

PISA 2009 TÜRKİYE ÖRNEKLEMİNDE ÖĞRENME STRATEJİLERİ MODELİNİN FARKLI GRUPLARDA ÖLÇME DEĞİŞMEZLİĞİNİN İNCELENMESİ*

Şeyma UYAR

Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi, syuksel@mehmetakif.edu.tr

Nuri DOĞAN

Hacettepe Üniversitesi Eğitim Fakültesi, nurid@hacettepe.edu.tr

Özet

Bu araştırmada, PISA 2009 Türkiye örnekleminde öğrenci anketinde öğrenmeyi öğrenme bölümünde yer alan öğrenme stratejilerine dair bir model test edilmiş ve modelin cinsiyet, okul türü ve istatistiksel bölge (12 NUTS) gruplarında değişmezliği incelenmiştir. Araştırma, PISA 2009 Türkiye örnekleminde 4340 veri ile yürütülmüştür. Öğrenme stratejileri modeli çeşitli alt gruplarda farklı yorumlanabileceğinden bu araştırmada, cinsiyet, okul türü ve istatistiksel bölge değişkenlerine göre modelin değişmezliğine bakılmıştır. 12 istatistiksel bölge arasında Doğu Karadeniz ve Kuzeydoğu Anadolu bölgeleri model uygunluğu için kabul edilebilir aralıklarda yer almadığından araştırma kapsamının dışında bırakılmıştır. Değişmezlik testleri aşamalı şekilde yürütülmüştür ve sınırlandırılmamış model ile daha sınırlı bir modelden elde edilen değerler arasındaki farklara bakılmış ve parametrelerin değişmez olup olmadığı incelenmiştir. Öğrenme stratejileri modelinin cinsiyet ve okul türü gruplarında sadece yapısal ve metrik değişmezlik koşullarını yerine getirdiği, bölgelerde ise tüm değişmezlik koşullarını yerine getirdiği sonucuna varılmıştır.

Anahtar sözcükler: Ölçmelerin değişmezliği, Çoklu grup doğrulayıcı faktör analizi, Yapısal eşitlik modeli

AN INVESTIGATION OF MEASUREMENT INVARIANCE OF LEARNING STRATEGIES MODEL ACROSS DIFFERENT GROUPS IN PISA TURKEY SAMPLE

Abstract

In this study, a model on learning strategies in "Learning by strategies" part of PISA 2009 Student Survey was examined in respect of the model's invariance across gender, school types, and Nomenclature of Territorial Units for Statistics (12 NUTS) in Turkey sample. The study was carried out by means of 4340 data in PISA 2009 Turkey sample. Since the learning strategies model may be interpreted differently across sub-groups, the invariance of model in relation to gender, school types, and 12 NUTS was examined in this study. The Eastern Black Sea and Northern Anatolia regions were not included in the study since these regions were not within the range of acceptability for model suitability. The variables were examined using differences test between the more restrictive invariance form and the basic form to determine whether the parameter invariant across groups. It was concluded that while the model only provided configural and metric invariance conditions in the groups of gender and school types, it provided all invariance conditions among regions.

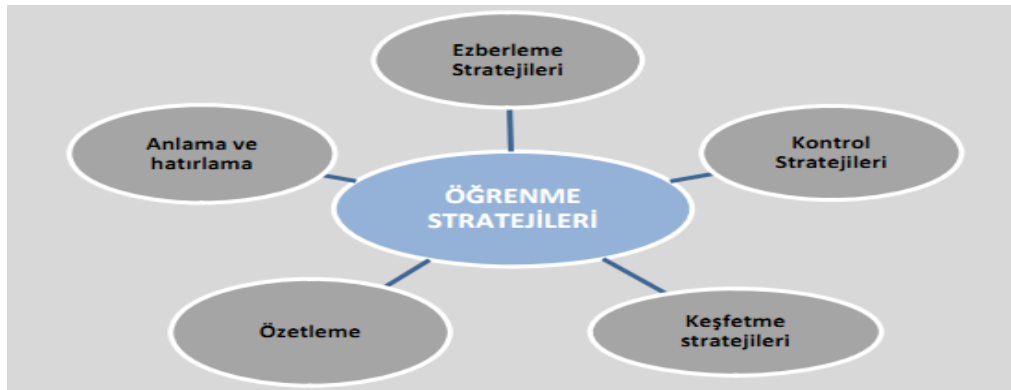
Key Words: Measurement Invariance, Multi Groups Confirmatory Factor Analysis, Structural Equation Model

GİRİŞ

Pek çok ülkede öğrencilere kazandırılacak davranışlar doğrultusunda sisteme ve öğelerine yönelik eksikliklerin belirlenmesi ve gereksinim duyulan alanların geliştirilebilmesi için bazı kararlar alınır. Bu kararlar çoğu zaman ulusal veya uluslar arası düzeyde çeşitli ölçme ve değerlendirme çalışmalarından elde edilen bulgulara dayanır.

Türkiye'nin uluslararası düzeyde katıldığı çalışmalardan biri, Ekonomik İşbirliği ve Kalkınma Örgütü'nün (OECD) kısa adı PISA olan Uluslar Arası Öğrenci Değerlendirme Projesidir (Programme for International Student Assessment). En son uygulaması 2009 yılında yapılan PISA değerlendirmesinde ağırlıklı olarak ele alınan konu alanı okuma becerileri olmuştur. Bu dönemde öğrencilerin okuma etkinliklerine katılımına, kendi okuma ve öğrenme stratejileri hakkındaki görüşlerine odaklanılmıştır (MEB, 2010). PISA 2009 öğrenci anketinde öğrenme stratejileri özellikle okuma becerileri çerçevesinde yoklanmaya çalışılmıştır.

PISA çerçevesinde ele alınan öğrenme stratejileri Şekil 1'de gösterildiği gibidir.



Şekil 1: PISA 2009 değerlendirmesinde tanımlanan öğrenme stratejileri [MEB, 2010]

Proje kapsamından elde edilen veriler ile öğrencilerin bu stratejileri kullanabilmeyi düzeyleri ve okuma becerileri arasındaki ilişki, cinsiyetlere göre stratejileri kullanmaları arasındaki farklılıkları ve Türkiye'deki öğrencilerin diğer OECD ülkelerinden farklılaşma düzeyleri özellikle psikolojik yapılar çerçevesinde incelenmeye çalışılmaktadır. Söz konusu gruplar arasında yapılan karşılaştırmalarda ulaşılan sonuçlar, psikolojik yapılar üzerinde ölçme değişmezliği çalışmalarının yapılmadığı durumlarda tartışmalı olabilmektedir; çünkü değişmezliğin, yapılan karşılaştırmaların gücünü arttırdığı ve derinlemesine bir inceleme yapmaya olanak verdiği düşünülmektedir.

Bireylerin demografik özelliklerinin etkisinin arındırılmadığı (cinsiyet, okul türü, bölge, gelir düzeyi vb.) ölçme araçlarının kullanılması ve bunlardan elde edilen sonuçların karşılaştırılmalı olarak yorumlanması yanlış sonuçlar verebilir (Reise, Widaman ve Pugh, 1993). Klasik test kuramına göre (KTK) yalnızca geçerlik ve güvenilirlik kapsamında hesaplanan test ve madde istatistikleri, gruba bağlı olarak değişebilmekte ve grubun özelliklerini yansıtmaktadır (Crocker ve Algina, 1986).

KTK'deki bu sınırlılığa dayalı olarak, ölçme değişmezliğinin temel sorunu gündeme gelmiştir. Başka bir ifade ile gözlenen değişkenlerin ölçme niteliklerinin ne gibi durumlarda ve düzeylerde, gruplar arasında geçerli ve genellenebilir olduğu yönündeki sorun ortaya atılmıştır. Bu durumun ardında daha çok, “farklı kültürlerde yer alan cevaplayıcılar aynı maddeyi benzer şekilde yorumlayabiliyorlar mı, cinsiyet ve yaş gibi bireysel farklılıklar cevaplama etki ediyor mu?” şeklindeki sorular yer almaktadır (Vandenberg ve Lance, 1998, 2000). Söz konusu sorulara cevap aramak adına ölçmelerin değişmezliği çalışmaları yapılmaktadır (Vandenberg ve Lance 2000; Millsap ve Kwok, 2004; Raykov ve Marcoulides, 2006).

Ölçme değişmezliği problemi temelde şu biçimde ele alınmaktadır:

Ölçmeye konu olan bireylere ilişkin “k” farklı grup tanımlanmakta ve bu gruplara göre bireylerin karşılaştırılması amaçlanmaktadır. “k” farklı grup için ifade edilen genel teorik yapı,

$$X_k = \tau_k + \Lambda_k \xi_k + \delta_k \quad (1)$$

biçimindedir.

X_k , px1 şeklinde gözlenen değişken matrisi; τ_k , px1 deseninde gözlenen değişkenler ile örtük yapılar arasındaki katsayı vektörü; Λ_k , pxr deseninde faktör yükleri matrisi; δ_k , bağımsız gözlenen değişkenlerin hata vektörü; ξ_k , örtük faktörler vektörü biçiminde tanımlanmaktadır. Ayrıca, örtük yapılar ve hatalara ilişkin kovaryanslar,

$$Cov(\xi_k \delta_k) = 0, \quad Cov(\xi_k) = \Phi_k \quad \text{ve} \quad Cov(\delta_k) = \theta_k$$

şeklinindedir. Örtük yapılar ve hatalara ilişkin ortalamaların ise,

$$E(\xi_k) = \kappa_k \quad \text{ve} \quad E(\delta_k) = 0$$

biçiminde olduğu varsayılmaktadır. Bu durumda verilere ilişkin kovaryans ve ortalamalar denklemi,

$$Cov(X_k) = \mu_{xk} = \tau_k + \Lambda_k \kappa_k \quad (2)$$

$$E(X_k) = \Sigma_k = \Lambda_k \Phi_k \Lambda_k' + \theta_k \quad (3)$$

biçiminde olmaktadır.

Ölçme değişmezliğindeki problem ise [τ_k , Λ_k , θ_k] parametrelerinin k tane grupta eşit olup olmadığı yönündedir. Başka bir ifade ile bu parametrelerin alt gruplar için serbest bırakılması, sınırlandırılması veya birleştirilmesi doğrultusunda modelin yapısının gruplar arasında farklılaşmamasını belirtmektedir (Millsap ve Meredith, 2007, s. 131-139; Millsap ve Meredith, 2004).

Ölçme değişmezliği genellikle aşamalı şekilde hipotezlerin test edilmesi ile araştırılmaktadır. Yapısal (configural) değişmezlik, en temel değişmezlik düzeyidir. Yapısal değişmezliğe ilişkin kanıt elde edilmesi, ölçme aracındaki maddelerin gruplar arasında aynı yapıyı test ettiği anlamına gelmektedir (Vandenberg ve Lance, 1998). Bu aşamada teorisi kurulan modele, gruplar arasında aynı olma sınırlaması getirilmektedir (Wu, Li ve Zumbo, 2007; s. 7). Yapısal değişmezlik daha sınırlayıcı modellerin karşılaştırılabilmesi açısından temel bir yer teşkil etmektedir ve metrik değişmezliğin test edilebilmesi için sağlanması gereklidir (Cheung ve Rensvold, 2002; Little, 1997). Metrik değişmezlik, faktör yüklerinin

gruplar arasında değişmez olup olmadığını test etmektedir. Başka bir deyişle eğer metrik değişmezlik sağlanır ise, farklı gruplardaki bireylerin maddeleri benzer biçimde yorumladığı söylenebilmektedir. Bu değişmezliğin sağlanmadığı durumlarda, bir grupta bazı maddelerin diğer gruptakilere göre daha önemli veya dikkat çekici olduğu anlamı ortaya çıkabilmektedir (Byrne, 1998). Diğer bir aşama olan skalar değişmezlik aşaması sağlanır ise “gözlenen değişkenlerdeki ortalama farklılıkları, örtük yapıların ortalamalarından kaynaklanmaktadır” şeklinde bir yorum yapmaya imkan vermektedir. Katı değişmezlik madde eğilimlerinin ve sabitlerinin yanında, gruplar arasında madde artıklarının varyanslarının eşitliğini gerektirmektedir. Katı değişmezlik, çok sınırlandırılmış bir modeldir ve pratikte karşılanması zordur. Ayrıca, skalar değişmezliğin sağlanmadığı bir durum katı değişmezliğin de sağlanmayacağı anlamına gelmektedir.

Pratikte, ölçme aracının gruplar arasında tam değişmezliği sağlamanın her zaman mümkün olmadığı, ancak bu durumlarda bile en azından kısmi ölçme değişmezliğine yer verilmesi gerektiği belirtilmektedir (Steenkamp ve Baumgartner, 1998). Kısmi ölçme değişmezliğinde modele, parametrelerin eklenmesi, çıkarılması ya da parametrelerin birleştirilmesi söz konusudur (Reise ve ark., 1993). Tam değişmezlik ve kısmi değişmezlik çalışmaları ölçme araçlarının yapı geçerliliğini güçlendirdiği düşünülen ve araştırılması önemli olan bir özelliktir. Ölçme araçlarının değişmezliğini kontrol etmenin en etkili tekniği yapısal eşitlik modellemesi (YEM)’ dir (Byrne ve ark., 1989). Bu uygulamalarda, araştırmacılar genellikle varsayımları sağlayan bir ölçme aracı ile ölçülen örtük değişkenin boyutunu veya boyutlarını değerlendirmektedirler. Söz konusu yaklaşım, bir ölçme aracının psikometrik özellikleri incelendiğinde kullanılmaktadır (Raykov ve Marcoulides, 2006, s. 6).

Literatürde çeşitli ölçme araçlarının bazı demografik gruplarda ölçme değişmezliğini sağlayıp sağlamadığına ilişkin çalışmalar bulunmaktadır. Tucker, Ozer, Lyubomirsky ve Boehm (2005), araştırmalarında yaşam doyumu ölçeğininin değişmezliğini (The Satisfaction with Life Scale) Rus ve Kuzey Amerika örneklemelerini karşılaştırarak incelemişler ve ölçeğin, değişmezlik koşullarını sağlamadığı, dolayısı ile puanların gruplar arasında karşılaştırılmasının doğru olmayacağını belirtmişlerdir. Stein, Lee ve Jones (2006), araştırmalarında “The Sense of Coherence Scale of Antonovsky (Antonovsky uyum duygusu ölçeği) (1993)” için kültürler arası değişmezliği Çinliler, Japonlar ve beyazlar grupları arasında inceleyerek, ölçeğin etnik gruplar arasında farklılıklar gösterdiğini ortaya koymuşlardır.

Öğretmen (2006), araştırmasında PIRLS 2001 verilerini kullanmış ve testlerin yapılarının Amerika Birleşik Devletleri ve Türkiye örneklemelerinde ölçme eşdeğerliğine sahip olmadığını göstermiştir. Hagger, Aşçı, Lindwall, Hein, Ball, Tarrant, Ruiz ve Sell (2007), “sosyal fizik kaygısı ölçeğinin (social physique anxiety scale-SPAS)” kültürlerarası eşdeğerliğini çok gruplu doğrulayıcı faktör analizi ile incelemiş ve faktör yükleri ve hata varyanslarının sınırlandırılması ile elde edilen modeli gruplar arasında eşdeğer bulmamıştır. Bir diğer araştırmada PIRLS 2001 testinin faktör yapısının uygulamaya katılan tüm ülkelerde (35 ülke) birbirine orta düzeyde denk sayılabileceği bulgusunu elde edilmiştir (Akyıldız, 2009). Uzun (2008), öğrencilerin fen başarısına etki eden faktörlerin modellenmesini ve modelde yer alan değişkenlerin cinsiyet grupları arasında ölçme değişmezliğini araştırmıştır. Modele alınan tüm değişkenlerin cinsiyet gruplarında metrik değişmezlik koşullarını sağladığını, ancak katı değişmezlik koşulunu hiçbir değişkenin sağlamadığını ortaya koymuştur. Bu durumda, modelin cinsiyet grupları arasında karşılaştırılmasının pek doğru olmayacağını belirtmiştir.

Amaç

Aynı zamanda farklı gruplardan elde edilen ölçümlerin birbirine denk olup olmadığının belirlenmesi test geliştirme sürecinde oldukça önemlidir. Bir ölçme aracından elde edilen puanların herhangi bir durum için karşılaştırılması, ölçme aracının farklı gruplarda aynı özelliği ölçebildiği durumlarda sağlanmakta ve o zaman anlamlı yorumlar yapılmasına imkan vermektedir. Bu çalışmanın amacı PISA 2009 değerlendirmesinde, öğrenme stratejileri modelinin ve öğrenme stratejileri ile okuma becerileri arasında tanımlanan modelin farklı gruplarda (cinsiyet, okul türü ve istatistiksel bölge) aynı yapıyı ölçüp ölçmediğini belirlemektir.

YÖNTEM

Evren-Örnekleme

Araştırmanın evrenini Türkiye’de eğitime devam eden 15 yaş öğrencileri oluşturmaktadır. Türkiye bu uygulamada, 12 istatistiksel bölge biriminden (NUTS) 56 il ve bu illerdeki okul türlerine göre tabakalandırılarak PISA ulusal merkez tarafından seçkisiz yöntemle belirlenen 170 okuldan 4996 öğrenci ile yer almıştır. Bu araştırma 4996 öğrenci üzerinden yürütülmüş ve anket maddelerine eksik yanıt veren öğrenciler değerlendirmeye alınmamıştır. Ayrıca analiz türlerindeki varsayımlarının incelenmesi sonucunda, koşulları sağlamayan bireyler araştırma dışında bırakılmıştır. Son durumda araştırma 4340 veri ile yürütülmüştür.

Araştırmaya 2162 kız öğrenci ve 2178 erkek öğrenci alınmıştır. Araştırma kapsamına alınan öğrenciler okul türlerine göre genel liseler, Anadolu ve fen liseleri ve teknik ve mesleki liseler olarak gruplandırılmıştır. MEB (2010) raporuna göre ilköğretim öğrencilerinin puanları tüm lise türlerindeki öğrencilerin puanlarından anlamlı bir şekilde düşüktür ve çalışmalarda ayrı değerlendirilmesi önerilmektedir. Anadolu ve fen lisesi öğrencilerinin okuma becerileri performanslarına göre başarı puanları birbirine oldukça yakındır ve puanlar arasında anlamlı bir fark yoktur. Genel liseye devam eden öğrencilerin başarı puanları Anadolu ve fen lisesi öğrencilerinin puanlarından anlamlı bir şekilde düşüktür. Teknik ve mesleki liselerdeki öğrencilerin puanları ise genel lise öğrencilerinin başarı puanlarından düşüktür. Teknik ve mesleki lise okul türleri kapsamına güzel sanatlar lisesi, çok programlı liseler ile Anadolu teknik ve mesleki liseler de dahil edilmiştir. Bu okulların başarı puanları teknik ve mesleki lisedeki öğrencilerin puanlarına yakın olduğu için araştırma kapsamında tek grup olarak birleştirilmesi uygun görülmüştür. Araştırmaya alınan öğrencilerin 1645’i (% 37.9) genel liseden, 813’ü (% 18.7) Anadolu, Fen ve Anadolu Öğretmen Liselerinden ve 1789’u (% 41.3) meslek lisesi ve teknik liselerden oluşmaktadır.

Araştırmaya İstanbul Bölgesi’nden 689 (% 15.9), Batı Marmara Bölgesi’nden 224 (% 5.2), Ege Bölgesi’nden 546 (% 12.6), Doğu Marmara Bölgesi’nden 446 (% 10.3) ve Batı Anadolu Bölgesi’nden 417 (% 9.6) öğrenci alınmıştır. Ayrıca, Akdeniz Bölgesi’nden 564 (% 13.0), Orta Anadolu Bölgesi’nden 265 (% 6.1), Batı Karadeniz Bölgesi’nden 311 (% 7.2), Doğu Karadeniz Bölgesi’nden 181 (% 4.2), Kuzeydoğu Anadolu Bölgesi’nden 120 (% 2.8), Ortadoğu Anadolu Bölgesi’nden 193 (% 4.4) ve son olarak Güneydoğu Anadolu Bölgesi’nden 384 (% 8.8) öğrenci araştırma grubunu oluşturmaktadır.

Veri Toplama Araçları

Araştırmada, Uluslar Arası Eğitim Başarısını Belirleme Kuruluşunun (IEA) PISA 2009 değerlendirmesinde Türkiye örnekleme uygulanan öğrenci anketinde yer alan maddelere verdikleri yanıtlar veri olarak kullanılmıştır. Veriler 2010 yılı Aralık ayında OECD resmi web sitesinde yayımlanmış olan rapordan alınmıştır. Ankette yer alan öğrenmeyi öğrenme bölümünden, öğrenme stratejilerine ilişkin maddelere verilen yanıtlar ve öğrencilerin kişisel bilgilerine ilişkin cinsiyet, okul türü ve katıldıkları bölge dikkate alınmıştır.

Verilerin analizi

Araştırma verilerinin analizi için SPSS 15.0 ve LISREL 8.80 for Windows programlarından yararlanılmıştır. SPSS 15.0 programı verilerin dağılımını belirlemek ve açıklayıcı faktör analizi yapmak amacı ile kullanılmıştır. LISREL programı ise araştırma problemi doğrultusunda, öğrencilerin okuma becerilerini etkileyen faktörlere ilişkin model oluşturma ve modelin gruplar arasında değişmezliğini incelemek için kullanılmıştır. Modelin değişmezliği çoklu gruplu doğrulayıcı faktör analizi (Multi Group Confirmatory Factor Analysis-MGCF) ile incelenmiştir (Jöreskog ve Sörbom, 1993; Byrne, 1998, s: 291-341; Schmacher ve Lomax, 2004; Raykov ve Marcoulides, 2006).

Verilerden doğru bir sonuç çıkartılabilmesi açısından analizlere başlamadan önce veri seti, veri yapısı ve verilerin analizlere ilişkin varsayımları karşılayıp karşılamadığı incelenmelidir. Analiz sonuçlarının kalitesi verilerin niteliğine bağlı olabileceğinden bu çalışmada veriler aşağıdaki adımlarla incelenmiştir.

Analiz sonuçlarında büyük değişikliklere yol açabileceğinden öncelikle kayıp veriler araştırılmıştır.

Uç Değerler: Kayıp verilerin analiz dışında bırakılmasının ardından geriye kalan veriler üzerinden tek değişkenli aykırı değer bulunup bulunmadığı incelenmiştir. Bu amaçla hesaplanan z değerlerinden ± 3 değerini aşanlar analiz dışında bırakılmıştır. Doğrulayıcı faktör analizi için bir ön koşul olan çok değişkenli artık değerlerden arınlık durumu Mahalanobis uzaklıkları (Mahalanobis Distance- D^2) hesaplanarak test edilmiştir. Sonuçlar, verilerde çok değişkenli aykırı değer bulunmadığını göstermiştir. Kayıp değer ve uç değerlerin çıkarılmasının ardından araştırma 4340 veri ile yürütülmüştür. Verilerin normal olup olmadığına karar vermede araştırmada kullanılan bağımsız değişkenlerin her birine ilişkin basıklık ve çarpıklık değerleri ile ortalamalarının standart sapmalarına oranları (bağıl değişim katsayısı) dikkate alınmıştır. Sonuçlar, değişkenlere ilişkin verilerin normal dağılıma uygun olduğunu göstermiştir. Maddelerin birbiri ile ilişkileri ve maddeler arasında çoklu bağlantı problemi olup olmadığı incelenmiştir. Çoklu bağlantı problemi öğrenme stratejilerini oluşturduğu düşünülen beş faktör için ayrı ayrı hesaplanmıştır. Her bir faktörde maddelerin birbiri ile ilişkisinin oldukça düşük olduğu, tolerans değerlerinin istenilen düzeyde olduğu, varyans şişkinlik faktör (variance inflation factor-VIF) değerlerinin 10'dan çok düşük değerler verdiği ve koşul (durum) indeksi (Condition index-CI) değerlerinin 30'dan çok düşük olduğu gözlenmiştir. Bu sonuçlar, maddeler arasında çoklu bağlantı problemi bulunmadığının kanıtı olarak kullanılmıştır (Gujarati, 2004, s: 348-359; Kline, 2005, s: 57; Stevens, 2009, s: 74-80; Büyüköztürk, 2010).

Yapılan incelemeler sonucunda verilerin çok değişkenli istatistiklere uygun olduğuna ve normal dağılım gösterdiğine karar verilmiştir. Bu sonuçtan yola çıkarak, araştırmada

tahmin yöntemi olarak en çok olabilirlik yöntemi (Maximum Likelihood Method) kullanılmıştır.

Öğrenme stratejileri için oluşturulan ölçme modeli LISREL 8.80 programı aracılığı ile doğrulayıcı faktör analizi tekniğinden yararlanılarak doğrulanmaya çalışılmıştır. Modele dair uyum katsayıları hesaplanmıştır ve modelin veri ile uyumuna karar verebilmek için kestirimin Ki-Kare, RMSEA, CFI, ECVI, NFI ve GFI değerleri dikkate alınmıştır, ancak Ki-kare değeri örneklem büyüklüğünden etkilenebileceğinden modelin uyumunun değerlendirilmesinde daha çok diğer uyum ölçütleri kullanılmıştır (Jöreskog ve Sörbom, 1993). Modelin veri ile uyum gösterdiği durumlarda değişmezlik testlerine geçilmiştir.

Değişmezlik aşamalarında cinsiyet için kızlar, okul türü için Anadolu ve fen liseleri, istatistiksel bölgeler için İstanbul bölgesi referans grup olarak alınmıştır. Diğer gruplara getirilen sınırlandırmalar referans gruplar doğrultusundadır.

Model değişmezliği testi için aşamalı olarak elde edilen modellerde uyum indekslerin (Ki-Kare, RMSEA, GFI, CFI) kabul edilebilir aralıkta olup olmadığı incelenmiştir. Modelin veri ile uyum gösterdiği durumlarda yapısal değişmezlik testlerinden elde edilen NNFI (TLI) ve CFI değerleri ile daha sınırlı modellerden elde edilen TLI ve CFI değerleri arasındaki farka bakılmıştır. Bu değerler arasındaki farkın çok küçük olması beklenmektedir. (Cheung ve Rensvold, 2000; Vandenberg ve Lance, 2000). İç içe geçmiş (nested) modellerin karşılaştırılmasında daha çok CFI ya da TLI fark değerlerinin (ΔCFI ve ΔTLI) kullanılması önerilmektedir. (Cheung ve Rensvold, 2002; Vandenberg ve Lance, 2000; Hu ve Bentler, 1998). Hesaplanan farkların $-0.01 \leq \Delta CFI \leq 0.01$ ve $-0.01 \leq \Delta TLI \leq 0.01$ aralığı içerisinde bulunmadığı durumlarda ilgili aşamada değişmezlik koşulunun sağlanmadığına karar verilmiştir (Cheung and Rensvold, 2002; Brannick, 1995; Kelloway, 1995; aktaran: Rusticus ve Hubley, 2006).

Araştırmada Kullanılan Değişkenler

Ezberleme Stratejileri (ezb)

ST27Q01: Ders çalışırken metinde yer alan her şeyi ezberlemeyi denerim.

ST27Q03: Ders çalışırken mümkün olduğu kadar fazla ayrıntıyı ezberlemeyi denerim.

ST27Q05: Ders çalışırken konuyu ezbere anlatabilecek kadar çok defa okurum.

Keşfetme Stratejileri (kesfetme)

ST27Q08: Ders çalışırken bu bilginin okul dışında ne kadar faydalı olabileceğini düşünürüm.

ST27Q10: Ders çalışırken, bu dökümanla kendi yaşantım arasında bağ kurarak dökümanı daha iyi anlamayı denerim.

ST27Q12: Ders çalışırken, metindeki bilginin, gerçek yaşamla ne kadar uyduğunu anlarım.

Kontrol Stratejileri (kontrol)

ST27Q02: Ders çalışmaya tam olarak ne öğrenemem gerektiğini düşünerek başlarım.

ST27Q04: Ders çalışırken başka konulardan edindiğim önceki bilgilerimle, yeni bilgiler arasında bağlantı kurmayı denerim.

ST27Q06: Ders çalışırken okuduğum metni anlayıp anlamadığımı kontrol ederim.

ST27Q09: Ders çalışırken, hala anlamadığım kavramların ne olduğunu bulmaya çalışırım.

ST27Q11: Ders çalışırken, metinde yer alan önemli noktaları hatırladığımdan emin olurum.

ST27Q13: Ders çalışırken, anlamadığım bir şey olursa, bunu açıklığa kavuşturmak için ek bilgiler ararım.

Anlama ve Bilgiyi Hatırlama Stratejileri (bilhat)

ST41Q03: Metni okuduktan sonra metnin içeriğini diğer insanlarla tartışırım.

ST41Q04: Metnin önemli kısımlarının altını çizerim.

ST41Q05: Metni kendi kelimelerimi kullanarak özetlerim.

ST41Q06: Metni sesli bir şekilde, bir başkasına okurum.

Bilgiyi Özetleme Stratejileri (bilozet)

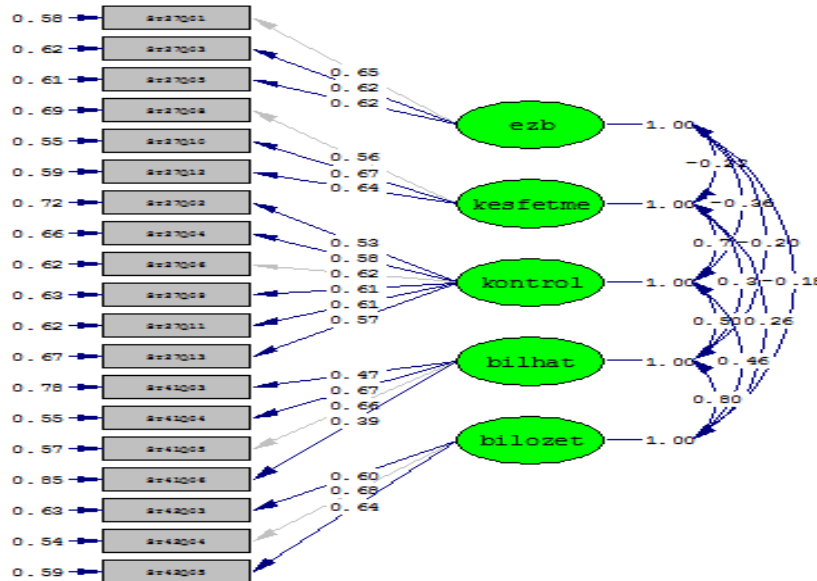
ST42Q03: Özeti yazmadan önce, metni mümkün olduğunca çok defa okurum.

ST42Q04: Metinde yer alan önemli gerçeklerin, benim özetimde yer alıp almadıklarını kontrol ederim.

ST42Q05: Metnin genelini okurum, en önemli cümlelerin altını çizerim. Bu cümleleri, kendi kelimelerimi kullanarak özet olarak yazarım.

BULGULAR VE YORUM

PISA öğrenci anketinde yer alan ezberleme, keşfetme, kontrol, anlama-bilgiyi hatırlama ve bilgiyi özetleme stratejileri faktörlerinin öğrenme stratejileri modelini oluşturup oluşturmadığını belirlemek için yapılan doğrulayıcı faktör analizinden elde edilen path diyagramı Şekil 2'de verilmiştir.



Şekil 2. PISA 2009 Türkiye örnekleminde öğrencilerin öğrenme stratejilerine ilişkin ölçme modeli

Tablo 1: Öğrenme Stratejileri Modelinin Uyum Katsayıları

| | χ^2 | df | $\frac{\chi^2}{df}$ | <i>RMSEA</i> | <i>CFI</i> | <i>GFI</i> | <i>AGFI</i> |
|----------------------------------|---------------|------------|---------------------|--------------|-------------|-------------|-------------|
| Öğrenme Stratejileri Modeli | 1164.39 | 142 | 8.19 | 0.041 | 0.97 | 0.97 | 0.96 |
| Kızlar | 733.97 | 142 | 5.17 | 0.044 | 0.96 | 0.97 | 0.95 |
| Erkekler | 576.16 | 142 | 4.06 | 0.037 | 0.98 | 0.97 | 0.96 |
| Anadolu ve Fen Lisesi | 329.73 | 142 | 2.32 | 0.040 | 0.96 | 0.96 | 0.95 |
| Genel Lise | 578.10 | 142 | 4.07 | 0.043 | 0.97 | 0.96 | 0.95 |
| Teknik ve Mesleki Liseler | 478.78 | 142 | 3.37 | 0.036 | 0.98 | 0.97 | 0.96 |
| Batı Marmara Bölgesi | 183.86 | 142 | 1.29 | 0.036 | 0.96 | 0.92 | 0.89 |
| İstanbul Bölgesi | 302.59 | 142 | 2.13 | 0.041 | 0.97 | 0.96 | 0.94 |
| Ortadoğu Anadolu Bölgesi | 199.43 | 142 | 1.40 | 0.046 | 0.95 | 0.90 | 0.87 |
| Güneydoğu Anadolu Bölgesi | 264.19 | 142 | 1.86 | 0.047 | 0.97 | 0.93 | 0.91 |
| Ege Bölgesi | 290.78 | 142 | 2.05 | 0.044 | 0.97 | 0.95 | 0.93 |
| Batı Karadeniz Bölgesi | 227.93 | 142 | 1.61 | 0.044 | 0.96 | 0.93 | 0.90 |
| Doğu Marmara Bölgesi | 221.68 | 142 | 1.56 | 0.036 | 0.98 | 0.95 | 0.93 |
| Doğu Karadeniz Bölgesi | 205.81 | 142 | 1.45 | 0.050 | 0.97 | 0.89 | 0.86 |
| Kuzeydoğu Anadolu Bölgesi | 218.29 | 142 | 1.54 | 0.067 | 0.85 | 0.84 | 0.78 |
| Batı Anadolu Bölgesi | 268.36 | 142 | 1.89 | 0.046 | 0.96 | 0.94 | 0.91 |
| Orta Anadolu Bölgesi | 257.03 | 142 | 1.81 | 0.055 | 0.94 | 0.91 | 0.88 |
| Akdeniz Bölgesi | 282.18 | 142 | 1.99 | 0.042 | 0.97 | 0.95 | 0.93 |

İlgili modele ait uyum istatistikleri incelendiğinde, model uyumunun değerlendirilmesinde kullanılan ölçütlerin kabul edilebilir aralıklarda yer aldığı söylenebilir ($RMSEA < 0.08$, $CFI > 0.90$, $GFI > 0.90$ ve $AGFI > 0.85$). Modelin ayrıca, her bir grupta veri ile uyumu incelenmelidir. Bu amaçla alt gruplarda modele ilişkin uyum ölçütlerine bakıldığında yalnızca Doğu Karadeniz Bölgesi ve Kuzeydoğu Anadolu Bölgesi'nde, Tablo 1'de görüldüğü gibi *CFI*, *GFI* ve *AGFI* değerlerinin model uyumu için sınır değerlerin hemen altında olduğu ifade edilebilir. Bu bölgeler programın çalıştırılması esnasında sorun çıkardığından ve kovaryans matrisinin negatif çıkmasına sebep olduklarından, modelin bölgeler arası değişmezliği incelenirken söz konusu bölgelerin analiz dışında bırakılmasına karar verilmiştir. Son durumda öğrenme stratejileri modelinin 19 madde ve beş boyuttan oluşan bir model şeklinde doğrulandığı söylenebilir. Model veri uyumunu sağlayan gruplar için değişmezlik analizleri yapılmıştır. Cinsiyet için yapılan değişmezlik analizlerine ilişkin sonuçlar Tablo 2'de verilmiştir.

Tablo 2: Cinsiyet Gruplarında Değişmezlik Analiz Sonuçlarına Göre Uyum Katsayıları

| Değişmezlik | χ^2 | df | <i>RMSEA</i> | <i>ECVI</i> | <i>CFI</i> | <i>GFI</i> | <i>NFI</i> | <i>TLI</i> | ΔTLI | ΔCFI |
|-------------|----------|------|--------------|-------------|------------|------------|------------|------------|--------------|--------------|
| Yapısal | 1310.12 | 284 | 0.041 | 0.34 | 0.97 | 0.97 | 0.96 | .96 | - | - |
| Metrik | 1358.38 | 298 | 0.041 | 0.35 | 0.97 | 0.97 | 0.96 | .96 | 0 | 0 |
| Skalar | 2128.89 | 331 | 0.050 | 0.57 | 0.95 | 0.97 | 0.94 | .94 | 0.02 | 0.02 |
| Katı | 2355.50 | 365 | 0.050 | 0.57 | 0.94 | 0.96 | 0.93 | .93 | 0.03 | 0.03 |

Tablo 2'ye göre model uyumunun değerlendirilmesinde kullanılan uyum ölçütlerinin yapısal değişmezlik aşamasında kabul edilebilir sınırlar içerisinde olduğu söylenebilir ($RMSEA < 0.08$, $CFI > 0.90$, $GFI > 0.90$ ve $TLI > 0.90$). Elde edilen bulgular ışığında, bu aşamada yalnızca modelin yapısının her iki grupta aynı olduğu söylenebilir. Metrik

değişmezlik aşamasında faktör yüklerine gruplar arasında eşit olma sınırlılığı getirilmesinin ardından elde edilen uyum ölçütlerine göre modelin veri ile iyi uyum gösterdiği sonucuna ulaşılabilir. Bu adımdan sonra yapısal değişmezlik ile metrik değişmezlik arasındaki ΔCFI ve ΔTLI fark değerleri incelendiğinde metrik değişmezliğin sağlandığı söylenebilir ($-0.01 \leq \Delta CFI \leq 0.01$, $-0.01 \leq \Delta TLI \leq 0.01$). Elde edilen bu bulgu, kız ve erkeklerde ölçülen özellikler ile ölçeğin boyutları arasındaki ilişkilerin benzer olduğu yorumunu yapmaya imkan verebilir. Bu durumda, kızlar ve erkeklerin maddelere benzer biçimde yanıt verdiği ve gruplardan elde edilen puanların karşılaştırılmasının anlamlı olduğunu söylemek mümkün olabilir. Tablo 2 incelendiğinde skalar değişmezlik aşamasında uyum indekslerinin kabul edilebilir sınırlar içerisinde yer aldığı görülebilir. Yapısal değişmezlikten elde edilen CFI ve TLI ile skalar değişmezlik için elde edilen CFI ve TLI değerleri arasındaki farklar incelendiğinde modelin skalar değişmezliği sağlamadığı ifade edilebilir ($\Delta TLI = 0.02$, $\Delta TLI = 0.02$). Elde edilen bulgular ışığında maddelere ilişkin faktör yükleri cinsiyetler arasında değişmez iken, faktörler arası korelasyonların değişmezlik sergileyemediklerini söylemek mümkün olabilir. Bu durum, örtük yapıların ortalamalarının gruplar arasında farklılaştığı ve kız ve erkeklere ilişkin ortalamaların karşılaştırılmasının anlamlı olmayacağı şeklinde yorumlanabilir. Aşamalı olarak yapılan değişmezlik çalışmalarında bir değişmezlik koşulunun yerine getirilmediği durumda diğer bir aşamaya geçilmesinin söz konusu olmadığı belirtilmektedir. Modelin, kız ve erkek gruplarında skalar değişmezlik testini sağladığına ilişkin bir kanıt bulunmadığı düşünüldüğünden, bu gruplarda katı değişmezliği incelemenin doğru olmayacağı söylenebilir. Tablo 7'ye göre katı değişmezlik aşamasındaki ΔCFI ve ΔTLI değerleri 0.03 olarak bulunmuştur. Bu durum skalar değişmezlik koşulunu yerine getirmeyen bir model için katı değişmezliğin de sağlanmayacağı bir göstergesi olarak düşünülebilir.

Okul türleri için yapılan değişmezlik analizlerine ilişkin sonuçlar Tablo 3'te verilmiştir.

Tablo 3. Okul Türlerinde Değişmezlik Analiz Sonuçlarına Göre Uyum Katsayıları

| | χ^2 | df | RMSEA | ECVI | CFI | GFI | NFI | TLI | ΔTLI | ΔCFI |
|---------|----------|-----|-------|------|------|------|------|-----|--------------|--------------|
| Yapısal | 1386.60 | 426 | 0.040 | 0.39 | 0.97 | 0.96 | 0.96 | .97 | - | - |
| Metrik | 1419.74 | 454 | 0.039 | 0.39 | 0.97 | 0.96 | 0.96 | .97 | 0 | 0 |
| Skalar | 2229.99 | 506 | 0.049 | 0.58 | 0.95 | 0.96 | 0.94 | .95 | 0.02 | 0.02 |
| Katı | 2572.06 | 574 | 0.050 | 0.63 | 0.95 | 0.95 | 0.93 | .95 | 0.02 | 0.02 |

Tablo 3'teki bilgilere göre, yapısal değişmezlik aşamasında uyum indekslerinin kabul edilebilir sınırlar içerisinde yer aldığı ve bu sayede alt gruplar için yapısal değişmezliğin sağlandığı söylenebilir (RMSEA<0.05, CFI>0.90, GFI>0.90, NNFI>0.90). Bu durum, örtük yapıların gruplar arasında değişmediği, Anadolu ve fen lisesi öğrencileri, genel lisedeki öğrenciler ve teknik ve mesleki lisedeki öğrenciler için modeldeki örtük değişkenlerin yapısının benzer olduğu şeklinde yorumlanabilir. Metrik değişmezliğe ilişkin analiz sonuçlarına göre uyum indekslerinin kabul edilebilir sınırlar içerisinde olduğu (RMSEA < 0.05, CFI > 0.90, GFI > 0.90, NNFI > 0.90) ve metrik değişmezliğin sağlandığı ($-0.01 \leq \Delta CFI \leq 0.01$ ve $-0.01 \leq \Delta TLI \leq 0.01$) söylenebilir. Bu bulgudan hareket ederek metrik değişmezlik koşulunu, okul türleri için sağlayan modelde, örtük değişkenler ile bunlara ait maddeler arasındaki ilişkiler alt gruplarda aynıdır yorumunu yapmak mümkün olabilir. Tablo incelendiğinde okul türü grupları için skalar değişmezlik aşamasında uyum ölçütlerinin kabul edilebilir sınırlar içerisinde olduğu söylenebilir. Bu adımdan sonra yapısal değişmezlik ile skalar değişmezlik arasındaki fark değerleri dikkate alındığında, modelin skalar değişmezliği sağlamadığı görülebilir ($\Delta CFI = 0.02$, $\Delta TLI = 0.02$). Skalar değişmezlik koşulu sağlanmayan

bir modelin katı değişmezliği sağlamayacağı tablo aracılığı ile görülebilir ($\Delta CFI = 0.02$, $\Delta TLI = 0.02$).

İstatistiksel bölge türleri için yapılan değişmezlik analizlerine ilişkin sonuçlar Tablo 4'te verilmiştir.

Tablo 4: İstatistiksel Bölgelerde Değişmezlik Analiz Sonuçlarına Göre Uyum Katsayıları

| | χ^2 | df | RMSEA | ECVI | CFI | GFI | NFI | TLI | ΔTLI | ΔCFI |
|---------|----------|------|-------|------|------|------|------|------|--------------|--------------|
| Yapısal | 2511.93 | 1435 | 0.043 | 0.85 | 0.96 | 0.94 | 0.92 | 0.96 | - | - |
| Metrik | 2654.64 | 1561 | 0.042 | 0.83 | 0.96 | 0.94 | 0.92 | 0.96 | 0 | 0 |
| Skalar | 3143.70 | 1746 | 0.045 | 0.95 | 0.96 | 0.94 | 0.91 | 0.96 | 0 | 0 |
| Katı | 3474.18 | 2037 | 0.042 | 0.89 | 0.96 | 0.94 | 0.90 | 0.96 | 0 | 0 |

Tablo incelendiğinde, tüm değişmezlik testi aşamalarında uyum indekslerinin kabul edilebilir aralıklarda yer aldığı ve modelin veri ile iyi uyum gösterdiği söylenebilir (RMSEA < 0.05, CFI > 0.90, GFI > 0.90, NNFI > 0.90). Modelin bölgeler arasında yapısal değişmezliğinin testinden elde edilen CFI ve TLI değerleri ile daha sınırlı modellerden elde edilen bu değerler arasındaki farka bakıldığında ölçeğin alt gruplarda tüm değişmezlik aşamalarını sağladığı söylenebilir ($\Delta TLI=0.00$ ve $\Delta TLI=0.00$).

SONUÇ VE TARTIŞMA

Çoğu araştırmalarda çeşitli gruplar arasındaki farklılıklar ortaya konmaya çalışılmaktadır. Bu farklılıkların ölçme aracından meydana gelip gelmediği ise araştırma sonuçlarına gölge düşüren bir durum olabilmektedir.

Öğretmen (2006) PIRLS okuma testlerinin Türkiye ve Amerika Birleşik Devletleri'nde eşdeğer olmadığını belirtmiştir. Bu araştırma, Öğretmen (2006) ve Uzun (2008)'in bulgularını destekleyerek uluslar arası sınavlarda kullanılan ölçme araçlarının çeşitli alt gruplarda farklı yorumlanabileceğini göstermektedir.

Bu çalışmada 5 faktör ve 19 madde şeklinde doğrulanan öğrenme stratejileri modelinde cinsiyete dayalı olarak yapılan değişmezlik testi sonuçları, modelin tam değişmezlik koşullarını yerine getirmediğini göstermiştir. Model kız ve erkek gruplarında yapısal ve metrik değişmezlik koşullarını sağlamıştır. Teorisi kurulan modelin cinsiyet gruplarında yapısı değişmezdir ve örtük değişkenler ile gözlenen değişkenler arasındaki ilişkiler alt gruplarda eşittir, ancak öğrenme stratejileri modeli skalar ve katı değişmezlik koşullarını sağlamamıştır. PISA 2009 Türkiye raporunda, cinsiyet gruplarında öğrenme stratejileri kullanma durumlarının farklılık gösterdiği, kız öğrencilerin ezberleme ve keşfetme stratejilerini, erkek öğrencilerin ise kontrol stratejilerini daha sık kullandıkları belirtilmiştir. Bireyler arasında örtük yapılar verilen yanıtlar, ölçülen özellik aynı olmasına rağmen, örtük yapıların gruplar arasında aynı yorumlanamamasından dolayı farklı çıkmış olabilir.

Öğrenme stratejileri modeli, okul türüne göre alt gruplarda yalnızca yapısal ve metrik değişmezlik koşullarını sağlamıştır. Model üzerinde modifikasyon yapılmadan parametrelere getirilen sınırlamalar doğrultusunda modelin alt gruplarda skalar değişmezlik ve katı değişmezlik koşullarını yerine getirdiğine dair kanıt bulunamamıştır. Öğrenme stratejileri modelinin okul türü gruplarında farklılık göstermesi, eğitim seviyelerinin farklılığından, derslerde gördükleri davranışlar kapsamında, eğitim ortamından vb. kaynaklanıyor olabilir.

PISA 2009 okuma becerileri alanında Türkiye örneklemini oluşturan öğrenciler arasında en düşük ortalamaya sahip öğrencilerin, meslek lisesi öğrencilerinin olduğu, teknik lise öğrencilerinin de onlara yakın bir puan ortalamasına sahip olduğu belirtilmektedir (MEB, 2010). Okullardaki başarı farklılıkları, öğrenme stratejileri modelinin bu gruplar arasında farklı yorumlanmasına neden olmuş olabilir. Okul türleri arasındaki böylesine büyük farklar OECD ülkeleri içerisinde en çok Türkiye’de gözlenmektedir (EARGED, 2004). Eğitim olanaklarının okul türlerine göre eşit dağıtılamaması ve bireylerin bu olanaklardan aynı ölçüden yararlanamaması bu sonucu doğurmuş olabilir (Berberoğlu ve Kalender, 2005).

Model istatistiksel bölge birimleri arasında, parametrelere sınırlandırma getirilerek elde edilen değişmezlik aşamalarında yapısal, metrik, skalar ve katı değişmezlik koşullarını yerine getirmiştir. Model istatistiksel bölge alt gruplarında eşittir ve gruplar arasında öğrenme stratejileri yönündeki herhangi bir farklılık ölçekten kaynaklanmamaktadır. Wu ve arkadaşları (2007), TIMSS 1999 değerlendirilmesinde kullanılan matematik testlerinin farklı kültürlerde sadece yapısal ve metrik değişmezlik koşullarının en iyi şekilde karşıladığını göstermişlerdir. Aynı kültürlerde ise ölçeğin katı değişmezliği sağladığını belirtmişlerdir. Bu bulgu Türkiye’de aynı kültür altında bulunup farklı istatistiksel bölgelerden alınan öğrencilerde, öğrenme stratejileri modeli katı değişmezlik koşulunu sağladığı bulgusu ile tutarlılık göstermektedir.

ÖNERİLER

Araştırmada sonuçlarına göre ortaya çıkan öneriler şu şekildedir:

Farklı ülkelerde cinsiyet grupları arasında modelin yapısının nasıl olduğuna ilişkin Türkiye ile karşılaştırılmalı olarak bir çalışma yapılabilir. Yine bu ülkelerdeki okulların türlerine göre (eğitim seviyesi daha düşük vb.) modelin değişmezliği incelenebilir. Cinsiyet ve okul türü gruplarında modelin ölçme değişmezliğinin sağlanmamasına neden olan maddelerin belirlenmesi, gruplar arasında maddelerin DMF (değişen madde fonksiyonu) gösterip göstermediği belirlenebilir. Bu araştırmada istatistiksel bölge birimlerinden Doğu Karadeniz ve Kuzeydoğu Anadolu Bölgeleri, model veri ile iyi uyum göstermediğinden araştırma kapsamından çıkarılmıştır. Bu gruplarda modelin kötü uyum göstermesinin nedenlerini saptamak amacı ile uyumsuzluğa yol açan maddelerin incelenmesi önerilebilir. Araştırma tam değişmezlik aşamaları kullanılarak yürütülmüştür. Tam değişmezlik koşulunu sağlamayan modellerin en azından kısmi değişmezlik yönünden incelenmesi, modelde teoriye uygun olacak şekilde modifikasyon yapılması önerilebilir. Öğrencilerin öğrenmelerini kalıcı kılmak amacı ile sıklıkla yararlandığı yöntemlerin belirlenmesi ve bu tür tekniklerin okullara ve ülke geneline yaygınlaştırılması sağlanabilir. Öğrencilerin bu stratejileri kullandıklarına dair farkındalıklarının artırılmasına yönelik çalışmalar yapılabilir.

KAYNAKÇA

- AKYILDIZ, M. (2009) PIRLS 2001 Testinin Yapı Geçerliliğinin Ülkelerarası Karşılaştırılması. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 6(1), 18-47.
- BERBEROĞLU, G. ve KALENDER, İ. (2005) “Öğrenci Başarısının Yıllara, Okul Türlerine, Bölgelere Göre İncelenmesi: ÖSS ve PISA Analizi” *Eğitim Bilimleri ve Uygulama*, 4(7), 21-35
- BYRNE, B. M. (1998). *Structural equation modeling with lisrel, prelis, and simplis: Basic concepts, applications, and programming*. USA: Lawrence Erlbaum Associates.
- CHEUNG, G. W. & RENSVDL, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255.
- CROCKER, L. & ALGINA, J. (1986). *Introduction to classical and modern test theory*. Philadelphia: Harcourt Brace Jovanovich College Publishers:
- HAGGER, M. S., AŞCI, F., LINDWALL, M., HEIN, V., BALL, M. O., TARRANT, M. & colleagues (2007) “Cross-Cultural validity and measurement invariance of the social physique anxiety scale in five european nations” *Scandinavian Journal of Medicine & Science in Sports*. Signapore, 17: 703-719.
- HU, L. & BENTLER, P. (1995). Evaluating model fit. R. Hoyle (Ed). *Structural Equation Modeling Concept, Issues and Application* (pp. 76-99). Thousand Oaks: Sage Publications.
- JORESKOG, K. G. & SORBOM, D. (1993). *LISREL 8: Structural equation modeling with the simplis command language*. USA: Scientific Software International, Inc.
- KLINE, R. B. (2005). *Principles and practice of structural equation modeling* (2nd Edt.). New York: The Guilford Press.
- LITTLE, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: practical and theoretical issues. *Multivariate Behavioral Research*, 32(1), 53-76.
- MİLLİ EĞİTİM BAKANLIĞI, (2009). *Türkiye Raporu*, Ankara: MEB
- MILLSAP, R. E & MEREDITH, W. (May-2004). Factorial invariance: historical trends and new developments [Bildiri]. *Factor Analysis at 100 Conference*, Thurstone Psychometric Laboratory, University of North Carolina
- MILLSAP, R. E & MEREDITH, W. (2007). Factor invariance: historical perspectives and new problems. R. Cudeck & R. C. MacCallum (Eds). *Factor analysis at 100 historical developments and future directions* (pp. 131-152). Mahwah, NJ: Lawrence Erlbaum.
- ÖĞRETMEN, T. (2006). “Uluslar Arası Okuma Becerilerinde Gelişim Projesi (PIRLS) Testinin Psikometrik Özelliklerinin İncelenmesi: Türkiye-Amerika Birleşik Devletleri Örneği.” *Doktora Tezi*, Hacettepe Üniversitesi, Ankara.

- RAYKOV, T. & MARCOULİDES, G. A (2006). A first course in structural equation modeling (2nd Edt). London: Lawrence Erlbaum Association.
- REISE, S. P., WİDAMAN, K. F. & Pugh, R. H. (1993) "Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance" *Psychological Bulletin*, 114(3), 552-566
- RUSTICUS, S. S & HUBLEY, M. A. (August 2006) "Measurement invariance of the Ası-R and Bıqlı across gender and age [Bildiri]" Annual Meeting of the American Psychological Association (APA), New Orleans, U.S.A.
- SCHUMACHER, R. E. & LOMAX, R. G. (2004). A beginner's guide to structural equation modeling (2nd Edt.). New Jersey: Lawrence Erlbaum Associates, Publishers
- STEENKAMP, E. M & BAUMGARTNER, H. (1998). Assessing measurement invariance in cross-national consumer research, *The Journal of Consumer Research*, 25(1), 78-90
- STEIN, A., LEE, J. W. & JONES, P. S. (2006) "Assessing cross-cultural differences through use of multiple-group invariance analyses, *Journal of Personality Assessment*" 87(3), 249 – 258.
- STEVENS, P. J. (2009). *Applied multivariate statistics for the social sciences* (5th Edt.) New York: Routledge Taylor and Francis Group.
- TUCKER, K. L., ÖZER, D. J, LYUBOMİRSK, S. & 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". Yüksek Lisans Tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- WU, D. A., LI, Z. & 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" *Practical Assessment, Research & Evaluation*, 12(3), 1-26.
- VANDENBERG, R. J. & LANCE, C.E. (1998). A summary of the issues underlying measurement equivalence and their implications for interpreting group differences. *Research Methods Forum*.
- VANDENBERG, R.J. & LANCE, C. E (2000). A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1) 4-70.