

Çok Boyutlu Test Yapılarında Alfa, Tabakalı Alfa ve Omega Güvenirlik Katsayılarının Karşılaştırılması¹

Sümevra SOYSAL 

¹Dr. Öğr. Üyesi, Necmettin Erbakan Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Konya, Türkiye

[sumevasoysal@hotmail.com](mailto:sumevrasoysal@hotmail.com)

Makale Bilgileri	ÖZ
<p>Makale Geçmişi Geliş: 05.10.2022 Kabul: 01.03.2023 Yayın: 31.03.2023</p> <p>Anahtar Kelimeler: Güvenirlik, Cronbach alfa, Tabakalı Alfa, Omega, Çok Boyutlu Test.</p>	<p>Bu çalışmada çok boyutlu ölçümlerde iç tutarlık anlamında güvenilirliğin değerlendirilmesinde Cronbach alfa katsayısı yerine daha uygun alternatiflerini kullanma konusunda farkındalık oluşmasına katkı sağlamak amaçlanmıştır. Bu amaç kapsamında örneklem büyüklüğü (100, 200 ve 400), boyutlar arası korelasyon (0,00 ve 0,50), maddeler arası korelasyon (0,30-0,50 ve $\geq 0,75$), test uzunluğu (10 ve 20) ve boyutlara düşen madde sayısı (5-5, 7-3, 10-10 ve 15-5) koşulları altında iki boyutlu basit ve karmaşık test yapılarında üretilen verilerde alfa, tabakalı alfa ve omega katsayılarının performansı incelenmiştir. Ayrıca araştırmacılara tabakalı alfa ve omega katsayılarının manuel olarak kolaylıkla hesaplanabildiğini göstermek amacıyla örnekler sunulmuştur. Veriler R programı ile psych ve sirt paketleri kullanılarak analiz edilmiştir. Araştırmanın sonucunda güvenilirlik kestirimleri üzerinde en fazla etkisi olan değişkenlerin sırasıyla maddeler arası korelasyon ve boyutlar arası korelasyon olduğu bulunmuştur. Test uzunluğunun ilişkisiz modellerde maddeler arası düşük korelasyon koşulu altında güvenilirlik kestirimleri üzerinde daha fazla etkiye sahip olduğu görülmüştür. Örneklem büyüklüğü ortalama güvenilirlik kestirimleri üzerinde etkili bulunmamıştır ancak kestirimlerin hatası üzerinde etkili bulunmuştur. Omega ve tabakalı alfa katsayılarının performansı benzer bulunurken alfa katsayısı ise bu katsayılardan daha düşük güvenilirlik değerleri kestirmiştir. Hem alanyazına hem de mevcut çalışmanın sonuçlarına dayalı olarak araştırmacılara, çok boyutlu ölçümlerde iç tutarlık anlamında güvenilirlik kanıtı olarak alfa katsayısını kullanmamaları ya da tek başına kullanmamaları; bunun yerine omega, tabakalı alfa gibi daha uygun katsayıları kullanıp raporlamaları önerilir.</p>

Comparison of Alpha, Stratified Alpha, and Omega Reliability Coefficients in Multidimensional Test Structures

Article Info	ABSTRACT
<p>Article History Received: 05.10.2022 Accepted: 01.03.2023 Published: 31.03.2023</p> <p>Keywords: Reliability, Cronbach Alpha, Stratified Alpha, Omega Multidimensional Test.</p>	<p>In this study, it is aimed to contribute to raising awareness about using more suitable alternatives instead of Cronbach's alpha coefficient in the evaluation of reliability in terms of internal consistency in multidimensional measurements. For this purpose, it is examined that how alpha, stratified alpha and omega coefficients are affected by the sample size (100, 200 and 400), correlation between dimensions (0.00 and 0.50), correlation between items (0.30-0.50 and ≥ 0.75), test length (10 and 20) and the number of items per dimensions (5-5, 7-3, 10-10, and 15-5) in the data simulated in two-dimensional simple and complex test structures. In addition, examples are presented to the researchers to show that the stratified alpha and omega coefficients can be easily calculated manually. Data were analyzed with R program using psych and sirt packages. As a result of the study, it is found that the variables that had the most effect on the reliability estimations were the correlation between items and correlation between dimensions, respectively. It is observed that test length has a greater effect on reliability estimates under the condition of low correlation between items in unrelated models. The sample size is not effective on the average of reliability estimations, but it is effective on the error of the estimations. While the performance of the omega and stratified alpha coefficients are quite similar, the alpha coefficient estimates lower reliability values than these coefficients. Based on both the literature and the results of the present study, it is suggested that the alpha coefficient should not be reported, or not be reported alone, as proof of reliability in terms of internal consistency in multidimensional measurements; instead, it is recommended to use and report more appropriate coefficients such as omega, stratified alpha.</p>

Atıf/Citation: Soysal, S. (2023). Çok Boyutlu Test Yapılarında Alfa, Tabakalı Alfa ve Omega Güvenirlik Katsayılarının Karşılaştırılması. *Ahmet Keleşoğlu Eğitim Fakültesi Dergisi (AKEF) Dergisi*, 5(1), 213-236.



"This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) (CC BY-NC 4.0)"

*Bu çalışmanın bazı koşullarına ait sonuçları 6. Uluslararası Eğitimde ve Psikolojide Ölçme ve Değerlendirme Kongresinde (6th International Congress on Measurement and Evaluation in Education and Psychology, 5-8 Eylül, Kosova) sözlü bildiri olarak sunulmuştur.

GİRİŞ

Sosyal ve davranış bilimlerinde araştırmacılar, bireylerin ilgi, kaygı, yetenek, başarı, tutum gibi örtük özellikleriyle ilgilenirler. Doğrudan gözlenemeyen bu özellikler ancak bireylerin ölçek, test veya anket maddelerine verdikleri tepkiler ile ölçülebilirler. Bu dolaylı ölçme işlemi, ölçme sonuçlarını tesadüfi hatalara daha açık hale getirir ve bu nedenle psikometrik araştırmalarda ölçme sonuçlarının güvenilirliği temel konulardan biridir. Genel olarak tutarlı koşullar altında bireyler için benzer ölçme sonuçları üreten bir ölçme aracının güvenilir olduğu kabul edilir. Bir ölçme aracının tutarlılığı ölçme hatasının kaynağına bağlı olmak üzere test-tekrar test güvenilirliği, paralel formlar güvenilirliği, iç tutarlılık, puanlayıcılar arası güvenilirlik olarak farklı yöntemlerle incelenebilir (Crocker ve Algina, 1986, s.132-143). Dunn ve diğerlerine (2014) göre American Psychological Association tarafından yayınlanan makalelerde en fazla raporlanan güvenilirlik türü, test-tekrar test veya paralel formlar güvenilirliğine kıyasla, iç tutarlılıktır. Bu durumun nedenlerinden biri birden fazla uygulama gerektiren test-tekrar test ve paralel formlar yöntemleri zahmetli, zaman alıcı veya maliyetli olarak kabul edildiğinden araştırmacıların tek uygulama gerektiren iç tutarlılık katsayılarına başvurmasıdır. İç tutarlılık katsayıları arasında en fazla raporlanan ise alfa ve Spearman-Brown eş değer yarılar katsayılarıdır (Cho, 2016; Warrens, 2016).

Felt ve Brennan (1989) klasik test kuramı çerçevesinde dört ölçme modeli tanımlamıştır: Paralel (parallel), eşdeğer (tau equivalent), temelde/özünde eşdeğer (essentially tau equivalent) ve konjenerik (congeneric). Paralel ölçme modellerinde maddelerin ortalaması, standart sapması, faktör yükleri, maddeler arası kovaryanslar ve hata terimlerinin eşit olduğu varsayılmaktadır. Paralel ölçmelerden farklı olarak eşdeğer ölçmelerde, maddelerin ortalaması ve maddeler arası kovaryansların eşit; temelde eşdeğer ölçmelerde ise madde kovaryanslarının eşit diğer özelliklerin eşit olmadığı varsayılır. Konjenerik ölçme modellerinde ise tüm madde parametrelerinin farklı olduğu varsayılır.

Alfa katsayısının güvenilirliğin yansız bir kestirici olması için ölçme aracının tek boyutlu, maddelerin en az temelde eşdeğer olması gereklidir (Raykov, 1998; Traub, 1994). Ancak gerçekte psikometrik ölçümler konjenerik ölçme modellerine daha uygundur. Ayrıca maddeler farklı parametrelere sahip olduğunda aynı örtük değişkenle veya faktörle aynı derecede ortak varyansa sahip olmayabilirler ve bu durum ölçümlerin tek boyuttan ziyade çok boyutlu yapı göstermesine neden olabilir (Soğan, 2000). Dolayısıyla hem konjenerik hem de çok boyutlu ölçme sonuçları, alfa kestirimlerinde önemli yanlışlıklara neden olabilmektedir (Cho, 2016; Cho ve Kim, 2015; Graham, 2006; Sijtsma, 2009). Alfa katsayısının bu sınırlılıklarına karşı, Cronbach ve diğerleri (1965) alt boyutlara sahip ölçme araçlarından elde edilen birleşik puanların güvenilirliği için tabakalı alfa (stratified alpha) katsayısının kullanımı önermiştir. Bazı araştırmacılar çok boyutlu yapılarda tabakalı alfa katsayısının gerçek güvenilirliğin iyi bir kestiricisi olduğunu raporlamışlardır (Kamata vd., 2003; Osburn, 2000, Revelle ve Zinbarg, 2009; Tang ve Cui, 2012). Konjenerik ölçme yapılarında güvenilirlik kestirimleri için önerilen bir diğer katsayı omegadır (McDonald, 1999). Omega katsayısının birçok araştırmacı tarafından hem alfa hem de diğer iç tutarlılık katsayılarıyla karşılaştırıldığında güvenilirliğin daha duyarlı bir kestiricisi

olduğu gösterilmiştir (Dunn vd., 2014; Graham, 2006; Hayes ve Coutts, 2020; Revelle ve Zinbarg, 2009; Zinbarg vd., 2005, 2006, 2007).

Alanyazın araştırmalarında alfa katsayısı, yukarıda sayılan varsayımlarına rağmen çok boyutlu ve konjenerik ölçme sonuçlarının güvenirliliğinin raporlanmasında hala sıkça kullanılmaktadır. Araştırmacılar toplanabilir özellikte alt boyutlara sahip ölçme araçlarından elde edilen ölçümlerin güvenirliliği için testte yer alan tüm maddeler üzerinden genel bir alfa raporlamayı tercih edebilmektedirler. Böyle bir yaklaşımın çok boyutlu yapılarda güvenirliliğin iyi bir kestiricisi olduğu belirtilen tabakalı alfa ve omega katsayılarına göre ne kadar doğru olduğu konusunun araştırılması gerektiği düşünülmüştür. Bu kapsamda alanyazında çok boyutlu veya bileşenli veriler üzerinde yapılan güvenirlilik analizi çalışmalarının sınırlı olduğu görülmektedir (Brunner ve SÜß, 2005; Cortina, 1993; Kamata vd., 2003; Raykov vd., 2015; Serbetar ve Sedlar, 2016; Viladrich vd., 2017; Watkins, 2017, Widhiarso ve Ravand, 2014). Bu araştırmalar arasında maddeler arası korelasyonu manipüle ederek araştırma yapan az sayıda çalışmayla (Cortina, 1993; Kamata vd., 2003; Osburn, 2000) karşılaşılmıştır. Maddeler arası korelasyon matrisinin çok boyutlu güvenirlilik kestirimleri üzerinde etkili olduğu bilinmekle birlikte bu etkinin korelasyon büyüklüklerinden nasıl etkilendiği de araştırılmak istenmiştir. Ayrıca bir maddenin yalnız tek boyutta yük verdiği çok boyutlu yapılara göre bir maddenin birden fazla boyuta yük vermesi durumunda güvenirliliğin nasıl etkilendiği de gözlenmek istenmiştir. Belirlenen ihtiyaca bağlı olarak bu araştırmada “örneklem büyüklüğü, boyutlar arası korelasyon, maddeler arası korelasyon, test uzunluğu ve boyutlara düşen madde sayısı değişkenlerinin farklı düzeyleri altında üretilen çok boyutlu basit ve karmaşık test yapılarında alfa, tabakalı alfa ve omega kestirimleri nasıl değişmektedir?” sorusuna yanıt aranmıştır.

YÖNTEM

Araştırma Modeli

Bu araştırmada farklı test koşulları altında üretilmiş veriler kullanılarak güvenirlilik katsayıları incelendiği için bu araştırma simülasyona dayalı deneysel bir araştırma özelliği taşımaktadır.

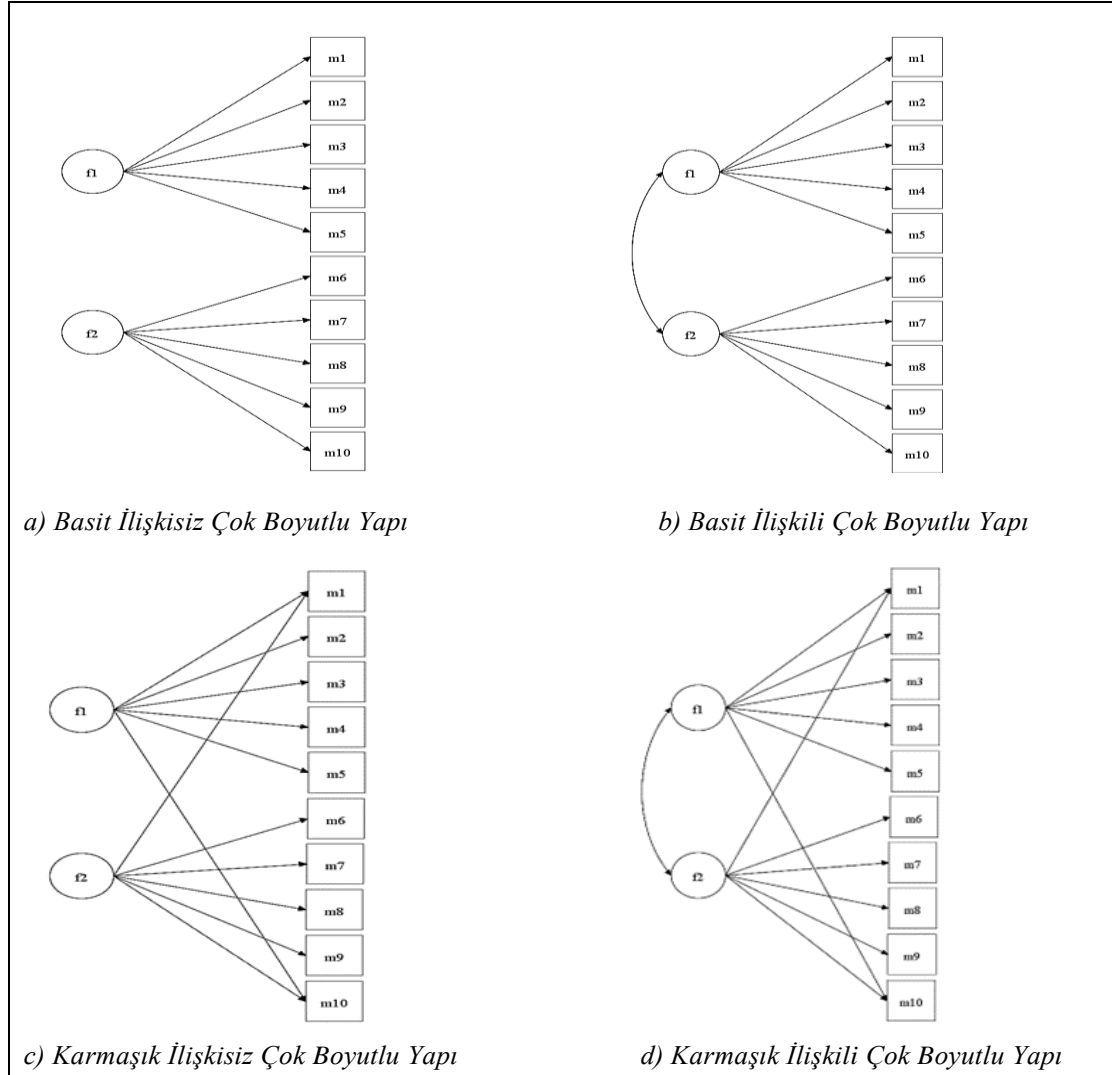
Araştırmanın Koşulları

Bu çalışmada ölçme modeli, örneklem büyüklüğü, boyutlar arası korelasyon, maddeler arası korelasyon, test uzunluğu ve boyutlara düşen madde sayısı değişkenleri bağımsız değişken olarak ele alınmıştır. Bağımsız değişkenlere ilişkin koşullar ve düzeyleri Tablo 1’de özetlenmiştir.

Tablo 1. *Simülasyon Koşulları*

Koşul	Düzye
Ölçme modeli	Basit ve karmaşık
Örneklem Büyüklüğü	100, 200 ve 400
Boyut	2
Boyutlar arası korelasyon	0.00 ve 0.50
Maddeler arası korelasyon	0.30-0.50 ve ≥ 0.75
Test uzunluğu	10 ve 20
Boyutlardaki madde sayısı	5-5; 7-3; 10-10 ve 15-5

Tablo 1’de özetlenen 96 çapraz koşulun ($2 \times 3 \times 2 \times 2 \times 2$) her biri için 1000 tekrar sayısı ile konjenerik ölçme modeline uygun veriler üretilmiş ve toplamda 96 000 veri seti analiz edilmiştir. Harwell ve diğerleri (1996) simülasyon çalışmalarında 25 tekrarın yeterli olabildiğini belirtmişlerdir. Bu araştırmada evren ortalamasına yakınlaşmak ve cevap örüntülerindeki çeşitliliği artırmak için 1000 tekrar sayısı tercih edilmiştir. Araştırmanın tüm verileri iki boyutlu ve beş dereceli likert ölçeğindedir.



Şekil 1. Araştırmanın Ölçme Modelleri

Araştırmanın koşullarından biri ölçme modelidir. Maddelerin yalnızca tek bir boyutta yük verdiği (bir diğer ifadeyle diğer boyutlarda ya hiç yük vermediği ya da çok düşük düzeyde yük verdiği) faktör analitik modellere “basit yapılı”, maddenin birden fazla boyutta önemli düzeyde yük verdiği modellere “karmaşık yapı” denilmektedir (Paek ve Cole, 2019; Zhang, 2007). Pratikteki ölçme uygulamalarında her iki yapıda ölçme sonuçlarıyla karşılaşılabilmekle birlikte bir maddenin birden fazla faktörde sahip olduğu yük değerleri arasındaki farkın 0,1’den küçük olması istenmektedir. Bu durumu da dikkate alarak bu araştırmada ölçme modelleri bir koşul olarak ele alınmıştır. Basit yapı için boyutlar arası korelasyon 0,00 ve 0,50 olacak şekilde manipüle edilmiştir ve bu ölçme modellerine sırasıyla basit ilişkisiz çok boyutlu yapı (Şekil 1a) ve basit ilişkili çok boyutlu yapı adı verilmiştir (Şekil 1b). Benzer şekilde karmaşık modeller için karmaşık ilişkisiz çok boyutlu yapı (Şekil 1c) ve karmaşık ilişkili çok boyutlu yapı tanımlaması yapılmıştır (Şekil 1d).

Araştırmada karmaşık yapı, her bir faktörün ilk maddesi bir diğer faktörde 0,30'dan büyük ancak diğer faktördeki yükünün minimum yarısı kadar yük verecek şekilde manipüle edilerek veriler üretilmiştir.

Alanyazında kullanılan ölçme araçları ve LGS gibi geniş ölçekli testler incelendiğinde sıklıkla 20 maddelik testler/alt testler kullanıldığı görülmektedir. Bu amaçla bu araştırmada toplam test uzunluğu olarak 20 ve daha kısa testleri temsilen (tipik performans testlerinin kısa formları gibi) 10 tercih edilmiştir. Pratikteki ölçme uygulamaları düşünülerek örneklem büyüklüğü N=100, 200 ve 400 olarak belirlenmiştir. Maddeler arası korelasyonun alfa katsayısının da formülünden hatırlanacağı gibi güvenirlilik kestirimleri üzerinde etkisi olduğu göz önüne alınarak bu araştırmanın koşullarından biri olarak seçilmiştir. Bu amaçla 0,30-0,50 ile düşük maddeler arası korelasyonu ve $\geq 0,75$ ile yüksek maddeler arası korelasyonu temsil edecek şekilde iki düzey belirlenmiştir. Maddeler arası korelasyonlar yalnızca aynı boyuttaki maddeler için düşük ve yüksek olarak manipüle edilmiştir. Boyutlara düşen madde sayısının güvenirlilik kestirimlerini etkileyebileceği düşünülerek eşit ve farklı olacak şekilde 10 maddelik test için 5-5 ve 7-3, 20 maddelik test için 10-10 ve 15-5 olarak bu koşulun düzeyleri belirlenmiştir.

Güvenirlilik katsayıları

Alfa

Kuder ve Richardson (1937) ikili puanlanan maddelerden oluşan ölçme araçlarından elde edilen ölçümlerin eşdeğer güvenirliliğini (tau-equivalent reliability) belirlemek üzere KR-20 ve KR-21 katsayılarını geliştirmiştir. Bazı araştırmacılar (Cronbach, 1947, 1951; Guttman, 1945; Hoyt, 1941) KR-20'nin likert tipi ölçümlere genişletilebileceğini göstermiş ve önermiştir. Alfa katsayısı, ölçümlerin tek boyutlu ve eşdeğer (tau-equivalent) olması durumunda güvenirliliği yansız kestirilebilmektedir. Günümüzde daha çok Cronbach alfa katsayısı olarak bilinmektedir ve aşağıdaki formül ile hesaplanmaktadır.

$$\alpha = \frac{n}{n-1} \left[\frac{\sum_{i=1}^n \sigma_i^2}{\sigma_x^2} \right]$$

Formülde yer alan n madde sayısını; σ_i^2 , i. maddenin varyansını; σ_x^2 ise testin varyansını temsil etmektedir. Maddeler standardize edilmişse, maddelerin korelasyon ortalaması veya varyans-kovaryanslarının ortalaması kullanılarak da katsayı hesaplanabilir (Carmines ve Zeller, 1982). Buna bağlı olarak, maddeler arasında negatif korelasyonların varlığı, bazı durumlarda negatif alfa katsayısının ortaya çıkmasına sebep olmaktadır.

Tabakalı Alfa (α_s)

Cronbach ve diğerleri (1965) tarafından birden fazla bileşen, boyut ya da alt ölçeklerden oluşan ölçme araçlarından elde edilen ölçümlerin güvenirliliğini belirlemek üzere alfa katsayısından geliştirmiştir. Tabakalı alfa katsayısı aşağıdaki formül ile hesaplanmaktadır:

$$\alpha_s = 1 - \frac{\sum_{j=1}^k \sigma_j^2 (1 - \alpha_j)}{\sigma_x^2}$$

Formülde yer alan α_j , j.alt boyutun Cronbach alfa güvenilirliğini; σ_j^2 , j. alt boyutun varyansını ve σ_x^2 ise toplam test varyansını temsil etmektedir. Bu araştırmada üretilen verilerden biri üzerinde tabakalı alfa katsayısının hesaplanmasına ilişkin bir örnek aşağıda paylaşılmıştır.

Tablo 2. Örnek Veri için Betimsel İstatistikler*

Boyut	Madde Sayısı	Varyans	Alfa	Boyutlar Arası Korelasyon	
				Boyut1	Boyut 2
Boyut 1	5	19.14	0.71	1	0,00
Boyut 2	5	18.80	0.69		1
Tüm test	10	38.26	0.62		

*Araştırma için üretilen verilerden seçilmiştir.

Tablo 2’de boyutlar arası korelasyonun sıfır, maddeler arası korelasyonun (aynı boyuttaki maddeler için, bkn. Ek Tablo1) 0,30-0,40 olduğu iki boyutlu bir veri setine ait betimsel istatistikler verilmiştir. Verilen değerler tabakalı alfa formülünde yerine yazılırsa,

$$\alpha_s = 1 - \frac{19,14 \times (1 - 0,71) + 18,80 \times (1 - 0,69)}{38,26} = 0,70$$

olarak tabakalı alfa kolaylıkla hesaplanır. Tablo 2’de görüldüğü gibi tüm test için alfa katsayısı 0,62 iken tabakalı alfa katsayısı ile güvenilirlik 0,70 olarak hesaplanır.

Omega (ω_g)

McDonald (1999) tarafından yapısal eşitlik modellemesinin bir parçası olarak faktör-analitik çerçeveye dayalı geliştirilmiş bir güvenilirlik katsayısıdır. Heise ve Bohrnstedt’in (1970) çok boyutlu ölçümlerin güvenilirliğini belirlemek için önerdiği omega katsayısının farklı bir versiyonudur. McDonald (1999), ölçme aracı ile ölçülen yapıyı gösterge değişkenlerin ne kadarını açıklayabildiğini ortak faktör modelleri (common factor models) aracılığıyla belirlemesi nedeniyle bu katsayıyı yapı güvenirligi (construct reliability) olarak da isimlendirmektedir. İki farklı omega katsayısı vardır: genel (ω_g) ve hiyerarşik (ω_h). Hiyerarşik omega, birden fazla boyutun olduğu ancak hala genel bir faktörün geçerli olduğu ölçümlerin bir güvenirligi olarak McDonald (1970) tarafından tanıtılmıştır. Çoğu durumda ω_g , ω_h ’a göre daha yüksek değerler vermektedir. Her bileşeni gerçek puan varyansı ile ağırlıklandırması nedeniyle ω_h , yapısal eşitlik modelleri kullanılarak değerlendirilen ölçümler için daha uygundur (Bacon vd., 1995). Her iki omega katsayısının formülü aşağıda verilmiştir. Ancak bu çalışmanın tüm koşulları için daha uygun olması nedeniyle bu araştırmada yalnızca ω_g katsayısının performansı raporlanmıştır.

$$\omega_g = \frac{\sum_{j=1}^k (\sum_{i=1}^p \lambda_{ij})^2}{\sum_{j=1}^k (\sum_{i=1}^p \lambda_{ij})^2 + \sum_{i=1}^p e_i}$$

$$\omega_h = \frac{(\sum_{i=1}^p \lambda_{ij})^2}{\sum_{j=1}^k (\sum_{i=1}^p \lambda_{ij})^2 + \sum_{i=1}^p e_i}$$

Formülde yer alan λ_{ij} , i. gösterge değişkenin (maddenin) j. faktördeki standardize faktör yükünü; e_i ise gösterge değişkenlerin tekil varyansını (hata) temsil etmektedir. $e_i = 1 - (\sum_{j=1}^p \lambda_{ij})^2$; yani bir gösterge değişkenin hata varyansı, maddenin standardize faktör yükünün karesinin birden farkına eşittir.

Omega, ortak faktör modeli olarak hem açımlayıcı faktör analizi (AFA) hem de doğrulayıcı faktör analizi (DFA) kullanılarak hesaplanabilir (Brunner ve Süß, 2005). Yukarıda tabakalı alfa katsayısının hesaplanmasında kullanılan aynı örnek veri ile omega katsayısının hesaplanması aşağıda paylaşılmıştır. Tablo 2’de bazı betimsel istatistikleri verilen örnek verinin DFA’dan elde edilen standardize faktör yükü ve hata varyansları ise Tablo 3’te özetlenmiştir.

Tablo 3. Örnek Verinin Standardize Faktör Yükü ile Hata Varyansı*

Boyut	Madde	$\lambda_{boyut 1}$	$\lambda_{boyut 2}$	e_i
Boyut 1	M1	0.50		0.75
	M2	0.50		0.75
	M3	0.58		0.66
	M4	0.62		0.62
	M5	0.60		0.64
Boyut 2	M6		0.51	0.74
	M7		0.63	0.60
	M8		0.61	0.63
	M9		0.50	0.75
	M10		0.62	0.62

*Araştırma için üretilen verilerden seçilmiştir.

Genel omega katsayısını hesaplamak için Tablo 3’teki değerler ω_g denkleminde yerine yazıldığında;

$$\sum_{j=1}^k (\sum_{i=1}^p \lambda_{ij})^2 = (0,50+0,50 +0,58+0,62+0,60)^2 + (0,51+0,63 +0,61+0,50+0,61)^2 = 16,08$$

$$\sum_{i=1}^p e_i = (0,75+0,75 +0,66+0,62+0,64+0,74+0,60 +0,63+0,75+0,62) = 6,76$$

$$\omega_g = \frac{16,08}{16,08+6,76} = 0,70 \text{ olarak bulunur.}$$

Verilerin Analizi

Üretilen verilerin yapısal ve madde korelasyon matrisi açısından doğruluğu R programında psych (Revelle ve Revelle, 2022) paketi kullanılarak ağırlıklandırılmamış en küçük kareler (unweighted least squares) kestirim ve oblimin döndürme yöntemi ile polikorik korelasyon matrisine dayalı olarak AFA ile kontrol edilmiştir. Basit yapılu üretilen verilerde bir madde yalnızca kendi boyutunda 0,30’dan büyük yük vermiştir, diğer boyutta verdiği yükün ise $\leq 0,05$ olduğu gözlenmiştir. Karmaşık yapılu üretilen veri setinde ise her bir boyutun ilk maddesinin bir diğer boyutta 0,30’dan büyük ancak diğer boyuttaki yükünün minimum yarısı kadar yük verdiği, geri kalan maddelerin ise yalnızca tek bir boyutta yük verdiği doğrulanmıştır. Maddeler arası korelasyonun sıfır olduğu koşullarda minimum faktör yükünün

0,50; boyutlar arası 0,50 olan koşullarda ise minimum faktör yükünün 0,85 olduğu gözlenmiştir. Araştırma kapsamında incelenen güvenirlilik katsayılarından alfa ve tabakalı alfa kestirimleri R programında *sirt* paketi (Robitzsch, 2022), omega kestirimleri ise *psych* (Revelle ve Revelle, 2022) paketi kullanılarak hesaplanmıştır. Araştırmanın sonuçları 1000 tekrardan kestirilen güvenirlilik katsayılarının ortalaması üzerinden yorumlanmıştır. Verilerin doğruluğunu kontrol etmede olduğu gibi omega kestirimlerinde faktör analizi için ağırlıklandırılmamış en küçük kareler ve oblimin döndürme yöntemi kullanılmıştır.

Etik

Araştırmanın verilerinin bilgisayar ortamında üretilmesi nedeniyle etik kurul onayı gerekmemektedir.

BULGULAR

Araştırmanın bulguları ölçme modellerine ve örneklem büyüklüğü koşulunun düzeylerine bağlı olarak sırayla sunulmuştur.

Basit Ölçme Modellerine İlişkin Sonuçlar

Basit ölçme modeline ilişkin sonuçlar Tablo 4-Tablo 6 arasında özetlenmiştir. Omega katsayısının konjenerik ölçmelerde gerçek güvenirliliğin daha yansız bir kestiricisi olduğunu belirten çalışmalar (Dunn vd., 2014; Graham, 2006; Hayes ve Coutts, 2020; Revelle ve Zinbarg, 2009; Zinbarg vd., 2005, 2006, 2007) referans alınarak omega katsayısına göre alfa ve tabakalı alfa katsayılarının ortalama yanlılığı Tablo 7’de özetlenmiştir.

Tablo 4. Basit Ölçme Modelinde N=100 için Kestirilen Güvenirlilik Değerleri

B.kor.	M.kor.	Madde Sayısı	Alfa			Tabakalı Alfa			Omega		
			Min.	Maks.	Ort.	Min.	Maks.	Ort.	Min.	Maks.	Ort.
0.00	0.30-0.50	5-5	0.35	0.74	0.62	0.50	0.80	0.70	0.61	0.81	0.72
		7-3	0.43	0.79	0.67	0.54	0.82	0.73	0.62	0.83	0.75
		10-10	0.66	0.86	0.77	0.74	0.89	0.82	0.78	0.89	0.83
		15-5	0.72	0.90	0.83	0.74	0.91	0.86	0.75	0.91	0.86
	≥0.75	5-5	0.72	0.90	0.83	0.74	0.91	0.86	0.75	0.91	0.86
		7-3	0.79	0.92	0.88	0.93	0.97	0.95	0.93	0.97	0.96
		10-10	0.87	0.95	0.92	0.96	0.98	0.97	0.96	0.98	0.97
		15-5	0.93	0.97	0.95	0.96	0.98	0.98	0.97	0.98	0.98
0.50	0.30-0.50	5-5	0.62	0.84	0.75	0.67	0.85	0.78	0.68	0.86	0.79
		7-3	0.63	0.85	0.76	0.66	0.85	0.78	0.67	0.87	0.79
		10-10	0.78	0.90	0.86	0.79	0.91	0.87	0.80	0.91	0.87
		15-5	0.81	0.92	0.87	0.82	0.92	0.88	0.82	0.93	0.88
	≥0.75	5-5	0.86	0.95	0.92	0.94	0.97	0.96	0.94	0.97	0.96
		7-3	0.90	0.96	0.93	0.95	0.98	0.96	0.95	0.98	0.96
		10-10	0.93	0.97	0.96	0.97	0.99	0.98	0.97	0.99	0.98
		15-5	0.95	0.98	0.97	0.97	0.99	0.98	0.97	0.99	0.98

*B.kor: Boyutlar arası korelasyon; M.kor: Maddeler arası korelasyon; Min:1000 tekrarda minimum kestirilen değer; Maks: 1000 tekrarda maksimum kestirilen değer; Ort:1000 tekrarın ortalaması

Tablo 4 incelendiğinde, alfa katsayısı hemen hemen tüm koşullarda omega ve tabakalı alfa katsayılarından daha düşük güvenirlilik değerleri kestirmiştir. Özellikle boyutların ilişkisiz olduğu ($B.kor=0,00$) ölçme modelinde kestirim ortalamaları açısından omega ile alfa katsayısı arasında minimum 0,10 kadar fark olduğu görülmektedir. Alanyazında ve pratikte bir ölçme aracından elde edilen ölçümlerin minimum 0,70 güvenirlilikte olması yaygın kabul edilen kuraldır. Bu bağlamda, basit ilişkisiz modelde maddeler arası korelasyon 0,30-0,50 koşulu altında ortalamaya göre tabakalı alfa ve omega tüm durumlarda kabul edilebilir ($>0,70$) güvenirlilik kestirimleri üretebilirken alfa katsayısı ancak toplam test uzunluğu 20 madde iken kabul edilebilir kestirimler yapabirmiştir. 1000 tekrarın tamamı değerlendirildiğinde alfa katsayısının 0,35 gibi oldukça düşük güvenirlilik değeri üretebildiği dikkat çekmektedir. Alfa ile güvenirliliği 0,35 olarak kestirilen veri setinde tabakalı alfa 0,50 ve omega 0,76'dır. Dolayısıyla bazı koşullarda alfa ile omega arasındaki fark 0,40'a çıkmıştır. Tabakalı alfa ve omega katsayıları ise 10 maddelik verilerde 0,70'in altında sonuçlar vermiştir. Tabaklı alfa, daha kısa testlerde omega katsayısından daha düşük kestirimler yapmıştır.

Boyutlar arası korelasyonun artması durumunda yani basit ilişkisiz modele ($B.kor=0,00$) göre basit ilişkili modelde ($B.kor=0,50$) tüm koşullarda güvenirliliğin daha yüksek kestirildiği gözlenmektedir. Cortina'nın (1993) alfa katsayısının ilişkisiz çok boyutlu kestirimleri üzerinde yaptığı çalışmasıyla tutarlı olarak basit ilişkisiz modelde ve maddeler arası korelasyonun $\geq 0,75$ olduğu koşullarda üç katsayı da 0,70'in üzerinde kestirim yapmıştır. Dolayısıyla maddeler arası korelasyonun boyutlar arası korelasyondan daha fazla güvenirlilik katsayıları üzerinde etkisi vardır. Bu beklenen bir durumdur. Çünkü tabakalı alfa ve özellikle alfa katsayıları madde korelasyonlarının ortalamasının bir fonksiyonu olduğu için maddeler arası ilişkiler kuvvetlendikçe daha yüksek güvenirlilik değerleri elde edilir. Faktör analitik modellerde maddeler arası korelasyon arttığında daha yüksek faktör yük değerlerinin ortaya çıkması beklenir. Nitekim hem ilişkili hem de ilişkisiz modelde maddeler, maddeler arası korelasyonun 0,30-0,50 olduğu koşullarda 0,50-0,70 arası; maddeler arası korelasyon $\geq 0,75$ iken $>0,85$ faktör yükü üretmiştir. Omega katsayısı da faktör analitik modele dayandığı için maddeler arası korelasyonun yükselmesi güvenirliliği yükseltmiştir.

Tabakalı alfa ile omega katsayısı iki koşul haricinde birbiriyle oldukça tutarlı sonuçlar vermiştir. İlişkisiz modellerde boyutlar arası korelasyonun 0,30-0,50 olduğu koşul ile test uzunluğunun 10 madde olduğu koşullarda tabakalı alfa kestirimleri omega kestirimlerine göre daha düşüktür; diğer koşullarda iki katsayı arasındaki fark 0,01'i aşmamıştır. Test uzunluğunun artmasıyla güvenirliliğin arttığı bilinmektedir (Cortina, 1993) ve bulgular bunu desteklemektedir. Bu bağlamda dikkat çeken bir diğer bulgu, eşit uzunluklardaki testlerde boyutlardan biri diğerinden daha fazla maddeye sahip olan testlerin (7-3 gibi) eşit uzunluktaki boyutlara sahip testlerden (5-5 gibi) daha yüksek güvenirliliğe sahip olmasıdır.

Tablo 5, $N=200$ için basit ölçme modelinde kestirilen güvenirlilik değerlerini özetlemektedir. Tablo 4 ile karşılaştırıldığında alfa, tabakalı alfa ve omega katsayılarının ortalamasının örneklem büyüklüğünden etkilenmediği görülmektedir. Örneklem büyüklüğünün ortalamalar üzerinde etkisi gözlenmemiştir. Ancak aynı parametreler üzerinden üretilmiş daha geniş veri setlerinde kestirilen minimum güvenirlilik değerlerinin yükseldiği gözlenmektedir. Dolayısıyla bu durum minimum değerler açısından alfa ile diğer iki katsayı arasındaki farkın da düşmesini sağlamıştır. Örneğin, alfa katsayısının 0,44 kestirildiği veri setinde tabakalı

alfa 0,57 ve omega 0,65 olarak elde edilmiştir. Dolayısıyla 100 örneklem büyüklüğü için bazı koşullarda alfa ile omega arasında ortaya çıkan 0,40 birimlik fark 0,30'a kadar düşmüştür.

Tablo 5. Basit Ölçme Modelinde N=200 için Kestirilen Güvenirlik Değerleri

B.kor.	M.kor.	Madde Sayısı	Alfa			Tabakalı Alfa			Omega		
			Min.	Maks.	Ort.	Min.	Maks.	Ort.	Min.	Maks.	Ort.
0.00	0.30-0.50	5-5	0.44	0.72	0.62	0.57	0.77	0.70	0.65	0.79	0.72
		7-3	0.55	0.75	0.67	0.63	0.81	0.73	0.66	0.82	0.75
		10-10	0.68	0.83	0.77	0.74	0.86	0.82	0.79	0.86	0.83
		15-5	0.79	0.87	0.84	0.82	0.89	0.86	0.82	0.89	0.86
	≥0.75	5-5	0.79	0.87	0.84	0.82	0.89	0.86	0.82	0.89	0.86
		7-3	0.84	0.91	0.87	0.94	0.96	0.95	0.94	0.97	0.95
		10-10	0.89	0.94	0.92	0.96	0.98	0.97	0.97	0.98	0.97
		15-5	0.93	0.96	0.95	0.97	0.98	0.98	0.97	0.98	0.98
0.50	0.30-0.50	5-5	0.67	0.81	0.75	0.71	0.83	0.78	0.71	0.83	0.79
		7-3	0.65	0.82	0.76	0.69	0.84	0.79	0.70	0.84	0.79
		10-10	0.81	0.90	0.86	0.83	0.90	0.87	0.84	0.91	0.87
		15-5	0.83	0.91	0.87	0.84	0.91	0.88	0.84	0.91	0.88
	≥0.75	5-5	0.89	0.94	0.92	0.94	0.97	0.96	0.95	0.97	0.96
		7-3	0.91	0.95	0.93	0.95	0.97	0.96	0.95	0.97	0.96
		10-10	0.94	0.97	0.96	0.97	0.98	0.98	0.97	0.98	0.98
		15-5	0.95	0.98	0.97	0.97	0.99	0.98	0.97	0.99	0.98

*B.kor: Boyutlar arası korelasyon; M.kor: Maddeler arası korelasyon; Min: 1000 tekrarda minimum kestirilen değer; Maks: 1000 tekrarda maksimum kestirilen değer; Ort:1000 tekrarin ortalamasi

Tablo 6. Basit Ölçme Modelinde N=400 için Kestirilen Güvenirlik Değerleri

B.kor.	M.kor.	Madde Sayısı	Alfa			Tabakalı Alfa			Omega		
			Min.	Maks.	Ort.	Min.	Maks.	Ort.	Min.	Maks.	Ort.
0.00	0.30-0.50	5-5	0.55	0.70	0.63	0.64	0.76	0.70	0.67	0.76	0.72
		7-3	0.59	0.73	0.67	0.67	0.78	0.73	0.70	0.78	0.75
		10-10	0.72	0.82	0.78	0.78	0.85	0.82	0.79	0.85	0.83
		15-5	0.79	0.86	0.84	0.82	0.88	0.86	0.83	0.88	0.86
	≥0.75	5-5	0.79	0.86	0.84	0.82	0.88	0.86	0.83	0.88	0.86
		7-3	0.84	0.90	0.88	0.94	0.96	0.95	0.94	0.96	0.95
		10-10	0.90	0.93	0.92	0.96	0.98	0.97	0.97	0.98	0.97
		15-5	0.94	0.96	0.95	0.97	0.98	0.98	0.97	0.98	0.98
0.50	0.30-0.50	5-5	0.69	0.81	0.75	0.72	0.83	0.78	0.72	0.83	0.79
		7-3	0.71	0.82	0.77	0.74	0.83	0.79	0.75	0.84	0.79
		10-10	0.82	0.88	0.86	0.84	0.90	0.87	0.85	0.90	0.87
		15-5	0.86	0.89	0.88	0.86	0.90	0.89	0.86	0.90	0.88
	≥0.75	5-5	0.90	0.94	0.92	0.95	0.97	0.96	0.95	0.97	0.96
		7-3	0.91	0.94	0.93	0.95	0.97	0.96	0.95	0.97	0.96
		10-10	0.95	0.97	0.96	0.97	0.98	0.98	0.97	0.98	0.98
		15-5	0.96	0.97	0.97	0.98	0.98	0.98	0.98	0.98	0.98

*B.kor: Boyutlar arası korelasyon; M.kor: Maddeler arası korelasyon; Min: 1000 tekrarda minimum kestirilen değer; Maks: 1000 tekrarda maksimum kestirilen değer; Ort:1000 tekrarin ortalamasi

Tablo 6, N=400 için basit ölçme modelinde kestirilen güvenilirlik değerlerini özetlemektedir. Tablo 4 ve Tablo 5 ile karşılaştırıldığında alfa, tabakalı alfa ve omega katsayılarının ortalamasının örneklem büyüklüğünden etkilenmediği görülmektedir. Tablo 5’te gözlemlendiği gibi örneklem büyüklüğünün artmasıyla kestirilen minimum güvenilirlik değerleri daha da yükselmiş, alfa ile omega arasında ortaya çıkan maksimum fark ise daha da düşmüştür. Alfa katsayısının 0,55 kestirildiği veri setinde tabakalı alfa 0,64 ve omega 0,67 olarak elde edilmiştir. Dolayısıyla 100 örneklem büyüklüğü için belirlenen alfa ile omega arasındaki maksimum fark 0,17’ye kadar düşmüştür.

Tablo 7. Basit Ölçme Modelinde Örneklemelere Göre Yanlılık Değerleri

B.Kor.	M.Kor.	Madde Sayısı	100		200		400	
			$\omega_g - \alpha$	$\omega_g - \alpha_s$	$\omega_g - \alpha$	$\omega_g - \alpha_s$	$\omega_g - \alpha$	$\omega_g - \alpha_s$
0.00	0.30-0.50	5-5	0.110	0.032	0.099	0.022	0.090	0.015
		7-3	0.089	0.029	0.079	0.020	0.073	0.014
		10-10	0.062	0.018	0.057	0.013	0.052	0.009
		15-5	0.028	0.005	0.024	0.002	0.022	-0.001
	≥ 0.75	5-5	0.110	0.005	0.108	0.003	0.108	0.003
		7-3	0.080	0.004	0.079	0.003	0.078	0.002
		10-10	0.054	0.003	0.053	0.002	0.053	0.001
		15-5	0.031	0.001	0.030	0.001	0.030	0.000
0.50	0.30-0.50	5-5	0.038	0.011	0.034	0.007	0.032	0.005
		7-3	0.035	0.012	0.030	0.008	0.028	0.005
		10-10	0.019	0.003	0.017	0.002	0.017	0.002
		15-5	0.011	0.002	0.010	0.001	0.009	0.000
	≥ 0.75	5-5	0.043	0.001	0.042	0.001	0.042	0.000
		7-3	0.033	0.001	0.033	0.001	0.033	0.001
		10-10	0.020	0.000	0.020	0.000	0.020	0.000
		15-5	0.014	0.000	0.013	0.000	0.013	0.000

*B.kor: Boyutlar arası korelasyon; M.kor: Maddeler arası korelasyon; Min:1000 tekrarda minimum kestirilen değer; Maks: 1000 tekrarda maksimum kestirilen değer; Ort:1000 tekrarın ortalaması

Tablo 7’de, basit yapıli modeller için alfa ve tabakalı alfa katsayılarının omega katsayısından ortalama ne kadar uzak kestirim yaptığını gösteren yanlılık değerleri özetlenmiştir. Alfa ile omega katsayılarının dramatik olarak birbirinden farklı kestirim yaptığı koşul, ilişkisiz çok boyutlu ölçme modelinde test uzunluğunun en kısa (10 madde) olduğu koşuldur. Bu fark maddeler arası korelasyon 0,30-0,50 arasında iken boyutlardaki madde sayıları farklılaştığında azalmaktadır. Benzer durum örneklem büyüklüğü arttığında da gözlenmektedir. Tabakalı alfa ile omega ise neredeyse tüm koşullarda benzer performans göstermiştir. Alfa katsayısı, ilişkişel çok boyutlu yapılarda daha yansız sonuçlar vermiştir.

Karmaşık Ölçme Modeline İlişkin Sonuçlar

Karmaşık ölçme modeline ilişkin sonuçlar Tablo 8-Tablo 10 arasında özetlenmiştir. Omega katsayısının konjenerik ölçmelerde gerçek güvenilirliğin daha yansız bir kestiricisi olduğunu belirten çalışmalar (Dunn vd., 2014; Graham, 2006; Hayes ve Coutts, 2020; Revelle ve Zinbarg, 2009; Zinbarg vd., 2005, 2006, 2007) referans alınarak omega katsayısına göre alfa ve tabakalı alfa katsayılarının ortalama yanlılığı Tablo 11’de özetlenmiştir.

Tablo 8. Karmaşık Ölçme Modelinde $N=100$ için Kestirilen Güvenirlik Değerleri

B.kor.	M.kor.	Madde Sayısı	Alfa			Tabakalı Alfa			Omega		
			Min.	Maks.	Ort.	Min.	Maks.	Ort.	Min.	Maks.	Ort.
0.00	0.30-0.50	5-5	0.46	0.80	0.69	0.55	0.83	0.73	0.62	0.84	0.75
		7-3	0.53	0.81	0.72	0.59	0.83	0.75	0.64	0.85	0.77
		10-10	0.68	0.85	0.79	0.73	0.88	0.82	0.75	0.89	0.84
		15-5	0.77	0.89	0.85	0.79	0.91	0.87	0.80	0.91	0.87
	≥ 0.75	5-5	0.77	0.89	0.85	0.79	0.91	0.87	0.80	0.91	0.87
		7-3	0.83	0.93	0.90	0.92	0.96	0.94	0.93	0.97	0.95
		10-10	0.88	0.96	0.93	0.95	0.98	0.97	0.96	0.98	0.97
		15-5	0.93	0.97	0.95	0.96	0.98	0.97	0.97	0.98	0.98
0.50	0.30-0.50	5-5	0.63	0.86	0.78	0.67	0.87	0.80	0.69	0.88	0.82
		7-3	0.65	0.88	0.80	0.67	0.88	0.81	0.69	0.89	0.83
		10-10	0.81	0.91	0.87	0.83	0.92	0.88	0.84	0.92	0.88
		15-5	0.81	0.92	0.88	0.82	0.92	0.89	0.82	0.93	0.89
	≥ 0.75	5-5	0.89	0.96	0.93	0.94	0.98	0.96	0.94	0.98	0.97
		7-3	0.90	0.96	0.94	0.93	0.97	0.96	0.94	0.98	0.97
		10-10	0.94	0.98	0.96	0.97	0.99	0.98	0.97	0.99	0.98
		15-5	0.96	0.98	0.97	0.97	0.99	0.98	0.97	0.99	0.98

*B.kor: Boyutlar arası korelasyon; M.kor: Maddeler arası korelasyon; Min:1000 tekrarda minimum kestirilen değer; Maks: 1000 tekrarda maksimum kestirilen değer; Ort:1000 tekrarın ortalaması

Tablo 8 incelendiğinde, basit yapıli verilerde olduđu gibi karmaşık yapıli verilerde de alfa katsayısı hemen hemen tüm koşullarda omega ve tabakalı alfa katsayılarından daha düşük güvenirlik değeri kestirmiştir. Ancak basit yapıli verilere kıyasla ortalama güvenirlik değeri daha yüksek olduđu dikkat çekmektedir. Bunun nedenlerinden biri iki madde her iki boyutta 0,30'dan yüksek yük verecek şekilde ilişkilendirildiğinden doğal olarak bazı replikasyonlarda boyutlar arası korelasyonlar 0,10 düzeyine kadar yükselmiştir. Ayrıca maddelerin farklı boyutlarda yük vermesi maddeler arası korelasyonları da basit modele kıyasla bir miktar yükselteceği (0,50 ye daha yakın) için alfa katsayısının da bir miktar yükselmesi beklenir. Karmaşık ilişkisiz modelde neredeyse tüm koşullarda üç katsayı da minimum 0,70 düzeyinde güvenirlik değeri üretebilmektedir. 1000 tekrarın tamamı değerlendirildiğinde alfa katsayısının 0,46 gibi oldukça düşük güvenirlik değeri kestirebildiği dikkat çekmektedir. Alfa ile güvenirliği 0,46 olarak kestirilen veri setinde tabakalı alfa 0,55 ve omega 0,62'dir. Dolayısıyla, basit yapıli verideki eş koşullara kıyasla daha düşük olmakla birlikte bazı koşullarda alfa ile omega arasındaki fark 0,22'ye, tabakalı alfa ile omega arasındaki fark ise 0,11'e kadar çıkmıştır. Boyutlar arası korelasyonun artması ile tüm koşullarda güvenirliğin daha yüksek kestirildiği gözlenmektedir. Karmaşık yapıli hem ilişkili hem de ilişkisiz modelde maddeler arası korelasyonun $\geq 0,75$ olduđu koşullarda, verilerin tamamında üç katsayı da 0,70'in üzerine kesitim yapmıştır. Dolayısıyla maddeler arası korelasyonun boyutlar arası korelasyondan daha fazla güvenirlik katsayıları üzerinde etkisi olduđu karmaşık model de gözlenmiştir. Bu beklenen bir durumdur. Tablo 4'ün altındaki paragrafta açıklaması yapılmıştır. Yine basit yapıli modele ilişkin sonuçlarda da gözlendiği üzere eşit uzunlukta olmayan boyutlardaki testlerden edilen güvenirlik farklı uzunlukta boyutlardan elde edilen güvenirlikten daha düşüktür.

Tablo 9. Karmaşık Ölçme Modelinde N=200 için Kestirilen Güvenirlik Değerleri

B.kor.	M.kor.	Madde Sayısı	Alfa			Tabakalı Alfa			Omega		
			Min.	Maks.	Ort.	Min.	Maks.	Ort.	Min.	Maks.	Ort.
0.00	0.30-0.50	5-5	0.59	0.78	0.69	0.65	0.82	0.74	0.62	0.83	0.75
		7-3	0.60	0.80	0.72	0.65	0.83	0.75	0.66	0.83	0.76
		10-10	0.72	0.84	0.79	0.78	0.86	0.83	0.78	0.87	0.84
		15-5	0.81	0.89	0.85	0.83	0.90	0.87	0.83	0.90	0.87
	≥0.75	5-5	0.81	0.89	0.85	0.83	0.90	0.87	0.83	0.90	0.87
		7-3	0.87	0.93	0.90	0.93	0.96	0.94	0.94	0.97	0.95
		10-10	0.89	0.95	0.93	0.96	0.98	0.97	0.96	0.98	0.97
		15-5	0.93	0.96	0.95	0.96	0.98	0.97	0.97	0.98	0.98
0.50	0.30-0.50	5-5	0.68	0.85	0.78	0.71	0.86	0.80	0.73	0.87	0.81
		7-3	0.74	0.86	0.81	0.75	0.87	0.82	0.76	0.88	0.83
		10-10	0.82	0.90	0.87	0.84	0.91	0.88	0.84	0.91	0.88
		15-5	0.84	0.91	0.88	0.85	0.91	0.89	0.85	0.92	0.89
	≥0.75	5-5	0.91	0.95	0.93	0.95	0.97	0.96	0.95	0.97	0.97
		7-3	0.92	0.96	0.94	0.95	0.97	0.96	0.95	0.98	0.97
		10-10	0.95	0.97	0.96	0.97	0.98	0.98	0.97	0.99	0.98
		15-5	0.96	0.98	0.97	0.98	0.99	0.98	0.98	0.99	0.98

*B.kor: Boyutlar arası korelasyon; M.kor: Maddeler arası korelasyon; Min:1000 tekrarda minimum kestirilen değer; Maks: 1000 tekrarda maksimum kestirilen değer; Ort:1000 tekrarin ortalaması

Tablo 10. Karmaşık Ölçme Modelinde N=400 için Kestirilen Güvenirlik Değerleri

B.kor.	M.kor.	Madde Sayısı	Alfa			Tabakalı Alfa			Omega		
			Min.	Maks.	Ort.	Min.	Maks.	Ort.	Min.	Maks.	Ort.
0.00	0.30-0.50	5-5	0.62	0.75	0.69	0.69	0.79	0.74	0.70	0.80	0.75
		7-3	0.62	0.79	0.72	0.67	0.81	0.76	0.68	0.82	0.77
		10-10	0.73	0.83	0.79	0.78	0.86	0.83	0.79	0.87	0.84
		15-5	0.81	0.88	0.85	0.83	0.89	0.87	0.83	0.89	0.87
	≥0.75	5-5	0.81	0.88	0.85	0.83	0.89	0.87	0.83	0.89	0.87
		7-3	0.88	0.92	0.90	0.93	0.95	0.94	0.94	0.96	0.95
		10-10	0.91	0.94	0.93	0.96	0.97	0.97	0.97	0.98	0.97
		15-5	0.94	0.96	0.95	0.97	0.98	0.97	0.97	0.98	0.98
0.50	0.30-0.50	5-5	0.73	0.83	0.78	0.76	0.84	0.80	0.77	0.85	0.81
		7-3	0.75	0.84	0.81	0.77	0.85	0.82	0.78	0.86	0.82
		10-10	0.83	0.89	0.87	0.85	0.90	0.88	0.85	0.90	0.88
		15-5	0.86	0.90	0.88	0.87	0.91	0.89	0.87	0.91	0.89
	≥0.75	5-5	0.91	0.95	0.93	0.95	0.97	0.96	0.96	0.97	0.97
		7-3	0.93	0.95	0.94	0.95	0.97	0.96	0.96	0.97	0.97
		10-10	0.95	0.97	0.96	0.97	0.98	0.98	0.98	0.98	0.98
		15-5	0.96	0.98	0.97	0.98	0.98	0.98	0.98	0.99	0.98

*B.kor: Boyutlar arası korelasyon; M.kor: Maddeler arası korelasyon; Min:1000 tekrarda minimum kestirilen değer; Maks: 1000 tekrarda maksimum kestirilen değer; Ort:1000 tekrarin ortalaması

Tablo 9, N=200 için karmaşık ölçme modelinde kestirilen güvenirlik değerlerini özetlemektedir. Tablo 8 ile karşılaştırıldığında alfa, tabakalı alfa ve omega katsayılarının ortalamasının örneklem büyüklüğünden etkilenmediği görülmektedir. Ancak veri setlerinden kestirilen minimum güvenirlik

değerleri yükselmiştir. Hatta alfa katsayısı 20 maddelik testlerin tamamında 0,70'ün üzerinde değerler vermiştir. Aynı zamanda Tablo 8'e göre alfa ile omega arasında ortaya çıkan maksimum farkın da azaldığı gözlenmiştir. Örneğin ilişkisiz modelde maddeler arası korelasyonun 0,30-0,50 olduğu koşulda alfa ile omega arasındaki maksimum fark 0,11'e; tabakalı alfa ile omega arasındaki fark ise 0,03'e kadar düşmüştür.

Tablo 10, N=400 için karmaşık ölçme modelinde kestirilen güvenirlilik değerlerini özetlemektedir. Tablo 8 ve Tablo 9 ile karşılaştırıldığında alfa, tabakalı alfa ve omega katsayılarının ortalamasının örneklem büyüklüğünden etkilenmediği görülmektedir. Neredeyse tüm koşullarda üç katsayı da 0,70 üzerinde güvenirlilik değerleri kestirmiştir. Tablo 9'a kıyasla alfa ile omega arasında ortaya çıkan maksimum farkın da azaldığı gözlenmiştir. Örneğin ilişkisiz modelde maddeler arası korelasyonun 0,30-0,50 olduğu koşulda alfa ile omega arasındaki maksimum fark 0,08'e; tabakalı alfa ile omega arasındaki fark ise 0,01'e kadar düşmüştür.

Tablo 11. Karmaşık Ölçme Modelinde Örneklemelere Göre Yanlılık Değerleri

B.Kor.	M.Kor.	Madde Sayısı	100		200		400	
			$\omega_g - \alpha$	$\omega_g - \alpha_s$	$\omega_g - \alpha$	$\omega_g - \alpha_s$	$\omega_g - \alpha$	$\omega_g - \alpha_s$
0.00	0.30-0.50	5-5	0.084	0.015	0.073	0.010	0.057	0.009
		7-3	0.051	0.014	0.047	0.010	0.045	0.008
		10-10	0.051	0.012	0.045	0.007	0.043	0.006
		15-5	0.017	0.001	0.016	-0.001	0.015	-0.002
	≥0.75	5-5	0.074	0.012	0.069	0.012	0.058	0.011
		7-3	0.055	0.010	0.055	0.009	0.055	0.009
		10-10	0.046	0.004	0.045	0.003	0.044	0.003
		15-5	0.025	0.002	0.025	0.002	0.025	0.002
0.50	0.30-0.50	5-5	0.036	0.014	0.032	0.011	0.030	0.009
		7-3	0.024	0.014	0.020	0.010	0.017	0.008
		10-10	0.018	0.004	0.017	0.003	0.016	0.003
		15-5	0.009	0.002	0.008	0.001	0.007	0.000
	≥0.75	5-5	0.034	0.006	0.033	0.006	0.033	0.005
		7-3	0.024	0.005	0.024	0.005	0.024	0.005
		10-10	0.018	0.002	0.018	0.001	0.018	0.001
		15-5	0.011	0.001	0.011	0.001	0.011	0.001

*B.kor: Boyutlar arası korelasyon; M.kor: Maddeler arası korelasyon; Min: 1000 tekrarda minimum kestirilen değer; Maks: 1000 tekrarda maksimum kestirilen değer; Ort:1000 tekrarin ortalaması

Tablo 11'de, karmaşık yapıli modeller için alfa ve tabakalı alfa katsayılarının omega katsayısından ortalama ne kadar uzak kestirim yaptığını gösteren yanlılık değerleri özetlenmiştir. Basit yapıli verilere kıyasla daha düşük olmakla birlikte alfa ile omega arasındaki farkın en yüksek olduğu koşul, ilişkisiz çok boyutlu yapı ve test uzunluğunun en kısa (10 madde) olduğu koşuldur. Bu fark maddeler arası korelasyon 0,30-0,50 arasında iken boyutlardaki madde sayıları farklılaştığında azalmaktadır. Benzer durum örneklem büyüklüğü arttığında da gözlenmektedir. Tabakalı alfa ile omega ise neredeyse tüm koşullarda benzer sonuçlar göstermiştir. Alfa katsayısı, ilişkisiz çok boyutlu yapılarda daha yansız sonuçlar vermiştir.

TARTIŞMA, SONUÇ ve ÖNERİLER

Geliştirildiği dönemden günümüze varsayımları sağlanmadığında gerçek güvenirligi yanlış kestirdiğini belirten araştırmalara rağmen hala birçok bilimsel çalışmada alfa katsayısının yanlış kullanımının devam ettiği belirtilmektedir (örn., Cho ve Kim, 2015; Cortina, 1993; Schmitt, 1996; Sijtsma, 2009; Sijtsma ve Pfadt, 2021). Alanyazında özellikle çok boyutlu ölçme araçlarından elde edilen sonuçların güvenirliginde omega, tabakalı alfa, maksimal güvenirlik gibi daha iyi alternatiflerinden ziyade alfa katsayısının raporlandığı çalışmalar ile sıklıkla karşılaşmaktadır. Bu durum, araştırmacıların çeşitli yayınlarda veya çalışmalarda omega, tabakalı alfa gibi iç tutarlılığı test eden farklı güvenirlik katsayılarına kıyasla alfa katsayısı ile çok daha fazla karşılaşmalarından ve SPSS gibi paket programlarda alfa katsayısının kolaylıkla hesaplanabilmesinden kaynaklanabilir. Cho ve Kim'in (2015, s. 225) ifade ettiği gibi "alfayı daha üstün bir alternatifle değiştirmek yalnızca kişisel bir tercih meselesi değil, aynı zamanda konuya bilinçli bir şekilde yanıt veren bir akademi meselesi" olarak görülmelidir. Bu nedenle ilgili alanyazındaki çalışmalara ek olarak mevcut araştırma ile pratikte daha fazla karşılaşılan çok boyutlu yapılarda alfa yerine daha uygun güvenirlik katsayılarının kullanımı konusunda farkındalık oluşmasına katkı sağlamak amaçlanmıştır. Bu amaca bağlı olarak çok boyutlu ölçme sonuçları üzerinde alfa, tabakalı alfa ve omega katsayılarının ölçme modeli, test uzunluğu, boyutlara düşen madde sayısı, maddeler arası korelasyon ve faktörler arası korelasyondan nasıl etkilendiği incelenmiştir. Ayrıca araştırmacılara tabakalı alfa ve omega katsayılarının manuel olarak kolaylıkla hesaplanabildiğini göstermek amacıyla örnekler sunulmuştur. Bu kolaylığın anlaşılması araştırmacıları, alfa yerine ya da alfa ile birlikte bu katsayıları kullanmaya daha fazla teşvik edebilir.

Bu araştırmada alfa, tabakalı alfa ve omega katsayılarının çok boyutlu test yapılarındaki doğasını, performansını daha iyi anlayabilmek için simülasyon verisi kullanılmıştır. Araştırmada incelenen üç katsayı için güvenirlik kestirimleri üzerinde en fazla etkisi olan değişkenin maddeler arası korelasyonun olduğu gözlenmiştir. Üstelik boyutlara arası korelasyonun etkisi test uzunluğu kıaldıkça daha fazla ortaya çıkmaktadır. Cortina'nın (1993) çalışmasında basit ilişkisiz çok boyutlu testler üzerinde alfa katsayısı için benzer sonuçlar raporlanmıştır. Alfa katsayısı için daha belirgin olmakla birlikte boyutlar arası korelasyon ile güvenirlik kestirimleri arasında pozitif yönlü bir ilişki bulunmuştur. Üstelik yanlılık değerlerine (omega katsayısına uzaklık) göre güvenirlik kestirimleri üzerinde boyutlar arası korelasyon daha etkili bir değişkendir.

Test uzunluğu ve testin basit ya da karmaşık yapıda olması da güvenirlik üzerinde etkilidir. Test uzunluğu arttıkça güvenirlik kestirimlerinin yükseldiği gözlenmiştir. Ancak test uzunluğunun güvenirlik üzerindeki etkisini tek başına yorumlamak yanıltıcı olabilir; test uzunluğunun etkisi özellikle maddeler arası korelasyonlara ve boyutlara arası ilişkilere dikkat ederek yorumlanmalıdır (Cortina, 1993; Wells ve Wollack, 2003). Çünkü maddeler arası korelasyonun daha yüksek olduğu diğerine göre daha kısa bir test (örn. 5-5 madde), maddeler arası korelasyonun daha düşük olduğu daha uzun bir teste (örn. 10-10 madde) göre daha güvenilir olabilir. Benzer durum boyutlara arası korelasyon için de geçerlidir. Mevcut araştırmanın dikkat çekici bir diğer sonucu ilişkisiz modellerde boyutlara düşen madde sayısının da güvenirlik üzerinde etkili olmasıdır. Yine maddeler arası korelasyonun düşük olduğu koşullarda boyutların farklı uzunlukta olması

güvenirliğin yükselmesine neden olmuştur. Bu durum alfa katsayısı için daha belirgin şekilde gözlenmektedir. Boyutlar arası ilişkilerin varlığı bu etkiyi azaltmaktadır.

Örneklem büyüklüğü güvenilirlik kestirimlerinin ortalamasını etkilememiştir, yalnızca daha homojen kestirimler sağlamıştır. Minimum ve maksimum kestirim değerleri üzerinde gözlenebileceği gibi Ek Tablo 2 ve Ek Tablo 3'te 1000 tekrar için standart sapma değerleri incelendiğinde örneklem büyüklüğü ile kestirimlerin standart sapması (hatası) arasında negatif ilişki ortaya çıkmaktadır. Bu sonuç Kamata ve diğerlerinin (2003) alfa, tabakalı alfa, maksimal güvenilirlik ve omega katsayıları için elde ettiği sonuçlar ile oldukça tutarlıdır.

Bu araştırmanın bulgularından çıkarılan en önemli sonuç hem basit hem de karmaşık çok boyutlu testlerde boyutlar arasında orta düzeyde ilişki varsa ve maddeler arası korelasyon da yüksek ise omega ve tabakalı alfa katsayısı benzer sonuçlar vermektedir; alfa katsayısı da bu iki katsayıya yakınsayan değerler kestirmektedir. Ancak pratikte maddeler arası korelasyonların bu kadar yüksek olduğu veri setleriyle karşılaşmak düşük olasılıklıdır. Genellikle orta ve daha düşük maddeler arası korelasyonlarla karşılaşmaktadır. Bu durumda bu çalışmanın düşük maddeler arası korelasyon koşulları dikkate alındığında hem ilişkili hem de ilişkisiz hem basit hem de karmaşık modellerde omega ve tabakalı alfa katsayıları 20 maddelik testlerde en az 100 kişilik örneklemelerden makul düzeyde ($>0,70$) güvenilirlik değerleri elde edilebilirken 200 ve üzeri örneklemelerden daha kısa testler için makul düzeyde güvenilirlik kestirimleri yapılabilmektedir.

Alfa katsayısı boyutların ilişkisiz olduğu koşullarda en düşük değerleri vermesinin yanı sıra bu modellerden elde edilen sonuçlar bazı araştırmacıların (örn., Cortina, 1993; Green vd., 1977; Revelle, 1979; Schmitt, 1996; Zinbarg vd., 2006) “alfa katsayısının tek başına tek boyutluluğun bir ölçüsü olamayacağı” iddiasını destekler niteliktedir. Çünkü ilişkisiz çok boyutluluğa ve maddeler arası düşük korelasyona rağmen 0,83 veya 0,90 gibi alfa katsayısı ile yüksek düzeyde güvenilirlik değerleri elde edilebilmektedir. Mevcut araştırmada alfa katsayısı ile özellikle ilişkisiz modellerde makul düzeyde güvenilirlik değerleri elde edebilmek için Cortina'nın (1993) belirttiği gibi testin minimum 20 maddeden oluşması gerekir. Boyutlar arası ilişkiler arttığında ve model basitten karmaşığa doğru ilerlediğinde daha kısa testlerden de alfa ile makul düzeyde güvenilir sonuçlar elde etmek mümkündür. Ancak bu bulgular dikkatli yorumlanmalıdır. Bazı araştırmacılar (örn., Cortina, 1993; Hair vd., 2010; Nunnally, 1978; Kaplan ve Saccuzzo, 2009) alfa katsayısı için kabul edilebilir çeşitli kesme noktaları tanımlamıştır. Bunlar arasından en yaygın kabul edilen ve kullanılan sınır 0,70'tir. Güvenirlik için bu değeri sınır kabul eden ve mevcut çalışmayı inceleyen araştırmacılar arasından ortalama güvenilirlik kestirimlerine göre alfa katsayısı nasılsa 0,70'in üzerinde sonuçlar vermektedir diye düşünenler ve dolayısıyla alfa katsayısının varsayımları dışında kullanılmasında büyük sakınca görmeyenler olabilir. Ancak unutulmamalıdır ki bu sonuçlar 1000 veri setinin ortalamasıdır ve oldukça steril ve kontrollü üretilmiş bu 1000 farklı veri seti içerisinde dahi 0,40'ın altında alfa katsayısı hesaplanmıştır. Hatta koşulların fazlalığı nedeniyle makalede yer verilmemiştir ancak 50 örneklem büyüklüğünde omega katsayısının 0,60 olarak kestirildiği durumlarda alfa katsayısının 0,18 olarak kestirildiği veri setleri ile de karşılaşılmıştır. Bu örnekler, gerçek hayatta buna benzer sonuçlar verebilecek veri setleriyle karşılaşılabileceğimizin bir kanıtıdır

ve aslında güvenilir olan bir ölçme sonucunun alfa katsayısı ile güvenilir olmadığı yönünde yanlış karar verilmesine neden olabileceği ihtimalini de ortaya koymaktadır.

Bir ölçme aracından elde edilen ölçümlerin makul düzeyde güvenilir olması gerekir ancak araştırmacıların tek odak noktası bu olmamalıdır. Alfa katsayısının eşdeğer olmayan ölçmelerde gerçek güvenilirliğin alt sınırını verdiği (Osburn, 2000) bilgisine dayanarak varsayımlarının karşılanmadığı durumlarda bu katsayı, yalnızca 0,70 sınır değerinin üzerinde kestirim yaptığı için raporlanmamalıdır. Burada odaklanılması gereken en önemli nokta ölçümlerin güvenilirliğinin gerçek değerinin veya gerçeğe en yakın değerinin raporlanması olmalıdır. Kamata ve diğerlerinin (2003) simülatif veriler üzerinde yaptıkları çalışmada hem basit hem de karmaşık modellerde gerçek güvenilirliği en yansız kestiren katsayılar tabakalı alfa ve omega, en uzak kestiren katsayı ise alfa olarak bulunmuştur. Yine aynı çalışmada tabakalı alfa ve omega katsayılarının performansı benzer bulunmakla birlikte 500 ve üzeri örneklemelerde tabakalı alfa katsayısının omega katsayısından daha yansız sonuçlar verdiği raporlanmıştır. Gerçek veri ile yapılan araştırmalarda da benzer şekilde konjenerik ölçme yapılarında omega katsayısının daha kesin ve doğru kestirimler yapabildiği belirtilmiştir (Brunner ve Süß, 2005; Dang ve Chan, 2017). Mevcut araştırma, alfa katsayısının ancak maddeler arası korelasyonlar $>0,75$ ve boyutlar arası minimum orta düzeyde ilişkili bulunan çok boyutlu modellerde omega ve tabakalı alfaya en fazla yakınsayabildiği ortaya çıkarmıştır. Ancak daha önce belirtildiği gibi bu koşulların gerçek verilerde ortaya çıkması oldukça düşük olasılıklıdır. Bu nedenle hem alanyazına hem de mevcut çalışmanın sonuçlarına dayalı olarak araştırmacılara, **çok boyutlu konjenerik ölçümlerde iç tutarlılık anlamında güvenirlilik kanıtı olarak alfa katsayısını kullanmamaları ya da tek başına kullanmamaları; bunun yerine omega, tabakalı alfa gibi daha uygun katsayıları kullanıp raporlamaları** önerilir.

Gelecek araştırmalar için mevcut araştırmada ele alınan koşulların benzer ya da farklı düzeylerinin bu çalışmada ele alınmayan iç tutarlılık katsayıları veya güvenilirliğin farklı anlamlarını taşıyan farklı güvenirlilik katsayıları üzerindeki etkisinin incelenmesi önerilebilir. Mevcut araştırmada karmaşık modelleri temsilen her boyuttan yalnızca bir madde diğer boyutta yük verecek şekilde veriler üretilmişti. Gelecek çalışmalar için araştırmacılar daha karmaşık ölçme modelleri üzerinde güvenirlilik katsayılarının performansını inceleyebilir. Gerçek veri matrislerine dayalı olarak da benzer çalışmalar yapılabilir.

KAYNAKÇA

Bacon, D. R., Sauer, P. L., & Young, M. (1995). Composite Reliability in Structural Equations Modeling. *Educational and Psychological Measurement*, 55(3), 394-406.

Brunner, M., & Süß, H. M. (2005). Analyzing the reliability of multidimensional measures: An example from intelligence research. *Educational and Psychological Measurement*, 65(2), 227-240.

Carmines, E. G., & Zeller, R. A. (1982). Reliability and validity assessment. Beverly Hills: Sage Publications.

- Cho, E., & Kim, S. (2015). Cronbach's coefficient alpha: Well-known but poorly understood. *Organizational Research Methods, 18*, 207-230.
- Cho, E. (2016). Making reliability reliable: A systematic approach to reliability coefficients. *Organizational Research Methods, 19*(4), 651-682.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart & Winston.
- Cronbach, L. J. (1947). Test "reliability": Its meaning and determination. *Psychometrika, 12*(1), 1-16.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.
- Cronbach, L. J., Schonemann, P., & McKie, D. (1965). Alpha coefficients for stratified-parallel tests. *Educational and Psychological Measurement, 25*, 291-312.
- Cortina, J. M (1993). What is coefficient alpha? An examination of theory and application. *Journal of Applied Psychology, 78*, 98-104.
- Deng, L., & Chan, W. (2017). Testing the difference between reliability coefficients alpha and omega. *Educational and psychological measurement, 77*(2), 185-203.
- Dunn, T. J., Baguley, T., & Brunsdon, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British journal of psychology, 105*(3), 399-412.
- Feldt, L. S. , Brennan, R. L. (1989) Reliability. In Linn RL (ed), *Educational measurement* (3rd ed., pp. 105-146), Macmillan.
- Graham, J. M. (2006). Congeneric and (Essentially) Tau-Equivalent Estimates of Score Reliability. *Educational and Psychological Measurement, 66*(6), 930-944.
- Green, S. B., Lissitz, R.W., and Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement, 37*, 827-838.
- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika, 10*, 255-282.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis*. Pearson College Division.
- Harwell, M., Stone, C. A., Hsu, T. C., & Kirisci, L. (1996). Monte Carlo studies in item response theory. *Applied Psychological Measurement, 20*(2), 101-125.
- Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But.... *Communication Methods and Measures, 14*(1), 1-24.
- Hoyt, C. (1941). Test reliability estimated by analysis of variance. *Psychometrika, 6*(3), 153-160
- Kamata, A., Turhan, A., & Darandari, E. (2003, April). *Estimating reliability for multidimensional composite scale scores*. In annual meeting of American Educational Research Association, Chicago, IL.

- Kaplan, R. M., & Saccuzzo, D. P. (2009). *Psychological testing: Principles, applications, and issues* (7. basım). CA: Thompson Wadsworth.
- McDonald, R. P. (1999). *Test theory: Unified treatment*. Lawrence Erlbaum Associates.
- McDonald, R. P. (1970). Theoretical foundations of principal factor analysis and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology*, 23, 1-21
- Nunnally, J. C. (1978). *Psychometric theory* (2. Basım). New York : McGraw-Hill.
- Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods*, 5, 343-355.
- Paek, I., and Cole, I. (2019). *Using R for Item Response Theory Applications*. New York: Routledge.
- Raykov, T. (1998). Coefficient alpha and composite reliability with interrelated nonhomogeneous items. *Applied Psychological Measurement*, 22, 375-385.
- Raykov, T., West, B. T., & Traynor, A. (2015). Evaluation of coefficient alpha for multiple-component measuring instruments in complex sample designs. *Structural Equation Modeling: A Multidisciplinary Journal*, 22(3), 429-438.
- Revelle, W. (1979). Hierarchical cluster analysis and the internal structure of tests. *Multivariate Behavioral Research*, 14, 57-74.
- Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, Omega, and the glb: Comments on Sijtsma. *Psychometrika*, 74, 145-154.
- Revelle, W., & Revelle, M. W. (2022). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Version R package version 2.2.9. <https://cran.r-project.org/package=psych>
- Robitzsch, A. (2022). *sirt: Supplementary Item Response Theory Models*. Version R package version 3.12-66. <https://cran.r-project.org/package=sirt>
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, 8, 350-353.
- Serbetar, I., & Sedlar, I. (2016). Assessing reliability of a multi-dimensional scale by coefficient alpha. *Revija za Elementarno Izobraževanje*, 9(1/2), 189.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74, 107-120.
- Sijtsma, K., & Pfadt, J. M. (2021). Part II: On the use, the misuse, and the very limited usefulness of Cronbach's alpha: Discussing lower bounds and correlated errors. *Psychometrika*, 86(4), 843-860.
- Sočan, G. (2000). Assessment of reliability when test items are not essentially tau-equivalent. *Advances in Methodology and Statistics*, 15, 23-35.
- Tang, W., & Cui, Y. (2012, April). *A simulation study for comparing three lower bounds to reliability*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, Canada

Traub, E. R. (1994). *Reliability for the social sciences: Theory and Applications. Measurement methods for the social sciences*. London: Sage Publications.

Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). A journey around alpha and Omega to estimate internal consistency reliability. *Anales de Psicología/Annals of Psychology*, 33(3), 755-782.

Warrens, M. J. (2016). A comparison of reliability coefficients for psychometric tests that consist of two parts. *Advances in Data Analysis and Classification*, 10(1), 71-84. <https://10.1007/s11634-015-0198-6>

Watkins, M. W. (2017). The reliability of multidimensional neuropsychological measures: from alpha to Omega. *The Clinical Neuropsychologist*, 31(6-7), 1113-1126.

Wells, C. S., & Wollack, J. A. (2003). *An instructor's guide to understanding test reliability*. Testing & Evaluation Services University of Wisconsin.

Widhiarso, W., & Ravand, H. (2014). Estimating reliability coefficient for multidimensional measures: A pedagogical illustration. *Review of Psychology*, 21(2), 111-121.

Zhang, J. (2007). Conditional covariance theory and DETECT for polytomous items. *Psychometrika* 72:69.

Zinbarg, R., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's, Revelle's, and McDonald's: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, 70, 123-133.

Zinbarg, R., Yovel, I., Revelle, W., & McDonald, R. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for estimators for ω_h . *Applied Psychological Measurement*, 30, 121-144

Zinbarg, R. E., Revelle, W., & Yovel, I. (2007). Estimating for structures containing two group factors: Perils and prospects. *Applied Psychological Measurement*, 31, 135-157.

Ek Tablo 1. Örnek Verinin Maddeler Arası Korelasyon Tablosu

Maddeler	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
M1	1.00	0.33	0.38	0.38	0.40	0.04	0.01	0.00	0.03	0.03
M2	0.33	1.00	0.39	0.39	0.36	0.04	-0.01	0.01	0.00	0.01
M3	0.38	0.39	1.00	0.43	0.43	-0.02	0.00	0.01	-0.01	-0.01
M4	0.38	0.39	0.43	1.00	0.42	0.00	0.00	-0.02	0.00	0.01
M5	0.40	0.36	0.43	0.42	1.00	0.00	0.00	0.01	0.03	0.00
M6	0.04	0.04	-0.02	0.00	0.00	1.00	0.35	0.35	0.35	0.35
M7	0.01	-0.01	0.00	0.00	0.00	0.35	1.00	0.45	0.38	0.44
M8	0.00	0.01	0.01	-0.02	0.01	0.35	0.45	1.00	0.35	0.45
M9	0.03	0.00	-0.01	0.00	0.03	0.35	0.38	0.35	1.00	0.37
M10	0.03	0.01	-0.01	0.01	0.00	0.35	0.44	0.45	0.37	1.00

Ek Tablo 2. Basit Ölçme Modelindeki Koşullara Göre 1000 Tekrar için Kestirimlerin Standart Sapması

B.Kor.	M.Kor.	Madde Sayısı	100			200			400		
			α	α_s	ω_g	α	α_s	ω_g	α	α_s	ω_g
0.00	0.30-0.50	5-5	0.057	0.044	0.034	0.040	0.030	0.024	0.028	0.021	0.017
		7-3	0.049	0.040	0.031	0.034	0.028	0.023	0.024	0.019	0.016
		10-10	0.032	0.025	0.019	0.022	0.017	0.013	0.015	0.012	0.010
		15-5	0.023	0.020	0.018	0.014	0.013	0.012	0.010	0.009	0.008
	≥ 0.75	5-5	0.029	0.009	0.006	0.019	0.006	0.004	0.013	0.004	0.003
		7-3	0.019	0.007	0.006	0.013	0.005	0.004	0.009	0.003	0.003
		10-10	0.014	0.004	0.003	0.009	0.003	0.002	0.006	0.002	0.002
		15-5	0.007	0.003	0.003	0.005	0.002	0.002	0.003	0.002	0.001
0.50	0.30-0.50	5-5	0.037	0.033	0.032	0.024	0.021	0.021	0.018	0.016	0.016
		7-3	0.037	0.033	0.031	0.024	0.022	0.021	0.016	0.015	0.015
		10-10	0.019	0.017	0.017	0.014	0.012	0.012	0.009	0.008	0.008
		15-5	0.018	0.016	0.017	0.012	0.011	0.011	0.007	0.007	0.007
	≥ 0.75	5-5	0.014	0.006	0.006	0.009	0.004	0.004	0.006	0.003	0.003
		7-3	0.010	0.005	0.005	0.008	0.004	0.004	0.005	0.003	0.003
		10-10	0.007	0.003	0.003	0.004	0.002	0.002	0.003	0.001	0.001
		15-5	0.005	0.003	0.003	0.003	0.002	0.002	0.002	0.001	0.001

α : Cronbach alfa katsayısı; α_s : Tabakalı alfa katsayısı, ω_g : omega katsayısı

Ek Tablo 3. Karmaşık Ölçme Modelindeki Koşullara Göre 1000 Tekrar için Kestirimlerin Standart Sapması

B.Kor.	M.Kor.	Madde Sayısı	100			200			400		
			α	α_s	ω_g	α	α_s	ω_g	α	α_s	ω_g
0.00	0.30-0.50	5-5	0.048	0.039	0.036	0.033	0.027	0.027	0.022	0.018	0.018
		7-3	0.044	0.038	0.035	0.029	0.026	0.025	0.021	0.018	0.017
		10-10	0.033	0.026	0.022	0.021	0.016	0.015	0.015	0.012	0.011
		15-5	0.021	0.019	0.019	0.013	0.012	0.012	0.009	0.008	0.009
	≥ 0.75	5-5	0.022	0.009	0.007	0.015	0.006	0.005	0.010	0.004	0.003
		7-3	0.015	0.008	0.006	0.010	0.005	0.004	0.007	0.004	0.003
		10-10	0.013	0.005	0.003	0.009	0.003	0.003	0.006	0.002	0.002
		15-5	0.006	0.003	0.003	0.005	0.002	0.002	0.003	0.002	0.001
0.50	0.30-0.50	5-5	0.033	0.030	0.027	0.022	0.019	0.018	0.015	0.013	0.013
		7-3	0.028	0.026	0.025	0.019	0.017	0.017	0.014	0.013	0.013
		10-10	0.019	0.017	0.016	0.013	0.012	0.012	0.009	0.008	0.008
		15-5	0.015	0.014	0.015	0.011	0.011	0.011	0.008	0.007	0.007
	≥ 0.75	5-5	0.012	0.006	0.005	0.007	0.004	0.003	0.005	0.003	0.002
		7-3	0.010	0.006	0.005	0.006	0.004	0.003	0.004	0.003	0.002
		10-10	0.006	0.003	0.003	0.004	0.002	0.002	0.003	0.001	0.001
		15-5	0.004	0.003	0.002	0.003	0.002	0.002	0.002	0.001	0.001

α : Cronbach alfa katsayısı; α_s : Tabakalı alfa katsayısı, ω_g : omega katsayısı

EXTENDED ABSTRACT

Introduction: Despite the violation of the assumptions of the alpha coefficient in the literature, it is still frequently used in reporting the reliability of multidimensional and congeneric measurement results. Researchers may prefer a general alpha reporting over all items in the test for the reliability of the measurements obtained from measurement tools with additive sub-dimensions. It is thought that it is necessary to investigate how accurate such an approach is according to the stratified alpha and omega coefficients, which are stated to be a good predictor of reliability in multidimensional structures. In addition, it is seen in the literature that reliability analysis studies on multidimensional data are limited (Brunner & Süß, 2005; Cortina, 1993; Kamata et al., 2003; Raykov et al., 2015; Serbetar & Sedlar, 2016; Viladrich et al., 2017; Watkins, 2017; Widhiarso and Ravand, 2014). Among these studies, there was only one study (Kamata et al., 2003) that conducted research by manipulating the correlation between items. It was also desired to investigate how effective the size of the correlation between the items would be on the multidimensional reliability estimations. In addition, it was also wanted to observe how reliability is affected from type of the measurement model. Depending on the identified need, in this study, "How do alpha, stratified alpha and omega estimations change in the multidimensional simple and complex test structures simulated under different levels of sample size, correlation between dimensions, correlation between items, test length and number of items per dimension?" search for an answer to the question.

Materials and Methods: In line with the purpose of the research, the data for multidimensional simple and complex test structures simulated under sample size (100, 200 and 400), correlation between dimensions (0.00 and 0.50), correlation between items (0.30-0.50 and ≥ 0.75), test length (10 and 20) and the number of items per dimensions (5-5, 7-3, 10-10, and 15-5) conditions, with 1000 repetitions. The accuracy of the simulated data in terms of structural and item correlation matrix was checked by DFA based on polychoric correlation matrix with

unweighted least squares estimation and oblimin rotation method using the psych (Revelle and Revelle, 2022) package in the R program. Among the reliability coefficients examined within the scope of the research, alpha and stratified alpha estimations were calculated in the R program using sirt package (Robitzsch, 2022), and omega estimations were calculated using the psych (Revelle and Revelle, 2022) package.

Findings: Alpha coefficient had estimated lower reliability values than omega and stratified alpha coefficients in all conditions. It was seen that at least 0.10 differences occur between omega and alpha coefficients in terms of averages of estimations, especially in the uncorrelated measurement models. While the correlation between items was 0.30-0.50 in the perfect simple unrelated model, stratified alpha and omega were acceptable by the standards of many (>0.70) in all cases, while the alpha coefficient was rather high and acceptable only when the total test length is 20 items. When the entire 1000 repetitions were evaluated, it was noteworthy that the alpha estimated the reliability with a quite low value of 0.35 (but the stratified alpha was 0.50 and omega was 0.76). Moreover, in some conditions, the difference between alpha and omega coefficients increased to 0.40. The stratified alpha and omega coefficients were less than 0.70 in the 10-item data. The stratified alpha yielded lower estimations than omega coefficient in shorter tests, but the performance of these two coefficients nearly remained stable in the other conditions. As the correlation between dimensions increased, it was observed that the reliability was estimated higher in all conditions. In the perfect simple uncorrelated model, when the correlation between the items was ≥ 0.75 , all three coefficients were above 0.70. Therefore, the inter-item correlation had a greater effect on the reliability coefficients than the interdimensional correlation. This is an expected situation. Because the stratified alpha and especially the alpha coefficients are a function of the mean of the item correlations, higher reliability estimations values can be obtained as the inter-item relationships become stronger. Also, when the correlation between items increases in factor analytical models, it is expected to yield higher loadings. The stratified alpha and omega coefficients yielded fairly consistent results with the exception of two conditions, the difference between the two coefficients never exceed 0.01. This coefficient gave more biased results than the stratified alpha estimations when the correlation between dimensions was 0.30-0.50 in the uncorrelated models and the test length was 10 items. It is known that reliability increases with increasing test length, and the findings supported this. Whether the model is complex or simple had an effect on reliability. The complex models provided higher estimations for all coefficients than the perfect simple models.

Discussion, Conclusion and Suggestions: In the study, it is found that the variables having the most effect on the reliability estimations are the correlation between items and correlation between dimensions, respectively. Test length and the type of structure (simple and complex) also effect on reliability. The sample size is not effective on the average of reliability estimations, but it is effective on the error of the estimations (only providing more homogeneous estimates). The most important conclusion drawn from the findings of this study is that, in both perfect simple and complex multidimensional tests, if there is a moderate relationship between the dimensions and a high correlation between the items, the omega and stratified alpha coefficients give similar results, and the alpha coefficient can give values that converge to these two coefficients. However, in practice, it is unlikely to encounter data sets with such high correlations between items, but medium and lower correlations are. In this case, the conditions with the low correlation between items of this study are more important and should be considered. especially in the uncorrelated models, these conditions yields omega and stratified alpha that are quite high and acceptable by convention than alfa. Maybe, among the researchers who examined the current study, there may be some who think that the alpha coefficient gives

results above 0.70 (according to the average reliability estimations), and therefore using the alpha coefficient outside of their assumptions yields no significant drawback. However, it should be remembered that these results are the average of 1000 data sets, and alpha coefficients below 0.40 were calculated even in these 1000 datasets simulated in a sterile and controlled manner. In fact, it was not included in the article due to the excess of conditions, but datasets with a sample size of 50 were encountered, while omega was 0.60 and the alpha coefficient was 0.18. These examples are proof that we can encounter data sets that can give similar results in real life, and they also reveal the possibility that a measurement result that is actually reliable may lead to a wrong decision that it is not reliable with the alpha coefficient. Therefore, based on both the literature and the results of the present study, it is suggested that the alpha coefficient should not be reported, or not be reported alone, as proof of reliability in terms of internal consistency in multidimensional measurements; instead, it is recommended to use and report more appropriate coefficients such as omega, stratified alpha.