



## Developing a Self Efficacy Scale toward Using Rubric \*

Miraç YILMAZ<sup>1</sup> Perihan GÜNEŞ<sup>2</sup> Belma TÜRKER BİBER<sup>3</sup>

**ABSTRACT:** The aim of this study is to develop a valid and reliable scale that can be used to measure pre-service teachers' self-belief in using rubrics. Self-efficacy scale toward using rubric was applied to 402 pre-service teachers from 3 state universities in Turkey. For the structure validity of the scale, firstly, an exploratory factor analysis was performed, and for the uni-dimensional structure that was seen to be meaningful in the first analysis, a principal component analysis was performed. In the last step of the analyses, the model, which was formed according to the basic components and factor analysis results, was tested through a confirmatory factor analysis. The scale is composed of 13 items, and the reliability coefficient was calculated to be Cronbach alpha= .91. As a result, it can be said that the developed scale is a valid and reliable measurement tool to be used in measuring self-efficacy toward using rubric.

**Keywords:** Pre-service teachers, rubric, self-efficacy, scale development.

---

\* This paper was presented as oral presentation at the European Conference on Educational Research (ECER) Congress in Porto (Portugall), September 1-5, 2014

<sup>1</sup> Assist. Prof. Dr. Hacettepe University, Faculty of Education, Ankara, Turkey. yilmazmirac@gmail.com

<sup>2</sup> Assist. Prof. Dr. Aksaray University, Faculty of Education, Aksaray, Turkey. perimensen@gmail.com

<sup>3</sup> Res. Asst., Hacettepe University, Faculty of Education, Ankara, Turkey. bturker@hacettepe.edu.tr

## SUMMARY

**Purpose and Significance:** The aim of this study is to develop a valid and reliable scale to be used in measuring pre-service teachers' self-efficacy toward using rubric beliefs.

**Method:** In order to develop a self-efficacy toward using rubric (SESUR) measurement tool, 402 pre-service teachers from three state universities. Pre-service teachers have been studying at biology education, science education, mathematics, counseling and primary education departments. 72% (288) of the pre-service teachers were female, and 28% (114) of them were male. All of them have taken the measurement-evaluation course at least once, and 99% of them have successfully completed the class. In preparing the scale, first of all literature, theses, already-developed scales and other basic sources on rubric and self-efficacy were consulted. After this exploration, a 20 Likert-type items were prepared. Then, a draft form with 14 items selected out of these was formed, and this form was examined by experts in terms of its appropriateness, comprehensibility, length, and its proficiency in measuring using rubrics self-efficacy. Then, the scale edited again and were organized as a 5-Likert type scale.

**Results:** As a result of the PCA, a single component was found. This approach indicates that the scale is unidimensional. In addition, scree plot results showed that one component are statistically significant. In relation to PCA, Kaiser-Meyer-Olkin (KMO) was calculated to be .92, and Barlett's Test of Sphericity was calculated to be  $(151) = 869.025, p = .00$ . These values provides enough proof for the appropriateness of the obtained data for the sampling group and the fact that matrix R can be factored (Pallant 2005). EFA was done and it was determined that the average variance the single dimension explains is 47.87%. Factor loadings of items are between in .60-.76 interval. In the second phase of the analysis process, an unidimensional CFA was conducted in order to determine to what extent the unidimensional structure can explain the relationship among the data obtained from other pieces of the data set (Jöreskog & Sörbom 1993). According to the results, standardized weights of the items are in the .58-.75 interval. Overall compliance coefficients related to the model are at sufficient levels ( $\chi^2_{251} = 187.17; p = .00$ ; CFI = .97; SRMR = .05; RMSEA = .08) (Hair, Anderson, Tatham & Black, 2005). The RMSEA value of the proposed model is .08. It can be seen that the proposed model's RMSEA value meets the critical value  $\leq .08$ , which is an acceptable one (Hair et al. 2005; Kline, 2011; Schumacker & Lomax, 2004). According to the obtained results, it can be said that the model is a good one. Internal consistency reliability coefficient (Cronbach Alpha) is  $\alpha = .91$ .

**Discussion and Conclusions:** According to the results of the explanatory and affirmative factor analyses, reliability of the uni-dimensional SESUR can be said to be high proficient. Cronbach Alpha internal consistency coefficient value of the 13-item SESUR is  $\alpha=.91$ . This can be considered a rather high value. According to the obtained results, it is possible to claim that this is a good model. According to this, validity results related to the SESUR are met according to CFA. Developed in the light of the explained findings, it seems possible to claim that “self-efficacy toward using rubric” scale is a valid and reliable measurement tool.

Studies and applications in which self-efficacy toward using rubric is determined can be applied to large teacher and pre-service teacher groups, and thus, examining whether additional activities related to developing their proficiencies are needed, and in relation, whether students’ success is affected or not can also be suggested.



## Rubrik Kullanma Özyeterlik Ölçeği Geliştirme Çalışması<sup>†</sup>

Miraç YILMAZ<sup>1</sup> Perihan GÜNEŞ<sup>2</sup> Belma Türker BİBER<sup>3</sup>

**ÖZET:** Bu çalışmada, öğretmen adaylarının rubriğe ilişkin özyeterliklerini belirlemede kullanılacak geçerli ve güvenilir ölçümler sunan bir ölçme aracının geliştirilmesi amaçlanmıştır. Uzman görüşleri doğrultusunda hazırlanan taslak “Rubrik Kullanma Özyeterliği Ölçeği” üç devlet üniversitesinde beş ayrı bölümde okumakta olan toplam 402 öğretmen adayına uygulanmıştır. Ölçekten elde edilen ölçümlerin yapı geçerliği faktör analizi kullanılarak araştırılmıştır. Araştırmanın ilk aşamasında temel bileşenler ve açımlayıcı faktör analizi, son aşamasında ise doğrulayıcı faktör analizi uygulanmıştır. Toplam 13 maddeden oluşan ölçeğin iç tutarlık güvenilirlik katsayısı Cronbach alfa= .91 olarak hesaplanmıştır. Araştırma sonuçları, rubrik kullanma özyeterliği ile ilgili geliştirilen ölçeğin geçerli ve güvenilir ölçümler sağlayan bir araç olarak kullanılabilmesi göstermektedir.

**Anahtar Kelimeler:** Öğretmen adayları, rubrik, özyeterlik, ölçek geliştirme.

\* Bu makale 1-5 Eylül 2014 tarihlerinde Porto /Portekiz’de yapılan ECER’14 Kongresi’nde sözlü bildiri olarak sunulmuştur.

<sup>1</sup> Yard. Doç. Dr. Hacettepe Üniversitesi, Eğitim Fakültesi, Ankara, Türkiye. yilmazmirac@gmail.com

<sup>2</sup> Yard. Doç. Dr. Aksaray Üniversitesi, Eğitim Fakültesi, Aksaray, Türkiye. periensen@gmail.com

<sup>3</sup> Araş. Gör. Hacettepe Üniversitesi, Eğitim Fakültesi, Ankara, Türkiye. bturker@hacettepe.edu.tr

## **GİRİŞ**

Son yıllarda Türkiye’de Milli Eğitim Bakanlığınca hazırlanan yeni öğretim programlarında, öğrenme-öğretme ortamına daha etkin katılan, araştıran, sorgulayan, bilgiyi kullanan ve paylaşan bireylerin yetiştirilmesine önem verilmektedir (Milli Eğitim Bakanlığı [MEB], 2013a, 2013b, 2013c, 2013d, 2013e). Öğretim programlarındaki bu anlayışın sonucunda okullarda ezber bilgi, derinliği olmadan çok sayıda kavram ve prensibi öğrenme, yazılı materyallerdeki bilgileri aynen aktarma gibi öğrenme çıktılarının öneminin azaldığı; aksine eleştirel düşünme, yorumlama, analiz edebilme, sonuç çıkarabilme, gözlemler yapabilme, günlük hayatta karşılığı olan problemleri çözebilme, araştırma yapma gibi üst düzey bilişsel becerilerin önemli hale geldiği görülmektedir (Berberoğlu, 2006). Sonuçta tüm bu yeni eğilimler ölçme ve değerlendirme anlayışının da değişmesine neden olmuş, sınıf içinde sıklıkla kullanılan yöntemlerin yanı sıra tamamlayıcı ölçme değerlendirme uygulamalarının da ön plana çıkmasını sağlamıştır. Öğretim programlarında ürünün yanı sıra sürecin de değerlendirildiği performansa dayalı ölçme ve değerlendirme anlayışının altı çizilmektedir (MEB 2013a, 2013b, 2013c, 2013d, 2013e).

Performansa dayalı değerlendirme önceden belirlenmiş seçenekler arasından seçim yapmaktan ziyade, öğrencinin cevabı oluşturması esasına dayanmaktadır (Baker, O’Neal & Linn, 1993). Performansa dayalı değerlendirme basit düzeyde bir cevabı yapılandırmaktan, bir aktiviteyi sergilemek ya da bir ürün ortaya koymaya kadar geniş bir perspektifi kapsamaktadır. Bir çalışmanın performansa dayalı olabilmesi için üst düzey zihinsel çabayı içermesi gerekmektedir (Aschbacher 1991; Bahar, Nartgün, Durmuş ve Bıçak, 2012; Baker et al., 1993; Kutlu, Doğan ve Karakaya, 2008). Performansa dayalı değerlendirme, öğrencilerin bir konudaki bilgi ve becerilerini yansıtmalarına fırsat verecek bir aktivite (örneğin model yapmak) ya da orijinal yanıtlar üretebilecek bir durumu (bir matematik probleminin çözümünü açıklamak, kompozisyon yazmak gibi) içerir (Madaus & O’Dwyer, 1999; Stiggins, 1987). Bunların yanı sıra performans değerlendirmenin öğrencilere gerçek yaşam durumlarında uygulanabilecek problemleri sunan, çoklu çözüm yolları ve stratejileri içeren, öğrencilerin ön bilgilere sahip olmasını gerektiren, aynı zamanda geniş bir zaman diliminde ölçülebilen özellikleri de bulunmaktadır (Aschbacher, 1991; Bahar vd., 2012; Madaus & O’Dwyer 1999, Stiggins 1987). Öğrencilerin ürün ya da süreç boyunca göze çarpan becerilerinin direk ölçülmesi (Frederiksen, 1984), aynı zamanda ölçümünün şeffaf olması (Frederiksen & Collins 1989) performans değerlendirmeyi popüler yapan gerekçelerdir.

Performansa dayalı değerlendirme, performans görevleri ve rubrik (dereceli puanlama anahtarı) olmak üzere iki bölümden oluşur (Popham, 2007). Performans görevleri eleştirel düşünme, problem çözme, yaratıcılık, araştırma yapma gibi öğrencilerin üst düzey zihinsel becerilerinin gelişmesini ve ölçülmesini sağlayan etkinliklerdir. Performansın ölçülebilmesi için, öğrencilere performanslarını ortaya çıkaracak performans görevlerinin verilmesi gereklidir. Rubrikler, performans görevlerini değerlendirmede yaygın olarak kullanılan tamamlayıcı ölçme ve değerlendirme araçlarından biridir (Popham, 2007).

Goodrich (1997) ve Andrade, Wang, Du ve Akawi (2009), rubriği “çalışmalarda yer alması istenilen ölçütlerin listelendiği ve bu ölçütlerin niteliğinin iyiden kötüye doğru detaylı bir şekilde tanımlandığı dokümanlar” olarak ifade etmişlerdir. Picket ve Dodge (2007) rubriklerin, performans veya nitelik gibi belirlenmiş bir hedefi ölçmesi gerektiğini vurgulamışlardır. Ayrıca araştırmacılar rubriklerin, performansın düzeylerini ve tanımlarını da içermesi gerektiğini belirtmişlerdir. Rubrikler kullanım amacına göre holistik (bütüncül) ve analitik olmak üzere iki farklı şekilde oluşturulabilirler (Bahar vd., 2012; Picket & Dodge, 2007). Holistik rubrik, öğretmenin ürünü bir bütün olarak dikkate aldığı puanlamayı içermektedir. Analitik rubrik ise, performansın bileşenlerini ayrı ayrı değerlendirmektedir. Ayrıca analitik rubrik, süreç boyunca ortaya çıkan performansa odaklanır (Atılğan, Kan ve Doğan, 2007).

Öğretim sürecinde öğrencilerin ezbere ve hatırlamaya dayanan bilgilerinin yanı sıra, üst düzey zihinsel becerilerini de kullanmaları gerekmektedir. Bu gereklilik üst düzey zihinsel becerilerin ölçülmesini ve değerlendirilmesini beraberinde getirmektedir. Rubrikler araştırma yapma, eleştirel düşünme, bilimsel düşünme gibi hem ürün hem de sürece odaklanan üst düzey becerilerin ölçülmesi için ön plana çıkmaktadır.

Eğitim sistemlerinde gerçekleştirilen yeniliklerin amacına ulaşabilmesinde yeniliklere açık ve onları uygulayabilen öğretmenlere ihtiyaç olduğunun altı önemle çizilmektedir (Bahar vd., 2012). Bu bağlamda öğretmenlerin sıklıkla kullanılan ölçme ve değerlendirme yöntemleri kadar tamamlayıcı ölçme ve değerlendirme yöntemlerine de önem vermeleri, bu yöntemleri kullanmayı bilmeleri ve bu konuda istekli olmaları önemlidir. Ancak bazı araştırmalarda, tamamlayıcı ölçme ve değerlendirme araçları ile ilgili öğretmenlerin bilgi düzeylerinin yetersiz olduğu bildirilmektedir (Fidan ve Sak, 2012; Sağlam-Arslan, Devcioğlu-Kaymakçı ve Arslan, 2009; Yeşilyurt, 2012). Gelbal ve Kelecioğlu (2007)'de yaptıkları çalışmada, ölçme ve değerlendirme tekniklerinin kullanımı ve hazırlanması konusunda öğretmenlerin eğitime ihtiyaçları olduğunu vurgulamaktadırlar. Ayrıca

Yeşilyurt (2012) eğitim fakülteleri öğretmen yetiştirme programlarının, eğitim anlayışındaki yeni gelişmelerle paralellik taşıma konusundaki yetersizlikleri olduğunu da belirtmektedir. Tüm bunların yanı sıra öğretmenlerin tamamlayıcı ölçme ve değerlendirme araçları hakkındaki bilgi düzeyleri, bu araçlara yönelik tutumları ve bu araçları hazırlayıp, uygulayabilmelerine ilişkin özyeterlikleri de merak konusudur.

Özyeterlik, bireyin belli bir performansı başarılı olarak yapıp yapmayacağına konusunda kendisine duyduğu inançtır. Özyeterlik yeteneklerimize ilişkin inancımıza dayanmaktadır ve belirli bir davranışı gerçekleştirmek için özyeterliğin gerekli olduğu bildirilmektedir (Bandura 1977, 1994). Bireyin olumlu özyeterlik inancının, bireyin o konuya ilişkin etkinliklerde motivasyonunu arttırdığı, zor ve ilk defa karşılaşıldığı görevlerde bireyi çaba göstermeye istekli kılarak bu görevlerle başa çıkabilmesini sağladığı bildirilmektedir. Herhangi bir konuya ilişkin sahip olunan olumsuz özyeterlik inancının ise kişinin kendi iradesiyle ile bir işi yapmamasına ya da sonuçlandırmamasına neden olduğu bildirilmektedir (Jerusalem, 2002). Pajares'e göre (1997) özyeterlik inançları göreve ve konuya özel inançlardır. Schwarzer ise (1993) özyeterliğin eğitim, sosyal, gelişim ve sağlık gibi pek çok özel alanla ilişkili olabileceğini bildirmektedir. Bu durumlara bağlı olarak özyeterlik daha iyi sağlık, daha yüksek başarı ve daha fazla sosyal ilişki şeklinde kendini göstermektedir (Schwarzer, 1993). Benzer şekilde özyeterlik inancının, öğretmen davranışlarını anlama ve geliştirmede katkılar sağlayacağı düşünülmektedir (Tschannen-Moran, Woolfolk-Hoy & Hoy, 1998; Tschannen-Moran & Woolfolk-Hoy, 2001). Nitekim özyeterlik inancı öğrenme-öğretme ve performansla ilgili araştırmalarda sıklıkla kullanılmaktadır (Özkan, Tekkaya ve Çakıroğlu, 2002; Riggs & Enochs, 1990; Tschannen-Moran & Woolfolk-Hoy, 2001). Bu bağlamda kişide tamamlayıcı ölçme değerlendirme araçlarını kullanabileceğine dair var olan inançların incelenmesi de önem taşımaktadır. Gelbal ve Kelececioğlu (2007) çalışmalarında öğretmenlerin kendilerini yeterli olarak algıladıkları ölçme yöntemlerini daha sık kullandıklarını tespit etmişlerdir. Bu durumda konuya göre farklılaşan özyeterlik algılarının, öğretmenlerin ölçme değerlendirme etkinliklerini uygulayabilmeleri açısından da farklılaşabileceği ve onların etkinliklerini değiştirebileceği düşünülebilir. Buna göre ölçme değerlendirme inançlarının ayrıntılı olarak ölçülüp incelenmesi, eğitim öğretim etkinliklerinin veriminin artırılmasına yönelik önlemler alınmasını sağlayabilir. Bu durum öğretmen ve öğretmen adaylarının ölçme ve değerlendirme inançlarının önceden tahmin edilip düzeltilmesini sağlayabileceğinden, yeni eğitim anlayışına uygun ve nitelikli öğretmenler yetiştirmek açısından önem taşımaktadır. Bu açılarından bakıldığında “rubrik kullanma özyeterliği” öğretmenlerin mesleki performansları için çok önemli bir yer tutan tamamlayıcı ölçme ve değerlendirme araçlarını kullanabileceklerine ilişkin inançlarının ne düzeyde

olduğu ve ileride bu konuda ne kadar çaba göstereceklerine ilişkin tahminler yürütülebilir. Ayrıca öğretmen adaylarının gelecekte tamamlayıcı ölçme ve değerlendirme araçlarını ne denli etkili ve iyi kullanabileceklerine dair tahminler yapılması, bu konuda önlemler alınması ve programlarda iyileştirmeler yapılmasında yol gösterici olabilir. Buna göre öğretmen ve öğretmen adaylarının, öğrencinin üst düzey bilişsel becerilerini değerlendirmek için rubrik gibi tamamlayıcı ölçme ve değerlendirme araçlarını kullanabilmeleri konusundaki inançları tespit edilerek, bunları amaca göre ve uygun şekliyle hazırlayıp uygulayabileceklerine dair gösterecekleri çabalar tahmin edilebilir.

Özyeterlik inancına dayanan “rubrik kullanma özyeterliği inancı” kavramı ile öğretmen ve öğretmen adaylarının, rubrik gibi önemli tamamlayıcı ölçme değerlendirme aracını hazırlayıp uygulayabilmeleri hakkındaki kendilerine ilişkin yeterlik beklentileri ele alınmaktadır. Bu konudaki özyeterlik, rubrik kullanma konusundaki inançları ölçmek üzere hazırlanmış ölçeklerle sorgulanabilir. Böylelikle öğretmen ve öğretmen adaylarının mesleki alan yeteneklerinin tespiti yapılabilir ve bu konudaki yeterliklerinin eğitim veya hizmet yaşantıları içinde geliştirilmesi sağlanabilir. Tüm bu nedenlerle geçerli ve güvenilir Türkçe “Rubrik Kullanma Özyeterliği Ölçeği”ne (RKÖÖ) ihtiyaç duyulduğu açıktır. Bu nedenlerle özellikle öğretmen adaylarının rubrik kullanma ile ilgili olarak daha sonra yapabileceklerinin tahmin edilmesi ve mesleki gelişim düzeylerinin geliştirilmesi açısından bakıldığında, rubrik kullanımına yönelik özyeterliği ölçeğin geçerli ve güvenilir bir aracın varlığının önemi anlaşılabilir.

Bu araştırmanın amacı, öğretmen adaylarının rubrik gibi tamamlayıcı ölçme araçlarını hazırlayıp uygulayabilmelerine, başka bir deyişle bu ölçme aracını kullanabilmelerine ilişkin kendilerine olan inançlarını ölçmede kullanılabilecek geçerli ve güvenilir bir ölçme aracı geliştirmektir.

## YÖNTEM

### Araştırmanın Modeli

Bu çalışma, tarama modelinde planlanmış bir ölçek geliştirme çalışmasıdır.

### Çalışma Grubu

Rubrik kullanma özyeterliğine yönelik ölçme aracını geliştirmek amacıyla 2012-2013 öğretim yılında üç devlet üniversitesinde öğrenim görmekte olan toplam 402 öğretmen adayı ile çalışılmıştır. Araştırma üç farklı çalışma grubu üzerinde yürütülmüştür. Açıklayıcı faktör analizi (AFA) verilerinin elde edildiği birinci çalışma grubu 116’sı (%76.8) kadın ve 34’ü (%22.5) erkek olmak üzere 151 kişiden oluşmaktadır. Bu gruptaki öğretmen



adaylarının % 19,9'u (30 kişi) Biyoloji Eğitimi Anabilim Dalı'na; %21.9'u (33 kişi) Fen Eğitimi Anabilim Dalı'na; %48.3'ü (73 kişi) Matematik Eğitimi Anabilim Dalı'na; %7.3'ü (11 kişi) Psikolojik Danışma ve Rehberlik Anabilim Dalı'na ve %2.6'sı (4 kişi) Sınıf Öğretmenliği Anabilim Dalı'na kayıtlı bulunmaktadır. Doğrulayıcı Faktör Analizi (DFA) verilerinin elde edildiği ikinci çalışma grubu 171'i (%68.1) kadın ve 80'i (%31.9) erkek olmak üzere 251 kişiden oluşmaktadır. Bu gruptaki öğretmen adaylarının % 15.5'i (39 kişi) Biyoloji Eğitimi Anabilim Dalı'na; %4'ü (10 kişi) Fen Eğitimi Anabilim Dalı'na; %20.7'si (52 kişi) Matematik Eğitimi Anabilim Dalı'na; %29.5'i (74 kişi) Psikolojik Danışma ve Rehberlik Anabilim Dalı'na ve %30.3'ü (76 kişi) Sınıf Öğretmenliği Anabilim Dalı'na kayıtlı bulunmaktadır. Ayrıca üçüncü çalışma grubunu, araştırmanın birinci ve ikinci çalışma gruplarının toplamı olan 402 kişi oluşturmaktadır. Toplam çalışma grubu verileri üzerinden Cronbach Alfa güvenirliği hesaplanmıştır. Araştırmanın yürütüldüğü 402 kişiden oluşan toplam çalışma grubuna ilişkin bilgiler aşağıdaki tablolarda (Tablo1-4) ayrıntılı şekilde verilmiştir.

**Tablo 1.** Çalışmaya alınan öğretmen adaylarının öğrenim gördükleri anabilim dalları

Uygulama Yapılan Anabilim Dalları	Öğretmen Adayı Sayısı	%
Matematik Öğretmenliği	125	31
Fen Bilgisi Öğretmenliği	43	11
Sınıf Öğretmenliği	80	20
Psikolojik Danışma ve Rehberlik	85	21
Biyoloji Öğretmenliği	69	17
<b>Toplam</b>	<b>402</b>	<b>100</b>

Tablo 1'e bakıldığında, araştırma kapsamında ele alınan çalışma grubunun % 17'sini (69 kişi) Biyoloji Eğitimi Anabilim Dalı'na; %11'ini (43 kişi) Fen Eğitimi Anabilim Dalı'na; %31'ini (125 kişi) Matematik Eğitimi Anabilim Dalı'na; %21'ini (85 kişi) Psikolojik Danışma ve Rehberlik Anabilim Dalı'na; %20'sini (80 kişi) Sınıf Öğretmenliği Anabilim Dalı'na kayıtlı, araştırmaya gönüllü olarak katılmak isteyen toplam 402 öğretmen adayı oluşturmaktadır.

**Tablo 2.** Çalışmaya alınan öğretmen adaylarının cinsiyet dağılımı

Öğretmen Adaylarının Cinsiyeti	Öğretmen Adayı Sayısı	%
Kadın	288	72
Erkek	114	28

Tablo 2'ye bakıldığında, tümünün ölçme değerlendirme dersini en az bir kez almış ve %99'u Ölçme ve Değerlendirme dersini başarıyla alıp geçmiş olduğu tespit edilen öğretmen adaylarının %72'si (288 kişi) kadın, %28'i (114 kişi) ise erkektir.

**Tablo 3.** Çalışmaya alınan öğretmen adaylarının yaşa göre dağılımları

Öğretmen Adaylarının Yaş Aralıkları	%
19-20	20
21-22	62
23-24	15
25 ve üzeri	3

Öğretmen adaylarının yaşlara göre dağılımlarına bakıldığında (Tablo 3) %20'sinin 19-20 yaş; %62'sinin 21-22 yaş; %15'inin 23-24 yaş ve %3'ünün 25 yaş ve üzeri yaş aralıklarında bulunduğu görülmektedir.

**Tablo 4.** Çalışmaya alınan öğretmen adaylarının sınıf düzeylerine göre dağılımları

Öğretmen Adaylarının Sınıf Düzeyleri	Öğretmen Adayı Sayısı	%
2. Sınıf	107	27
3. Sınıf	193	48
4. Sınıf	89	22
5. Sınıf	13	3

Tablo 4'e bakıldığında %27'sini 2. sınıf (107 kişi), %48'ünü 3. sınıf (193 kişi), %22'sini 4. sınıf (89 kişi) ve %3'ünü ise 5. sınıf (13 kişi) öğrencilerinin oluşturduğu görülmektedir.

### Ölçme Aracının Geliştirilmesi

Ölçeğin hazırlanmasında ilk aşama olarak rubrik ve özyeterlik ile ilgili literatür, tez, geliştirilmiş ölçekler ve diğer temel kaynaklardan faydalanılmıştır (Bahar vd., 2012; Kutlu, 2004; Kutlu vd., 2008; Schmitz & Schwarzer, 2000; Tschannen-Moran et al. 1998; Tschannen-Moran & Woolfolk-Hoy, 2001). Bu incelemelerden sonra, öğretmen adaylarının rubrik kullanma özyeterliklerini ölçmek üzere aracın seçimi, hazırlanması ve uygulanması konularını kapsayan ve özyeterliğe ilişkin inançları ifade eden, tek boyutlu Likert tipi 20 maddelik bir havuz oluşturulmuştur. Daha sonra, bu havuzdaki 20 madde ile bir taslak form hazırlanmıştır. Hazırlanan bu taslak form, ikisi özyeterlik ve biri rubriklerle ilgili alanda çalışmaları olan 3 uzman kişi tarafından gözden geçirilmiştir. Taslak form uzmanlar tarafından konuya uygunluk, anlaşılabilirlik, uzunluk ve rubrik kullanma özyeterlikleri, ölçme yeterliliği açılarından incelenmiştir. Uzmanların görüşleri sonucunda 6

maddenin ölçekten çıkarılmasına karar verilmiştir. Nihai ölçekte yer alan 14 madde “kesinlikle katılıyorum”, “katılıyorum”, “kararsızım”, “katılmıyorum”, “kesinlikle katılmıyorum” şeklinde belirtilen 5’li Likert tipinde düzenlenmiştir.

### **Verilerin Analizi**

RKÖÖ taslak formundan elde edilen ölçümlerin yapı geçerliği faktör analizi kullanılarak araştırılmıştır. Ölçümlerin güvenilirliği Cronbach Alfa, kotalı yöntemi kullanılarak test edilmiştir. Veri çözümleme sürecinin ilk aşamasında, veri setinde yer alan gözlemler kotalı örnekleme yoluyla ikiye bölünmüştür. 151 kişiden oluşan ilk grupta temel bileşenler analizi (TBA) ve açımlayıcı faktör analizi (AFA); 251 kişiden oluşan ikinci grupta ise doğrulayıcı faktör analizi (DFA) uygulanmıştır.

Temel bileşenler ve açımlayıcı faktör analizinin yapıldığı ilk çalışmada, veri girişi ve sayıtlılar (aykırı değer, çoklu bağlantı, doğrusallık ve normallik) gözden geçirilmiş ve uygunlukları belirlenmiştir. Verilerin faktör analizine uygun olup olmadıklarının belirlenmesi amacıyla Kaiser-Meyer-Olkin (KMO) değeri incelenmiş ve Barlett’in Küresellik testi yapılmıştır. KMO değerinin .60’tan yüksek ve Barlett’in Küresellik testinin istatistiksel olarak anlamlı olması verilerin faktör analizine uygun olduğuna işaret etmektedir (Büyüköztürk, 2006). AFA sonuçlarının yorumlanmasında, faktör yükü .30’un üzerinde olan maddelerin ölçekte kalmasına dikkate edilmiştir (Büyüköztürk, 2006).

Araştırmada AFA ve Cronbach alfa için SPSS 20.0 paket programından yararlanırken, DFA için ise LISREL 8.80 paket programından yararlanılmıştır. Polychoric korelasyonlar üzerinde temel bileşenler analizi (TBA) yönteminin yanı sıra, faktör sayısının belirlenmesinde özdeğerlerin büyüklüğü ve özdeğer grafiği (scree plot) yaklaşımları kullanılmıştır.

DFA sonucunda elde edilen  $\chi^2$  (Ki-Kare) değerinin anlamlı olduğu tespit edilmiştir.  $\chi^2$  değerinin örneklem büyüklüğünden etkilenebilmesi ve örneklem büyüklüğü arttıkça istatistiksel olarak anlamlı çıkabilmesi (Bentler & Bonett, 1980; Kline, 2011; Schermelleh-Engel, Moosbrugger & Müller, 2003) nedeniyle, modelin verilere uyumlu olup olmadığına karar verilmeden önce,  $\chi^2/sd$ ’nin örneklem büyüklüğü değerine ve diğer uyum indekslerine bakılması gerektiği önerilmektedir (Hu ve Bentler, 1999). Çalışmada 2.87 olarak bulunmuş  $\chi^2/sd$  değerinin 3’ten küçük olması modelin kabul edilebilir bir uyumu sahip olduğunu göstermektedir (Kline, 2011; Marsh, Hau, Artelt, Baumert & Peschar, 2006; Schermelleh-Engel et al., 2003). Uyum indeksleri gözden geçirilirken iyilik uyum indeksi (GFI), karşılaştırmalı uyum indeksi (CFI), normlaştırılmamış uyum indeksi (NNFI), normlaştırılmış uyum indeksi (NFI), fazlalık uyum indeksi (IFI) gibi uyum indeksleri ele alınmıştır. Bu

uyum indekslerinin uygunluğu için .90 ve üzeri kabul edilebilir bir düzeyi göstermektedir (Bentler & Bonett, 1980; Kline, 2011). Modelde düzeltilmiş iyilik uyum indeksi (AGFI) değeri bakılmış ve .85 ve üzeri uygunluk düzeyi olarak kabul edilerek incelenmiştir (Kline, 2011; Marsh et al., 2006). Bunların yanısıra DFA analizi için tahmin hatalarının ortalamasının karekökü (RMSEA), standartlaştırılmış hata kareleri ortalamasının karekökü (SRMR), sıkı iyilik uyum indeksi (PGFI) ve sıkı normlaştırılmış uyum indeksi (PNFI) değerlerine bakılmıştır. RMSEA için .08'den SRMR için .10'dan küçük olan değerler, PGFI ve PNFI için ise .50'den büyük olan değerler kabul edilebilir uyum ölçütü olarak dikkate alınmıştır (Byrne, 2010; Kline, 2011)

### **BULGULAR**

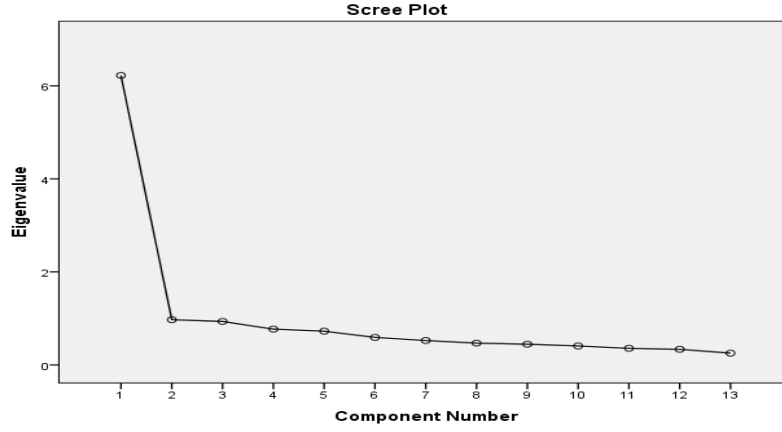
Bu bölümde RKÖÖ'nden elde edilen ölçümlerin geçerlik (açımlayıcı ve doğrulayıcı faktör analizleri) ve güvenilirliklerine (iç tutarlık) ilişkin bulgular aşağıda verilmiştir.

#### **Açımlayıcı Faktör Analizi (AFA)**

Araştırmada öncelikle çalışmanın birinci grubunda açımlayıcı faktör analizine yer verilmiştir. Analizlerin birinci aşamasında uzman görüşlerince ölçekte yer almasına karar verilen 14 maddenin madde toplam korelasyonlarına bakılmıştır. Madde toplam test puanı korelasyonu 0.20 den düşük olan bir madde (Madde 3) ölçekten çıkarılmıştır (Büyüköztürk, 2006).

Araştırmada KMO değeri .92 bulunmuş ve Barlett'in Küresellik testinin istatistiki olarak anlamlı olduğu ( $\chi^2 = 869.025$ ,  $sd=78$   $p= .00$ ) belirlenmiştir (Büyüköztürk, 2006; Pallant 2005). Bu değerler ölçme aracından elde edilen verilerin örneklem grubuna uygunluğu ve R matrisinin faktörlenebilirliği, bir başka deyişle verilerin faktör analizi için uygunluğu hakkında yeterli kanıtları sağlamaktadır.

Bu tespitin ardından AFA'de maddelerin boyutlarda ne oranda yer aldığını görmek için temel bileşenler analizi (TBA) yapılmıştır. TBA sonucunda özdeğeri 1'den büyük tek bileşen bulunmuştur. Döndürülmemiş temel bileşenler analizi sonucunda özdeğer grafiğinde (scree plot) boyuta ait özdeğerde ani bir düşüşün gözlenmesi, ölçeğin tek boyutlu bir yapıya sahip olduğunu destekler nitelikte bulunmuştur (Şekil 1) (Büyüköztürk, 2006).



Şekil 1. Özdeğer grafiği (scree plot)

Analiz sonucunda toplam varyansın %47.87'ini açıklayan tek faktörlü ölçeğin kuramsal yapıyla uyduğu tespit edilmiştir. AFA sonucunda elde edilen faktör analizi sonuçları Tablo 5'te verilmiştir. Tablo 5'e bakıldığında, ölçek maddelerinin faktör yüklerinin .30 ve üzerinde olduğu belirlenmiştir. Tek boyutta yer alan ve birlikte "Rubrik Kullanmaya İlişkin Özyeterlik" inancını yoklayan maddelerin faktör yükleri .60-.76 arasında değişmektedir.

Tablo 5. RKÖÖ'nin maddeleri ve faktör yükleri

Maddeler	Faktör Yükleri
1. Eğer uğraşırsam, derste kullanacağım bir rubriği (dereceli puanlama anahtarını) hazırlayabileceğime inanıyorum.	.76
2. Rubrik hazırlarken uygun hedef davranışları sağlayan ölçütlere kolayca karar verebileceğime inanıyorum.	.69
3. Ölçüt tanımlarını yaparken uygun dil ve terminoloji kullanabileceğime inanıyorum.	.60
4. Rubrik hazırlarken bir sorunla karşılaşırsam nasıl halledeceğimi bilirim.	.65
5. Öğrenciler daha önce rubrik kullanmamış olsa bile rubrikleri sınıfta uygulayabileceğime inanıyorum.	.68
6. Öğrencilerin seviyeleri yakın olsa bile rubriklerle puanlama yapabileceğime eminim.	.70
7. Eğer farklı rubrik türleri arasında karar vermem gerekirse, uygun rubriği seçebileceğime eminim.	.71
8. Eğer uğraşırsam rubriği öğrenci düzeyine uygun hale getirebileceğime inanıyorum.	.76
9. Uğraşırsam, rubriklerle objektif bir puanlama yapabileceğime eminim.	.76
10. Konu soyut ve zor olsa bile bir rubrik hazırlayabilirim.	.65
11. Eğer uğraşırsam, öğrencilere rubrik kullanmayı benimseteceğime inanıyorum.	.66
12. Az bir zamanım olsa bile, derste kullanmak için bir rubrik hazırlayabilirim.	.65
13. Rubrikleri öğrencilere detaylı geri bildirim vermek için kolayca kullanabileceğime inanıyorum.	.70
Açıklanan Toplam Varyans %	47.873

### Doğrulayıcı Faktör Analizi (DFA)

Analiz sürecinin ikinci aşamasında ikinci çalışma grubundan toplanan ve tek boyutlu yapının, veri setinin diğer parçasından elde edilen veriler arasındaki ilişkileri ne ölçüde açıkladığını incelemek amacıyla tek boyut doğrulayıcı faktör analizi (DFA) yapılmıştır (Jöreskog & Sörbom, 1993). RKÖÖ'ne ilişkin uyum indeksi değerleri  $\chi^2/sd=2.87$ , GFI=.90, AGFI= .86, CFI=.97, NFI= .95, NNFI= .96, IFI= .97, RMSEA= .087, SRMR= .05, PNFI.80 ve PGFI=.64 olarak bulunmuştur. Sınanan modelin uyum indekslerine ilişkin uyum değerleri ile DFA'dan elde edilen uyum indeksi değerleri tablo 6'da gösterilmiştir. Tablo 6'daki uyum indekslerine bakıldığında, DFA sonucunda elde edilen tek faktörlü modelin uyum düzeyinin kabul edilebilir düzeyde olduğu ortaya çıkmaktadır (Byrne, 2010; Kline, 2011; Meyers, Gamst & Guarino, 2006; Schermelleh-Engel et al., 2003)

**Tablo 6.** Uyum İndeksi Değerleri İle DFA'dan Elde Edilen Uyum İndeksi Değerleri

İncelenen Uyum İndeksleri	Mükemmel Uyum Ölçütleri	Kabul Edilebilir Uyum Ölçütleri	Elde Edilen Uyum İndeksleri	Sonuç
$\chi^2/sd$	$0 \leq \chi^2/sd \leq 2$	$2 \leq \chi^2/sd \leq 3$	2.87	Kabul Edilebilir Uyum
GFI	$.95 \leq GFI \leq 1.00$	$.90 \leq GFI \leq .95$	.90	Kabul Edilebilir Uyum
AGFI	$.90 \leq AGFI \leq 1.00$	$.85 \leq AGFI \leq .90$	.86	Kabul Edilebilir Uyum
CFI	$.95 \leq CFI \leq 1.00$	$.90 \leq CFI \leq .95$	.97	Mükemmel Uyum
NFI	$.95 \leq NFI \leq 1.00$	$.90 \leq NFI \leq .95$	.95	Mükemmel Uyum
NNFI	$.95 \leq NNFI \leq 1.00$	$.90 \leq NNFI \leq .95$	.96	Mükemmel Uyum
IFI	$.95 \leq IFI \leq 1.00$	$.90 \leq IFI \leq .95$	.97	Mükemmel Uyum
RMSEA	$.00 \leq RMSEA \leq .05$	$.05 \leq RMSEA \leq .08$	.087	Kabul Edilebilir Uyum
SRMR	$.00 \leq SRMR \leq .05$	$.05 \leq SRMR \leq .10$	.05	Mükemmel Uyum
PNFI	$.95 \leq PNFI \leq 1.00$	$.50 \leq PNFI \leq .95$	.80	Kabul Edilebilir Uyum
PGFI	$.95 \leq PGFI \leq 1.00$	$.50 \leq PGFI \leq .95$	.64	Kabul Edilebilir Uyum

$\chi^2 187.17$ ,  $sd = 65$ , RMSEA için %90 olasılıklı güven aralığı = 0.072 ; 0.10

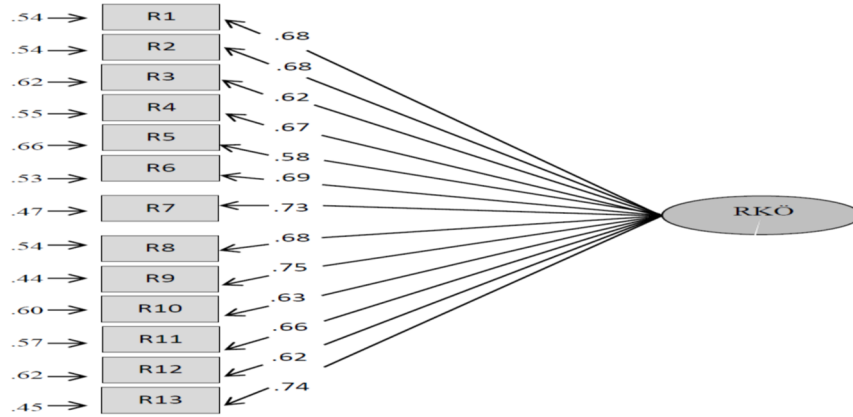
DFA sonucunda elde edilen tek boyutlu modele ilişkin t-testi değerleri ise Tablo 7'de sunulmuştur. Tablo 7'e bakıldığında, rubrik kullanma özyeterliği boyutunda yer alan maddelerin t-testi değerleri 9.73-13.44 aralığında değiştiği görülmektedir. Hesaplanan t değerlerinin 1.96'dan büyük olması .05 düzeyinde; 2.58'den büyük olması ise .01 düzeyinde anlamlı olduğunu ifade etmektedir (Jöreskog & Sörbom,1993; Kline, 2011). Çalışmada, ölçekte yer alan 13 maddenin tümünün DFA'dan elde edilen t testi değerlerinin .01 düzeyinde anlamlı olduğu belirlenmiştir. Kline (2011)'de, anlamsız t testi sonuçları içeren maddelerin modelden çıkarılması ya da katılımcı sayısının artırılması ile anlamlı değerler elde edileceğini vurgulamıştır. Bu çalışmadaki değerlerin anlamlı oluşu araştırmadaki katılımcı sayısının uygunluğunu ve modelden çıkarılması gereken madde olmadığını ortaya koymaktadır.

**Tablo 7.** RKÖÖ için DFA'dan Elde Edilen t-Testi Değerleri

Madde No	t
1	11.80*
2	11.74*
3	10.45*
4	11.68*
5	9.73*
6	12.03*
7	12.93*
8	11.85*
9	13.44*
10	10.79*
11	11.32*
12	10.47*
13	13.41*

\*p&lt;.01 düzeyinde anlamlı

DFA sonucu elde edilen tek boyutlu modelin faktör yükleri ise Şekil 2'de sunulmuştur. Şekil 2 incelendiğinde, rubrik kullanma özyeterliği boyutunda yer alan maddelerin faktör yüklerinin .58-.75 aralığında değiştiği görülmektedir.

**Şekil 2.** Rubrik kullanma özyeterliği ölçeği'ne (RKÖÖ) ilişkin ölçüm modeli

### Güvenirlilik

RKÖÖ'den elde edilen ölçümlerin güvenirliliği için Cronbach Alfa iç tutarlık katsayısı hesaplanmıştır. Ölçümlerin Cronbach Alfa iç tutarlık

katsayısı .91 olarak hesaplanan ölçeğin güvenilirlik katsayısının yeterli olduğu söylenebilir (Büyüköztürk, 2006). Ölçekten alınabilecek en düşük puan 13 ve en yüksek puan 65'dir.

Sonuç olarak DFA, RKÖÖ'nün öğretmen adaylarının rubrik kullanma özyeterliklerini ölçebildiğine kanıt oluşturmaktadır. Bir başka deyişle AFA'nda ortaya koyulan tek boyutlu yapı doğrulanmıştır.

### SONUÇ VE ÖNERİLER

Öğretmen ve öğretmen adaylarının rubrik kullanabilmeleri konusundaki özyeterlik inançları tespit edilerek, öğrenci kazanımlarını değerlendirmeye uygun rubrik gibi ölçme ve değerlendirme araçlarını kullanabilmeleri konusundaki inançları tespit edilebilir. Bunun sonucunda rubrik gibi tamamlayıcı ölçme değerlendirme araçlarını uygun amaçlar için ve uygun şekliyle hazırlayıp uygulayabileceklerine dair gösterecekleri çabalar tahmin edilebilir. O halde "rubrik kullanma" konusundaki inançları ölçmek üzere hazırlanmış ölçeklerle bu konuda sahip oldukları inançlar sorgulanabilir. Bu amaçla çalışmada öğretmen adaylarının rubrik gibi tamamlayıcı ölçme araçlarını hazırlayıp uygulayabilmelerine başka bir deyişle bu ölçme araçlarını kullanabilmelerine ilişkin kendilerine olan inançlarını ölçmeye yarayan geçerli ve güvenilir bir ölçek geliştirilmeye çalışılmıştır.

RKÖÖ taslak formundan elde edilen ölçümlerin yapı geçerliği AFA ve DFA kullanılarak araştırılmıştır. Çalışmanın birinci grubunda AFA sonucunda toplam varyansın %47.87'sini açıklayan 13 maddeden oluşan tek boyutlu bir ölçek elde edilmiştir. İkinci çalışma grubunda tek boyutlu ölçeğin, veri setinin diğer parçasından elde edilen veriler tarafından doğrulanıp doğrulanmadığı DFA ile test edilmiştir. DFA sonucunda, elde edilen uyum indekslerinin yeterli olduğu ve tek faktörlü modelin uyum düzeyinin kabul edilebilir olduğu ortaya çıkmaktadır. AFA sonucunda tek faktörlü ölçeklerde açıklanan varyansın %30 ve daha fazla olması, az sayıda madde içeren ölçeklerde faktör yüklerinin .30 üzerinde olması (Büyüköztürk, 2006) ve DFA'da elde edilen uyum indekslerinin yeterli olması (Byrne, 2010; Kline, 2011; Meyers, Gamst & Guarino, 2006; Schermelleh-Engel et al., 2003), tek faktörlü RKÖÖ'nden elde edilen ölçümlerin yapı geçerliğinin sağlandığının kanıtı olarak gösterilebilir.

RKÖÖ'nden elde edilen ölçümlerin güvenirliliği, Cronbach Alfa yöntemiyle incelenmiştir. Ölçümlerin Cronbach Alfa güvenirliliği .91 olarak bulunmuştur. Cronbach Alfa değerinin .70 ve üzerinde olması ölçümlerin güvenirliliği için yeterli görülmektedir. (Büyüköztürk, 2006). Buna göre Cronbach Alfa yöntemiyle hesaplanan güvenilirlik katsayısı RKÖÖ'nden elde edilen ölçümlerin güvenirliliğinin sağlandığının kanıtı olarak sunulabilir.



Araştırmada ulaşılan bulgular, öğretmen adaylarının rubrik kullanımına ilişkin özyeterlik inançlarını belirlemek için geliştirilen RKÖÖ'nin, geçerli ve güvenilir ölçümler sağlayan bir araç olarak kullanılabilceğini ortaya koymaktadır.

Bu araştırmada geliştirilen ölçeğin, farklı örneklerle yeni geçerlik ve güvenilirlik çalışmalarının yapılması önerilebilir. Bu bağlamda RKÖÖ'nin kullanılacağı yeni araştırmaların yürütülmesinin, ölçeğin ölçme gücüne katkı sağlayacağı düşünülebilir. Ayrıca rubrik kullanma özyeterliliğinin belirlendiği yeni çalışmaların, geniş öğretmen ve öğretmen adayı kitleleri üzerinde yapılarak, onların yeterliklerinin geliştirilmesi ile ilgili olarak rubrik hazırlama etkinliklerine ihtiyaç duyulup duyulmadığı belirlenebilir. Bu bağlamda öğretmenlerin öğrencileri değerlendirirken kullanabileceği tamamlayıcı değerlendirme ve öğrenme araçları olan rubriklerin ilişkili olduğu değişkenlerin (örneğin öğrenci başarıları gibi) nasıl etkilendiğine dair yeni araştırmaların da yapılması önerilebilir.

#### KAYNAKÇA

- Andrade, H. G., Wang, X., Du, Y., & Akawi, R. L. (2009). Rubric-referenced self assessment and self efficacy for writing. *The Journal of Educational Research, 102*(4), 287-302
- Aschbacher, P. R. (1991). Performance assessment: State activity, interest and concerns. *Applied Measurement in Education, 4*(4), 275-288.
- Atılgan, H., Kan, A. ve Doğan, N. (2007). *Eğitimde ölçme ve değerlendirme*. Ankara: Anı Yayıncılık.
- Bahar, M., Nartgün, Z., Durmuş, S. ve Bıçak, B. (2012). *Geleneksel-tamamlayıcı ölçme ve değerlendirme teknikleri öğretmen el kitabı*. Ankara: Pegem A Yayıncılık.
- Baker, E. L., O'Neil, H. F. & Linn, R. L. (1993). Policy and validity prospects for performance based assessment. *American psychologist, 48*(12), 1210-1218.
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review, 84* (2), 191-215.
- Bandura, A. (1994). Self-efficacy, In V.S. Ramachandran (Ed.), *Encyclopedia of human behavior*, (4), 71-81. New York: Academic Press.
- Bentler, P. M. & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588-606.
- Berberoğlu, G. (2006). *Sınıf içi ölçme ve değerlendirme teknikleri*. İstanbul: Morpha.
- Büyüköztürk, Ş. (2006). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem A Yayıncılık.
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications and programming*. New York, NY: Taylor and Francis Group.
- Fidan, M. ve Sak, İ. M. (2012). İlköğretim öğretmenlerinin tamamlayıcı ölçme değerlendirme teknikleri hakkında görüşleri. *Bartın Üniversitesi Eğitim Fakültesi Dergisi, 1*(1), 174-189.

- Frederiksen, N. (1984). The real test bias: Influences of testing on teaching and learning. *American Psychologist*, 39, 193-202.
- Frederiksen, J. R. & Collins, A. (1989). A systems approach to educational testing. *Educational Researcher*, 18 (9), 27-32.
- Gelbal, S. ve Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlik algıları ve karşılaştıkları sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 33, 135-145.
- Goodrich, H. G. (1997). Understanding rubrics. *Educational Leadership*, 54 (4).
- Hair, J. F. Jr., Anderson, R. E., Tatham, R. L., & Black, W.C. (2005). *Multivariate data analysis*, 6th edition, Upper Saddle River, NJ: Prentice Hall.
- Hu, L. & Bentler, P. M. (1999). *Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Structural Equation Modeling*, 6(1), 1-55.
- Jerusalem, M. (2002). Theoretischer Teil - Einleitung I, *Zeitschrift für Pädagogik*, 44, 8-12.
- Jöreskog, K. G. & Sörbom, D., (1993). *Lisrel 8: Structural equation modeling with the simplis command language*, Hillsdale, NJ, Lawrence Erlbaum Associates Publishers.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York: The Guilford Press.
- Kutlu, Ö. (2004). Tek Soru İle Öğrenci Performansının Belirlenmesi. [Online]: <http://www.erg.sabanciuniv.edu/iok2004/bildiriler> adresinden 13 Kasım 2010 tarihinde indirilmiştir.
- Kutlu, Ö., Doğan, C. D. ve Karakaya, İ. (2008). *Öğrenci Başarısının Belirlenmesi: Performansa ve Portfolyaya Dayalı Durum Belirleme*. Ankara: Pegem Akademi.
- Madaus, G. F., & O'Dwyer, L. M. (1999). A short history of performance assessments. *Phi Delta Kappan*, 80 (9), 688-696.
- Marsh, H. W., Hau, K. T., Artelt, C., Baumert, J. & Peschar, J. L. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing*, 6(4), 311-360.
- Meyers, L. S., Gamst, G. & Guarino, A. J. (2006). *Applied multivariate research: Design and interpretation*. London: SAGE Publications.
- Milli Eğitim Bakanlığı (MEB) (2013a). Ortaokul matematik dersi öğretim programı (5, 6, 7 ve 8. sınıflar). [Online]: <http://ttkb.meb.gov.tr/> adresinden 10 Eylül 2015 tarihinde indirilmiştir.
- Milli Eğitim Bakanlığı (2013b). İlköğretim kurumları (ilkokullar ve ortaokullar) fen bilimleri dersi (3, 4, 5, 6, 7 ve 8. sınıflar) öğretim programı. [Online]: <http://ttkb.meb.gov.tr/> adresinden 10 Eylül 2015 tarihinde indirilmiştir.
- Milli Eğitim Bakanlığı (2013c). Ortaöğretim biyoloji dersi (9, 10, 11 ve 12. sınıflar) öğretim programı. [Online]: <http://ttkb.meb.gov.tr/> adresinden 10 Eylül 2015 tarihinde indirilmiştir.
- Milli Eğitim Bakanlığı (2013d). Ortaöğretim fizik dersi (9, 10, 11 ve 12. sınıflar) öğretim programı. [Online]: <http://ttkb.meb.gov.tr/> adresinden 10 Eylül 2015 tarihinde indirilmiştir.

- Milli Eğitim Bakanlığı (2013e). Ortaöğretim kimya dersi (9, 10, 11 ve 12. sınıflar) öğretim programı. [Online]: <http://ttkb.meb.gov.tr/> adresinden 10 Eylül 2015 tarihinde indirilmiştir.
- Özkan, Ö., Tekkaya, C. ve Çakıroğlu, J. (2002). Fen bilgisi aday öğretmenlerin fen kavramlarının anlama düzeyleri, fen öğretimine yönelik tutum ve öz-yeterlik inançları. *V. Ulusal Fen Bilimleri Eğitimi Kongresi*, ODTÜ, Ankara.
- Pallant, J. (2005). SPSS survival manual: A step by step guide to data analysis using SPSS for windows. Australia: Australian Copyright.
- Pajares, F. (1997). *Current directions in self-efficacy research*. In M. Maehr ve P.R. Pintrich (Eds.), *Advances in motivation and achievement*. (Vol. 10, pp. 1-49). Greenwich, CT: JAI Press.
- Pickett, N. & Dodge. B. (2007). Rubrics for web lessons. [Online] Retrieved on 10-September-2010, at URL: <http://webquest.sdsu.edu/rubrics/weblessons.htm>.
- Popham. J. W. (2007). *Classroom Assessment: What Teachers Need to Know*. Pearson Education. 5th Edition. USA.
- Riggs, I. M. & Enochs L. G. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument, *Science Education*, 74(6), 625-637.
- Sağlam-Arslan, A., Devocioğlu-Kaymakçı, 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.
- Schmitz, G.S. ve Schwarzer, R. (2000). Selbstwirksamkeitserwartung von Lehrern: Längsschnitt befunde mit einem neuen Instrument. *Zeitschrift für Pädagogische Psychologie*, 14(1),12-25.
- Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research-Online*, 8(2), 23-74.
- Schumacker, R. E., Lomax, R. G. (2004). *A Beginner's Guide to Structural Equation Modeling*. Second Edition. New Jersey: Lawrence Erlbaum Associates.
- Schwarzer R. (1993). General Perceived Self-Efficacy in 14 Cultures. [Online] Retrieved on 5-June-2007, at URL <http://Web.Fu-Berlin.De/Gesund/Publicat/Ehpscd/Health/World14.Htm>.
- Stiggins, R. J. (1987). Design and development of performance assessments. *Educational measurement: Issues and Practice*, 6 (1). 33-42.
- Tschannen-Moran, M., Woolfolk Hoy, A. & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, 68, 202-248.
- Tschannen-Moran, M. & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing and elusive construct. *Teaching and Teacher Education*, 17, 783-805.
- Yeşilyurt, E. (2012). Fen ve teknoloji dersinde kullanılan ölçme değerlendirme yöntemleri ve karşılaşılan güçlükler. *Turkish Studies - International Periodical For The Languages, Literature And History Of Turkish Or Turkic*, 7(2), 1183-1205.

