

## Fen Başarısı ile İlgili Bazı Değişkenlerin TIMSS-R Türkiye Örnekleminde Cinsiyete Göre Ölçme Değişmezliğinin Değerlendirilmesi

### Assessing the Measurement Invariance of Factors that are Related to Students' Science Achievement across Gender in TIMSS-R Turkey Sample

Bilge UZUN\*  
Hacettepe Üniversitesi

Tuncay ÖĞRETMEN\*\*  
Ege Üniversitesi

#### Öz

Bu çalışmanın amacı, öğrencilerin fen başarısına etki eden duyuşsal faktörlerin neler olduğunu belirleyerek bu değişkenlerin ölçmelerin değişmezliği koşulunu sağlayıp sağlamadığını belirlemektir. Bu çalışmada örneklemini 7841 büyüklüğünde olan TIMSS Türkiye verileri, Yapısal Eşitlik Modellemesi (YEM) tekniği kullanılarak analiz edilmiştir. Modelde yer alan örtük değişkenler (özyeterlik, önem, tutum ve sınıf içi öğrenci etkinlikleri) için aşamalı bir şekilde değişmezlik testleri yapılmıştır. Analizlerde hiçbir sınırlandırmanın yapılmadığı (biçimsel değişmezlik) durum ile daha fazla sınırların şart konulduğu diğer değişmezlik testleri arasındaki CFI karşılaştırmalı uyum iyiliği kriterleri farkına bakılmıştır. Ölçmelerin değişmezliği çalışmalarından modele alınan tüm değişkenlerin metrik değişmezlik koşulunu yerine getirdiği, ancak cinsiyet gruplarında katı değişmezlik koşulunu hiçbirinin sağlayamadığı görülmüştür.

*Anahtar Sözcükler:* Ölçmelerin değişmezliği, Çok Örneklemli Doğrulayıcı Faktör Analizi (MGCFI), Yapısal Eşitlik Modeli (YEM), TIMSS, cinsiyet.

#### Abstract

The purpose of this study is to investigate the factors that are related to students' science achievement and assessing the measurement invariance of these factors across gender. This study examined the TIMSS data for Turkish students with the sample size of 7841 through the analysis of Structural Equation Modeling (SEM). First, it is designed a model over the factors from student questionnaire affecting student's science achievement with Structural equation modeling. It was understood that the designed model's goodness fit index output is acceptable. Then four forms of invariance was assessed for each latent variables progressively and assessed MI. The analyses basically begin by fitting a proposed model to the data for each sample considered separately with none of the parameters constrained to be equal across groups(configural invariance). This unconstrained model serves as the baseline model. Subsequently, in a hierarchical fashion, more stringent constraints are placed on the model by specifying the parameters of interest to be constrained across the groups (e.g., factor loadings, factor intercorrelations, error variances). The variables were examined using comparative fit index (CFI) difference test between the more restrictive invariance form and the basic form to determine whether the model and the individual parameter estimates are invariant across the samples. The MI results revealed that there were invariance problems across gender in TIMSS. All latent variables at least have metric invariance, but none of them have strict invariance across groups.

\* Arş. Gör. Bilge UZUN, Hacettepe Üniversitesi, Tıp Fakültesi, Tıp Eğitimi Anabilim Dalı.

\*\* Yrd. Doç. Dr. Tuncay ÖĞRETMEN, Ege Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü.

*Keywords:* Measurement Invariance, MGCFI, Structural equation modeling (SEM) Third International Mathematics And Science Study (TIMSS), Gender

## Summary

### *Purpose*

The purpose of this study is to investigate the factors that are related to students' science achievement and assessing the measurement invariance of these factors across gender. Invariance testing has been developed to address the question of how we can be sure that a test or survey measures the same trait dimension, in the same way, when administered to two or more qualitatively distinct groups. Through invariance testing we are able to determine the structure of TIMSS student questionnaire's items were measuring the same trait in the groups.

### *Result*

This study examined the TIMSS data for Turkish students with the sample size of 7841 through the analysis of Structural Equation Modeling (SEM). Measurement invariance was tested using Multi-group Confirmatory Factor Analysis (MG-CFI), which examines the change in the comparative-of-fit index (CFI) when cross-group constraints are imposed on a measurement model.

First, it is designed a model over the factors from student questionnaire affecting student's science achievement with Structural equation modeling. We determine 5 factors after the analysis of AFA. And four of them were used for the structural model. It was understood from the outputs that the designed model's goodness fit index output was good. Then four forms of invariance were assessed for each latent variable progressively and assessed MI. The analyses basically begin by fitting a proposed model to the data for each sample considered separately with none of the parameters constrained to be equal across groups (configural invariance). This unconstrained model serves as the baseline model. Subsequently, in a hierarchical fashion, more stringent constraints are placed on the model by specifying the parameters of interest to be constrained across the groups (e.g., factor loadings, factor intercorrelations, error variances). The variables were examined using comparative fit index (CFI) difference test between the more restrictive invariance form and the basic form to determine whether the model and the individual parameter estimates are invariant across the samples. Only the latent variable TUTUM comparisons of means were significant. Because the comparison of means makes sense only when scalar invariance holds. The MI results revealed that there were invariance problems across gender in TIMSS. All latent variables at least have metric invariance, but none of them have strict invariance across groups.

### *Discussion*

For any use of scale scores, there is a critical assumption that the scale is measuring the same trait in all of the groups. If that assumption holds, then comparisons and analyses of those scores are acceptable and yield meaningful interpretations. If that assumption is not true, then such comparisons and analyses do not yield meaningful results. We designed a model over the factors from student questionnaire affecting student's science achievement and tested invariance of these factors of measurement model. All latent variables at least have metric invariance, this means that the same scale units of the latent variable(s) have been established and differences within separate groups may be compared, However, even metric invariance is not the ultimate prerequisite for mean comparisons across cultures for additive biases may still be present. The comparison of means makes sense only when scalar invariance holds. Hence, only the comparison according to TUTUM is meaningful.

### *Conclusion*

The designed model had good fit index. But for gender groups it works different from each other.

## Giriş

Toplumu oluşturan ve aralarında işlevsel bağ bulunan kurumların her biri toplumsal ilerlemenin, değişimin, bütünlüğü ya da gereksinimlerin sağlanmasında kendine düşen bazı önemli işlevleri vardır. Söz konusu önemli kurumlardan biri, eğitim ve öğretim işlevlerinin gerçekleştirildiği ve sürdürüldüğü yer olan her düzeyde okullar olup amaçları ve işlevleri, bazı dönemlerde sorgulanmakta, nitelik ve verimlilik yönünden belirlenen birtakım ölçütlere göre çok yönlü incelenmesi ve karşılaştırılması gerekmektedir. Bu bağlamda, TIMSS etkinlikleri, söz konusu karşılaştırmalı eğitim araştırmaları yönünden birtakım gelenekleri oluşmuş, saygınlık kazanmış örnek bir çalışmadır.

Farklı grupların aynı ölçme aracı ile karşılaştırıldığı çalışmalarda, “ölçme aracından kaynaklanabilecek farklılıklar yoktur” biçimindeki bir kabul, karşılaştırmanın ve yorumların eksik kalmasına neden olabilmektedir. Bir ölçme aracı modeli, ‘farklı gruplarda da aynı özelliği ölçer’ kabulü ile hazırlanır. Eğer bu kabul onay alabilirse, yapılan karşılaştırma ve puanlara ait analizlerin doğruluğu anlamlı olacaktır. Yapılan bir ölçmedeki ölçme sonuçları bireylerin farklı özellikler taşıması nedeniyle farklı olabilir. Ancak bu farklılığın sadece birey özelliklerine bağlanması ve bu biçimde açıklanması çoğu zaman doğru değildir. Ölçme sonuçlarındaki farklılık aynı zamanda ölçme aracının kendinden kaynaklanıyor olabilir.

Bir ölçme modelinin birden fazla grupta aynı yapıya sahip olması demek, söz konusu ölçüğün maddelerinin faktör yüklerinin, faktörler arası korelasyonların ve hata varyanslarının aynı olması demektir (Byrne, 1998; Jöreskog & Sörbom, 1993). Değişmezlik çalışmaları ile yeni bir ölçme aracı geliştirilmez. Ölçmelerin değişmezliği gruplar arası karşılaştırmalarda bu karşılaştırmanın anlamlılığında bir koşul olarak ortaya çıkar (Bollen, 1989; aktaran: Cheung & Rensvold, 2000). Yani bu testler ölçme aracının elde edilen ölçümlerde eşit yapıyı ortaya koyup koymadığını garantilemeye çalışır.

Ölçmelerin değişmezliği aşamalılık içerir. Test edilmesi gereken dört aşama vardır (Meredith, 1993). Bunlar;

I) Biçimsel değişmezlik (Configural invariance): Kurulan teoriye bağlı kavramsal yapı alt gruplarda aynıdır ve sadece bu temel yapının gruplar arası karşılaştırması mümkündür.

II) Metrik değişmezlik (Weak factorial invariance = Metric invariance): Kurulan bir modelde yer alan yapısal ilişkiler aynıdır ve yapının değişkenler açısından karşılaştırılması ve araştırılması alt gruplarda uygun ve anlamlı olabilir.

III) Skalar değişmezlik (Strong invariance = Scalar invariance): Kavramsal yapı, yapısal ilişkilendirme ve hata kaynakları alt gruplarda aynı olduğu için örtük değişkenlerin ortalamalarının karşılaştırılması anlamlıdır.

IV) Katı değişmezlik (Strict invariance): Kavramsal yapı, yapısal ilişkilendirme, hata kaynakları ve madde artıklarının varyansları gruplarda aynı olduğu için örtük ortalamaların karşılaştırılması anlamlıdır.

Bu değişmezlik formları hiyerarşik bir sıra takip edilerek ve birçok model uyum katsayısı kullanılarak test edilir. Her bir değişmezlik aşaması kendinden önce gelen aşamanın koşulunu da barındırmaktadır. Metrik değişmezlik koşulunu sağlamayan bir modellemenin neden skalar değişmezlik koşulunu da sağlamayacağı bu biçimde açıklanır.

Gruplar arasındaki farklılıklar değerlendirilirken geleneksel yaklaşımlar kullanmak hatalı sonuçlar doğurabilir. Değişmezlik testleri bir tür kovaryans yapı analizi olup farklı gruplarda belli bir yapının ölçülmesi üzerine tasarlanmıştır. İstatistiksel olarak değişmezlik araştırılırken en çok ve yaygın kullanılan metot Çok Örneklemli Doğrulayıcı Faktör Analizi (MG-CFI)'dir (Jöreskog & Sörbom, 1999). MG-CFI en az sınırlı modelden en sınırlı modele kadar modellemeler arası karşılaştırma yaparak gruplar arası parametrelerin değişmezliğini sağlamaya çalışır (Horn

& McArdle, 1992). Ölçmelerde değişmezlikle ilgili kanıtların bulunmadığı çalışmanın zayıf bir çalışma olacağı belirtilmiştir (Horn, 1992).

Ülkemizde fen bilimleri başarısını etkileyen yapılar üzerine çalışmalar bulunmasına rağmen, bu yapıların farklı gruplarda ölçme değişmezlikleri ile ilgili bir çalışmaya rastlanmamıştır. Fen bilimleri eğitiminin gelişmesine katkıda bulunmak için öğrencilerin öğrenme düzeylerinin derinlemesine incelenmesi faydalı olabilir. Araştırmalara bakıldığında, değişmezlik çalışmaları grupların karşılaştırılmasında daha doğru ve güvenilir sonuçlar vermektedir. Yapılan karşılaştırmaların anlamlı ve doğru olması için ölçmelerin değişmezlik konusunu çalışmak, Türkiye örnekleminin yapısının anlaşılması açısından açıklayıcı olabilir.

Katılımcı ülkelerin kendi eğitim sistemlerini sorgulamayı amaçlayan ve elde edilen sonuçlardan yola çıkarak mevcut sistemde gereken düzeltme ve yeniliklerin yapılmasına olanak veren TIMSS; PIRLS ve PISA gibi güvenilir verilerle çalışmak, eğitim sisteminde yer alan temel sorunların nedenlerine ulaşmamızı kolaylaştıracaktır. Bu çalışmada TIMSS-R (Third International Mathematics and Science Study-Repeat ) adlı araştırma çalışmasının verileri kullanılmıştır. TIMSS 1999-R verilerine göre öğrencilerin fen bilimleri başarısını etkileyen modelde yer alan değişkenlerin cinsiyet grupları için karşılaştırılmasının anlamlı olup olmadığına bakılmaktadır. Araştırma bulgularının literatüre katkı yapabileceği, elde edilen araştırma bulgularının Türk eğitim sistemindeki değişim ve gelişim amaçlı yapılan deneysel çalışmalar için önemli bilgiler sağlayabileceği, farklı gruplarda bazı değişkenlerin değişmezlik çalışmaları yapıldıktan sonra karşılaştırılmasının, araştırılan değişkenler açısından Türkiye'nin sosyal yapısını açıklamaya yardımcı olabileceği düşünülmektedir.

#### Yöntem

Bu araştırma, 1999 TIMSS-R'ye katılan öğrencilerden toplanan verilerle öğrencilerin fen bilimleri başarısını açıkladığı düşünülen cinsiyetlere göre değişmez bir model ortaya koyarak modelde yer alan değişkenler ile başarı arasındaki ilişkiyi cinsiyetler bazında test etmeye yönelik tasarlanmış betimsel bir araştırmadır.

TIMSS-99 (R)'ye katılan öğrenciler, ulusal düzeyde iki basamaklı bir işlem içeren süreçle belirlenmiştir. İlk basamakta okullar, ikinci basamakta da aynı okuldaki sınıflar rastgele seçilmiştir. Böylece, araştırmaya katılan her ülkede yaklaşık 150 okul rasgele seçilmiş, bu okullarla ilgili bazı bilgiler TIMSS Araştırma Merkezi'ne gönderilmiştir. Her ilköğretim/temel eğitim okulunda 8. sınıfta bir veya iki şube olmak üzere değerlendirmeye katılan ülkeden yaklaşık 3500 öğrencinin ortalama başarı durumunun incelenmesi planlanmıştır. Bu çalışmada anlatılan biçimde seçilen öğrencilerden Türkiye'de sınava katılan öğrencilere ait veriler kullanılmıştır. Çalışma grubu 2204 okulun 7841 8. sınıf öğrencisinden oluşmaktadır.

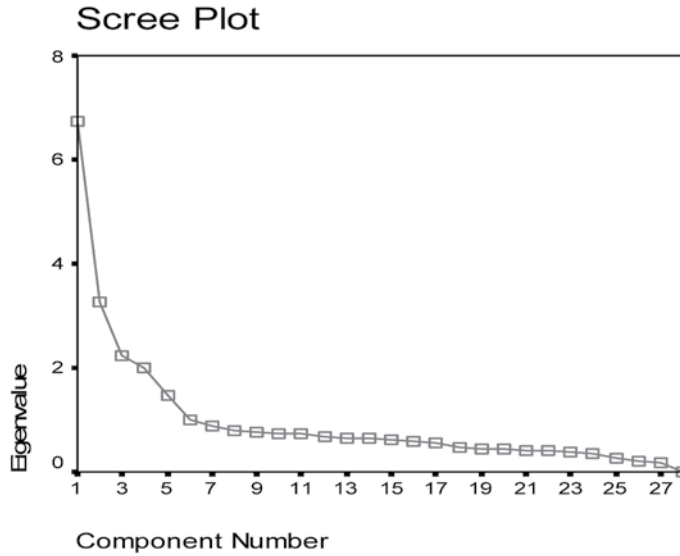
Araştırma, Uluslararası Eğitim Başarılarını Belirleme Kuruluşunun (IEA) 1999 TIMSS-R'ye katılan öğrencilere uygulamış olduğu öğrenci anketinde yer alan fen ile ilgili duyuşsal özellikleri ölçen sorulara verilen cevaplar veri olarak kullanılmıştır.

Bu çalışmada temel olarak üç istatistiksel analiz yöntemi kullanılmıştır. Bunlardan birincisi, açımlayıcı faktör analizidir (AFA). İkincisi, Yapısal Eşitlik Modeli'dir (YEM). Yapısal Eşitlik Modeli (YEM) ile regresyon modelindeki değişkenler arasındaki yordayıcı yapısal ilişki ve faktör analizindeki gizil faktör yapılarını kapsamlı bir analizde birleştirmek amaçlanmaktadır. Üçüncü olarak ise ölçme aracının cinsiyetler bakımından değişmezliğinin sağlanmasına yönelik çok örneklemlili doğrulayıcı faktör analizi (MGCFI) tekniği kullanılmıştır. Bu çalışmada doğrulayıcı uyum katsayıları (CFI) arasındaki farklar kullanılmıştır. Belirlenen temel model çerçevesinde değişmezlik testi olarak sırasıyla; biçimsel değişmezlik sonucu elde edilen CFI ile, metrik değişmezlik için bulunan CFI, skalar değişmezlik için bulunan CFI ve katı değişmezlik için bulunan CFI farkları incelenmiştir. CFI farklarının incelenmesinin nedeni ise uyum katsayılarının örtük skorlarla gözlenen skorlar arasındaki ilişkiye dair bilgi vermesinden kaynaklanır. Diğer

uyum katsayıları da ölçmelerin derğişmezliđi çalışmalarında kullanılabilir. Uyum katsayılarının eş olduđu durumda maddelere ait skorların gruplar arasında anlamlı biçimde karşılaştırılması yapılabilir. Bu çalışmada veri ile uyum gösteren MGCFI çalışma dosyaları için  $0.01 \geq \Delta CFI \geq -0.01$  olan derğişmezlik koşullarının sağlanıp sağlanmadığına bakılmıştır.

### Bulgular ve Tartışma

TIMSS öğrenci anketinin fen bilgisi ile ilgili olan maddelerine açımlayıcı faktör analizi yapılarak teoriye konu olacak bileşenlerin neler olduğuna karar verilmiştir. Açımlayıcı faktör analizi (AFA), maksimum olabilirlik, varimax dik döndürme tekniđi kullanılarak yapılan analiz sonucunda ölçekte yer alan 77 maddenin faktörler altındaki yükleri incelenmiştir. Faktör yükleri 0.30'un altında olan maddeler alınmamıştır. Faktörlerin özdeğerlerine dayalı olarak çizilen çizgi grafiđi aşağıda verilmiştir.



Şekil 1: Faktörlerin Özdeğerlerine Dayalı Olarak Çizilen Çizgi Grafiđi

Şekil 1'e göre Özdeğeri 1' in üzerinde olan faktör sayısı 5 olarak belirlenmiştir. Analiz sonucu elde edilen faktörlerin özdeğerleri ve açıkladıkları varyans miktarları ise, birinci faktör için 6,740 ve %24,073; ikinci faktör için 3,279 ve %11,711; üçüncü faktör için 2,227 ve %7,954; dördüncü faktör için 1,996 ve %7,127; beşinci faktör için 1,484 ve %5,299 bulunmuştur. Ancak bu çalışmada sosyoekonomik seviye faktörü olarak düşünölen derğişkene ilişkin maddeler modelin ve derğişmezlik çalışmasının dışında bırakılmıştır. Yapılan açımlayıcı faktör analizi (AFA) sonunda 4 temel faktör olduğu sonucuna ulaşılmıştır. ÖZYETERLİK gizil derğişkeninin yordadığı 5 gösterge derğişken, TUTUM gizil derğişkeninin yordadığı 4 gösterge derğişken, sınıf içi öğrenci etkinlikleri (SIOE) gizil derğişkeninin yordadığı 6 gösterge derğişken, ÖNEM gizil derğişkeninin yordadığı 4 gösterge derğişken bulunmaktadır.

Ölçmede derğişmezlik dosyalarının çıktıları ve CFI katsayı farklar bulunarak anlamlılık tespit edilmiştir. Aşamalı bir biçimde, her bir derğişken için biçimsel derğişmezlik ile metrik derğişmezlik, skalar derğişmezlik ve katı derğişmezlik arasındaki CFI fark değeri yorumlanmıştır. Biçimsel derğişmezlik koşullarını sağlayan derğişkenleri belirlemek için alt gruplara ait hiçbir parametreye sınır getirilmemiştir. Her iki grup için tüm parametreler (faktör yükleri, faktörler arası korelasyon, hata varyansları) serbest bırakılmış ve farklı değerler almalarına izin verilmiştir.

Tablo 1.

*Biçimsel Değişmezlik İçin Araştırılan Çalışma Dosyalarına Ait Uyum İstatistikleri*

	df	X <sup>2</sup>	ECVI	RMSEA	CFI	GFI	NFI
ÖZYETERLİK	11	251,17	0037	0,075	0,99	0,99	0,99
ÖNEM	4	23,5	0,007	0,035	1,0	1,0	1,0
TUTUM	4	60,49	0,012	0,061	0,99	0,99	0,99
SIOE	18	389,2	0,06	0,07	1,0	1,0	1,0

Bu Tablo 1’de yer alan *ECVI*, *RMSEA*, *CFI*, *GFI* ve *NFI* kabul edilir aralık içerisinde (Meredith, 1993). Hatta elde edilen değerler için “alt gruplarda bu modellemeler veri ile mükemmel uyum sergilemektedir” demek yanlış olmayacaktır. Bu uyum istatistiklerinden her bir değişkene ilişkin ölçme modelinin alt gruplar için anlamlı olduğu sonucu çıkar. Yani örtük değişkenler hem kız hem de erkekler için aynı yapıdadır. Başka bir deyişle tüm örtük değişkenler gruplar için değişmezdir.

Metrik değişmezlik koşullarını sağlayan değişkenlerin hangileri olduğuna karar vermek için gruplar arasında faktör yüklerinin aynı olması sınırlılığına yer verilmektedir. Tablo 2’ de faktör yüklerinin eşit olduğu koşulu ile hazırlanan çalışma dosyalarına ilişkin uyum katsayıları yer almaktadır.

Tablo 2.

*Metrik Değişmezlik İçin Araştırılan Çalışma Dosyasına Ait Uyum İstatistikleri*

	df	X <sup>2</sup>	ECVI	RMSEA	CFI	GFI	NFI
ÖZYETERLİK	15	301,40	0,042	0,07	0,98	0,99	0,98
ÖNEM	7	29,5	0,007	0,03	1,0	1,0	1,0
TUTUM	7	69,8	0,01	0,05	1,0	1,0	1,0
SIOE	23	414	0,06	0,07	1,0	1,0	1,0

Tablo 2’ de yer alan *ECVI*, *RMSEA*, *CFI*, *GFI*, *NFI* değerleri veri ile mükemmel uyum sağlamıştır. Bu uyum istatistiklerinden her bir değişkene ilişkin ölçme modelinin alt gruplar için anlamlı olduğu sonucu çıkar. Değişmezlik koşulunu sağlayan değişkenlere karar vermek için metrik model ile biçimsel model kıyaslanarak bu iki model arasındaki ( $\Delta CFI$ ) uyum katsayıları fark değerleri incelenmiş ve Tablo 3’te yer alan değerler elde edilmiştir.

Tablo 3.

*Biçimsel Değişmezlik ve Metrik Değişmezlik İçin Araştırılan Çalışma Dosyasına Ait Fark Testleri Tablosu Uyum İstatistikleri*

	$\Delta CFI$ $CFI_{BİÇİMSEL} - CFI_{METRİK}$
ÖZYETERLİK	0,01
ÖNEM	0,00
TUTUM	0,01
SIOE	0,00

Fark değerinin  $-0,01 \leq \Delta CFI \leq 0,01$  aralığında olması tüm değişkenler için metrik değişmezlik koşulunun mevcut olduğunun işaretidir. Yani grupların maddelere aynı biçimde cevap verdiğini ve böylece iki gruptan elde edilen puanların karşılaştırmada anlamlı olduğu sonucuna ulaşılmıştır.

Değişmezliğin üçüncü aşaması olan skalar değişmezlik, diğer değişmezlik testlerinde yer alan sınırlandırmaların yanı sıra benzer maddelerin gruplar arası eş olduğu sınırlandırılması yapılarak test edilebilir. Yani alt gruplarda hem faktör yükleri hem de faktörler arası korelasyonların eşit olma sınırlılığı bulunmaktadır. Metrik bakımdan değişmez olan çalışma dosyası, skalar değişmezlik sınırlandırılması eklenerek test edildiğinde elde edilen sonuçlar ve uyum katsayıları aşağıdaki gibi çıkmıştır.

Tablo 4.

*Skalar Değişmezlik İçin Araştırılan Çalışma Dosyasına Ait Uyum İstatistikleri*

	df	$\chi^2$	p	ECVI	RMSEA	CFI	GFI	NFI
ÖZYETERLİK	20	2578,36	0.00	0,34	0,18	0,85	0,88	0.85
ÖNEM	14	688.7	0.00	0,092	0,11	0,90	0,96	0.90
TUTUM	14	301	0.00	0,04	0,07	1,0	1,0	1,0
SIOE	34	1197,15	0.00	0,16	0,095	0.91	0.96	0.91

Tablo 5 incelendiğinde, ÖZYETERLİK değişkeninin uyum katsayılarının kabul edilebilir sınır aralığında bulunmadığı görülmektedir. ÖNEM değişkeni de RMSEA katsayısı için belirtilen aralıkta yer almamaktadır. Alt gruplarda ÖZYETERLİK ve ÖNEM değişkeni için kurulan modellemeler veri ile uyum göstermemektedir. ÖZYETERLİK dışındaki TUTUM ve SIOE değişkenlerine ait uyum istatistikleri incelendiğinde, kabul edilir aralıkta yer aldığı için veri ile uyum göstermektedir.

Elde edilen uyum istatistikleri ve biçimsel değişmezlik dosyalarına ait uyum istatistikleri incelenerek skalar değişmezliğin bulunup bulunmadığına yönelik yapılan CFI fark testine ait sonuçlar aşağıdaki tabloda gösterilmiştir.

Tablo 5.

*Biçimsel Değişmezlik ve Skalar Değişmezlik İçin**Araştırılan Çalışma Dosyasına Ait Fark Testleri Tablosu Uyum İstatistikleri*

	$\Delta CFI$ $CFI_{\text{BİÇİMSEL}} - CFI_{\text{SKALAR}}$
ÖZYETERLİK	0,14
ÖNEM	0.1
TUTUM	0.01
SIOE	0.09

Tablo 6'da fark değerinin  $-0,01 \leq \Delta CFI \leq 0,01$  aralığında olan tek değişken TUTUM değişkenidir. Yani kız ve erkek öğrenci grupları için TUTUM dışındaki diğer örtük değişkenler için skalar değişmezlik sağlanamamıştır. Bu, örtük yapı ortalamalarının skalar değişmezliğin sağlanmadığı değişkenlerde, gruplar arası fark olduğu anlamına gelmektedir. Sadece TUTUM değişkeni için gözlenen maddeler üzerindeki grup farklılıkları bakımından cinsiyetler arası bir karşılaştırma anlamlı olabilecektir.

Son değişmezlik kademesinde yer alan katı değişmezlik için bundan önce skalar değişmezliğin test edildiği çalışma dosyasına hata varyansların gruplar arası denklik sınırlandırılması eklenmiştir. Burada önemli olan TUTUM değişkeni için yapılacak fark testidir. Çünkü değişmezlik aşamalık gerektirmektedir ve bir aşama sağlanmadan diğerine geçiş söz konusu değildir.

Tablo 6.

*Katı Değişmezlik İçin Araştırılan Çalışma Dosyasına Ait Uyum İstatistikleri*

	<i>df</i>	<i>X<sup>2</sup></i>	<i>p</i>	<i>ECVI</i>	<i>RMSEA</i>	<i>CFI</i>	<i>GFI</i>	<i>NFI</i>
TUTUM	19	514,09	0.00	0,068	0,082	0,94	0,98	0,94

TUTUM değişkeninin veri ile uyum gösterdiği görülmektedir. Katı değişmezliğin sağlanıp sağlanmadığına ilişkin yapılan doğrulayıcı uyum katsayıları arasındaki farkın belirtilen düzeyde olmadığı sonucuna ulaşılmıştır. CFI uyum katsayıları arasındaki farkın 0,05 olması aranan koşulu sağlamamaktadır. Bu sonuca ilişkin tablo değerleri aşağıda verilmiştir.

Tablo 7.

*Biçimsel Değişmezlik ve Katı Değişmezlik İçin*

*Araştırılan Çalışma Dosyasına Ait Fark Testleri Tablosu Uyum İstatistikleri*

	$\Delta CFI$ $CFI_{BİÇİMSEL} - CFI_{KATI}$
TUTUM	0.06

Belli modifikasyonlar yapılmadan görülmüştür ki modeldeki hiçbir değişken kız ve erkekler için katı değişmezlik koşulunu sağlamamaktadır.

Yapılan karşılaştırmaların anlamlılığı, değişkenlerin değişmezlik aşamalarını sağlayıp sağlayamadığı ile ilgilidir. Aslında alt gruplarda yapılacak karşılaştırmanın ne anlamda yapılacağı da bu durumda önem kazanmaktadır. Test edilen değişmezlik aşamaları ve bu aşamalarda belirlenen koşulları sağlayıp sağlamadıkları aşağıdaki tabloda gösterilmiştir.

Tablo 8.

*Araştırılan Değişkenlere İlişkin Ölçmelerin Değişmezlik Aşamaları Sonuçları*

	<i>BİÇİMSEL DEĞİŞMEZLİK</i>	<i>METRİK DEĞİŞMEZLİK</i>	<i>SKALAR DEĞİŞMEZLİK</i>	<i>KATI DEĞİŞMEZLİK</i>
ÖZYETERLİK	VAR	VAR	YOK	YOK
TUTUM	VAR	VAR	VAR	YOK
ÖNEM	VAR	VAR	YOK	YOK
SIOE	VAR	VAR	YOK	YOK

Tablo 8'e bakılarak tüm değişkenlere ilişkin metrik bakımdan değişmezliğin mevcut olduğu sonucu çıkarılmıştır. Bu çalışmada yer alan tüm değişkenlerle yapılacak bir modellemede de maddelere ait puanların gruplar arasında anlamlı biçimde karşılaştırılmasının mümkün olacağı anlamına gelmektedir. Metrik değişmezlik koşullarını sağlayan değişkenler için cinsiyetler arası yapılacak karşılaştırmaların anlamlı olacağı ve alt gruplarda meydana gelen farklılığın ölçme aracından kaynaklanmadığı söylenebilir. Tablo 8'de kızlar ve erkekler için sağlanmış olan değişmezlik testleri ile teste ait maddelerin faktör yük değerleri eşittir ve örtük değişkenlerin hata varyanslarına ait değerler her iki grupta da farklılık göstermektedir. Tüm değişkenlerin kullanılarak oluşturulduğu bir modellemede (kız öğrenciler ve erkek öğrenciler için metrik değişmezlik dosyasına ait standardize edilmiş değerler) madde puanlarına yönelik yapılan karşılaştırmalar anlamlı olabilecektir. Çünkü tüm değişkenlerin karşıladığı değişmezlik koşulu metrik değişmezliktir. Ancak bu, tüm faktör yapısına göre yapılan karşılaştırmaların anlamlılığı anlamına gelmemelidir.



Görüldüğü üzere katı değişmezlik testini sağlayabilen bir örtük değişkenin olmayışı, bu karşılaştırmaların anlamlı olmayacağı sonucuna bizi götürecektir. Bu durumda kızlar ve erkekler başarıyı açıklayan değişkenler bakımından farklıdır yorumunu yapmak mümkündür. Elde edilen bu değerlerin farklılık boyutu ve değişikliği Türkiye'nin sosyal yapısını ortaya koymaya yardımcı bir kanıt niteliği taşıyabilir.

### Sonuçlar ve Yorum

- Biçimsel değişmezlik: Teorisi sunulan modele alınan değişkenler ile kurulan modellemelerin uyum istatistikleri uygundur ve MG-CFI ile analizi yapılan değişkenler yapısal bakımdan değişmezdir. Her iki alt grubun almış olduğu testte yer alan dört değişkene ilişkin kavramsal yapı değişmezdir.
- Metrik değişmezlik: Sınırlandırılmamış modelle aralarında yapılan CFI fark testi sonucunda modelde yer alan tüm gözlenen örtük değişkenler bu sınırlandırmayı karşılamıştır. Yani modele alınan her değişken metrik bakımdan alt gruplar için değişmez bulunmuştur.
- Skalar değişmezlik: Modelin MGCFI ile analizi sonunda elde edilen uyum katsayıları istenilen aralıkta bulunmadığı gibi sınırlandırılmamış model ile yapılan CFI fark testinde de elde edilen değer 0,01 den büyük olduğu için ancak belli modifikasyonlarla kısmi (partial) değişmezlik söz konusu olabilir. Skalar değişmezlik için yapılan sınırlandırılmanın tek karşılandığı değişken TUTUM değişkenidir.
- Modele alınan değişkenlerden skalar değişkenliği sağlamayan üç değişken (SIOE; ÖNEM, ÖZYETERLİK) için yapılacak sınırlandırmanın değişmezlik çalışmalarında aşamalılık ilkesinden ötürü anlamsız olduğu belirtilmiştir. Öte yandan skalar değişmezlik koşulunu sağlayan TUTUM değişkeni için yazılan katı değişmezlik çalışma dosyası uyum istatistikleri ve biçimsel model ile elde edilen uyum istatistikleri arasındaki fark testi anlamlı çıkmamıştır. Hiçbir değişken için katı değişmezlik söz konusu değildir.
- Yapılan testler sonucu tüm değişkenler ile kurulup karşılaştırmanın anlamlı olacağı aşama metrik değişmezlik aşamasıdır.

Sonuç olarak, bu çalışmanın asıl odağı olan “TIMSS uygulamasının gruplar arası karşılaştırılmasının anlamlı olup olmadığı” sorusu şu biçimde cevaplanabilir: Tüm değişkenler en azından metrik değişmezliğe sahipler; ancak ölçmelerin değişmezliğinde katı değişmezliğe kadar sağlanması gereken koşulları modifikasyon yapmadan sağlamadıkları ve uygulama bu değişkenler bakımından ölçmelerin değişmezliği koşullarını tam olarak ve aşamalı biçimde sonuna kadar yerine getirmediği için araştırma problemi 2 bulgularına bakılarak yapılan karşılaştırmaların kesinliğinden ve anlamlılığından söz edilemez.

Bu çalışmada modele dahil edilmeyen değişkenlerin bağımsız değişkenle ilişkileri nedensellik ilişkisini etkilemiş olabilir. Bu nedenle dikkate alınmayan değişkenler açısından da değişmezlik testleri ile karşılaştırmalar yapılmasına ihtiyaç olduğu düşünülmektedir.

Gelecek çalışmalarda öğretmen ve okullara ilişkin anketler kullanılarak benzer araştırmalar yapılması ve değişmezlik testleri ile zenginleştirilmesi, mesleki eğitim ve bölge gelişimi açısından önemli bilgiler sağlayabilir.

Ülkemizin bütününe yönelik anlamlı genellemeler yapılabilmesi için değişmezlik çalışması ile farklı il ve bölgeler çapında karşılaştırmalar yapılması, sosyal yapının ve konu alanına yönelik durumun açıklanmasında yararlı olacaktır. Böylece öğrencilerin fen bilgisi dersinden beklentileri daha açık şekilde ortaya konulabilecektir.

Kaynakça

- Byrne, B. M. (1998). *Structual equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. New Jersey: Lawrence Erlbaum.
- Cheung, G. W., Rensvold, R. B. (2000). Assessing Extreme And Acquiescence Response Sets İn Cross-Cultural Research Using Structural Equations Modeling. *Journal Of Cross-Cultural Psychology*, 31(2), 187-212.
- Horn & Mcardle.(1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, 18 (3) ,117 – 144
- International Study Center (2000b) TIMSS 1999 (TIMSS-R) Released science items. <http://isc.bc.edu/timss1999i/study.html> (Ağustos, 2002)
- Jöreskog, K. G. (1993). *Testing Structural Equation Models*. In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models* (Pp. 294–316). Newbury Park, CA: Sage.
- MEB. (2003). *1999 TIMSS Türkiye Raporu, Ankara*.
- Meredith, W. (1993). Measurement İnvariance, Factor Analysis, And Factorial İnvariance. *Psychometrika*, 58, 525-543.
- TIMSS 1999 (TIMSS-R) *International science report*. <http://isc.bc.edu/timss1999i/publications.html> (Ağustos, 2002)