

Çoklu-Grup Örtük Sınıf Analizi ve Ölçme Eşdeğerliği

Duygu Güngör
İzmir Üniversitesi

Mediha Korkmaz
Ege Üniversitesi

Oya Somer
Gediz Üniversitesi

Özet

Sosyal bilimler alanında çoğu görgül araştırma grupların karşılaştırılmasına odaklanmaktadır. Örtük değişkenler üzerinden yapılan karşılaştırmaların geçerli olması için ölçme eşdeğerliğinin sağlanmış olması gerekmektedir. Gözlenen ve örtük değişkenlerin süresiz yapıda olduğu durumlarda; grup karşılaştırmaları ve ölçme eşdeğerliği çalışmalarında örtük sınıf modelleri kullanılmaktadır. Örtük sınıf analizi (ÖSA) ile yapılan çoklu-grup çalışmalarında grup karşılaştırması öncesi ölçme eşdeğerliğinin varlığını incelemek amacıyla homojen, kısmi homojen ve heterojen olarak adlandırılan 3 temel model test edilmektedir. Homojen model, ölçme eşdeğerliğinin varlığına işaret etmektedir. Karşılaştırmalı olarak yapılan analizlerde en iyi model seçiminde bilgi kriterlerinden yararlanılmaktadır. Eylem Değerleri Ölçeği'nin, sevgi kapasitesi boyutu kullanılarak yapılan bu çalışmanın katılımcılarını 496 kadın, 237 erkek toplam 733 üniversite öğrencisi oluşturmaktadır. Tüm grup, kadın ve erkekler için 2 örtük sınıf tespit edilmiştir. En iyi model olarak homojen modelin kabul edildiği araştırmada koşullu olasılıklar sunulmuş ve yorumlanmıştır. Ayrıca kadın ve erkeklerin örtük sınıf olasılıklarının anlamlı şekilde farklılaştığı tespit edilmiştir. Bu durum ÖSA alanyazınında tamamen homojen modelin karşılanamadığı anlamına gelmektedir.

Anahtar kelimeler: Çoklu-grup örtük sınıf analizi; ölçme eşdeğerliği; süresiz örtük değişken

Abstract

Social and behavioral research relies heavily on group comparison. Empirical demonstration of measurement equivalence is necessary when the attributes are compared. Latent class models are used when attributes are conceptualized as having discrete distributions and when discrete observed variables are used to estimate those distributions. This study illustrates the measurement equivalence in multi-group latent class analysis. Heterogeneous model imposes the same number of latent classes on two or more groups without imposing between-group parameter constraints. Homogeneous model is obtained by constraining the between-group conditional probabilities of the heterogeneous model equal. Measurement equivalence is established by demonstrating that the homogenous model fits the data as good as, if not better than, the heterogeneous model. Otherwise, the analysis moves to the exploratory mode to identify the offending equality constraints. The resultant model, if any, is called partial homogenous model. Using the Love Capacity dimension of Values in Action Inventory, the measurement invariance for latent class analysis is illustrated with a dataset involving 496 female and 237 male college students. A two-class model is chosen to represent both gender groups. Homogenous model was found to be the best fitting model. There was evidence for the equality of conditional probabilities between the groups, however, the equality constraint on unconditional probabilities was not supported, which led to the conclusion that the latent classes have the same meaning in both of the gender groups, i.e., measurement equivalence, but that the prevalence rates are gender specific.

Key words: Multi-group latent class analysis; measurement equivalence, discrete latent variable

Yazışma Adresi: Yrd. Doç. Dr. Duygu Güngör, İzmir Üniversitesi Fen-Edebiyat Fakültesi Psikoloji Bölümü, Gürsel Aksel Bulvarı No:14, İzmir

E-posta: duygu.gungor@izmir.edu.tr

Yazar Notu: Bu çalışma, Ege Üniversitesi tarafından, 2011 / EDB / 006 numaralı Bilimsel Araştırma Projesi kapsamında desteklenmiştir. Çalışmanın bir kısmı 3. Eğitimde ve Psikolojide Ölçme Değerlendirme kongresinde sözlü bildiri olarak sunulmuştur.

Psikolojik ölçümlerin yanlılık incelemelerini ele alan istatistiksel yöntemleri anlamak için öncelikli olarak örtük değişken modelleri hakkında bazı önemli bilgilere ihtiyaç vardır. Bir örtük değişken modeliyle, test / ölçek puanı gibi gözlenen değişkenler ile bir ya da daha fazla sayıdaki örtük değişken arasındaki olası ilişkiler ifade edilmektedir (Millsap, 2011). Literatürde tanımlanmış dört farklı örtük değişken modeli söz konusudur ve gözlenen ve örtük değişkenlerin sürekli ya da süreksiz/kesikli olması biçimine göre adlandırılmaktadırlar (Bartholomew, Knott ve Moustaki, 2011). Bu modeller şunlardır: (i) Faktör analizi modelleri: Bu modelde, hem örtük hem de gözlenen değişken sürekli bir formdadır ve normal dağılım sayıltısı geçerlidir. (ii) Örtük özellik (madde tepki kuramı) modeli: Bu modelde, örtük özellik sürekli bir formda ve normal dağılım sayıltısı geçerli iken, gözlenen değişken ise süreksiz bir yapıda olup multinomial bir dağılım vardır. (iii) Örtük profil modeli: Bu modelde, örtük değişken süreksiz, buna karşın gözlenen değişken sürekli biçimdedir. (iv) Örtük sınıf modeli: Hem örtük hem de gözlenen değişkenler süreksiz biçimdedir. Bu modeller içerisinde araştırmacılar tarafından en çok bilinen ve kullanılan model ise doğrusal ilişkileri temel alan faktör analitik yöntemlerdir.

Sosyal bilimler alanında olduğu gibi psikolojide çoğu görgül araştırmada cinsiyet, etnik köken, yaş gibi deneklerin herhangi bir alt grup üyeliğini içeren değişkenler kullanılarak oluşturulan grupların karşılaştırılmasına odaklanılmaktadır. Bu karşılaştırmalar genellikle ya gözlenen puan ortalamaları ya da ortalama ve kovaryans yapıları göz önünde bulundurularak elde edilmiş örtük değişkenler üzerinden yapılmaktadır. Psikometrik anlamda geçerli grup karşılaştırmalarının varlığından söz edebilmek için madde ve test/ölçek düzeyinde gruplar arasında eşdeğerliğin diğer bir ifadeyle ölçme eşdeğerliğinin (measurement equivalence / invariance) sağlanmış olması gerekir. Genel anlamda ölçme eşdeğerliği, bir bireyin hangi karşılaştırma grubunda yer aldığından bağımsız olarak belli bir gözlenen puana sahip olması şeklinde tanımlanabilmektedir (Horn ve McArdle, 1992; Millsap, 2011). Örneğin, aynı hoşgörü düzeyine sahip kadın ve erkeklerin aynı gerçek puanlara sahip olacağı düşünülebilir. Eğer cinsiyet farkı aynı seviyede hoşgörü olan kişilerde farklı gerçek puanlar gözlenmesine neden oluyorsa, ölçme eşdeğerliğinin varlığından söz edilemez (Güngör-Culha, 2012).

Psikometrik teoriler içerisinde ölçme eşdeğerliği incelemelerinin çoğunluğu Madde Tepki Kuramı - MTK (Item Response Theory-IRT), Yapısal Eşitlik Modelleri (Structural Equating Modeling-SEM) ya da Klasik Test Kuramı (Classical Test Theory-CTT) çerçevesinde yapılmaktadır. Bu yöntemlerin ortak özelliği ise, gözlenen ve/veya örtük değişkenlerin genellikle sürekli bir yapıda olduğu, normal dağılım, homojenlik gibi sa-

yıltıların karşılandığı koşullarda kullanılabilmesidir. Oysa ki, psikolojik özelliklerin ölçümünde kullanılan ölçekleme tipine bağlı olarak araştırmaya konu edilen örtük ve gözlenen değişkenler çoğu zaman sürekli bir yapı göstermeyebilirler. Bu noktada hem gözlenen hem de örtük değişken ya da değişkenlerin süreksiz yapıda olduğu durumlar için kullanılan örtük sınıf modelleri (latent class models) iyi bir alternatif oluşturmaktadırlar. Örtük sınıf modelleri hem klasik test kuramından hem de doğrusal faktör analitik modellerden örtük değişken metriğinin kategorik ve kesikli-süreksiz (discrete) bir değişken olması nedeniyle özellikle farklılaşmaktadır. En genel tanımıyla örtük sınıf analizi (ÖSA); kategorik verilerle, gözlenemeyen kategorik grupları tanımlayan istatistiksel bir yöntemdir (Muthen ve Muthen, 1998-2007; Samuelson ve Dayton, 2010). Örtük sınıf modelleri kapsamında ise; örtük sınıf küme analizi, örtük sınıf açımlayıcı ve doğrulayıcı faktör analizleri, regresyon modelleri ve çoklu-grup örtük sınıf analizleri yapılabilmektedir. Bu modellerden çoklu-grup örtük sınıf analizleri ölçme eşdeğerliği ve grup karşılaştırmaları çalışmalarında kullanılmaktadır. Alanyazındaki bu konuyla ilgili örneklerin sayısı gün geçtikçe artmaktadır. Örnek olarak; Eid, Langeheine ve Diener'in (2003) üniversite öğrencilerinin yaşam doyumlarını Çin ve Amerika Birleşik Devletleri arasında; Moors ve Wennekers'in (2003) homoseksüellik, kürtaj, ötenazi gibi kavramlarla ilgili düşüncelerin 12 farklı ülke arasında, Kankaras ve Moors'un (2009) dayanışma ile ilgili düşüncelerin 33 farklı ülke arasında ölçme eşdeğerliğini incelediği çalışmalar verilebilir.

Geleneksel ve Çoklu-grup Örtük Sınıf Analizleri

Geleneksel örtük sınıf analizinin Goodman (2002) tarafından önerilen gösterimi X örtük değişken, A, B ve C gözlenen değişkenler olmak üzere $i = 1, 2, \dots, I$; $j = 1, 2, \dots, J$; $k = 1, 2, \dots, K$; $t = 1, 2, \dots, T$ için eşitlik 1'deki gibidir.

$$\pi_{ijkt}^{ABCX} = \pi_i^X \pi_{it}^{A/X} \pi_{jt}^{B/X} \pi_{kt}^{C/X} \quad (1)$$

Önerilen model, gözlenen değişkenlerden elde edilen veri setine uyum sağlıyorsa, en az bir örtük değişken vardır ve bu model yardımıyla T sayıda sınıf tanımlanabilir. ÖSA'da örtük sınıf olasılıkları (latent class probabilities/prevelances) ve koşullu olasılıklar (conditional probabilities/item response prevelances) olmak üzere iki temel parametre tahminlenmektedir (Collins ve Lanza, 2010; McCutcheon, 1987; Magidson ve Vermunt, 2004a). 1 numaralı eşitlikte π_i^X bir gözlemin/bireyin X 'in t sınıfında olma olasılığını, başka bir deyişle örtük sınıf olasılıklarını; $\pi_{jt}^{A/X}$ ise X 'in t sınıfında olan bir gözlemin/bireyin A 'nın i . kategorisinde olma koşul-

lu olasılığını temsil etmektedir. Örtük sınıf olasılıkları, sınıf sayısını ve bu sınıfların yaygınlığını/büyükliğini ifade ederken, her örtük X değişkeni için T sayıda sınıf ve sınıf sayısı kadar özellik ya da ölçek üzerinde tanımlanabilir sınıflar belirlenebilir. Örtük değişkenin en az iki sınıfı vardır, eğer tek sınıfla tanımlanabilir bir model elde edilirse gözlenen değişkenler istatistiksel olarak birbirinden bağımsızdır ve herhangi bir örtük değişken tanımlanamaz. Sınıfların büyüklüğü de bize evrene ait alt gruplar hakkında bilgi verir, iki sınıflı bir modelde bir sınıfın büyüklüğünü bilmek yeterlidir çünkü tüm örtük sınıf olasılıklarının toplamı 1'e eşittir.

Örtük sınıf analizinin bir diğer parametresi olan koşullu olasılıklar faktör analizindeki faktör yüklerine benzemektedir. Bu parametreler X örtük değişkeninin t sınıfında yer alan bir bireyin/gözlemin, gözlenen değişkenin belli bir seviyesinde olma koşullu olasılığını gösterir. Örtük sınıf olasılıkları gibi, koşullu olasılıkların da toplamı 1'e eşittir (McCutcheon,1987).

Çoklu-grup örtük sınıf analizi iki ya da daha fazla grubun karşılaştırılması için Clogg ve Goodman (1984; 1985) tarafından geliştirilmiştir (akt. Kankaras, Moors ve Vermunt, 2011). Geleneksel örtük sınıf analizinin gözlenen ve örtük değişkenlerine ek olarak yine gözlenen bir değişken olan grup değişkeninin de eklenmesiyle çoklu-grup örtük sınıf modeli eşitlik 2'deki gibi tanımlanmaktadır (Eid ve ark.,2003; Kankaras ve ark., 2011; McCutcheon,1987; 2002).

A kategori sayısı 1'den i'ye kadar, B kategori sayısı 1'den j'ye kadar, C de kategori sayısı 1'den k'ye kadar olan gözlenen değişkenler; X sınıf sayısı 1'den t'ye kadar olan örtük değişken, G grupları 1'den S'ye kadar olan grup değişkeni olsun.

$$\pi_{ijks}^{ABCX|G} = \pi_{ts}^{X|G} \pi_{its}^{A|XG} \pi_{jts}^{B|XG} \pi_{kts}^{C|XG} \quad (2)$$

$\pi_{ijks}^{ABCX|G}$ = G'nin s. grubuna ait olan bir gözlemin/bireyin, X'in t. sınıfında, A'nın i., B'nin j., C'nin k. kategorisinde olma birleşik olasılığı;

$\pi_{ts}^{X|G}$ = G'nin s. grubuna ait olan bir gözlemin/bireyin, X'in t. sınıfında olma koşullu olasılığı;

$\pi_{its}^{A|XG}$ = G'nin s. grubuna ve X'in t. sınıfına ait olan bir gözlemin/bireyin, A'nın i. kategorisinde olma koşullu olasılığı;

$\pi_{jts}^{B|XG}$ = G'nin s. grubuna ve X'in t. sınıfına ait olan bir gözlemin/bireyin, B'nin j. kategorisinde olma koşullu olasılığı;

$\pi_{kts}^{C|XG}$ = G'nin s. grubuna ve X'in t. sınıfına ait olan bir gözlemin/bireyin, C'nin k. kategorisinde olma koşullu olasılığını göstermektedir.

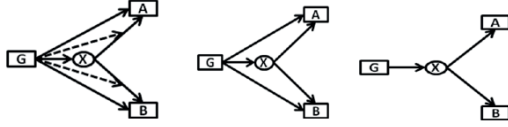
Bu eşitlikte koşullu olasılıklar ve örtük sınıf olasılıkları olmak üzere iki temel parametre olasılık cinsinden tanımlanmıştır ve her biri 0 ile 1 arasında değerler alırlar. Söz konusu olan çoklu-grup örtük sınıf analizleri olduğunda her bir parametrenin kendi tanım kümesi üzerinden ve böylece her alt grup için toplamları da 1'e eşit olacaktır. Örneğin iki örtük sınıfın tespit edildiği bir modelde bir grup için birinci örtük sınıf olasılığının .20 olduğu durumda, bu örtük sınıfta toplam grubun %20'sinin yer aldığı söylenebilmektedir. Kalan %80'lik bölüm ise ikinci örtük sınıfta yer alacaktır.

Çoklu-grup örtük sınıf analizleri lojistik ve log-lineer parametrelerle de çalışılabilmektedir. Lojistik terimler kullanıldığında $\alpha_{ts}^{A|G}$ gruba özgü sabit değerini (group-specific intercept), $\beta_{its}^{AX|G}$ gruba özgü eğim değerini (group-specific slope) temsil etmektedir. Log-linear terimler kullanıldığında sabit parametresi yerine grup değişkeninin örtük değişkene olan etkisini tanımlayan "doğrudan etki (direct effect)" ve eğim parametresi yerine grupla örtük değişkenin birlikte etkisini tanımlayan "ortak etki (interaction effect)" terimleri kullanılmaktadır (Kankaras ve ark., 2011).

Çoklu-grup örtük sınıf analizi sonucunda kabul edilen modelde pek çok olası çıktı vardır; gruplar tamamen eşdeğer olabilirler, grup farkları sadece örtük sınıf olasılığı ya da sadece koşullu olasılıklarda olabilir. Bu durumlardan örneğin, koşullu olasılıkların eşdeğer olduğu modellerde, örtük sınıf olasılıkları karşılaştırılabilir. Böylesi bir durumda gruplar arası fark niteliksel değil nicelikseldir (Collins ve Lanza, 2010). Millsap (2011) örtük sınıf modelleri çerçevesinde ölçme eşdeğerliğini, farklı gruplarda fakat aynı örtük sınıfta yer alan bireylerin aynı koşullu olasılıklara sahip olması olarak tanımlamaktadırlar.

Olası başka bir durum tüm parametrelerin gruba özgü olmasıdır, bu durumda da ölçme eşdeğerliğinden söz edilemeyeceği için gruplar arası karşılaştırma yapmak mümkün değildir. Collins ve Lanza (2010) çoklu grup örtük sınıf analizi yapan bir araştırmacının, karşılaştırma gruplarında eşit sayıda örtük sınıf tanımlanıp tanımlanamayacağı, koşullu olasılıkların gruplar arasında eşdeğer varsayıлып sayılamayacağı, örtük sınıf yapılarının gruplar arasında aynı olup olmadığı, eğer farklıysa bu farkların nasıl açıklanabileceği gibi pek çok soruyu cevaplama gerektiğini vurgulamaktadırlar. McCutcheon'a (1987) göre olası pek çok sonuç homojen, kısmi homojen, heterojen model olarak adlandırılan üç modelden birine uyum sağlayacaktır. Bu 3 model şekil 1'deki gibi gösterilebilir. G grup, X örtük değişken, A

ve B gözlenen değişkenler olmak üzere;



Şekil 1a. Heterojen Model

Şekil 1b. Kısmi Homojen Model

Şekil 1c. Homojen Model

Şekil 1a'da görülen heterojen model eşitlik 1'i temsil etmektedir. Bu modelde herhangi bir kısıt/sınırlandırma yoktur, eğim ve sabit parametreleri gruba özgü tanımlanmaktadır. Model seçim süreçlerinde temel model olarak sıklıkla -gruplar için ayrı ayrı sınıf sayısı belirlendikten sonra- heterojen model kullanılmaktadır (örn., Eid ve ark., 2003; Morren ve ark., basımda). Daha sonra Şekil 1b'de sunulan kısmi homojen modelde, eğim parametrelerinin gruplar için eşitlendiği durum ($\beta_{it}^{AX|G} = \beta_{it}^{AX}$) test edilmektedir. Son aşamada ise eğim parametrelerine ek olarak sabit parametreleri de ($\alpha_{is}^{AG} = \alpha_{is}^{AG} = \alpha_{ks}^{CG}$) eşitlenerek homojen model test edilmektedir. Araştırmaların bir kısmında da temel model olarak homojen model alınmakta, parametrelere konan kısıtlar serbest bırakılarak diğer modeller üretilmektedir (örn., Kankaras ve Moors, 2009). Ölçek/test maddelerinin ya da gözlenen değişkenlerin tamamına aynı anda kısıt koyma ya da serbest bırakma işlemi yapıldığında analizler ölçek düzeyinde değerlendirilmektedir. Yanı sıra maddelerin sadece birine ya da bir kısmına kısıt konarak madde düzeyinde de analizler yapmak mümkün olmaktadır.

Pratikte heterojen ve homojen modeller test edilerek homojen model uyum sağlarsa kısmi homojen modelleri madde düzeyinde test etmeye gerek kalmadan grup karşılaştırmaları yapılabilmektedir. Bununla birlikte özellikle yapı geçerliği çalışmalarında kullanıldığı durumda madde düzeyinde de analizlerin yapılması araştırmacılara daha derin bilgiler sağlamaktadır. Bazı simülasyon çalışmaları 5 ya da 10 maddeden sadece biri yanlı olduğunda kısmi homojen yerine homojen modelin iyi model olarak seçilebileceğini ve yanlı olan maddenin bu durumda tespit edilemeyebileceğini göstermiştir (Güngör-Culha, 2012).

Örtük sınıf modellerinde parametre tahmini için tekrarlı (iterative) yöntemlere ihtiyaç duyulmakta, genellikle beklenti-maksimizasyon (expectation-maximization) algoritması ya da Newton-Raphson algoritması veya bu iki algoritmanın bir kombinasyonu kullanılmaktadır (Collins ve Lanza, 2010; Vermunt ve Magidson, 2005a).

Model Seçimi

Örtük sınıf analizi model seçim sürecinde istatistiksel ölçütler olan sadelik (parsimony) ve yorumlanabilirlik göz önüne alınmalıdır (Collins ve Lanza, 2010; Silvia, Kaufman ve Pretz, 2009). Sade modeller, veriye uyum sağlayan yani en az sayıda örtük sınıfa sahip ve daha az parametre tahminlenen modeller tanımlanmaktadır. Bu koşulları sağlayan modeli bulmak için; sınıflandırma kalitesi ile ilgili bilgi veren entropi (entropy), temelinde gözlenen değerler ile beklenen değerlerin karşılaştırılması yatan Pearson Ki-Kare (χ^2), çok olasılık ki-kare oranı (likelihood ratio chi square)(L^2) gibi istatistikler ve Akaike bilgi kriteri (Akaike Information Criteria) (AIC), Bayesian bilgi kriteri (Bayesian Information Criteria) (BIC) gibi bilgi kriterleri kullanılabilir (McCutcheon, 2002; Silvia ve ark., 2009). Örtük sınıf modellerinde sınıf sayısı belirlenirken karşılaştırılmalı olarak seçim yapılmaktadır, amaç doğru modeli bulmaktan çok, daha fazla bilgi sağlayan modeli tespit etmektir (Moors ve Wennekers, 2003).

Model seçim sürecinde genel olarak analizlere gözlenen değişkenlerin bağımsızlığını ifade eden sınıf sayısının 1'e eşit olduğu modelle başlanır (Magidson ve Vermunt, 2004a). Eğer bu model kabul edilirse örtük sınıf analizine ihtiyaç kalmayacaktır. Ancak bu temel model verilere uyum sağlamazsa yani en az 2 örtük sınıf tanımlanabileceksa analizler her seferinde sınıf sayısı bir artırılarak serbestlik derecesinin izin verdiği sayıya kadar devam ettirilebilir.

Model seçimi için, AIC, BIC bilgi kriterlerinin kullanılması genel olarak kabul görmüştür. Bununla birlikte CAIC, AIC3, DBIC gibi türevleri de var olan kriterlerden hangisinin kullanılmasının uygun olduğuna dair bir görüş birliği bulunmamakta, alan yazında bu konu ile ilgili pek çok yayın bulunmaktadır (Bauer ve Curran, 2003; Dias, 2006; Lin, 2006; 2012; Nylund, Asparouhov ve Muthen, 2007; Vrieze, 2012; Yang ve Yang, 2007). Farklı simülasyon koşullarında Dias (2006), Güngör-Culha (2012) AIC3, Nylund ve arkadaşları (2007) BIC'nin daha iyi sonuçlar verdiğini bildirmişlerdir. Kankaras, Vermunt ve Moors (2011) ölçme eşdeğerliğinin farklı analiz yöntemleriyle (çoklu grup doğrulayıcı faktör analizi, madde cevap kuramı ve örtük sınıf faktör analizi) test edilmesini karşılaştırdıkları simülasyon çalışmasında, madde düzeyinde örneklem sayısı küçük olduğunda AIC'nin, örneklem sayısı arttığında ise BIC'nin AIC'den daha iyi sonuçlar verdiğini; ölçek düzeyinde ise AIC'nin daha iyi sonuçlar verdiğini göstermişlerdir. Morren, Gelissen ve Vermunt (2011; basımda) örtük sınıf faktör analizi kullanarak yaptıkları ölçme eşdeğerliği çalışmalarında BIC değerini kullanmayı tercih etmişlerdir. Bilgi kriterleriyle ilgili temel kural, düşük bilgi kriteri değerinin daha iyi modeli işaret ettiğidir (Moors ve Wennekers, 2003).

Bu çalışmanın amacı uygulamada araştırmacılara hem ölçme eşdeğerliği çalışmaları yaparken hem de grup karşılaştırmalarında, özellikle değişkenlerin süresiz yapıda olduğu ya da geleneksel yöntemlere ait sayıtların karşılanmadığı durumlar için iyi bir alternatif olan çoklu-grup örtük sınıf analizi ile bir uygulama örneği sunmaktır. Örtük sınıf analizi, tanımlanan bir psikolojik özellik etrafında sınıflayıcı ya da sıralayıcı ölçekleme düzeyinde oluşturulan test maddesi ya da gözlemlerin geleneksel faktör analizine benzer şekilde, o yapının varlığının kanıtı olarak kullanılmaktadır. Diğer bir ifadeyle bir yapı geçerliğinin test edilebilirliğini gösteren önemli bir psikometrik yöntemidir. (Bu noktada faktör analizi ile örtük sınıf analizi arasındaki ölçekleme seviyeleri dışındaki en büyük farkın örtük sınıf analizinin kişi-merkezli (person oriented) olması olduğunu belirtmekte yarar vardır (bkz. Collins ve Lanza, 2010.)

Yöntem

Katılımcılar

Araştırmanın katılımcılarını Ege Üniversitesinin 14 farklı fakültesinin yaklaşık olarak 50 kadar farklı bölümlerinde okuyan her sınıf düzeyinden 496 kadın, 237 erkek olmak üzere toplam 733 üniversite öğrencisi oluşturmaktadır. Katılımcıların %99'unun yaş aralığı 18-25 yaş olup, ortalaması ise 21.4 ve standart sapması da 2.2'dir.

Veri Toplama Araçları

Bu araştırmanın verileri, Peterson ve Seligman tarafından geliştirilen, pozitif psikoloji kapsamındaki 6 temel faktör ve 24 alt boyut ile temsil edilen, Eylem Değerleri Ölçeği (Values in Action - VIA, Park, Peterson ve Seligman, 2004)'ni kapsayan özelliklerin Türkçe Formu'nun üniversite öğrenci örnekleminde geçerlik ve güvenilirlik araştırmasının yürütüldüğü bir çalışmadan alınmıştır (Korkmaz, 2011). Eylem Değerleri olarak tanımlanan bu modelin alt boyutlarını temsil eden maddeler ise; Uluslar arası Kişilik Madde havuzundan (International Personality Item Pool-IPIP, Goldberg ve ark., 2006) alınmıştır. Kişilik modelleri çerçevesinde bir madde havuzu olan IPIP günümüzde yaklaşık olarak 2413 madde civarında büyük bir sayıya ulaşmıştır.

Bu çalışmada IPIP madde havuzundaki maddelerden Peterson ve Seligman'ın "Eylem Değerleri Ölçeği" ile örtüşen ve 24 alt boyut ile temsil edilen modelden "sevgi kapasitesi (capacity for love, intimacy) (<http://iPIP.ori.org/newVIAKey.htm>) alt boyutu seçilmiştir. Sevgi kapasitesi boyutu; insanlarla yakın ilişkileri önemseme, değer verme, paylaşma, insanlara yakın-sıcak davranma gibi özellikleri kapsamaktadır (Park ve ark., 2004). Bu alt boyutta toplam 10 madde bulunmaktadır.

Örtük sınıf analizleriyle 5'li Likert tipindeki maddelerin de analizi mümkün olmasına rağmen, bu çalışmada 5'li Likert tipindeki (1= tamamen uygun, 2 = biraz uygun, 3 = ne uygun, ne uygun değil, 4 = pek uygun değil, 5 = hiç uygun değil) yanıtlar ,1-2 ve 4-5 seçenekleri birleştirilerek 3 kategorili (uygun, ne uygun ne uygun değil ve uygun değil) hale dönüştürülmüştür. Buradaki temel amaç bir örnek uygulama niteliğinde olan bu çalışmada analiz sonuçlarının yorumlanmasını ve anlaşılabilirliğini kolaylaştırmaktır.

İşlem Yolu ve Analizler

Araştırmanın analizleri Vermunt ve Magidson (2005b) tarafından örtük sınıf modellerini test etmek üzere geliştirilen Latent GOLD 4.5 istatistik programı kullanılarak yapılmıştır.

Analizlere tüm grup, kadın ve erkekler için ayrı ayrı sınıf sayısı tespit edilerek başlanmıştır. Hem kadın hem erkeklerde 2 sınıflı modelin en iyi model olarak tespit edilmesinden sonra öncelikle tüm parametrelerin serbest olduğu, yani sabit ve eğitim parametrelerinin grubu özgü olduğu heterojen model birinci sıradaki temel model olarak oluşturulmuştur. İkinci modelde; eğitim parametreleri gruplar arasında eşitlenerek kısmi homojen modeller, üçüncü modelde; eğitim parametrelerinin yanısıra sabit parametrelerinin de eşitlendiği homojen model test edilmiştir. Örtük sınıf olasılıklarının karşılaştırılması amacıyla ise tamamen homojen model de analizlere eklenmiştir.

Giriş kısmında da değinildiği gibi pratikte homojen model kabul edilirse, kısmi homojen modelleri madde düzeyinde incelemeye gerek olmaksızın araştırmaya devam edilebilmektedir. Ancak bir örnek uygulama olması bakımından bu çalışmada madde düzeyindeki analizlerde uygulanıp sunulmuştur.

Bulgular

Tüm Grup, Kadın ve Erkekler İçin Örtük Sınıf Sayılarının Belirlenmesi

Çoklu grup karşılaştırmalarına başlamadan önce sınıf sayılarının gruplar için eşit olup olmadığı her biri ayrı ayrı analiz edilerek incelenmiştir. 1 sınıflı yerel bağımsızlık modeli ile başlanan analizler hiyerarşik olarak 6 sınıfa kadar çıkartılmıştır. Tablo 1'de de görüldüğü üzere 2 örtük sınıfın var olduğu model tüm grup, kadın ve erkekler için en düşük BIC ve CAIC değeri ile en iyi model olarak tespit edilmiştir. AIC ve AIC3 değerleri daha fazla sınıf sayısına işaret etmekle birlikte sadelik ve yorumlanabilirlik göz önünde bulundurulduğunda 2 sınıflı modelin araştırma için uygun olduğuna karar verilmiştir. Bu sonuç sevme kapasitesi örtük değişkeninin 2 seviyesinin olduğu şeklinde yorumlanabilir.

Tablo 1. 1-6 Sınıflı Modellere Bilgi Kriterleri Değerleri

	Tahminlenen Parametre Sayısı	L ²	BIC	AIC	AIC3	CAIC	sd
Tüm Grup							
1. Sınıf	20	2648.19	-2055.58	1222.19	509.19	-2768.58	713
2. Sınıf	31	2340.43	-2290.77	936.43	234.43	-2992.77	702
3. Sınıf	42	2272.32	-2286.31	890.32	199.32	-2977.31	691
4. Sınıf	53	2223.70	-2262.36	863.70	183.70	-2942.36	680
5. Sınıf	64	2189.70	-2223.80	851.70	182.70	-2892.80	669
6. Sınıf	75	2150.08	-2190.84	834.08	176.08	-2848.84	658
Kadın							
1. Sınıf	20	1871.75	-1082.58	919.75	443.75	-1558.58	476
2. Sınıf	31	1646.44	-1239.62	716.44	251.44	-1704.62	465
3. Sınıf	42	1594.20	-1223.58	686.20	232.20	-1677.58	454
4. Sınıf	53	1576.36	-1173.15	690.36	247.36	-1616.15	443
5. Sınıf	64	1538.10	-1143.14	674.10	242.10	-1575.14	432
6. Sınıf	75	1516.31	-1096.66	674.31	253.31	-1517.66	421
Erkek							
1. Sınıf	20	1273.98	87.41	839.98	622.98	-129.59	217
2. Sınıf	31	1184.45	58.03	772.45	566.45	-147.97	206
3. Sınıf	42	1133.67	67.40	743.67	548.67	-127.60	195
4. Sınıf	53	1085.52	79.40	717.52	533.52	-104.60	184
5. Sınıf	64	1054.04	108.07	708.04	535.04	-64.93	173
6. Sınıf	75	1020.67	134.84	696.67	534.67	-27.16	162

Ölçme Eşdeğerliğinin Ölçek ve Madde Düzeyinde Test Edilmesi

Sonraki adım olarak 2 örtük sınıflı modelde ölçek düzeyinde (i) örtük sınıf olasılıkları ve koşullu olasılıklarının tamamının gruba özgü olduğu, başka bir deyişle eğitim ve sabit parametrelerinin serbest bırakıldığı heterojen model; (ii) eğitim parametrelerinin her iki grup için eşit olduğu kısmi homojen model ve (iii) eğitim parametresine ek olarak sabit parametrelerinin gruplar için eşit olduğu homojen model oluşturulmuştur. Bu analizlere ek olarak maddelere ait eğitim değerlerinin eşitlendiği ve her seferinde yeni bir maddenin sabit değerlerinin de eşitlendiği madde düzeyindeki analizler yapılmış ve kısmi homojen modelle karşılaştırılmıştır. Bu durumda örneğin Tablo 2’de sunulan “+Madde 4” değeri ilk dört maddenin hem eğitim hem sabit değerlerinin, kalan altı maddenin ise sadece eğitim değerlerinin iki grup için eşitlendiği koşulu temsil etmektedir. Söz konusu analizler

“+Madde10” ulaştığında ise artık homojen modelle eşdeğer koşul oluşmuştur.

Tablo 2’de de görüleceği üzere ölçek düzeyinde homojen model BIC, AIC3 ve CAIC bilgi kriterlerince en iyi model olarak tespit edilmiştir. Daha derin bir inceleme adına yapılan madde düzeyindeki analizlerde BIC ve CAIC değerleri kullanıldığında modele eklenen sabit değeri eşitlenmiş her yeni maddenin bilgi kriteri değerlerinde düşüşe sebep olduğu gözlemlenmiştir. Ayrıca kadın ve erkek karşılaştırma gruplarının madde düzeyindeki (10 madde üzerinden) analizlerin hiçbirinde kısmi homojen modele ait değerden daha yüksek bir değere rastlanmadığı görülmektedir.

Hem eğitim, hem sabit parametrelerinin kadın ve erkek karşılaştırma grupları için eşdeğer olduğu homojen modelin kabul edilmesiyle, cevaplanması gereken son bir soru bulunmaktadır. Gruplar için yapının aynı olduğu bu koşulda kadın ve erkeklerin örtük sınıfları dağılı-

Tablo 2. İki Sınıflı Çoklu-Grup Örtük Sınıf Analizi Model Sonuçları

	Tahminlenen Parametre Sayısı	L ²	BIC	AIC	AIC3	CAIC	sd
Heterojen Model	62	2830.78	-1595.91	1488.78	817.78	-2266.91	671
Kısmi Homojen Model	52	2842.93	-1649.73	1480.93	799.93	-2330.73	681
+ Madde 1	50	2853.20	-1652.65	1487.20	804.20	-2335.65	683
+ Madde 2	48	2853.51	-1665.53	1483.51	798.51	-2350.53	685
+ Madde 3	46	2854.54	-1677.70	1480.54	793.54	-2364.70	687
+ Madde 4	44	2858.14	-1687.29	1480.14	791.14	-2376.29	689
+ Madde 5	42	2861.70	-1696.93	1479.70	788.70	-2387.93	691
+ Madde 6	40	2873.27	-1698.55	1487.27	794.27	-2391.55	693
+ Madde 7	38	2873.30	-1711.72	1483.30	788.30	-2406.72	695
+ Madde 8	36	2874.54	-1723.67	1480.54	783.54	-2420.67	697
+ Madde 9	34	2882.97	-1728.43	1484.97	785.97	-2427.43	699
+ Madde 10	32	2885.47	-1739.13	1483.47	782.47	-2440.13	701
Homojen Model	32	2885.47	-1739.13	1483.47	782.47	-2440.13	701
Tamamen Homojen Model	31	2894.62	-1736.58	1490.62	788.62	-2438.58	702

mının eşdeğer olup olmadığını tespit etmek için örtük sınıf olasılıklarının da gruplar için eşitlendiği tamamen homojen model ile karşılaştırılmıştır. Tablo 2'den de izlendiği üzere, modelde bir iyileşme gözlenmediği için gruplar arası örtük sınıf olasılıklarının anlamlı şekilde farklılaştığı söylenebilir.

Kabul edilen modele ait parametre tahminleri Tablo 3'te sunulmuştur. Koşullu olasılıklar faktör yüklerine benzemekte ve yorumlanabilmektedir. Genel olarak değerlendirdiğinde soru içerikleri sevme/sevilme kapasitesini temsil etmekte ve 1. örtük sınıfta yer alan bireyler sevme kapasiteleri yüksek, 2. örtük sınıfta yer alan bireyler ise sevme kapasiteleri düşük özellikler sergilemektedirler. Örnek olarak maddelerden bir kısmı şu şekilde yorumlanabilir;

Koşullu olasılıklara ayrıntılı olarak bakıldığında, “Bir insan olarak gerçekten önemseydiğim birileri var” maddesine “uygun değil” cevabı verenlerin tamamının 2. sınıfta yer alma olasılığı ile karakterize olmaktadır. “Hayatımda en az kendilerini önemsedikleri kadar beni de önemseyen insanlar olduğumu biliyorum” maddesinde “uygun” yanıtı verenlerin 1. sınıfta yer alma olasılığı .80 olarak bulunmuştur, 2. sınıfta yer alanların bu maddede “uygun değil” deme olasılıkları ise .92'dir. “Diğer insanlarla duygularımı kolaylıkla paylaşmam” maddesi için tüm cevap kategorilerindeki koşullu olasılıklar; 1. sınıf üyeleri için 2. sınıftakilerden daha yüksek olmakla

birlikte, “uygun değil” kategorisi için bu oran .81'lere ulaşmıştır. Bu madde “Bir ilişki kurmak uğruna gönüllü risk almayı isterim” maddesiyle benzer koşullu olasılık örüntüsü sergilemektedir. Madde içerikleri açısından da düşünüldüğünde iki maddenin de kişinin kendini diğerlerine karşı koruma ile ilgili olduğu söylenebilir. “Bir ilişki kurmak uğruna gönüllü risk almayı isterim” maddesine 1. sınıfta yer alan katılımcıların yanıtlardan herhangi birini verme olasılıkları .72 iken, bu oran 2. sınıfta yer alan katılımcılarda .28'e düşmüştür. Bu yapıyla madde iyi çalışmayan ve yapı ile ilgili bilgi vermeyen bir madde olarak görünmektedir.

Örtük sınıf olasılıkları incelendiğinde sevme kapasitesi yüksek olan 1. örtük sınıfta kadınların %76'sının, erkeklerin ise %63'ünün yer aldığı görülmektedir. Sevme kapasitesi düşük olan bireylerin oluşturduğu 2. örtük sınıfta ise kadınların %24'ü erkeklerin ise %37'si yer almaktadır. Örtük sınıf olasılıklarının gruplar arası eşdeğer varsayıldığı modelin daha iyi model olarak kabul edilmemesi bu dağılım farklarının istatistiksel olarak anlamlı olduğunun göstergesi olmaktadır. Özetle, hem erkekler hem de kadınlar için sevme kapasiteleri düşük ve yüksek iki eşdeğer sınıf bulunmakla birlikte, erkeklerin sevme kapasitesi düşük grupta yer alma olasılıkları kadınlara oranla; kadınların da sevme kapasitesi yüksek olan grupta yer alma olasılıkları erkeklere oranla daha yüksek bulunmuştur.

Tablo 3. Homojen Modele Ait Parametre Tahminleri

Bir ilişki kurmak uğruna gönüllü risk almayı isterim.					Bir insan olarak gerçekten önemseyemediğim birileri var.				
Koşullu Olasılıklar	Sınıf	Uygun	Ne uygun ne uygun değil	Uygun Değil	Koşullu Olasılıklar	Sınıf	Uygun	Ne uygun ne uygun değil	Uygun Değil
	1	.72	.72	.72		1	.76	.07	.00
	2	.28	.28	.28		2	.24	.93	1
Hayatımda en az kendilerini önemseydikleri kadar beni de önemseyen insanlar olduğunu biliyorum.					Diğer insanlarla duygularımı kolaylıkla paylaşmam.				
Koşullu Olasılıklar	Sınıf	Uygun	Ne uygun ne uygun değil	Uygun Değil	Koşullu Olasılıklar	Sınıf	Uygun	Ne uygun ne uygun değil	Uygun Değil
	1	.80	.30	.08		1	.65	.76	.81
	2	.20	.70	.92		2	.35	.24	.19
Bazı insanların kusurlarımı kabul ettiğini biliyorum.					Kendimi diğer insanlardan yalıtılmış hissediyorum.				
Koşullu Olasılıklar	Sınıf	Uygun	Ne uygun ne uygun değil	Uygun Değil	Koşullu Olasılıklar	Sınıf	Uygun	Ne uygun ne uygun değil	Uygun Değil
	1	.79	.57	.31		1	.39	.65	.85
	2	.21	.43	.69		2	.61	.35	.15
Bazılarının hayatındaki en önemli insan benim.					Başkalarının sevgisini kabullenmekte güçlük çekerim.				
Koşullu Olasılıklar	Sınıf	Uygun	Ne uygun ne uygun değil	Uygun Değil	Koşullu Olasılıklar	Sınıf	Uygun	Ne uygun ne uygun değil	Uygun Değil
	1	.81	.70	.35		1	.20	.44	.87
	2	.19	.30	.65		2	.80	.56	.13
Başka birine sevgimi belli edebilirim.					Ne yaparlarsa yapsınlar dostlarımı ve ailemi sevmekten vazgeçmem.				
Koşullu Olasılıklar	Sınıf	Uygun	Ne uygun ne uygun değil	Uygun Değil	Koşullu Olasılıklar	Sınıf	Uygun	Ne uygun ne uygun değil	Uygun Değil
	1	.80	.66	.34		1	.76	.62	.54
	2	.20	.34	.66		2	.24	.38	.46
Örtük Sınıf Olasılıkları									
Cinsiyet	Sınıf 1	Sınıf 2							
Kadın	.76	.24							
Erkek	.63	.37							

Tartışma

Sosyal bilimlerdeki araştırmaların pek çoğu grup karşılaştırmalarına odaklanmakta, ancak çok azı ölçme eşdeğerliğini test etmektedir. Ölçme eşdeğerliği çalışmalarında çeşitli teknikler arasında en popüler yaklaşım yapısal eşitlik modelleri kapsamındaki çoklu-grup doğrulayıcı faktör analizi (ÇGDFA) ve daha az olmakla birlikte ortalama ve kovaryans (mean and covariances

structures-MACS) yapıları ile ölçme eşdeğerliği çalışmalarıdır (Somer, Korkmaz, Dural ve Can, 2009). Diğer psikometrik popüler yaklaşım ise madde tepki kuramına (MTK) dayalı madde ve test/ölçek işlevsel farklılıklarını (differential item functions - DIF) inceleyen modellerdir. ÇGDFA ile madde ya da test/ölçek düzeyinde ölçme eşdeğerliği incelenirken gözlenen değişkenler ve örtük yapılar arasındaki ilişkilerin doğrusal olması gerekirken MTK ile bu ilişkilerin doğrusal olmayan modellemeleri

söz konusudur. Bununla birlikte gözlenen ve örtük değişkenlerin her ikisinin de kategorik olduğu yani sınıflama, sıralama ya da dereceli olduğu durumlarda bu iki yaklaşımın kullanılması mümkün değildir. Bu noktada örtük sınıf modelleri kapsamındaki çoklu-grup örtük sınıf analizi karşımıza çıkmaktadır. Çoklu-grup örtük sınıf analizinin ölçme eşdeğerliği çalışmalarında kullanılmaya başlaması oldukça yeni olmakla birlikte, Kankaras, Vermunt ve Moors (2011) bu üç yaklaşımı karşılaştırdıkları çalışmalarında bu yöntemin de oldukça iyi bir alternatif olduğuna değinmişlerdir.

Örtük sınıf modellerinin, yapısal eşitlik ve madde tepki kuramı modellerinden en büyük farkı değişkenlerin yapısından kaynaklanan normal dağılım sayılısının olmamasıdır. Ayrıca, hem örtük hem de gözlenen değişkenlerin kategorik olduğu koşullarda ölçme eşdeğerliği çalışmalarındaki tek alternatif model ÖSA'dır. Ölçme eşdeğerliği çalışmalarında örtük sınıf modelleri kullanmanın diğer modellerden daha avantajlı yönlerini şöyle sıralanabilir (Kankaras ve Moors, 2009): (i) ÖSA, diktomatik, sınıflayıcı, sıralayıcı ve Likert tipi değişkenlerde dahil olmak üzere ÇGDFA'nın aksine daha iyi uyum sağlamakta ve doğrusal ilişki, normal dağılım ve homojenlik gibi geleneksel model sayılıtlarına ihtiyaç duyulmamaktadır. (ii) ÖSA ile modeldeki eşdeğer olmayan madde sayısı çok olsa ve hatta karşılaştırma gruplarının faktör sayıları eşit olmasa bile model karşılaştırmalarını yapmak mümkündür. Buna karşın, ÇGDFA'da karşılaştırma gruplarının faktör yapılarında en az iki madde eşdeğer olmalı ya da en azından kısmi homojenlik aşaması sağlanmış olmalıdır. (iii) ÖSA'nın MTK'den farkı ise son yıllarda çok boyutlu madde tepki modelleri geliştirilmesine rağmen çok boyutlulukla daha kolay baş edebilmesi ve kullanıcı dostu programların varlığıdır. Ölçmeye konu olan bazı özellikler tek bir boyutta açıklanamayabilir. Böylesi durumlarda çok boyutlulukla baş edebilecek yöntemlerin tercih edilmesi gerekmektedir. Ölçme eşdeğerliği çalışmalarında da kimi maddeler bir boyut için yanlılık kaynağı iken diğer boyut için olmayabilirler (Millsap, 2011).

Örtük sınıf modellerinin avantajları yanında bazı sınırlılıkları da mevcuttur. Örneğin örtük değişkenin de süresiz varsayılması ile geleneksel faktör analizindeki faktör yükleri koşullu olasılıklara dönüşerek yorumlanması zorlaşmaktadır (Kankaras ve Moors, 2009). Magidson ve Vermunt (2004b), özellikle örtük sınıf faktör analizinin loglineer parametrelerle çalışıldığında faktör yüklerinin yorumlamanın güçlüğüne dikkat çekmişlerdir. Model seçim sürecinde çeşitli alternatiflerden hangisinin kullanılacağına ilişkin henüz kesin bir görüş birliği bulunmamaktadır. Özellikle sınıf sayısının belirlenmesine ilişkin çeşitli simülasyon çalışmalarından elde edilen farklı bulgular vardır (örn., Dias, 2006; Kankaras ve ark., 2011; Lin, 2012; Nylund ve ark., 2007; Yang ve Yang, 2007).

Sunulan örnekte, hem kadın hem erkek grupları için 2 sınıflı model kabul edilmiştir. İlgili modele ait bilgi kriteri değerleri incelendiğinde BIC ve CAIC değerlerinin bu modeli işaret ettiği görülmektedir. Dias (2006) ve Nylund ve arkadaşları (2007) yaptıkları çalışmalarda AIC'nin olması gerekenden fazla sınıf sayısı tespit etme eğilimini vurgulamaktadır.

Maddelere ait koşullu olasılıklar ve madde içerikleri incelenerek sınıflar sevgi kapasitesi yüksek ve sevgi kapasitesi düşük olarak adlandırılmıştır. Bu işlem faktör analizindeki faktör yüklerini yorumlamaya benzemektedir.

Bu çalışmada kabul edilen homojen model ölçme eşdeğerliğinin kanıtı olarak kullanılmakta ve gruplar arasında örtük sınıf olasılıklarının karşılaştırılmasına olanak sağlamaktadır. Hem ölçek hem madde düzeyindeki analizlere ilişkin sonuçlar incelendiğinde AIC'nin kadın ve erkek grupları için sabit değerlerinin farklı olduğu kısmi homojen modeli işaret ettiği görülmektedir. Diğer üç bilgi kriteri (BIC, CAIC, AIC3) ise homojen modeli işaret etmektedir. Bu bulgu AIC'nin eşdeğerlik varken de yanlılık tespit etme eğilimini belirten çalışmalarla paralellik göstermektedir (Güngör-Culha, 2012; Kankaras ve ark., 2011). Homojen model, yapısal eşitlik modellemelerindeki skalar eşdeğerlik modeline, madde cevap kuramındaki madde güçlük ve ayırt edicilik parametrelerinin fokal ve referans grup için eşdeğer olduğu modellere benzetilebilir. Kısmi homojen modeller kabul edildiğinde ise, bazı maddelerin sabit ya da eğitim parametrelerinin eşdeğer olmadığı bilinmektedir. Madde düzeyinde yapılan incelemeler sonucu, hangi maddelerin eşdeğerliği bozduğu tespit edilerek yorum yapmak mümkün olmaktadır. Bu çalışmada madde düzeyindeki analizler de yapılmış ve homojen modelin kabulüne ek bir kanıt olarak kullanılmıştır. Pratikte de araştırmacıların ölçek düzeyindeki analizlerle yetinmeyip, madde düzeyinde de incelemeler yapmalarının verilerin derinlemesine incelenmesine ve anlaşılmasına katkı sağlayacağı düşünülmektedir. Böylece yeni bir ölçme aracı geliştirme sürecinde hangi test maddelerinin ölçme aracında yer alması gerektiği kararının verilmesine geçerlik kanıtı sağlanabilecektir. Kankaras ve arkadaşları (2011), eşdeğerlik çalışmaları ya da grup karşılaştırmaları yaparken farklı yöntemlerin denenmesini önermektedir.

Hem uluslararası hem de ülkemiz alan yazınında çok az örneği bulunan örtük sınıf modelleri ile ölçme eşdeğerliği çalışmalarının görgül verilerle kullanımına ilişkin daha fazla örneğe ihtiyaç bulunmaktadır. Bunun yanı sıra örtük sınıf modelleri ile herhangi bir psikolojik örtük yapı etrafında kategorik, sıralayıcı ya da sınıflayıcı ölçekleme düzeylerinde oluşturulan ölçme araçları için açılımlayıcı ve doğrulayıcı örtük sınıf faktör analizlerinin yapılmasına imkan verdiği kadar, aynı zamanda örtük sınıf regresyon analizlerine de olanak sağlamaktadır. Bu yöntem, özellikle psikolojik yapılar hakkında eşit aralık-

lı ölçkleme düzeyinde test/ölçek geliştirmenin katı bir kural olarak uygulanamayacağı durumlarda test geliştiricilere ve araştırmacılara alternatif sunmaktadır.

Kaynaklar

- Bauer, D. J. ve Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods*, 8, 338-363.
- Bartholomew, P. M., Knott, M. ve Moustaki, I. (2011). *Latent variable models and factor analysis a unified approach (3. baskı)*. West Sussex: John Wiley & Sons, Ltd.
- Collins, L. M. ve Lanza, S. T. (2010). *Latent class and latent transition analysis with applications in the social, behavioral, and health sciences*. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Dias, J. G. (2006). Latent class analysis and model selection. M. R. Kruse, C. Borgelt, A. Nürberger ve W. Gaul, (Ed.), *From data and information analysis to knowledge engineering* içinde (95-102). Berlin: Springer-Verlag.
- Eid, M., Langeheine, R. ve Diener, E. (2003). Comparing typological structures across cultures by multigroup latent class analysis: A primer. *Journal of Cross-Cultural Psychology*, 34, 195-210.
- Goodman, A. L. (2002). Basic concepts and procedures in single-and multiple-group latent class analysis. J. A. Hage-naars ve A. L. McCutcheon, (Ed.), *Applied latent class analysis* içinde (3-55). New York: Cambridge University Press.
- Goldberg, L. R., Johnson, J. A, Eber, H. W, Hogan, R., Ashton, M. C., Cloninger, C. R. ve Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40, 84-96.
- Güngör-Culha, D. (2012). Örtük sınıf analizlerinde ölçme eşdeğerliğinin incelenmesi. *Yayılanmamış doktora tezi*, Ege Üniversitesi Sosyal Bilimler Enstitüsü, İzmir.
- Horn, J. L. ve McArdle, J. J. (1992) A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, 18, 117-144.
- Kankaras, M. ve Moors, G. (2009). Measurement equivalence in solidarity attitudes in Europe. Insights from a multiple group latent class factor approach. *International Sociology*, 24(4), 557-579.
- Kankaras, M., Moors, G. ve Vermunt, J. K. (2011). Testing for measurement invariance with latent class analysis. E. Davidov, P. Schmidt ve J. B. Billiet, (Ed.), *Cross-cultural analysis: methods and applications* içinde (359-384). New York: Routledge.
- Kankaras, M., Vermunt, J. K. ve Moors, G. (2011). Measurement equivalence of ordinal items: A comparison of factor analytic, item response theory, and latent class approaches. *Sociological Methods & Research*, 40(2), 279-310.
- Korkmaz, M (2011) Kategorik değişkenlerin analizinde güçlü bir yaklaşım: Örtük sınıf analizi. *E. Ü. Bilimsel Araştırma Projesi*, 2011/EDB/006, İzmir.
- Lin, T. H. (2006). A comparison of model selection indices for nested latent class models. *Monte Carlo Methods and Applications*, 12(3-4), 239-259.
- Lin, T. H. (2012). Model selection information criteria in latent class models with missing data and contingency question. *Journal of Statistical Computation and Simulation*, DOI: 10.1080/00949655.2012.698621.
- Magidson, J. ve Vermunt, J. (2004a). Latent class models. D. Kaplan, (Ed.), *Handbook of quantitative methods in social science research* içinde (175-198). Newbury Park, CA: Sage Publications.
- Magidson, J. ve Vermunt, J. K. (2004b). Comparing latent class factor analysis with the traditional approach in data mining. H. Bozdogan, (Ed.), *Statistical data mining and knowledge discovery* içinde (373-383). Boca Raton: Chapman & Hall/CRC.
- McCutcheon, A. L. (1987). *Latent class analysis*. Beverly Hills: Sage Publication.
- McCutcheon, A. L. (2002). Basic concepts and procedures in single-and multiple-group latent class analysis. J. A. Hage-naars ve A. L. McCutcheon, (Ed.), *Applied latent class analysis* içinde (56-89). New York: Cambridge University Press.
- Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York: Taylor & Francis Group.
- Moors, G. ve Wennekers, C. (2003). Comparing moral values in western european countries between 1981 and 1999. A multigroup latent-class factor approach. *International Journal of Comparative Sociology*, 44, 155-172.
- Morren, M., Gelissen, J. ve Vermunt, J. K. (2011). Dealing with extreme response style in cross-cultural research: A restricted latent class factor analysis. *Sociological Methodology*, 41, 13-47.
- Morren, M., Gelissen, J. ve Vermunt, J. K. (basımda). The impact of controlling for extreme responding on measurement equivalence in cross-cultural research. *Methodology*.
- Muthen, L. K. ve Muthen, B. O. (1998-2007). *Mplus user's guide fifth edition*. Los Angeles: CA: Muthen & Muthen.
- Nylund, K. L., Asparouhov, T. ve Muthen, B. O. (2007). Deciding on the number of classes in latent class analysis in latent class analysis and growth mixture modeling: a monte carlo simulation study. *Structural Equation Modeling*, 14(4), 535-569.
- Park, N., Peterson, C. ve Seligman, M. E. (2004). Strengths of character and well-being. *Journal of Social and Clinical Psychology*, 23(5), 603-619.
- Samuelsen, K. ve Dayton, M. (2010). Latent class analysis. G. R. Hancock ve R. O. Mueller, (Ed.), *The reviewer's guide to quantitative methods in the social sciences* içinde (173-184). London: Taylor & Francis Group.
- Silvia, P. J., Kaufman, J. C. ve Pretz, J. E. (2009). Is creativity domain-specific? Latent class models of creative accomplishments and creative self-descriptions. *Psychology of Aesthetics, Creativity, and the Arts*, 3(3), 139-148.
- Somer, O., Korkmaz, M., Dural, S. ve Can, S. (2009). Ölçme eşdeğerliğinin yapısal eşitlik modellemesi ve madde cevap kuramı kapsamında incelenmesi. *Türk Psikoloji Dergisi*, 24(64), 61-75.
- Vermunt, J. K. ve Magidson, J. (2005a). *Technical guide for Latent GOLD 4.0: Basic and advanced*. Belmont Massachusetts: Statistical Innovations Inc.
- Vermunt, J. K. ve Magidson, J. (2005b). *Latent GOLD 4.5 [Bilgisayar Yazılımı]*. Belmont, MA: Statistical Innovations Inc.
- Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion and the Bayesian information criterion. *Psychological Methods*, 17(2), 228-243.
- Yang, C. ve Yang, C. (2007). Separating latent classes by information criteria. *Journal of Classification*, 24, 183-203.

Summary

Multi-Group Latent Class Analysis and Measurement Equivalence

Duygu Güngör
İzmir University

Mediha Korkmaz
Ege University

Oya Somer
Gediz University

Most of the studies in social sciences focus on group comparisons based on participants' sub-group membership such as gender, ethnicity, and age. Measurement equivalence (ME) must be hold so as to make psychometrically valid comparisons. Traditionally, ME is described as having a true score independent from being a member of any group (Horn & McArdle, 1992; Millsap, 2011). Multi-group confirmatory factor analysis in the context of structural equation modeling and differential item function in the context of item response theory are the most popular methods in detecting ME. The common aspect of these two methods is the assumptions they hold such as normal distribution and homogeneity, and they both need continuous observed and/or latent variables. In psychology, however, latent and observed variables are not always continuous or normally distributed. In this point, latent class models -a latent variable model in which both observed and latent variables are discrete- are alternative methods.

Multi-group latent class analysis (MGLCA) first introduced by Clogg and Goodman in early 1980s (Kankaras, Moors, & Vermunt, 2011). By adding an observed group variable to traditional LCA, the model proposes the following. . For instance, assume that A is an observed variable having I categories ($I = 1, \dots, i$), B is an observed variable having J categories ($J = 1, \dots, j$), C is an observed variable having K categories ($K = 1, \dots, k$), X is a latent variable having T classes ($T = 1, \dots, t$) and G is an observed group variable having S subgroups ($S = 1, \dots, s$). Then,

$$\pi_{ij kts}^{ABCX|G} = \pi_{ts}^{X|G} \pi_{its}^{A|XG} \pi_{jts}^{B|XG} \pi_{kts}^{C|XG}$$

In this equation;

$\pi_{ij kts}^{ABCX|G}$ is the joint probability that an observation would be in category i on variable A , in category j on

variable B , in category k on variable C , given that the observation is in class t on variable X and in subgroup s on variable G . $\pi_{ts}^{X|G}$ is the conditional probability that an observation/individual is in class t on variable X , given that the observation is in group s on variable G . $\pi_{its}^{A|XG}$ is the conditional probability that an observation is in category i on variable A , providing that the observation in class t on variable X as well as in subgroup s on variable G . The latest description is exactly the same for $\pi_{jts}^{B|XG}$ and $\pi_{kts}^{C|XG}$.

There are two basic parameters in LCA, which describe conditional probabilities and latent class probabilities. Conditional probabilities are similar in terms of their factor loadings and those probabilities inform us about the structure. On the other hand, latent class probabilities are about class proportions. In unrestricted MGLCA, these parameters are estimated for all subgroups. Additionally, there are many possible results that can be gathered when studying MGLCA. McCutcheon (1987) mentioned that results will fit for three models; homogenous, partial homogenous, and heterogenous models.

Model selection also seems to be controversial in latent class analysis. There has been a growing body of simulation studies regarding this procedure. An alternative for them seems using the likelihood ratio test (LRT) statistics; however, there exists a consensus on the risks of using this method since it might not follow chi-squared distribution especially when the data is sparse (Collins & Lanza, 2010; Dias, 2006; Morren, Gelissen, & Vermunt, 2011; Yang & Yang, 2007). In addition to that, using information criteria like AIC, BIC, as well as the derivate of these methods are more common.

This study aims at introducing multi-group latent class analysis and presenting the researches an empirical example based on the data gathered from a personality test.

Method

Participants

Participants were 733 undergraduate students from 14 different faculties including 50 different departments in Ege University. Among 733 participants, 496 were female and 237 were male with the mean age of 21.4 ($SD = 2.2$).

Materials

Values in Action - VIA (Park, Peterson, & Seligman, 2004) Inventory has six factors and 24 sub-dimensions. Data of this study was taken from a research aiming to investigate the reliability and the validity of Turkish version of VIA (Korkmaz, 2011). Only 10 items of the "capacity of love" dimension were used as an example in order to introduce MGLCA. Although latent class analysis could be applied to the 5-point scales, we recoded the data by using three response categories [1 = much like me (Categories 1 and 2), 2 = neither like me nor unlike me (Category 3), 3 = much unlike me (Categories 4 and 5)]. The major reason for recoding the scales was to make fewer explanations to describe empirical results so that we could focus on the basic premise of MGLCA.

Analyses

Multi-group latent class analyses were employed by using Latent Gold 4.5 (Vermunt & Magidson, 2005b).

Heterogeneous model, several partial homogenous models, homogenous model, and completely homogenous model were tested respectively.

Results

Before conducting MGLCA, we first tested as to whether the class numbers are comparable between the group of males, the group of females, and the group as a whole. group as a whole, including both females and males. For this purpose 1 to 6-classes LCA models were tested separately. Lowest BIC and CAIC values pointed out a two-class structure for the both groups. Although AIC and AIC3 values indicated more classes, we decided to inspect two-class model due to parsimony and interpretability.

As a next step heterogeneous model, several partial homogenous models, homogenous model, and completely homogenous model were tested respectively for two-class model. As seen on Table 1, homogenous model was found as the best model based on BIC and CAIC.

Homogenous model showed us that the structure was equivalent for comparison groups. Completely heterogeneous model was also tested so as to see whether the prevalence of classed was equivalent (by restricting the latent class probabilities). Based on the results, the latent class probabilities are found significantly different for male and female groups (see Table 1).

Table 1. Two-Cluster Multi-Group Latent Class Analyses Results

	Number of Parameters	L ²	BIC	AIC	AIC3	CAIC	df
Homogenous Model	62	2830.78	-1595.91	1488.78	817.78	-2266.91	671
Partial Homogenous Model	52	2842.93	-1649.73	1480.93	799.93	-2330.73	681
+ Item 1	50	2853.20	-1652.65	1487.20	804.20	-2335.65	683
+ Item 2	48	2853.51	-1665.53	1483.51	798.51	-2350.53	685
+ Item 3	46	2854.54	-1677.70	1480.54	793.54	-2364.70	687
+ Item 4	44	2858.14	-1687.29	1480.14	791.14	-2376.29	689
+ Item 5	42	2861.70	-1696.93	1479.70	788.70	-2387.93	691
+ Item 6	40	2873.27	-1698.55	1487.27	794.27	-2391.55	693
+ Item 7	38	2873.30	-1711.72	1483.30	788.30	-2406.72	695
+ Item 8	36	2874.54	-1723.67	1480.54	783.54	-2420.67	697
+ Item 9	34	2882.97	-1728.43	1484.97	785.97	-2427.43	699
+ Item 10	32	2885.47	-1739.13	1483.47	782.47	-2440.13	701
Homogenous Model	32	2885.47	-1739.13	1483.47	782.47	-2440.13	701
Completely Homogenous Model	31	2894.62	-1736.58	1490.62	788.62	-2438.58	702

Table 2. Parameter Estimations of Homogenous Model

Am willing to take risks to establish a relationship.					Know someone whom I really care about as a person.						
Conditional Probabilities	Class	Much like me	Neither like me nor unlike me	Much unlike me	Conditional Probabilities	Class	Much like me	Neither like me nor unlike me	Much unlike me		
		1	.72	.72			.72	1	.76	.07	.00
		2	.28	.28			.28	2	.24	.93	1
Know that there are people in my life who care as much for me as for themselves.					Do not easily share my feelings with others.						
Conditional Probabilities	Class	Much like me	Neither like me nor unlike me	Much unlike me	Conditional Probabilities	Class	Much like me	Neither like me nor unlike me	Much unlike me		
		1	.80	.30			.08	1	.65	.76	.81
		2	.20	.70			.92	2	.35	.24	.19
Know that some others accept my shortcomings.					Feel isolated from other people.						
Conditional Probabilities	Class	Much like me	Neither like me nor unlike me	Much unlike me	Conditional Probabilities	Class	Much like me	Neither like me nor unlike me	Much unlike me		
		1	.79	.57			.31	1	.39	.65	.85
		2	.21	.43			.69	2	.61	.35	.15
Am the most important person in someone else's life.					Have difficulty accepting love from anyone.						
Conditional Probabilities	Class	Much like me	Neither like me nor unlike me	Much unlike me	Conditional Probabilities	Class	Much like me	Neither like me nor unlike me	Much unlike me		
		1	.81	.70			.35	1	.20	.44	.87
		2	.19	.30			.65	2	.80	.56	.13
Can express love to someone else.					Could never stop loving my family and close friends, no matter what they did.						
Conditional Probabilities	Class	Much like me	Neither like me nor unlike me	Much unlike me	Conditional Probabilities	Class	Much like me	Neither like me nor unlike me	Much unlike me		
		1	.80	.66			.34	1	.76	.62	.54
		2	.20	.34			.66	2	.24	.38	.46

Latent Class Probabilities											
Gender	Class 1	Class 2									
Female	.76	.24									
Male	.63	.37									

Parameter estimations of two-cluster homogenous model are shown in Table 2. Conditional probabilities are like the factor loadings so that they can be interpreted just like them.

In general item contents represent capacity of love; Class-1 can be labeled as *high love capacity* whereas

Class-2 can be labeled as *low love capacity*. When taking a closer look at the sex differences with respect to class assignment it is seen that the differences were not critical; males were slightly overrepresented in Class 2 ($m = 37\%$; $f = 24\%$); however, it was females who were overrepresented in Class 1 ($m = 63\%$; $f = 76\%$).

Discussion

Detecting measurement equivalence has been a critical focus for many studies until recently, however, the analyses of data in these studies has mainly based on the methodologies like multi-group confirmatory factor analysis or differential item functioning, although multi-group latent class analysis is comparatively a new method used in ME studies. The major difference of the latter method is that normal distribution is not required for the analysis. Additionally, when both latent and observed variables are discrete, only latent class analysis can be used.

The advantages of MGLCA are as follows: (i) The only assumption of the method is local independence. (ii) Model comparisons can be made even if class numbers are different between the comparisons groups and/or inequivalent item numbers are high. (iii) It can handle with multidimensionality. However, there are some limitations of latent class models. Both Kankaras and Moors (2009) and Magidson and Vermunt (2004b) underlined the difficulty of reporting factor loadings since they turn

out to be conditional probabilities due to the discrete variables.

To sum up, two-class model was found as the best model in this study based on BIC and CAIC, however AIC indicated more classes. Dias (2006) and Nylund et al. (2007) reported that AIC tends to yield more classes than the number of classes actually exists as results of their simulation studies.

Moreover, only AIC indicated partial homogenous model, whereas homogenous model was hold when the analyses based on CAIC, BIC and AIC3. This result is parallel with the simulation study results that reported AIC results were false positive in some conditions. Homogenous model is like scalar equivalence. When partial homogenous model accepted it is known that intercept and/or slope parameters are different for some items. In this study, item level analyses were also done and they were used as a supporting evidence for the homogenous model.

Although there is a growing body of research using latent class analysis, further empirical studies are still needed with latent class models detecting ME.