemektedir.

#### **4.1. Öneriler**

Sonuç olarak, elde edilen bulgular ışığında çalışmada geliştirilen ölçeğin ilgili alanyazındaki önemli bir boşluğu doldurduğu ve sahip olduğu psikometrik özelliklerle gelecek çalışmalarda kullanılabileceği belirtilebilir. Bunun yanında, geliştirilen bu ölçeğin program geliştirme uzmanlarına ve okul yöneticilerine de uyarlanarak, ölçeğin bu gruplardaki geçerlik ve güvenilirlik çalışmalarının test edilmesi önerilmektedir. Ayrıca, öğretmenlerin hazırlanan eğitim programlarında oynadıkları rolleri konu alan bir ölçeğin geliştirilmesi de burada önemli görülmektedir. Bu çalışma, göreceli olarak sınırlı sayıda bir çalışma grubu ile gerçekleştirilmiş olup, ölçeğe ilişkin ileri doğrulayıcı faktör analizi çalışmaları daha büyük örneklem grupları kullanılarak yapılabilir. Bunun dışında, hali hazırdaki çalışmada geliştirilmiş olan ölçek karma bir yaklaşımla desenlenerek, ölçekten elde edilecek nicel veriler nitel yollarla elde edilecek verilerle bütünleştirilerek, eğitim programlarının daha sağlıklı bir şekilde değerlendirilmesi sağlanabilir.

#### **4.2. Sınırlılıklar**

Öncelikle bu çalışma, küçük bir ilde görev yapmakta olan görece sınırlı sayıdaki öğretmenler üzerinde gerçekleştirilmiştir. Her ne kadar araştırmada dikkate alınan çalışma grubundaki öğretmen sayısı görece sınırlı olsa da, bu sayı alanyazında yeterli olarak değerlendirilmektedir (Kline, 1994). Ancak, bu araştırmada yalnızca açımlayıcı faktör analizi ile

güvenirlilik analizi çalışmalarına yer verilmiş olup, doğrulayıcı faktör analizi çalışmalarına yer verilmemiştir. Bu anlamda, ölçeğin daha büyük çalışma grupları kullanılarak doğrulayıcı faktör analizi çalışmaları gerçekleştirilebilir. Aynı zamanda, hali hazırdaki çalışmada geliştirilen ölçeğin ölçüt-dayanıklı geçerliği yoklanmamış; dolayısıyla, bu ölçekle birlikte benzer diğer ölçeklerden eş zamanlı olarak veri toplanmamıştır. Diğer yandan, hali hazırdaki ölçek Türkiye’de Milli Eğitim Bakanlığına (MEB) bağlı ilköğretim okulu ve liselerde görev yapmakta olan öğretmenler üzerinde uygulanarak geliştirilmiş olup, özel ve(ya) vakıf okullarında görev yapmakta olan öğretmenler çalışma dışında tutulmuştur. Son olarak, gerçekleştirilen bu ölçek geliştirme çalışması yalnızca il merkezinde görev yapmakta olan öğretmenler üzerinde gerçekleştirilmiş olup, kırsal (köy ve kasaba) bölgelerde görev yapmakta olan öğretmenler çalışma dışında tutulmuştur. Bu bağlamda, gelecek çalışmalarda gerek il merkezinde gerekse kırsal bölgelerde görev yapmakta olan öğretmenler dikkate alınarak ölçeğin doğrulayıcı çalışmalarına yer verilebilir.

#### **Kaynaklar**

- Acar, V., Bıçak, B., Tutkun, Ö. F., & Pektaş, S. (2009, Ekim). *Eğitim programları değerlendirme ölçeği (EPDÖ)*. 18. Eğitim bilimleri kongresinde sunulan bildiri, Ege Üniversitesi Eğitim Fakültesi, İzmir.
- Adıgüzel, O. C., & Özüdoğru, F. (2014). İlkokul 2. sınıf İngilizce öğretim programına yönelik aydınlatıcı değerlendirme modeline dayalı program değerlendirme ölçeği çalışması. *Trakya Üniversitesi Eğitim Fakültesi Dergisi*, 4(2), 124-136.
- Anderson L. W. (1988). Attitudes and their measurement. J. P. Keeves (Ed.), *Educational research, methodology and measurement: An international handbook* (pp. 885-895). New York: Pergamon Press.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme*. Ankara: ÖSYM Yayınları.
- Büyükkaragöz, S. S. (1997). *Program geliştirme*. Konya: Kuzucular Ofset.
- Büyükoztürk, Ş. (2007). *Sosyal bilimler için veri analizi el kitabı*. (12. baskı). Ankara: Pegem A Yayınları.

- Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Kuram ve Uygulamada Eğitim Yönetimi*, 32, 470-483.
- Cronbach L. J. (1990) *Essentials of psychological testing* (5th ed.). New York: Harper Collins Publishers.
- Cusimano, M. D., & Garg, M. (1996). A scale for curriculum evaluation and development in the health problems. *Evaluation & Health Problems*, 19(1), 91-103.
- Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2010). *Sosyal bilimler için çok değişkenli istatistik*. Ankara: Pegem Akademi Yayınları.
- Çubukçu, Z. (2008). Eğitim programı tasarımı ve geliştirilmesi. B. Duman (Ed.), *Öğretim ilke ve yöntemleri* (2. baskı, ss. 132-174). Ankara: Maya Akademi.
- Demir, S., & Fer, S. (2009). 2005 ilköğretim programlarını değerlendirme envanterinin geçerlik ve güvenilirlik çalışması. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 6(2), 207-231.
- Demirel, Ö. (2005). *Eğitimde program geliştirme: Kuramdan uygulamaya* (8. baskı). Ankara: Pegem A Yayıncılık.
- Doll, R. C. (1986). *Curriculum improvement: Decision making and process*. Boston, MA: Allyn and Bacon.
- Erden, M. (2000). *Eğitimde program değerlendirme* (3. baskı). Ankara: Anı Yayıncılık.
- Ertürk, S. (1972). *Eğitimde program geliştirme*. Ankara: Yelken Tepe Yayınları.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299.
- Ferguson, F., & Takane, Y. (1989). *Statistical analysis in psychology and education*. New York: McGraw-Hill.
- Fraser, B. J. (1977). Selection and validation of attitude scales for curriculum evaluation. *Science Education*, 61(3), 317-329.
- Göle, M. O., & Temel, Z. F. (2015, Haziran). *Program niteliği değerlendirme ölçeği geliştirme, geçerlik ve güvenilirlik çalışması*. II. Uluslararası Avrasya eğitim araştırmaları kongresinde sunulan bildiri, Hacettepe Üniversitesi Eğitim Fakültesi, Ankara.
- Gömleksiz, M. N., & Bulut, İ. (2006a). Yeni fen ve teknoloji dersi öğretim programına ilişkin öğretmen görüşleri. *Fırat Üniversitesi Sosyal Bilimler Dergisi*, 16(2), 173-192.

- Gömleksiz, M. N., & Bulut, İ. (2006b). Yeni sosyal bilgiler dersi öğretim programının uygulamadaki etkililiğinin değerlendirilmesi. *Kuram ve Uygulamada Eğitim Yönetimi*, 47, 393-421.
- Henson, K. T. (2006). *Curriculum planning*. Illinois: Waveland Press.
- Kline, P. (1994). *An easy guide to factor analysis*. London: Routledge.
- Korkmaz, İ. (2007). Eğitim programı: Tasarımı ve geliştirilmesi. A. Doğanay (Ed.), *Öğretim ilke ve yöntemleri* (2. baskı, ss. 2-34). Ankara: Pegem A Yayıncılık.
- Marsh, C. J., & Willis, G. (2007). *Curriculum: Alternative approaches, ongoing issues*. New Jersey: Prentice-Hall.
- McNeil, J. D. (2006). *Contemporary curriculum*. New Jersey: John Wiley and Sons.
- Murphy K. R., & Davidshofer (1991). *Psychological testing: Principles and applications*. New Jersey: Prentice-Hall.
- Oliva, P. F. (1988). *Developing the curriculum*. Boston: Scott, Foresman and Winston.
- Ornstein, A. C., & Hunkins, F. P. (1993). *Curriculum: Foundations, principles and issues* (2nd ed.). Needham Heights, MA: Allyn and Bacon.
- Öksüz, C. (2015). İlkokul matematik programını değerlendirme ölçeği. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 37(1), 21-33.
- Özçelik, D. A. (2010). *Eğitim programları ve öğretim* (2. baskı). Ankara: Pegem Akademi.
- Özdemir, S. M. (2009). Eğitimde program değerlendirme ve Türkiye’de eğitim programlarını değerlendirme çalışmalarının incelenmesi. *Yüüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 6(2), 126-149.
- Özdemir, S. M. (2007). Eğitimde program geliştirme ve program geliştirme sürecinin unsurları. G. Ocak (Ed.), *Öğretim ilke ve yöntemleri* (ss. 59-94). Ankara: Pegem A Yayıncılık.
- Özen, Y., Gülaçtı, F., & Kandemir, M. (2006). Eğitim bilimleri araştırmalarında geçerlik ve güvenilirlik sorunsalı. *Erzincan Eğitim Fakültesi Dergisi*, 8(1), 69-89.
- Reuterberg, S., & Gustafsson, J. E. (1992). Confirmatory factor analysis and reliability: Testing measurement model assumptions. *Educational and Psychological Measurement*, 52, 795-811.



- Saylor, J. G., Alexander, W. M., & Lewis, A. J. (1981). *Curriculum planning for better teaching and learning*. New York: Holt, Rinehart and Winston.
- Sönmez, V., & Alacapınar, F. G. (2015). *Örnekleriyle eğitimde program değerlendirme*. Ankara: Anı Yayıncılık.
- Sönmez, V. (2007). *Program geliştirme öğretmen el kitabı* (13. baskı). Ankara: Anı Yayıncılık.
- Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik*. Ankara: Seçkin Yayıncılık.
- Taba, H. (1962). *Curriculum development: Theory and practice*. New York: Harcourt, Brace and World.
- Tabachnick B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). MA: Allyn and Bacon.
- Tanner, D., & Tanner, L. (1995). *Curriculum development: Theory into practice*. New Jersey: Prentice-Hall.
- Tatlıdil, H. (2002). *Uygulamalı çok değişkenli istatistiksel analiz*. Ankara: Akademi Matbaası.
- Tavşancıl, E. (2005). *Tutumların ölçülmesi ve SPSS ile veri analizi*. Ankara: Nobel Yayın Dağıtım.
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington: American Psychological Association.
- Tyler, R. W. (1950). *Basic principles of curriculum and instruction*. Chicago, Illinois: The University of Chicago Press.
- Uşun, S. (2012). *Eğitimde program değerlendirme: Süreçler, yaklaşımlar ve modeller*. Ankara: Anı Yayıncılık.
- Varış, F. (1978). *Eğitimde program geliştirme: Teori ve teknikler* (3. baskı). Ankara: Ankara Üniversitesi Eğitim Fakültesi Yayınları.
- Yüksel, İ., & Sağlam, M. (2012). *Eğitimde program değerlendirme: Yaklaşımlar, modeller, standartlar*. Ankara: Pegem Akademi.
- Wiles, J., & Bondi, J. (1993). *Curriculum development: A guide to practice*. New York: MacMillan Publishing Company.
- Worthern, B. R., Sanders J. R., & Fitzpatrick J. L. (1997). *Program evaluation: Alternative approaches and practical guidelines*. New York: Longman.

### Extended Abstract

The purpose of this study was to develop a valid and reliable scale that could be used in evaluating curriculum in elementary and high schools by teachers. This study sought to improve the understanding of curriculum evaluation by teachers. Hence, the findings obtained in this study are believed to provide information for policymakers concerned with education administration, curriculum development, teacher education as well as insights that may be relevant to similar studies elsewhere.

The sample consisted of voluntary elementary and high school teachers ( $n = 200$ ) in the centre of Nigde province. In the light of scales obtained by the literature review, an item pool was formed by the researcher. The first form of the scale was prepared in the pool by the chosen items. This first form was presented into the views of a group of experts. According to the experts' directions, necessary changes were made and it was decided that 35 items would be placed in the scale. All of the items included in the scale were scored as 1 for totally disagree and 5 for totally agree. The measurement form was applied as a pilot study in elementary and high schools of Nigde province. For construct validity of the scale, exploratory factor analysis based on principal component analyses was applied.

In the exploratory factor analysis, 0.40 was accepted as the lowest factor loading limit in determining whether the items were included in the scale. Firstly, the validity reliability analyses of the Curriculum Evaluation Scale were applied. Kaiser-Meyer-Olkin (KMO) value and Barlett's test of sphericity were examined in order to eligibility of data for factor analysis. KMO results of the scale was found as 0.940 and the Bartlett result of the scale was calculated as  $\Delta\chi^2 = 3710.726/df = 595$ . In the study, KMO value was perceived as excellent and the Barlett's test of sphericity level was found out to be significant. According to the findings of the exploratory factor analysis, the scale consisted of 35 items and four sub-dimensions. These sub-dimensions were named as *i.* aims, *ii.* content, *iii.* teaching-learning process, and *iv.* measurement-evaluation. It was also seen that the factor loadings of the scale were between 0.715 and 0.457. It was understood that a variance of approximately 54.390% was explained when the variance that four different sub-dimensions. Besides, the Cronbach's Alpha value of the total scale was found as 0.92 and the Spearman-Brown split half correlation of the scale was calculated as 0.90 in the study. Besides these, it was found out that the sub-dimensions of the scale were calculated as ranging between 0.89 and 0.78.

Not: Bu çalışma, Necmettin Erbakan Üniversitesi, Ahmet Keleşoğlu Eğitim Fakültesi tarafından düzenlenen "Uluslararası Eğitimde Değişim ve Yeni Yönelimler Sempozyumu"nda (22-24 Kasım 2013) sözlü bildiri olarak sunulmuştur.

*EK-1. Eğitim Programlarını Değerlendirme Ölçeği*

Maddeler

- M1. Programda hedefler, eğitim durumlarını ve ölçme-değerlendirme boyutlarına yol gösterici niteliktedir.
- M2. Programda hedefler, öğretmenin değil, öğrencinin neleri yapması gerektiğini belirtmektedir.
- M3. Programda hedefler, ülkenin başar eğitim felsefesi ve toplumsal yapısı ile uyumludur.
- M4. Programın ölçme-değerlendirme boyutu, öğrencilerin yalnızca bilişsel öğrenmelerini değil, aynı zamanda duyuşsal ve psiko-motor öğrenmelerinin belirlenmesine de yol göstermektedir.
- M5. Programda hedefler, eğitim yolu ile gerçekleştirilebilir niteliktedir.
- M6. Programda hedefler, açık ve net olarak ortaya konulmuştur.
- M7. Programda hedefler, bireyin yaşamış olduğu çevreyi dikkate almaktadır.
- M8. Programın içerik boyutundaki bilgileri öğrenciler uygulayabilir ve kullanabilirler.
- M9. Programda hedefler birbirinin önkoşulu ve birbirlerini tamamlayacak bir şekilde düzenlenmiştir.
- M10. Programda hedefler yalnızca ürüne değil, aynı zamanda süreçte de önem vermektedir.
- M11. Programın ölçme-değerlendirme boyutu, hedeflerin gerçekleştirilip gerçekleştirilmediğini saptamaya hizmet etmektedir.
- M12. Programın içeriği, bireylerin ilgi ve ihtiyaçlarına cevap verir niteliktedir.
- M13. Programın içeriğindeki konular organize edilmiş bilgi anlamında anlamlıdır.
- M14. Programın içerik boyutundaki konular kullanılabilirliğe sahiptir.
- M15. Programın içerik boyutu, öğrencilerin önceki öğrenmeleriyle kaynaşıklık sağlamaktadır.
- M16. Hedefler, programın içerik boyutu ile tutarlı ve kaynaşıktır.
- M17. Programın içerik boyutunda sunulan bilgiler bilimsel açıdan doğrudur.
- M18. Programın içeriği, sosyal gerçeklerle tutarlı ve toplumsal ihtiyaçları karşılamaktadır.
- M19. Öğrenme-öğretme süreci, programın hedeflerinin gerçekleştirilebilmesinde ekonomiklik sağlamaktadır.
- M20. Programın ölçme-değerlendirme boyutu, öğrenci öğrenmelerindeki eksik ve yanlış öğrenmeler ile güçlükleri belirlemeye yardımcı olmaktadır.
- M21. Programın öğrenme-öğretme süreci boyutu, öğrencilerin düzeylerine ve ilgi ve ihtiyaçlarına uygun bir şekilde düzenlenmiştir.

M22. Öğrenme-öğretme süreci, programın içerik boyutunun öğrenilmesi ve öğretilmesi için uygun bir şekilde düzenlenmiştir.

M23. Programın içerik boyutu iyi organize edilmiş sistematik bir yapı sunmaktadır.

M24. Programın öğrenme-öğretme süreci boyutu, birden çok hedefi gerçekleştirebilir niteliğe sahiptir.

M25. Öğrenme-öğretme süreci, programın hedefleri ile tutarlıdır.

M26. Programın öğrenme-öğretme süreci boyutu, kolay uygulanabilir bir nitelik taşımaktadır.

M27. Programda hedefler, öğrencilerin genel özelliklerine ve ihtiyaçlarına uygundur.

M28. Programın içeriği, hedef boyutuyla tutarlı ve kaynaşıktır.

M29. Programın ölçme-değerlendirme boyutu, öğrencilere ve ortama göre değişebilecek alternatif ölçme-değerlendirme yaklaşımları sunmaktadır.

M30. Programın içerik boyutundaki konular günlük yaşamla ilgilidir.

M31. Öğrenme-öğretme süreci, programın hedeflerini gerçekleştirebilmeye dönüktür.

M32. Programın ölçme-değerlendirme boyutu, farklı öğrenme becerilerini (eleştirel düşünme, yaratıcı düşünme, problem çözme, vb.) değerlendirmeye olanak sağlamaktadır.

M33. Programın öğrenme-öğretme süreci, farklı öğrenme yollarına sahip olan öğrencileri gözetir niteliktedir.

M34. Programın öğrenme-öğretme süreci boyutu, öğrencilerin etkili öğrenmeler kazanabilmeleri için uygun bir şekilde düzenlenmiştir.

M35. Programın öğrenme-öğretme süreci boyutu, öğrencilerin ilgi ve ihtiyaçlarına cevap verir niteliktedir.

---