



## **Fen Bilimlerine Yönelik Öğrenme Modelinin Ölçme Değişmezliğinin İncelenmesi: PISA 2015 Örneği**

### **Investigating Measurement Invariance of Learning Model Towards Science: PISA 2015 Sample**

Şeyma UYAR<sup>1</sup>, Gülden KAYA UYANIK<sup>2</sup>

#### **Öz**

Bu çalışmada PISA 2015 öğrenci anketinde yer alan öğrenci özelliklerine göre fen bilimlerine yönelik kurulan öğrenme modelinin Türkiye ve Singapur örneklemelerinde dil ve kültür açısından değişmezliği, Türkiye örneğinde ise cinsiyetler arasında değişmezliği incelenmiştir. Bu amaçla yapısal eşitlik modeli temelinde yer alan çoklu grup doğrulayıcı faktör analizi LISREL 8.80 programında uygulanmıştır. Analiz sonuçları modelin Türkiye örneğinde cinsiyetler arasında şekilsel ve metrik değişmezliği sağladığını, ancak skaler (güçlü) değişmezliği sağlayamadığını göstermiştir. Türkiye ile Singapur arasında ise model yalnızca şekilsel değişmezlik koşulunu yerine getirmiştir.

**Anahtar Kelimeler:** fen bilimleri, ölçme değişmezliği, yapısal eşitlik modeli, PISA

#### **Abstract**

In this study, the learning model toward science in PISA 2015 was examined in terms of measurement invariance across Turkey and Singapur samples by cultural invariance and in Turkey sample by gender invariance. For this purpose, multiple group confirmatory factor analysis based on the structural equation model was applied in LISREL 8.80 program. Results of the analyze showed that the model provides configural and metric invariance between gender in Turkey sample but scalar invariance had failed to ensure. The model was ensure only configural invariance across Turkey and Singapore samples.

**Keywords:** science, measurement invariance, structural equation modeling, PISA

1. Mehmet Akif Ersoy Üniversitesi, Eğitim Fakültesi, Burdur, Türkiye; <https://orcid.org/0000-0002-8315-2637>

2. Sakarya Üniversitesi, Eğitim Fakültesi, Sakarya, Türkiye; <https://orcid.org/0000-0002-8100-6994>

**Atf / Citation:** Uyar, Ş., & Kaya Uyanık, G.(2019). Fen Bilimlerine Yönelik Öğrenme Modelinin Ölçme Değişmezliğinin İncelenmesi: PISA 2015 Örneği. *Kastamonu Education Journal*, 27(2), 297-507. doi:10.24106/kefdergi.2570

## Extended Summary

In the field of social sciences such as education and psychology a feature to be measured is being compared between different groups such as gender, culture, class level, socioeconomic status (Dimitrov, 2010; Hirschfeld and Brachel, 2014). However, the findings of the difference between the groups for the same trait can also be based on the psychometric properties of the scale or the difference between individuals (Başusta and Gelbal, 2015). To confirm the validity of obtained results it is necessary to have evidence that the scale measure equally in all groups. In the test development process it is very important to determine whether the scores obtained from different subgroups for similar features are equal (Önen, 2007; Rutkowski and Stevina, 2014). For this reason researchers can investigate measurement invariance, differential item functioning (DIF) or item bias. Measuring invariance refers to the fact that groups with the same level of measure are identical in raw score. Or it means that individuals in different groups are able to perceive and interpret the items in the same way (Bryne and Wakins, 2003; Flowers, Raju and Oshima, 2002). In other words, the property examined by a scale which is provided with invariance can be validly compared in different groups. It can be argued that when the scale is proved to be equal, the group difference does not arise from the measurement tool (Uzun and Öğretmen, 2010).

The interpretation based on the results can also be biased because the measurement tool provides an advantage to groups when the invariance of the measurement tool is not achieved. For this reason, more work needs to be given to studies related to measurement invariance. In particular, it is important that measurement tools used in international comparison in order to direct the education policies of the countries should not include errors.

PISA is one the most popular international assessment in the World. In PISA 2015 science literacy was weighted area. With those results there are many studies compared in terms of countries (cultures) or gender. It must be shown for meaningful comparisons that the means of between gender or between countries are equal. Therefore, in this study it is examined whether the Learning Model towards Science provide measurement invariance across Turkey and Singapore samples or across gender in Turkey sample.

In PISA 2015 assessment 5895 students participated in Turkey and 6115 students in Singapore sample. Following removal of the missing data and extreme values there were 4490 students in Turkey and 4988 students in Singapore sample. There were total 10 variables (adaption of instruction, perceived feedback, disciplinary climate in science classes, teacher-directed science instruction, teacher support in a science classes of students choice are in the first factor; epistemological beliefs, interest in broad science topics, enjoyment of science, instrumental motivation, science self-efficacy are in the second factor) in two factors. In this study, student characteristics model for science was formed by using explanatory factor analysis and confirmatory factor analysis (CFA). In the study, data were firstly examined for missing data, extreme values, single and multivariate normality, linearity and multicollinearity. The model validated by CFA were examined then for invariance across gender in Turkey sample and across culture in Turkey and Singapore samples. Multiple group confirmatory factor analysis was used to examine the measurement invariance. The relationship between the items and the factors has been tested hierarchically. The model was tested with CFA in terms of model-data fit across subgroups and whole data. When the student characteristics are examined it is seen in Turkey data RMSEA = 0.054, CFI = 0.98, the NNFI = 0.93, GFI = 0.98 and SRM = 0.04. In Singapore, RMSEA = 0.046, CFI = 0.98, NNFI = 0.97, GFI = 0.98 and SRMR = 0.032. And RMSEA = 0,058, CFI = 0.93, NNFI = 0.90, GFI = 0.98 for males, while  $\chi$ RMSEA = 0.048, CFI = 0.96, NNFI = 0.95, GFI = 0.98 and SRMR = 0.97 and SRMR = 0.046 for females.

Learning model towards science has provided only configural and metric invariance across gender in Turkey sample. It has provided only configural invariance between Turkey and Singapore samples.

## 1. Giriş

Eğitim ve psikoloji gibi alanlarda genellikle ölçülmek istenen bir özelliğin cinsiyet, kültür, sınıf düzeyi, sosyoekonomik durum gibi farklı gruplar arasında karşılaştırıldığı gözlenmektedir (Dimitrov, 2010; Hirschfeld ve Brachel, 2014). Bu karşılaştırmalardan elde edilen herhangi bir sonuç yorumlanırken ölçülen özelliğe ilişkin yapının farklı gruplarda benzer fonksiyon gösterdiği varsayılmaktadır (Cheung ve Rensvold, 1999; Vandenberg ve Lance, 2000). Oysa aynı özellik için gruplar arasında fark olduğu yönünde elde edilen sonuçlar bireyler arasındaki gerçek farklılıktan ya da kullanılan ölçme aracının psikometrik özelliğinden kaynaklanabilmektedir (Başusta ve Gelbal, 2015). Bu nedenle elde edilen sonuçların değerlendirilmeye alınması ve doğruluğunun onaylanması için, ölçme aracının konusu ne olursa olsun, ölçme aracıyla ölçülen özelliğin tüm gruplarda benzer şekilde anlaşıldığı ve aracın eşit ölçme yaptığına dair kanıtların olması gereklidir (Önen, 2007; Rutkowski ve Stevina, 2014). Bu kanıtlar ölçmenin yanlı olup olmadığı (Lord, 1980), değişen madde fonksiyonu (DMF) (Hambleton ve Rogers, 1989; Swaminathan ve Rogers, 1990) ya da ölçme değişmezliği (Meredith, 1993) incelemelerine yönelik olabilmektedir.

Ölçme değişmezliği (ya da ölçme eş değeri) ölçülen özellik bakımından eşit düzeyde olan grupların ölçme aracından aldıkları ham puanın aynı olması ya da farklı gruplarda yer alan bireylerin ölçme aracındaki maddeleri aynı şekilde algılama ve yorumlaması anlamına gelmektedir (Bryne ve Wakins, 2003; Flowers, Raju ve Oshima, 2002). Bir başka deyişle, değişmezliği sağlanan bir araçla incelenen özellik farklı gruplarda geçerli bir şekilde karşılaştırılabilmektedir. Ölçme aracının değişmez olduğu ortaya konulduğunda gruplar arasında gözlenen farklılık örtük özelliğe (ölçülen özelliğe) dayandırılabilir ve farklılığın en azından ölçme aracından kaynaklanmadığı söylenebilir (Uzun ve Öğretmen, 2010).

Yapılan birçok değişmezlik çalışmasında ölçme araçlarının dil, kültür, cinsiyet ya da bölge gruplarında ölçme değişmezliğini sağlamadığına rastlanmaktadır (Ayvalli, 2016; Du ve Tang, 2005; Kıbrıslıoğlu, 2015; Önen, 2007; Tucker, Ozer, Lyubomirsky ve Boehm, 2006, Uzun, 2008; Uyar ve Doğan, 2014; Wu, Li ve Zumbo, 2007). Bu durum "ölçülen yapının eşit olduğu" yönündeki varsayımın çoğu durumda yerine getirilmediğini göstermektedir (Başusta ve Gelbal, 2015). Ölçme aracının değişmezliği sağlanmadığında ölçme aracı gruplardan birine avantaj sağlayacağından sonuçlara dayalı olarak yapılan yorumlar da yanlı olabilmektedir. Bu puanlarla yapılan yorumlar çalışmaların niteliğini zedelerken bir yandan da alınan kritik kararlar bireylerin hayatına yanlış yön verebilmektedir. Özellikle uluslararası karşılaştırmalarda ülkelerin eğitim politikalarına yön verebilmeleri açısından kullanılan ölçme araçlarının hatalardan arınık olması ve karşılaştırılan sonuçların doğruluğu önem taşımaktadır.

Eğitim alanında, Uluslararası Matematik ve Fen Eğilimleri Araştırması (TIMSS) ve Uluslararası Öğrenci Değerlendirme Programı (PISA) öğrencilerin fen, matematik, okuma becerileri ve finansal okuryazarlık alanlarındaki başarılarını ve tutumlarını ölçmeye yönelik uluslararası büyük ölçekli değerlendirmelerdir. Bu uygulamalarda başarı ya da tutumlar hem ülke genelinde hem de ülkeler arasında karşılaştırılmaya olanak bulmaktadır. PISA'nın altıncı döngüsü olan PISA 2015 uygulaması, 35'i OECD üyesi olmak üzere 72 ülke ve ekonomideki yaklaşık 29 milyon öğrenciyi temsilen 540.000'e yakın öğrencinin katılımıyla 2015 yılı içerisinde gerçekleştirilmiştir. PISA 2015 uygulamasında ağırlıklı alan olarak fen okuryazarlığı ele alınmıştır (Milli Eğitim Bakanlığı [MEB], 2015).

PISA'da kullanılan fen okuryazarlık terimi ile öğrencilerin fen alanında bildiklerinin yanı sıra bunlarla ne yapabildiği ve bilimsel bilgiyi gerçek hayatta nasıl yaratıcı bir şekilde uygulayabildiği değerlendirilmektedir (MEB, 2015). PISA 2015 uygulamasında tüm alanlarda en başarılı ülke Singapur olmuştur. Asya ülkeleri de (Hong Kong, Tayvan, Japonya, Çin ve Güney Kore) başarı bakımından Singapur'u takip eden ülkelerdendir. Finlandiya, Estonya, Kanada ve İrlanda, Asya ülkeleri dışından olup PISA 2015'te ilk 10 sırada başarı gösteren ülkeler arasındadır. Türkiye ise OECD ortalama puanı olan 493 ve katılımcı tüm ülkelerin ortalama puanı olan 465'in altında kalarak 425 ortalama puan başarı göstermiştir (OECD, 2016). Öte yandan MEB (2015) ulusal raporunda fen okuryazarlığına yönelik duyuşsal özellikler bakımından Türkiye'deki öğrencilerin ilgi ve motivasyon düzeylerinin OECD ortalamasından daha yüksek olduğu, fen dersinden daha çok zevk aldıkları ve fen alanında kendilerini OECD ortalamasına göre daha yeterli gördükleri belirtilmiştir. Fen ile ilintili bir meslek sahibi olmayı bekleyen öğrenci oranının OECD ortalamasına göre daha yüksek olduğu ve öğrencilerin genel olarak fene yönelik olumlu bir tutuma sahip olduğu ortaya konmuştur (MEB, 2015).

PISA uygulamalarının duyuşsal birçok özelliği ölçmeye yönelik çabalarının olması buradan elde edilen verilerle cinsiyete, bölgelere ya da ülkelere göre karşılaştırma yapma olanağı sunmaktadır. Özellikle öğrenci başarısının duyuşsal özelliklerle olan ilişkisine yönelik kızların ve erkeklerin ya da ülkelerin performanslarının karşılaştırıldığı sonuçlar bu uygulamada pek çok kez rapor edilmektedir (OECD, 2016). Ancak, PISA raporları da dahil farklı gruplarda yer alan bireylerin karşılaştırıldığı sosyal bilimlere yönelik çalışmalarda kullanılan ölçme araçlarının geçerliliğine ilişkin verilen

bilgilerin sınırlı kaldığı söylenebilir. Özellikle ölçme araçlarının ölçme değişmezliğine ilişkin çalışmalara yer verilmemesi ortak bir sorun olarak ele alınabilir. Steenkamp ve Baumgartner (1998), yapılan çalışmalarda hayli önemli olan ölçme değişmezliğinin ihmal ediliyor olması hakkında bazı görüşler öne sürmüşlerdir. Ölçme değişmezliğinin aşamaları hakkında uzmanlar tarafından görüş birliğinin sağlanamamış olması, araştırmacıların bunların test edilmesine yönelik bilgiye ulaşamaması ya da kullanılan yöntemlerin karmaşık olması, kültürlerarası karşılaştırmalarda eşitliğin sağlanması ile ilgili ölçüt bulunmaması ve bunun için bir rehberin olmadığı yönünde ifadelerde bulunmuşlardır. Tüm bu nedenler ölçme değişmezliği ile ilgili çalışmalara daha fazla yer verilmesi gerektiğini bir kez daha gözler önüne sermektedir.

PISA gibi büyük ölçekli sınavlardan elde edilen sonuçlar ülkelerin eğitim sistemlerine yön vermede oldukça önemlidir. Bu nedenle ülkelerin başarılarının karşılaştırıldığı çalışmalara genellikle başarı bakımından üst sıralarda yer alan ülkelerin de dahil edildiği söylenebilir (Kıbrıslıoğlu, 2015). Farklı ülkelerde yer alan bireylerin dilleri ve kültürleri de farklılık gösterebilmektedir. Dolayısıyla bu çalışmalardan elde edilen sonuçlardan iki anlam çıkabilir. Birincisi puan farklılıkları gerçekten ülkeler arasındaki başarı farklılıklarından kaynaklanmaktadır. İkincisi, her ne kadar bireyler sınavı kendi diline uyarlanmış biçimde alsalar da ölçme aracındaki hata bu farklılığın sebebi olabilir (Steenkamp ve Baumgartner, 1998). Bulunan herhangi bir farkı cinsiyete, kültüre ya da farklı bir alt gruba dayandırabilmek ancak, ölçülen yapının ve uygulanan maddelerin gruplar arasında değişmez olduğunun test edilmesiyle mümkündür (Vandenberg ve Lance, 2000).

Bu nedenle ilgili çalışmada fen bilimlerine yönelik öğrenci anketinde yer alan öğrenci özellikleri yardımıyla kurulan modelin ülke ve cinsiyet grupları arasında değişmezliği incelenmiştir. Bu çalışmada ülke karşılaştırmasında temel olarak dikkat edilen nokta ülkelerin farklı dile ve kültüre sahip olmasıdır. Bunun yanında diğer bir ölçüt ise ülkelerin başarısıdır (OECD, 2016). Bu nedenle PISA 2015 uygulamasında tüm alanlarda birinci sırayı alarak en yüksek performansı sergileyen Singapur ile başarı bakımından ortalamanın altında kalan Türkiye arasında karşılaştırma yapmak uygun görülmüştür. Öte yandan hemen hemen bütün araştırmalara cinsiyet değişkeninin dahil edilmesi bu araştırmada kurulan modelin Türkiye örneğinde, cinsiyet grupları arasında değişmezliği sağlayıp sağlamadığı sorusunu akla getirmektedir. Bu ihtiyaç doğrultusunda araştırmada ele alınan temel problem “fen bilimlerine yönelik öğrenme modeli ülkeler arasında (Singapur-Türkiye) ve cinsiyetler arasında ölçme değişmezliğini sağlamakta mıdır? şeklindedir.

## 2. Yöntem

### Araştırmanın Türü

Araştırmada PISA 2015 öğrenci anketinde yer alan fen bilimlerine ilişkin anket maddelerinin kültür ve cinsiyet açısından eşdeğer olup olmadığı incelenmiştir. Bu yönüyle çalışma, PISA 2015 değerlendirmesinin geçerlik düzeyini saptamaya yönelik betimsel bir araştırma niteliğindedir.

### Evren ve Örneklem

PISA 2015 uygulaması 540.000'e yakın öğrencinin katılımıyla gerçekleştirilmiştir. İlk aşamada İstatistikî Bölge Birimleri Sınıflaması (İBBS) Düzey 1, eğitim türü, okul türü, okulların buldukları yer ve okulların idari biçimleri tabakaları kullanılarak okullar tabakalı seçkisiz örnekleme yöntemiyle belirlenmiştir, ikinci aşamada ise bu okullarda uygulamaya katılacak olan öğrenciler seçkisiz yöntemle belirlenmiştir. PISA 2015 uygulamasına Türkiye’de İBBS Düzey 1’e göre 12 bölgeyi temsil eden 61 ilden 187 okul ve 5895 öğrenci katılmıştır. Singapur örneğinde ise 6115 öğrenci yer almıştır. Ancak, analiz için uygun olmayan veriler çıkarıldıktan sonra Türkiye örneğinde 4490, Singapur örneğinde 4988 veri ile çalışma yürütülmüştür.

### Veri Toplama Araçları

PISA başarı testleri ile belirlemeye çalıştığı fen okuryazarlığının yanında öğrencilerin akademik başarıları ile ilişkili duyuşsal özelliklere ilişkin olarak da belirlemeler yapmaya çalışmaktadır. Bu bağlamda öğrencilerin fen öğrenme motivasyonları (fen öğrenmeyi öğrencilerin ne kadar ilginç ve eğlenceli buldukları), fen konularına ilgileri ve fen öğrenmeye araçsal güdülenmeleri (onların fen dersini gelecekteki eğitimlerinde ve kariyer planlarında yararlı olarak algılayıp algılamadıklarını), fen öz yeterliği (öğrencilerin kendi becerileriyle etkin bir şekilde fen ödevlerinin ve zorlukların üstesinden gelebilme düzeyleri) konularında öğrencilere anketler yoluyla sorular sorulmuştur (MEB, 2015). Çalışmada ele alınan öğrenci özellikleri Tablo 1’de yer almaktadır.

**Tablo 1. Fen bilimlerine yönelik öğrenci özellikleri**

Fene yönelik duyuşsal özellikler	Fen dersinde sınıf yönetimi - SINIFYÖN ( Disciplinary climate in science classes)
	Fen dersinde öğrencinin isteğine bağı öğretmen desteğı - ÖDESTEGİ (Teacher support in a science classes of students choice)
	Öğretmen merkezli anlatım - ÖMANLATI (Teacher-directed science instruction)
	Öğretime adaptasyon – ÖADAPT (Adaption of instruction)
	Algılanan geribildirim - ALGBILD ( Perceived Feedback)
	Fen öğrenmekten zevk alma - FENZEVK(Enjoyment of science)
	Geniş fen konularına ilgi – GFENILGI (Interest in broad science topics)
	Fen öğreniminde araçsal güdülenme- ARACSALG (Instrumental motivation)
	Fen özyeterliğı - FENÖZYET (Science self-efficacy)
	Epistemolojik inanç - EPIST (Epistemological beliefs)

Tablo 1’de göre OECD tarafından yapılandırılan ve PISA 2015 uygulamasında yer alan anketler aracılığıyla oluşturulan fen bilimlerine yönelik öğrenme modelinin öğretmene yönelik duyuşsal özellikler faktörü için beş, fene yönelik duyuşsal özellikler faktörü için beş tane öğrenci özelliğı ele alınmaktadır.

### Verilerin Analizi

Bu çalışmada fen bilimlerine yönelik öğrenci özellikleri modeli açımlayıcı faktör analizi (AFA) ve doğrulayıcı faktör analizi (DFA) yardımıyla oluşturulmuştur. Analizlerden doğru bir sonuç elde edebilmek amacıyla öncelikle veri seti ve verinin yapısı, AFA ve DFA analizlerinin gerektirdiğı varsayımlara uygunluk açısından incelenmiştir. Bu amaçla kayıp veri, uç değeri, tek ve çok değerişkenli normallik ile doğrusallık ve çoklu bağlantı incelemeleri yapılmıştır.

Doğrulayıcı faktör analizi ile doğrulanan model çoklu grup doğrulayıcı faktör analizi ile Türkiye örnekleminde cinsiyete göre, Türkiye ve Singapur arasında ise kültüre göre değerişmezlik bakımından incelenmiştir. Ölçme değerişmezliğini incelemek amacıyla sıklıkla kullanılan yöntem çoklu grup doğrulayıcı faktör analizidir (Joreskog, 1971; Wu et al., 2007). Bu yöntemle maddeler ve faktörler arasındaki ilişki hiyerarşik olarak dört adımda test edilebilmektedir. (Meredith & Teresi, 2006; Vandenberg and Lance, 2000). Tablo 2’de şekilsel, metrik, skaler ve katı değerişmezlik şeklinde isimlendirilen ve aşamalı olarak test edilen değerişmezlik adımları verilmektedir.

**Tablo 2. Ölçme değerişmezliğı aşamaları**

Şekilsel Değerişmezlik	Faktör yapısının gruplar arasında eşitliğı test edilir (her bir grupta kurulan model için kaç faktör var ve hangi maddeler hangi faktörlerle ilişkili)	Tüm parametrelerin gruplar arasında serbest bırakıldığı aşamadır. Kurulan modelin her iki grupta da yapı bakımından benzer olup olmadığı sonucuna varılır.
Metrik (zayıf) değerişmezlik	Faktör yüklerinin gruplar arasında eşitliğı test edilir.	Bu değerişmezlik aşaması sağlanırsa bir sonraki aşama test edilebilir. Eğer sağlanırsa grupların maddeleri benzer şekilde yanıtladığı yorumu yapılır.
Skalar (güçlü) değerişmezlik	Faktör yüklerinin yanında regresyon sabitlerinin gruplar arasında eşitliğı test edilir.	Bu aşamada değerişmezlik sağlanırsa bir sonraki adıma geçilebilir. Eğer sağlanırsa madde ortalamalarının gruplar arasında aynı olduğu ifade edilir.
Katı değerişmezlik	Hata varyanslarının eşitliğı test edilir.	Madde ortalamaları ve korelasyonlarının eşit olduğu ve maddeye verilen yanıtların aynı olduğu sonucuna varılır.

Gregorich, (2006).

Bu çalışmada ölçme değerişmezliğı, sınırlandırılmamış modelden elde edilen uyum istatistikleri ile daha sınırlı modelden elde edilen uyum istatistikleri arasındaki farka dayalı olarak incelenmiştir. Bu amaçla kurulan modeller arasındaki CFI değerişimleri arasındaki farklar incelenerek -0.01 değerişiminden küçük, ya da 0.01 değerişiminden büyük farklar dikkate alınarak modelin ilgili değerişmezlik koşulunu sağalamadığına karar verilmiştir.

Analize başlamadan önce Türkiye ve Singapur örneklemlerinde kayıp veriler incelenmiştir. Her bir değerişken için kayıp veri sayısının %5’ten fazla olması ve bunların rastgele dağılım göstermemesi nedeniyle kayıp veriler analiz dışında bırakılmıştır (Tabachnick & Fidell, 2007). Bu aşamadan sonra uç değerişimleri, z değerişimlerine (-3, +3) ve Mahalanobis uzaklıkları

ki-kare değeri ( $p < .001$ ) dikkate alınarak incelenmiştir. Bu aralığın dışında kalan veriler incelendiğinde Türkiye örnekleminden 239, Singapur örnekleminden 675 kişi uç değer olarak kabul edilmiştir.

Ölçme değişmezliği için çok gruplu doğrulayıcı faktör analizinin uygulanabilmesi için verilerin normalliği basıklık ve çarpıklık değerlerine ve histogramlarına bakılarak incelenmiş ve buna ilişkin sonuçlar Tablo 3'te gösterilmiştir. Her iki örnekleme de 10 tane değişken için basıklık ve çarpıklık değerleri -1 ile +1 arasında değişmektedir.

**Tablo 3. Değişkenlere ilişkin normallik sonuçları**

		Ortalama	Medyan	Mod	Çarpıklık	Basıklık
Türkiye	SINIFYÖN	-0,12	0,00	0,00	0,00	0,40
	ÖDESTEĞİ	0,21	0,08	1,45	-0,30	-0,42
	ÖMANLATI	-0,06	-0,01	-0,61	-0,09	0,73
	ALGBILD	0,33	0,27	0,01	0,12	0,06
	ÖADAPT	0,11	0,02	-0,38	0,01	0,22
	ZEVKFEN	0,15	0,51	0,51	-0,14	-0,25
	GFENILGI	0,72	0,72	0,73	0,61	0,38
	ARAÇSALG	-0,18	-0,19	-0,19	0,16	0,18
	FENÖZYET	0,40	0,37	0,37	-0,27	-0,06
	EPIST	0,42	0,32	0,32	0,77	0,90
Singapur	SINIFYÖN	0,24	0,16	0,00	0,12	-0,28
	ÖDESTEĞİ	0,32	0,19	1,45	-0,08	-0,85
	ÖMANLATI	0,28	0,26	0,45	0,01	0,39
	ALGBILD	0,32	0,17	0,01	0,10	0,29
	ÖADAPT	0,42	0,37	0,65	0,06	0,12
	ZEVKFEN	0,66	0,51	0,51	-0,03	0,19
	GFENILGI	0,39	0,42	0,53	0,33	0,85
	ARAÇSALG	0,57	0,37	0,37	0,00	-0,37
	FENÖZYET	-0,27	-0,30	-0,50	0,01	-0,31
	EPIST	0,27	-0,15	-0,19	0,73	0,26

Doğrusallık için saçılım grafiklerinden yararlanılmıştır. Grafikler incelendiğinde verilerin bu sayıltıyı karşıladığı görülmüştür. Son olarak tolerans, varyans şişkinlik (Variance inflation factor) ve koşul indeksi (Condition index) değerleri incelenerek çoklu bağlantı araştırılmış ve buna ilişkin sonuçlar Tablo 4'te verilmiştir. Tolerans değerinin .01'den büyük, varyans şişkinlik değerinin 10 dan küçük ve koşul indeksi 30'dan küçük olduğu için verilerde çoklu bağlantı bulunmadığı görülmüştür (Büyüköztürk, 2010; Tabachnick & Fidell, 2007).

**Tablo 4. Değişkenlere ilişkin çoklu bağlantı sonuçları**

	B	Std. Hata	Beta	Tolerans	Varyans Şişkinlik	Koşul indeksi	
Sabit	2420,22	25,26					
Türkiye	SINIFYÖN	-96,01	15,75	-0,04	-0,09	-0,04	1,17
	ÖDESTEĞİ	2,16	8,60	0,00	0,00	0,00	1,54
	ÖMANLATI	7,77	9,46	0,01	0,01	0,01	1,74
	ALGBILD	-31,13	9,81	-0,02	-0,05	-0,02	2,52
	ÖADAPT	-2,30	10,33	0,00	0,00	0,0	7,58
	ZEVKFEN	-47,77	7,93	-0,04	-0,09	-0,04	1,81
	GFENILGI	-1175,56	8,64	-0,89	-0,90	-0,81	1,96
	ARAÇSALG	1,25	7,24	0,00	0,00	0,00	2,30
	FENÖZYET	-14,71	9,26	-0,01	-0,02	-0,01	2,08
	EPIST	-49,11	6,97	-0,04	-0,11	-0,04	2,20

	B	Std. Hata	Beta	Tolerans	Varyans Şişkinlik	Koşul indeksi
Sabit	2731,24	10,29				
SINIFYÖN	-3,24	9,31	0,00	-0,01	0,00	1,87
ÖDESTEGİ	3,40	11,40	0,00	0,00	0,00	2,24
ÖMANLATI	-10,73	9,92	0,1	-0,02	0,01	2,30
ALGBILD	-19,63	9,74	-0,01	-0,03	-0,01	3,71
ÖADAPT	-21,13	11,90	-0,01	-0,03	-0,01	3,88
ZEVKFEN	-31,54	10,00	-0,02	-0,05	-0,02	2,56
GFENILGI	-19,32	11,46	-0,01	-0,02	-0,01	2,98
ARAÇSALG	39,20	9,90	0,02	0,06	0,02	3,53
FENÖZYET	-3,79	11,18	0,00	-0,01	0,00	3,04
EPIST	-1378,29	8,38	-0,93	-0,92	-0,83	3,17

### 3. Bulgular

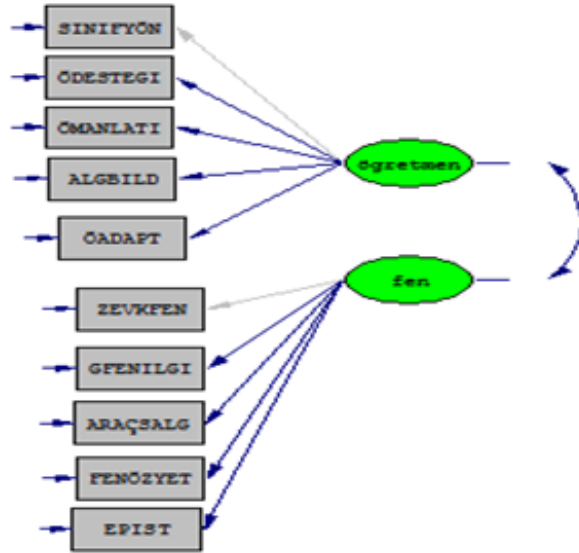
Bu araştırmada ele alınan temel problem fen bilimlerine yönelik öğrenme modelinin cinsiyet ve ülke gruplarında ölçme değişmezliğini sağlayıp sağlamadığı yönündedir. Bu amaçla varsayımların incelenmesinin ardından 10 değişken için verilerin açımlayıcı faktör analizine uygunluğu incelenmiştir. Sonuçlar Türkiye (KMO=0,749, Bartlett küresellik testi =5613.004, sd= 45 ve p<.01) ve Singapur (KMO=0,821, Bartlett küresellik testi =10660,85, sd= 45 ve p<.01) verilerinin analize uygun olduğunu göstermiştir AFA sonucunda verilerin her iki örnekleme iki boyutta toplandığı görülmüştür. Türkiye örnekleminde bu iki boyut birlikte toplam varyansın % 39'unu, Singapur verilerinde ise %48'ini açıklamaktadır. AFA sonucunda elde edilen faktör yükleri Türkiye ve Singapur örneklemleri için Tablo 5'te verilmektedir.

**Tablo 5. Açımlayıcı faktör analizi sonuçları**

	Türkiye		Singapur	
	Faktör			
	1	2	1	2
SINIFYÖN	0,805		0,834	
ÖDESTEGİ	0,743		0,785	
ÖMANLATI	0,698		0,749	
ALGBILD	0,469		0,599	
ÖADAPT	0,437		0,595	
ZEVKFEN		0,584		0,687
GFENILGI		0,332		0,311
ARAÇSALG		0,564		0,639
FENÖZYET		0,708		0,762
EPIST		0,628		0,685

Tablo 5'te yer alan değerlere göre, AFA sonucunda elde edilen ilk faktör öğretmene yönelik duyuşsal özellikler, ikinci faktör ise fen bilimlerine yönelik duyuşsal özellikler olarak adlandırılmıştır. Türkiye verilerinde birinci boyutta beş tane gözlenen değişken bulunmaktadır ve faktör yükleri 0,33 ile 0,81 arasında değişmektedir. İkinci alt boyutta faktör yükleri 0,43 ile 0,70 arasında değişen beş değişken yer almaktadır. Singapur verilerinde ise birinci boyutta 0,31 ile 0,83 arasında faktör yüküne sahip beş değişken bulunurken, ikinci boyutta faktör yükleri 0,76 ile 0,59 arasında değişen beş değişken yer almaktadır.

AFA sonucunda teorisi kurulan model Şekil 1'de verilmektedir.



**Şekil 1. Fen Bilimlerine yönelik öğrenci özelliklerine ilişkin teorisi kurulan model**

Teorisi kurulan model ölçme değişmezliği bakımından incelenmeden önce DFA ile tüm veride ve Türkiye, Singapur, kız ve erkek alt gruplarında model-veri uyumu bakımından test edilmiştir. Elde edilen uyum istatistikleri Tablo 6'da yer almaktadır.

**Tablo 6. DFA sonuçlarına göre alt gruplarda elde edilen uyum istatistikleri**

	RMSEA	SRMR	AGFI	GFI	NFI	CFI
Türkiye	0,054	0,040	0,98	0,97	0,93	0,94
Singapur	0,046	0,032	0,98	0,97	0,97	0,98
Kız	0,048	0,037	0,98	0,97	0,95	0,96
Erkek	0,058	0,046	0,97	0,96	0,90	0,93

Tablo 6'da fen bilimlerine yönelik öğrenci özellikleri modelinin uyum istatistikleri incelendiğinde, Türkiye verilerinde RMSEA = 0,054, CFI = 0,98, NNFI=0,93, GFI = 0,98 ve SRMR= 0,04 olarak bulunduğu görülmektedir. Singapur verilerinde RMSEA=0,046, CFI=0,98, NNFI=0,97, GFI = 0,98 ve SRMR = 0,032 olarak bulunmuştur. Kızlar alt grubunda RMSEA=0,048,CFI=0,96 NNFI=0,95, GFI=0,98 ve SRMR=0,037 olarak bulunurken, erkekler için RMSEA=0,058, CFI=0,93, NNFI=0,90, GFI = 0,97 ve SRMR = 0,046 olarak bulunmuştur. Bu değerler, modelin tüm gruplarda kabul edilebilir olduğunu göstermektedir. Ayrıca her bir grup için elde edilen standartlaştırılmış yol katsayılarının 0,22 ile 0,85 arasında değişen değerler aldığı, bu katsayıların t değerleri incelendiğinde istatistiksel olarak anlamlı olduğu ifade edilebilir.

Doğrulan fen bilimlerine yönelik öğrenme modelinin cinsiyet gruplarında değişmezliği çoklu grup doğrulayıcı faktör analizi kullanılarak incelenmiştir. Değişmezlik aşamalarından elde edilen uyum istatistikleri Tablo 7'de verilmektedir.

**Tablo 7. Cinsiyete göre değişmezlik sonuçları**

		$\chi^2$ katkısı (%)	GFI	$\chi^2$	df	RMSEA	CFI	NNFI	$\Delta$ CFI
Şekilsel	K	43,62	0,98	502,89	70	0,052	0,95	0,93	-
	E	56,38	0,97						
Metrik	K	43,63	0,98	585,53	78	0,054	0,94	0,93	0,01
	E	56,37	0,97						
Skalar	K	47,02	0,91	3237,23	97	0,120	0,66	0,69	0,28
	E	52,98	0,91						
Kati	K	47,27	0,91	3295,89	107	0,115	0,65	0,71	
	E	52,73	0,90						

K: Kız, E: Erkek



Tablo 7'ye göre modelin cinsiyet gruplarında şekilsel değişmezliği incelendiğinde uyum indekslerinin kabul edilebilir aralıkta olduğu görülmektedir (RMSEA < 0,08; CFI > 0,90; NNFI > 0,90). Buna durumda cinsiyete göre şekilsel değişmezliğin sağlandığı ve modelin yapısının her iki grupta benzer olduğu söylenebilir. Şekilsel değişmezliğin sağlanması metrik değişmezlik için bir önkoşuldur. Metrik değişmezlik aşamasında modele faktör yüklerinin gruplar arasında değişmez olduğu sınırlaması getirilmiştir. Uyum indeksleri incelendiğinde metrik değişmezlik aşamasında model veri uyumunun sağlandığı görülmektedir (RMSEA > 0,08; CFI > 0,90; NNFI > 0,90). Metrik değişmezlik ile sınırlandırılmamış model olan şekilsel değişmezlikten elde edilen CFI değerleri arasındaki fark ( $\Delta CFI = 0,01$ ) olduğundan bu koşulun sağlandığı söylenebilir. Buna göre grupların faktör yükleri de benzer şekildedir. Ancak bu durum gözlenen değişken ortalamalarının karşılaştırılabilirliğini sağlamamaktadır, yalnızca ölçek değişmezliği için bir önkoşuldur. Ölçek değişmezliği aşamasında madde sabitlerine gruplar arasında eşit olma sınırlaması getirilmiştir. Uyum indeksleri incelendiğinde model veri uyumunun sağlanmadığı görülmektedir (RMSEA > 0,08; CFI < 0,90; NNFI < 0,90). Buna göre gruplarda regresyonların paralellliği sağlanmamaktadır. Ölçek değişmezliğinin sağlanmaması katı değişmezliğin de sağlanmayacağını göstermektedir. Tablo 7 incelendiğinde katı değişmezlik aşamasında uyum indekslerinin kabul edilebilir aralığın dışında kaldığı görülmektedir (RMSEA > 0,08; CFI < 0,90; NNFI < 0,90).

Fen bilimlerine yönelik öğrenme modelinin ülkeler arasındaki değişmezliğine ilişkin bulgular Tablo 8'de verilmektedir.

**Tablo 8. Ülkelere göre değişmezlik sonuçları**

		$\chi^2$ katkısı (%)	GFI	$\chi^2$	df	RMSEA	CFI	NNFI	$\Delta CFI$
Şekilsel	T	45,01	0,98	927,70	70	0,051	0,97	0,96	-
	S	54,99	0,98						
Metrik	T	40,80	0,98	257,82	78	0,057	0,95	0,95	0,02
	S	59,20	0,97						
Skalar	T	30,96	0,97	605,48	97	0,109	0,81	0,83	
	S	69,04	0,90						
Katı	T	44,25	0,88	10174,17	107	0,141	0,68	0,73	
	S	55,75	0,85						

T: Türkiye, S: Singapur

Tablo 8'e göre fen bilimlerine yönelik öğrenci özellikleri modeli Türkiye ve Singapur ülkeleri arasında şekilsel değişmezliği sağlamıştır (RMSEA < 0,08; CFI > 0,90; NNFI > 0,90). Bu durum modelin yapısının her iki grupta aynı olduğunu göstermektedir. Şekilsel değişmezlik sağlandığı için metrik değişmezlik aşamasına geçilmiştir. Faktör yükleri gruplarda aynı olacak şekilde sınırlandırıldığında modelin yine veri ile uyumlu olduğu görülmektedir (RMSEA < 0,08; CFI > 0,90; NNFI > 0,90). Şekilsel değişmezlik ile metrik değişmezlik arasındaki CFI farkları incelendiğinde metrik değişmezliğin ülkeler arasında karşılanmadığı söylenebilir ( $-0,01 \leq \Delta CFI \leq 0,01$ ).

Buna göre Türkiye ve Singapur arasında ölçülen özellikler ile ölçeğin boyutları arasındaki ilişkilerin benzer olmadığı ve maddelere benzer biçimde yanıt vermedikleri yorumu yapılabilir. Bu durum gruplardan elde edilen puanların karşılaştırılmasının yanlı olabileceğini göstermektedir.

#### 4. Sonuçlar

Bu araştırmada, PISA 2015 öğrenci anketinde yer alan anketler yardımıyla oluşturulan "fen bilimlerine yönelik öğrenme modelinin" ölçme değişmezliği incelenmiştir. Model, Türkiye örneğinde cinsiyete göre yalnızca şekilsel ve metrik değişmezliği sağlamıştır. Güçlü (skalar) değişmezlik koşulu yerine getirilememiştir. Cinsiyetler arası ölçme değişmezliğinin yürütüldüğü çalışmalarda genellikle ölçeklerin en azından zayıf (metrik) ölçme değişmezliğini ya da kısmi değişmezliği sağladığı görülmüştür (Grouzet Otis ve Pelletier, 2006; Hirschfeld ve Brown, 2009; Uyar ve Dogan, 2014; Uzun ve Ogretmen, 2010). Öte yandan cinsiyetler arasında tam değişmezlik koşullarını yerine getiren çalışmalara da rastlanmaktadır (Ayvalli, 2016; Başusta ve Gelbal, 2010; Kıbrıslıoğlu, 2015). Çalışmalarda çıkan bu sonuçlar cinsiyet gruplarının benzer kültüre sahip olmasından ve kültür benzerliğinin cevaplama davranışları üzerinde etkili olmamasından kaynaklanıyor olabilmektedir (Wu, Li ve Zumbo, 2007).

Fen bilimlerine yönelik öğrenci özellikleri modeli Türkiye ve Singapur ülkeleri arasında yalnızca şekilsel değişmezliği

sağlamıştır. Byrne ve Watkins (2003), metrik değişmezliğin sağlanmadığı durumlarda madde yanlılığı olabileceğine dikkat çekmişlerdir. Bu sonuç, uygulanan anketlerin gruplardan birine yanlı davrandığı ya da maddelerin farklı kültürlerden gelen bireyler için farklı anlamlar ifade ettiğini göstermektedir. Özellikle büyük ölçekli sınavlarda kullanılan anket ya da testlerin farklı dillere çevrilmesi sınırlılık olarak ele alınmaktadır. Çünkü, bireylerin yaşadıkları kültür maddeleri algılama biçiminde etkili olabilmektedir. Bu konuda yapılan bazı çalışmalarda farklı dilleri konuşan bireyler arasında başka deyişle kültürler arası farklılıklar arttıkça kullanılan ölçek ya da testlerin ölçme değişmezliğini sağlanmadığı belirtilmiştir (Asil ve Gelbal, 2012; Ercikan ve Koh, 2005; Kıbrıslıoğlu, 2015; Marsh, Abduljabbar, Ebu-Hilal ve arkadaşları, 2013; Yandı, Köse ve Uysal, 2017). Kıbrıslıoğlu (2015) çalışmasında PISA 2012 matematik öğrenme algısı modelinin Türkiye, Çin-Şangay ve Endonezya örneklemelerinde ölçme değişmezliğini sağlayıp sağlamadığını incelemiştir. Araştırmanın sonuçları modelin ülkeler arasında yalnızca şekilsel değişmezliği sağladığını göstermektedir. Yandı, Köse, Uysal ve Oğul (2017) dilsel ve kültürel farklılaşmanın az olduğu durumlarda dahi ölçme araçlarında bulunan maddelerde DMF bulunduğunu belirtmişlerdir. Bu çalışmada elde edilen bulgular çeviri problemlerinin ve maddelerde bulunan kültürel öğelerin doğru uyarlanamaması durumunu bir kez daha alan yazındakilerle tutarlılık göstererek ortaya koymaktadır.

## Öneriler

Araştırmanın sonuçları doğrultusunda PISA çerçevesinde uygulanan anketlerin farklı kültür ya da cinsiyet açısından ölçme değişmezliğini sağlayamadığı görülmüştür. Bu bulgular farklı gruplara uygulanan ölçme araçlarının karşılaştırma yapmadan önce değişmezlik şartlarını sağlayıp sağlamadığının kontrol edilmesi gerekliliğini ortaya koymaktadır. Araştırmacıların ölçme aracının karşılaştırılan gruplar arasındaki geçerliliğine ilişkin kanıt sunması önemlidir. Bu nedenle ölçekler uygulama aşamasına geçilmeden önce yapı ve kapsam geçerliliğinin yanı sıra ölçme değişmezliği bakımından da gözden geçirilmelidir (Kıbrıslıoğlu, 2015).

Araştırmada kurulan model değişmezlik aşamalarının tamamını sağlayamamıştır. Bu durum ölçek maddelerinin bir kısmının gruplardan bir kısmına yanlı davrandığını göstermektedir. İleride yapılacak araştırmalarda yanlılık gösteren maddelerin belirlenmesi amacıyla yanlılık çalışmaları yapılabilir. Bunun yanında farklı ülkelerin karşılaştırıldığı çalışmalara ve farklı ölçme değişmezliği yöntemlerinin karşılaştırıldığı çalışmalara yer verilebilir.

## 5. Kaynakça

- Asil, M. ve Gelbal, S. (2012). PISA öğrenci anketinin kültürler arası eşdeğerliği. *Eğitim ve Bilim*, 166, 236-249.
- Ayvallı, M. (2016). *PISA 2012 matematik okuryazarlığı testinin ölçme değişmezliğinin incelenmesi*. (Yayımlanmamış Yüksek Lisans Tezi). Akdeniz Üniversitesi, Antalya.
- Başusta, N. B., ve Gelbal, S. (2015). Gruplararası Karşılaştırmalarda Ölçme Değişmezliğinin Test Edilmesi: PISA Öğrenci Anketi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 30(4), 80-90.
- Bryne, B. M. ve Watkins, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology*, 34(2), 155-175.
- Cheung, G.W. ve Rensvold, R.B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management*, 25, 1-27.
- Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*. 43(2), 121-149.
- Du, L. ve Li-Ping Tang, T. (2005). Measurement invariance across gender and major: The love of money among university students in people's republic of China. *Journal of Business Ethics*, 59, 281-293.
- Ercikan, K., ve Koh, K. (2005). Examining the construct comparability of the English and French version of TIMSS. *International Journal of Testing*, 5, 23-35.
- Flowers, C.P., Raju, N.S. ve Oshima, T. C. (2002). A comparison of measurement equivalence methods based on confirmatory factor analysis and item response theory. *Paper presented at the Annual Meeting of the National Council on Measurement in Education*, New Orleans.
- Gregorich, S.E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? : Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*, 44, 78-94
- Grouzet, F.M.E., Otis, N., ve Pelletier, L.G. (2006). Longitudinal cross-gender factorial invariance of the Academic Motivation Scale. *Structural Equation Modeling*, 13(1), 73-98.
- Hirschfeld, G. ve Brachel, R. (2014). Multiple-Group confirmatory factor analysis in R – A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research & Evaluation*. 19(7), 1531-7714.
- Kıbrıslıoğlu, N. (2015). *PISA 2012 matematik öğrenme modelinin kültürlere ve cinsiyete göre ölçme değişmezliğinin incelenmesi: Türkiye, Çin(Şangay)-Endonezya örneği*. (Yayımlanmış Yüksek lisans tezi). Hacettepe Üniversitesi, Ankara.

- Marsh, H. W., Abduljabbar, A. S., Ebu-Hilal, M. M. ve arkadaşları, (2013). Factorial, convergent, and discriminant validity of TIMSS math and science motivation measures: a comparison of arab and anglo-saxon countries. *Journal of Educational Psychology*, 105 (1), 108-128.
- MEB (2015). PISA 2015 Ulusal raporu.
- Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika*, 58, 525-543.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, N. J: Lawrence Erlbaum.
- OECD (2016). PISA 2015 results in focus
- Önen, E. (2007). Gruplar arası karşılaştırmalarda ölçme değişmezliğinin incelenmesi: epistemolojik inançlar envanteri üzerine bir çalışma. *Ege Eğitim Dergisi*, 8(2), 87–110.
- Rutkowski, L. ve Stevina, D. (2013). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement* 74(1) 31–57 .
- Steenkamp, Jan-Benedict, E. M. ve Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25(1), 78-90.
- Swaminathan, H., ve Rogers, H. J. (1990). Detecting item bias using logistic regression procedures. *Journal of Educational Measurement*, 27, 361-370.
- Tucker, K. L., Ozer, D. J, Lyubomirsk, S. ve Boehm, J. K. (2006). Testing for measurement invariance in the satisfaction with life scale: A comparison of Russians and North Americans. *Social Indicators Research*. 78: 341–360.
- Uzun, B. N. (2008). *TIMSS-R Türkiye örnekleminde fen başarısını etkileyen değişkenlerin cinsiyetler arası değişmezliğinin değerlendirilmesi*. (Yayımlanmamış Yüksek Lisans Tezi). Hacettepe Üniversitesi, Ankara.
- Uzun, B., ve Öğretmen, T. (2010). Fen başarısı ile ilgili bazı değişkenlerin TIMSS-R Türkiye örnekleminde cinsiyete göre ölçme değişmezliğinin değerlendirilmesi. *Eğitim ve Bilim*, 35(155).
- Vanderberg, R. J. ve Lance, C. E., (2000). A Review and synthesis of the measurement invariance literature: Suggestions practices, and recommendations for organizational research. *Organizational Research Methods*, 3 (4), 4-70.
- Yandı, A.; Köse, İ. A. ve Uysal, Ö. (2017). Farklı yöntemlerle ölçme değişmezliğinin incelenmesi: PISA 2012 örneği. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 13(1), 243-253.
- Yandı, A.; Köse, İ. A.; Uysal, Ö. ve Oğul, G. (2017). Küreselleşen dünyada eğitim. (PISA 2015 öğrenci anketinin (st094q01na-st094q05na) ölçme değişmezliğinin farklı yöntemlerle incelenmesi). Pegem, Ankara
- Wu, D. A., Li, Z. ve Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. [Elektronik Sürüm]. *Practical Assessment, Research & Evaluation*, 12(3), 1-26.