

Likert tipi ölçek geliştirme uygulamaları üzerine¹

Fatih Yılmaz², Ahmet Pekgör³

DOI 10.5281/zenodo.7774553

Özet

Sosyal bilimlerde gerçekleştirilen deneysel çalışmaların can damarını oluşturan ölçme araçlarının geliştirilmesi uzun ve yorucu bir süreçtir. Bu tür çalışmalarda çeşitli ölçme araçları kullanılsa da Likert ölçeklerin çeşitli gerekçelerle yaygın kullanımı yadsınamaz bir gerçektir. Her ne kadar Likert ölçek kullanımı literatürde oldukça yaygın olsa da bu ölçeklerin de bazı artıları ve eksileri bulunmaktadır. Bu bağlamda Likert ölçeklere yöneltilen çeşitli eleştiriler de bu çalışma kapsamında ele alınan başlıklardan biridir. Literatür taraması şeklinde gerçekleştirilen bu çalışmada Likert ölçek geliştirme sürecinde araştırmacıların dikkatle üzerinde durmaları gereken bazı hususlara değinilmiştir. Madde havuzunun oluşturulmasından başlayarak ön uygulamanın yapılması, geçerlik ve güvenilirlik çalışmalarının gerçekleştirilmesi ve araştırmanın yürütülmesi sürecinde araştırmacının göz önünde bulundurması gereken hususları irdeleyen bu çalışma; ölçek geliştirme konusunda araştırmacılar için bir rehber niteliğindedir. Ayrıca bu çalışmada değinilen hususlar ile araştırmacıların bizzat geliştirdiği ya da çeşitli kültürlerde veya çeşitli yabancı dillerde geliştirilen ölçeklerin uyarlanmasında karşılaşılabilecek problemlere ve bu problemlerin çözümlerine yönelik çeşitli yaklaşım ve önerilere de yer verilmektedir.

Anahtar Kelimeler: geçerlik, güvenilirlik, likert ölçek, ölçek geliştirme, psikometrik ölçekler

About Likert type scale development applications

Abstract

The development of measurement tools, which are the heart of experimental studies in social sciences, is a long and exhausting process. Although various measurement tools are used in such studies, it is undeniable that Likert scales are widely used for various reasons. Although the use of Likert scales is quite common in the literature, they also have pros and cons. In this context, various criticisms of the Likert scale are some of the issues covered in this study. In this study, which is a literature review, some issues that researchers should keep in mind in the Likert scale development process have been studied. This research, focusing on the issues that researchers should bear in mind during the pilot study, validity and reliability studies, and during the research, starting from the creation of the item pool, is like a guide for researchers in developing Likert-type scales. Moreover, this study also includes some suggestions for the problems that may be encountered by the researchers in the development or adaptation of scales from various cultures or languages.

Keywords: validity, reliability, likert scales, scale development, psychometric scales

¹ Bu çalışma yüksek lisans semineri olarak sunulmuştur.

² Öğretmen, Milli eğitim Bakanlığı, yilmazoglu@gmail.com, ORCID: 0000-0001-7852-6756

³ Dr. Öğr. Üyesi, Necmettin Erbakan Üniversitesi, istatistik@hotmail.com, ORCID: 0000-0001-9446-7960

Giriş

Ölçekler; ölçülmek istenilen çeşitli özelliklerin sınıflanabilmesi, sıralanabilmesi veya miktarının belirlenebilmesi için gereken hususları belirleyen araçlar olarak ifade edilebilir (Karakoç & Dönmez, 2014: 40). Psikometrik ölçekler ise doğrudan gözlenmesi ve ölçülmesi mümkün olmayan -gizil- özellikler ve yapılar hakkında bilgi edinmek amacı ile kullanılan ölçme araçlarıdır. Ancak sağlıklı ve güçlü bir ölçek geliştirme; araştırmacılar için oldukça yorucu, bilgi birikiminin yanında zaman ve emek gerektiren bir süreçtir (Schmitt & Klimoski, 1991).

Çeşitli disiplinlerde çeşitli ölçeklerden yararlanılmaktadır. Sosyal bilimlerde ise en sık kullanılan -Rensis Likert tarafından ortaya atılan- Likert tipi ölçek başta olmak üzere, Bogardus toplumsal uzaklık ölçeği, Guttman ölçeği, Edward ve Kilpatrick ölçeği ve Thurstone ölçekleri gibi ölçekler ile karşılaşabilmektedir. Bu ölçek türleri arasında Likert ölçeği sosyal bilimlerde başta olmak üzere birçok disiplinde ölçeğin geliştirilme süreci, uygulama aşaması ve veri analizinde sağladığı çeşitli kolaylıklar nedeni ile tarama çalışmalarında en sık kullanılan ölçek türü haline gelmiş görünmektedir (Li, 2013: 1609). Öyle ki anketlerin, diğer bir ifade ile ölçeklerin saha araştırmalarında sıklıkla kullanılması nedeni ile tarama niteliğindeki araştırmaların kalbi olduğunu ifade eden araştırmacılara rastlanmaktadır (Krosnick & Presser, 2010: 263).

Psikometrik bir ölçme aracı geliştirmenin ilk aşaması yukarıda da belirtilen ölçek türlerinden hangisinin geliştirileceğine karar verilmesidir. Bu çalışmada ise sosyal bilimlerde en sık kullanılan (Joshi vd., 2015) Likert tipinde ölçek geliştirme sürecine yer verilecektir.

1932 yılında Rensis Likert'in "Bir Tutum Ölçüm Tekniği" isimli çalışması ile literatüre giren Likert tipi ölçek (Batterton & Hale, 2017: 32) gerçek veya kuramsal bir konu hakkındaki bir dizi ifadenin katılımcıya yöneltilerek onların konu hakkındaki görüşlerinin toplanmasını içeren ölçme araçları olarak ifade edilebilir (Joshi vd., 2015: 397). Diğer bir ifadeyle psikometrik bir ölçme aracı olan Likert tipi ölçekler; katılımcıların görüşlerini, tutumlarını ya da duygularını yansıtabilecekleri cevap seçenekleri sunan araçlardır (Nemoto & Beglar, 2014). Literatürde Likert tipi ölçek olarak adlandırılan bu ölçme araçları doğrudan gözlenemeyen yapıların ölçülmesinde oldukça kullanışlı olup (Jebb vd., 2021; Li, 2013) bu tür ölçeklerdeki katılımcı ifadelerinin eşit aralıklarla derecelendirildiği kabul edilmektedir (Chyung vd., 2017).

Likert tipi ölçekler tutumların iki ucu bulunan, pozitiften negatife uzanan, ne pozitif ne de negatif bir tutum ifade eden tarafsız bir orta noktası bulunan doğrusal bir yapı üzerinde yer aldığı varsayımına dayanmakta (Symeonaki vd., 2015: 739) ve birçok disiplinde yapılan araştırmalarda sıklıkla tercih edilmektedir.

Diğer taraftan Likert tipi ölçekler genellikle Likert'in (1932) de kullandığı (Edmonson, 2005: 128) 5'li Likert diye adlandırılan "Kesinlikle katılıyorum, Katılıyorum, Kararsızım, Katılmıyorum, Kesinlikle katılmıyorum" şeklinde 5 cevap seçeneğinden oluşsa da (Jamieson, 2004) literatürde cevap seçeneğinin sınırlandırılabilmesine ya da sınırlandırılması gerektiğine ilişkin kaynak bulunmamaktadır. Bu bağlamda ölçeğin kaç cevap seçeneğinden oluşacağı katılımcıların özelliklerine, araştırmanın amacına ve araştırmacının araştırma konusuna yaklaşımına bırakılmış gibi görünmektedir.

Bu bağlamda literatürde çok geniş bir yelpazede cevap seçenekleri sunan Likert tipte ölçeklerle karşılaşmak mümkündür (Chakrabartty, 2014; Preston & Colman, 2000; Krosnick, 2018). Ancak yapılan araştırmalarda edinilen deneyimler Likert tipi ölçeklerde yer alan cevap seçenekleri arttıkça katılımcıların dikkatlerinin dağıldığını ve sonuçta ölçek sorularına verilen cevapların katılımcıların gerçek duygu, düşünce ya da tutumlarını yansıtmaktan uzaklaştığını ya da katılımcıların ölçme aracındaki tüm sorulara aynı cevap seçeneğini işaretlediklerini göstermektedir. Ayrıca katılımcıların özellikleri ve araştırma konusunun hassasiyeti de cevap

seçenekleri üzerinde etkili olabilmektedir. Örneğin ileri yaşlardaki katılımcılarla ya da küçük yaşlardaki katılımcılarla yapılan araştırmalarda 3 cevap seçeneğinden oluşan ölçeklerin uygun olacağı düşünülebilir (Bora Semiz & Altunışık, 2016, 587).

Ayrıca, ölçeklerde yer alan cevap seçeneğinin artmasıyla birlikte bazı katılımcılarda ölçekleri özensizce doldurma sonucu tüm maddelere cevap seçeneklerinden ilk seçeneği işaretleme, tüm maddelere aynı yanıtı verme, ya da tüm maddelere “Karasızım, Nötrüm” gibi orta noktada yer alan cevap seçeneğini işaretleme eğilimleri de gözlenebilmektedir (Li, 2013: 1610).

Katılımcı ifadelerinin yönünü ve şiddetini içeren cevap seçenekleri barındıran Likert tipi ölçeklerde araştırma sorularına “Katılıyorum” ya da “Katılmıyorum” ifadeleri katılımcı tutumunu, “Kesinlikle” ifadesi ise tutumun şiddetini belirtmektedir (Albaum, 1997: 332). Tablo 1’de Likert ölçeklerde yer alabilecek cevap seçeneklerine ilişkin bazı örnekler verilmiştir.

Tablo 1: Likert ölçeklerde kullanılacak cevap seçenekleri

3'lü Cevap Seçeneği	4'lü Cevap Seçeneği	5'li Cevap Seçeneği
Katılıyorum	Tamamen Katılıyorum	Tamamen Katılıyorum
Ne Katılıyorum Ne Katılmıyorum	Katılıyorum	Katılıyorum
Katılmıyorum	Katılmıyorum	Ne Katılıyorum Ne Katılmıyorum
	Hiç Katılmıyorum	Katılmıyorum
		Hiç Katılmıyorum
6'lı Cevap Seçeneği	7'li Cevap Seçeneği	
Tamamen Katılıyorum	Tamamen Katılıyorum	
Katılıyorum	Katılıyorum	
Biraz Katılıyorum	Biraz Katılıyorum	
Biraz Katılmıyorum	Ne Katılıyorum Ne Katılmıyorum	
Katılmıyorum	Biraz Katılmıyorum	
Hiç Katılmıyorum	Katılmıyorum	
	Hiç Katılmıyorum	

Vagias, 2006

Bu çalışma ile ise araştırmacılara Likert tipi ölçeklerin geliştirilmesi sürecine yönelik özet bilgiler sunularak araştırmacıların ölçek geliştirme sürecinde dikkate almaları gereken bazı hususlara değinmenin yanında karşılaşılabilecekleri çeşitli zorluklara ilişkin bilgiler sunulması amaçlanmıştır.

Yöntem

Likert tipi ölçek geliştirme uygulamalarını ele alan bu çalışma nitel araştırma yöntemlerinden -birçok disiplinde de yararlanılan- doküman analizi yönteminden yararlanılarak hazırlanmıştır. Doküman analizi; fikir edinmek, yorumlamak ve çıkarımda bulunmak amacıyla basılı ve/veya elektronik dokümanları sistematik biçimde incelemek olarak ifade edilebilir (Bowen, 2009: 27). Bu bağlamda bu çalışmada, sırasıyla ölçek hazırlama aşamaları içerisinde yer alan madde havuzunun oluşturulması, pilot uygulamanın yapılması, kapsam geçerliği, yapı geçerliği, güvenilirlik, ölçek geliştirme sürecinde dikkat edilmesi gereken hususlar ve Likert tipi ölçeklere yöneltilen eleştirilere yer verilmiştir.

Ölçek Hazırlama Aşamaları

Ölçme aracının geliştirilmesi süreci, özen gerektiren bir dizi aşama içermektedir. Bu aşamalara kısaca aşağıda değinilmiştir:

Madde Havuzunun Oluşturulması

Ölçek hazırlama safhasının ilk basamağı madde havuzunun oluşturulması ile başlar. Bu süreçte araştırmacı tarafından araştırmanın dayanağını oluşturan kuramsal çerçeve göz önünde bulundurularak kapsamlı bir literatür taraması gerçekleştirilir. Diğer bir ifade ile literatür taraması yapılır. Bu aşamada konu ile ilgili yapılmış araştırmaların araştırmacıya yol gösterici etkisi büyüktür. Önceden yapılmış araştırmalar araştırmacıya hangi konuların çalışılıp hangi konuların çalışılmadığını, çalışılan konularda ise hangi değişkenlerin önceden irdelendiğini, örneklem grubun özelliklerini, söz konusu örneklemden elde edilen veriler ile hangi sonuçlara ulaşıldığını, ulaşılan sonuçların nasıl yorumlandığını göstermesi bakımından son derece önemlidir. Bir bakıma önceki araştırmalar, yeni araştırmanın gerçekleştirilmesi sürecinde yolu aydınlatan bir fener gibidir.

Kapsamlı bir literatür taraması yapmadan araştırmaya başlayan bir araştırmacının ilerleyen süreçlerde seçmiş olduğu araştırma konusunun önceden çalışılmış olduğu sonucuna ulaşabilmesi muhtemeldir. Bunun yanında araştırmanın ilerleyen süreçlerinde ölçek maddelerine ilişkin bazı maddelerin eksikliği ya da bazı maddelerin gereksizliğini de hissedebilir. O nedenle araştırmacının ölçme aracının belirlenmesi ya da oluşturulması sürecinde sabırlı ve titiz davranması, ilerleyen süreçlerde yaşanabilecek eksiklik duygusunun önüne geçebilecektir.

Literatür taraması sürecinde sık yapılan hatalardan biri yalnızca konu ile ilgili önceden yapılmış araştırmaların taranmasıdır. Halbuki araştırma konusu hakkındaki bir köşe yazısından bile ölçme aracı için bir madde yazmak hatta bireyler arasında gerçekleşen konuya ilişkin bir diyalogdan bile bir madde yazmak mümkündür. İyi ölçek maddesi yazma konusunda çok çeşitli görüşler bulunsa da ölçek maddelerinin aşağıda yer alan özellikleri taşıması önemli görülmektedir. Bunlar:

- İfadeler kısa ve kolay anlaşılır olmalı, teknik ifadelerden ve jargonlardan kaçınılmalıdır.
- Söz dizimi oldukça basit olmalıdır.
- Muğlak ifadelerden kaçınılmalıdır.
- Katılımcıları belli bir cevap seçeneğine yönlendirebilecek ifadelerden kaçınılmalıdır.
- Bir madde ile yalnızca bir soru sorulmalıdır.
- Ölçekte yer alan tüm ifadeler aynı zamanda kullanılmalıdır.
- Ölçekteki ifadelerle verilebilecek cevap seçenekleri ile soru kökleri arasında çelişki bulunmamalıdır.
- Bir ölçek maddesinde birden fazla olumsuzluk eki yer almamalıdır.
- Aynı şeyi soran birden fazla sayıda soru yazmaktan kaçınılmalıdır.
- Ölçek maddelerinden birkaç tanesinin ters puanlanacak şekilde yazılmasının katılımcıların yanıtlarını denetlemede kolaylık sağlayabileceği düşünülebilir.
- Çok boyutlu ölçeklerde aynı boyutta yer alan sorular aynı grupta yer almalıdır.
- Ölçek maddelerinin katılımcılar tarafından yanlış anlaşılmasının önüne geçmek için araştırmacı tarafından gerekli tedbirler alınmalıdır (Hoşgörür, 1997: 354; Krosnick, 2018: 450).

Madde havuzu oluşturulması sürecinde araştırmanın dayanağını oluşturacak kavramsal çerçeve göz önünde bulundurularak olabildiğince fazla sayıda soru yazmak hem literatürü kapsamı açısından hem de yapılacak analizlerde belli bir madde sayısının korunabilmesi

açısından önemli görülmektedir. Bu süreçte literatürde yer alan yabancı kaynakların taranması da önem arz etmektedir. Ancak yabancı dilde yazılmış kaynakların taranmasında ya da yabancı dilde oluşturulan ölçme araçlarından madde alınması durumunda araştırmacının yabancı dile, konuya ve jargona hakimiyeti son derece önemlidir. Aksi halde madde aktarımı esnasında ciddi dil problemleri gündeme gelebilecektir.

Bazı araştırmacılar katılımcıların araştırma sorularına içtenlikle cevap verip vermediklerini anlayabilmek amacıyla özellikle madde sayısı fazla olan ölçeklere aynı soruyu tekrar yazmak gibi kontrol maddeleri yazma yoluna gitmektedirler. Ancak bu yaklaşım bazen uygulamada sorunlara yol açabilmektedir. Araştırma sorularını içtenlikle yanıtlayan katılımcılar da kontrol sorularını gözden kaçırabilmekte, bu durumda söz konusu katılımcının cevapları geçersiz sayılarak araştırmaya dâhil edilmemektedir. Bazı durumlarda ise bu tür kontrol soruları katılımcıları sinirlendirmekte ve de katılımcılar tepki olarak bu sorulara kasıtlı olarak farklı cevaplar vermekte, sonuç olarak ise ölçek analize dâhil edilmeyebilmektedir.

Ölçek maddelerinin hazırlanması esnasında dikkat edilmesi gereken bir diğer husus ise katılımcıların bireysel, toplumsal ya da dini hassasiyetleridir. Bu bağlamda araştırmacı, katılımcıların dini hassasiyetleri başta olmak üzere katılımcıları rahatsız edebilecek araştırma sorularından kaçınmalıdır. Bu tür hassasiyetlere dikkat edilmeden ölçek maddeleri hazırlandığında veri toplama aşamasında araştırmacı ciddi tepkiler ile karşılaşabilmekte, araştırmacının motivasyonu kırılmakta hatta bazı uç durumlarda araştırmaya son verilebilmektedir.

Diğer taraftan Likert tipte hazırlanmış ölçme araçlarında ölçek sorularına verilen cevap seçenekleri, diğer bir ifadeyle etiketleme seçenekleri de son derece önemlidir. Likert ölçeklerde araştırmacılar araştırma sorularına cevap olarak uygun olabileceğini düşündükleri cevabı, cevap seçeneği olarak belirleyebilmektedir. Bu nedenle araştırmacılar geliştirdikleri ya da uyarladıkları Likert tipi ölçekler için cevap seçeneklerini ve soru köklerini çok iyi değerlendirmeli, uygulama öncesi mutlaka ölçek geliştirme konusunda uzman kişilere danışmalıdır.

Pilot Uygulamanın Yapılması

Pilot uygulama, ölçek geliştirmenin önemli aşamalarından biri olarak kabul edilebilir. Pilot uygulama; madde havuzuna yazılan maddelerin katılımcılar tarafından nasıl algılandığını, araştırmacı tarafından kurgulandığı şekilde algılanıp algılanmadığını ortaya çıkararak araştırmacıya gerekli düzenlemeleri yapma fırsatı sunar. Pilot uygulamada dikkat edilmesi gereken hususlardan biri, pilot uygulamanın yapılacağı katılımcı grubun araştırmanın yürütüleceği evrenden benzer özellikleri taşıyan bireyler olmasıdır (Büyüköztürk, 2002: 478). Örneklemek gerekirse hemşirelerin motivasyon düzeylerini ölçmeye yönelik hazırlanmış bir ölçme aracının pilot uygulamasını hemşirelik öğrencileri gibi erişimi daha kolay ve kısa sürede daha fazla sayıda bireye erişilebilecek bir grup üzerinden yapmak doğru sonuçlar vermeyecektir. Diğer bir ifadeyle pilot uygulama ile araştırmanın yapılacağı örneklemin özelliklerinin farklı olması ölçeğin faktör yapısının doğru biçimde belirlenmesine olanak sağlamayabilecektir.

Pilot uygulamanın ardından geçerlik türlerinden ve geçerlik çalışmalarından söz etmek yerinde olacaktır. Geçerlik en basit ifade ile, Kelley'in tanımladığı şekilde, bir testin ölçmeyi amaçladığı şeyi ölçebilmesi ile ilgilidir (Kelley, 1927: 14). Ölçme aracının geçerli olabilmesi, hedeflenen faktörler dışındaki değişkenleri ölçüme karıştırmaması olarak da düşünülebilir (Thanasegaran, 2009). Literatürde geçerliğe ilişkin yordama geçerliği, eş zamanlı geçerlik, yapı geçerliği ve kapsam geçerliği (Cronbach & Meehl, 1955: 281) gibi çeşitli geçerlik türleri yer alsa da bu

çalışmada kapsam geçerliği ve yapı geçerliği başlıkları altında irdelenecek buna ilaveten güvenilirlik konusu da ele alınacaktır.

Kapsam Geçerliği

Hazırlanan testin ölçülmek istenen davranış özelliklerini ölçüp ölçmediğini anlamının yollarından biri testin kapsam geçerliğini irdelemektir. Bu bağlamda başvurulan yollardan biri uzman görüşü almaktır. Bu amaçla araştırmacı; hazırlamış olduğu madde havuzunda yer alan maddelerin araştırmacının amacına uygunluğu konusunda uzman görüşleri alarak pilot uygulama öncesi, uzman görüşleri doğrultusunda gerekli düzenlemeleri gerçekleştirir. Bu aşamada uzman görüşleri doğrultusunda ölçüğe yeni maddeler eklenebileceği gibi ölçüğe uygun olmadığı değerlendirilen maddelerin ölçekten çıkarılması da mümkündür. Ayrıca birden fazla sayıda uzmanın görüşüne başvurulması, özellikle madde sayısı fazla olan ölçüklere, gözden kaçabilecek hataların elenmesinde yararlı olacaktır.

Yapı Geçerliği

Yapı geçerliği bir testin dayandırıldığı yapıyı ölçme derecesi olarak düşünülebilir (Huot, 1990: 206). Yapı geçerliliği, ilgili yapının göstergeleri olduğu düşünülen veya teorik olarak ilgili olduğu varsayılan diğer değişkenlerle korelasyon ilişkilerine dayalı olarak yapının “doğrulanması” olarak ifade edilebilir (Locke, 2012). Daha anlaşılır bir ifade ile bir ölçme aracında yer alan maddeler arasındaki ilişkileri ve ölçme aracının bütünü ile olan ilişkiyi ifade eder.

Sosyal bilimler, eğitim bilimleri gibi birçok disiplindeki saha araştırmalarında araştırmacılar çok sayıda gözlemden veri toplamaktadır. Elde edilen bu veri seti içerisindeki gözlemlenemeyen yapıların anlaşılmasında açılımlı faktör analizinden yararlanılmaktadır. Faktör analizinin kökenleri 1900’lü yıllarda Charles Spearman’ın “İki-Faktör Teorisi”ne dayanmaktadır (Yong & Pearce, 2013: 79). Faktör analizi yaklaşımına göre değişkenler arasındaki gözlemlenemeyen ilişkiler, faktör analizi yoluyla faktör olarak adlandırılan birimler altında toplanır (Cudeck, 2000: 265; Fabrigar & Wegener, 2011). Diğer bir ifadeyle faktör analizi n sayıdaki değişkeni bir ya da birkaç faktör altında toplamayı amaçlar.

Esasen faktör analizi geliştirilen psikometrik bir ölçme aracının gözlemlenemeyen yapılarının nasıl etkilendiğini, gözlemlenebilen değişkenler aracılığıyla ortaya çıkarılmasını amaçlayan bir dizi yöntemlere verilen isimdir (DeCoster, 1998: 1; DiStefano vd., 2009; Yong & Pearce, 2013: 80). Bu yaklaşım ölçekte yer alan maddeler arasında ilişki bulunduğu varsayımına dayanmaktadır (Shrestha, 2021: 4).

Geliştirilen bir ölçme aracının faktör yapılarını belirlemek için açılımlı faktör analizi yönteminden yararlanılırken bu yöntemle belirlenen faktör yapılarının teyidi için ise doğrulayıcı faktör analizinden yararlanılabilir. Faktör analizi kulağa hoş bir terim gibi gelse de matematiksel olarak bir dizi karmaşık işlemler içeren bir analizdir. Bu analizi gerçekleştiren bir araştırmacının özdeğer, köşegen, faktör yükü, döndürme, toplam varyans gibi kavramlara aşina olması beklenir (Rummel, 1967: 444).

Güvenilirlik

Herhangi bir ölçme aracının geçerli olabilmesi için aynı zamanda güvenilir olması gerekmektedir (Voss vd., 2000: 177). Güvenilirlik -kısaca- tekrarlayan ölçümlerden aynı sonuçların alınması, ölçümlerin tutarlı olması şeklinde ifade edilebilir (Bravo & Potvin, 1991: 381). Literatürde psikometrik ölçme araçlarının güvenilirliklerinin belirlenmesinde kullanılan “omega, KR20, test tekrar test” gibi çeşitli güvenilirlik hesaplama alternatifleri yer alsa da ölçme araçlarının birden fazla madde içermesi durumunda maddeler arasındaki iç tutarlığın

hesaplanmasında yararlanılan (Bonett & Wright, 2014: 3; Christmann & Aelst, 2006: 1661) alfa katsayısı tekniği; psikometrik ölçme araçlarının güvenilirlik düzeylerini belirlemede en sık kullanılan yöntemdir (Bonett & Wright, 2014; Iacobucci & Duhachek, 2003: 478). İç tutarlılık kısaca katılımcıların bir ölçme aracıyla yer alan farklı maddelere tek bir veri toplama süreciyle tutarlı bir biçimde yanıt vermelerinin belirlenmesi olarak ifade edilebilir (Spiliotopoulou, 2009: 150).

Alfa katsayısı Gutmann (1945) tarafından "lambda3" olarak literatüre kazandırılrsa da popüler kullanımını Cronbach'ın 1951 yılında yayınlanan çalışmasıyla kazanmıştır (Warrens, 2015: 128). Söz konusu çalışma psikometrik ölçme araçlarını kullanan araştırmacılar ve test geliştiriciler için bir dönüm noktası olmuş, o günden bu yana psikometrik ölçme araçları kullanılarak yapılan neredeyse tüm araştırmalarda Cronbach'ın alfa değeri raporlanmadan yayınlanan bir çalışma son derece az hale gelmiştir (Sijtsma, 2009: 108).

Ölçme araçlarının güvenilirliğinin matematiksel bir ölçütü olarak düşünülebilen alfa katsayısı 0 ve 1 arasında bir değer alır. Alfa katsayısı 1'e yaklaştıkça ölçme aracının güvenilirlik düzeyi de artmaktadır. Kabul edilmesi gereken minimum alfa katsayısının ne olması gerektiğine ilişkin kabul edilen evrensel kurallar bulunamamakla birlikte (Bonett & Wright, 2014) bazı araştırmacılar minimum kabul edilebilir alfa katsayısının 0,70 - 0,75 arasında olması gerektiğini (Christmann & Aelst, 2006: 1661), bazı araştırmacılar ise 0,65 – 0,80 arasındaki alfa değerlerinin kabul edilebilir olduğunu (Vaske vd., 2017: 165), bazı araştırmacılar alfa değerinin 0,80'in üzerinde olması gerektiğini ifade etse de (Goodboy & Martin, 2020, s.423) George ve Mallery'den (2003) aktarılan aşağıdaki tablonun alfa değerinin yorumlanmasında yararlı olabileceği düşünülmektedir.

Tablo 2: α değerlendirme ölçütleri

α	Değerlendirme
$\alpha > ,9$	Mükemmel
$\alpha > ,8$	İyi
$\alpha > ,7$	Kabul edilebilir
$\alpha > ,6$	İdare eder
$\alpha > ,5$	Zayıf
$\alpha < ,5$	Kabul edilemez

Akt: Gliem & Gliem, 2003: 87

Alfa katsayısı, ölçme aracıyla ya da çok faktörlü ölçeklerde ilgili faktörde yer alan madde sayısından doğrudan etkilenmektedir (Spiliotopoulou, 2009: 153; Vaske vd., 2017: 165). Bu nedenle alfa katsayısı düşük çıkan ölçeklerde ya da ölçeklerin belirli faktörlerinde madde sayısını artırmak, ölçme araçlarında güvenilirliği artırmanın bir yolu olarak araştırmacıların sıklıkla başvurduğu bir yöntem olarak karşımıza çıkmaktadır. Ancak bu yöntem, oldukça uzun olabilen bazı veri toplama araçları için var olan veri toplama aracını daha da uzatacağından her zaman kullanışlı bir yöntem olarak düşünülmemelidir (Spiliotopoulou, 2009: 152).

Ölçek Geliştirme Sürecinde Dikkat Edilmesi Gereken Hususlar

Hazırlanan madde havuzunda yer alan maddeler üzerinde faktör analizi çalışmalarının yapılmasına geçmeden önce verilerin analiz için uygunluğunun gözden geçirilmesi gerekmektedir. İlk olarak gözlem sayısından başlamak gerekirse, faktör analizi çalışmalarındaki "daha fazla katılımcı daha iyi sonuç" inancı yerini daha net bazı yaklaşımlara bırakmıştır.

Örneğin faktör analizinde gözlem sayısının, madde sayısının 4 ya da 5 katı olması gerektiği; minimum gözlem sayısının 200 olması gerektiği; 100 gözlemden az olmamak koşuluyla gözlem sayısının madde sayısının 5 katı olması gerektiği; toplam gözlem sayısının 100 gözlemden az olduğu durumlarda ise gözlem sayısının madde sayısının 10 katı olması gerektiği ya da analize dahil edilmesi gereken gözlem sayısının faktör yük değerlerine bağlı olduğunu benimseyen yaklaşımlar literatürde yer almaktadır (Floyd & Widaman, 1995: 289-290). Diğer taraftan Tabachnick kuralına göre ise faktör analizi için en az 300 gözlem gerekli iken Hair ve arkadaşları gözlem sayısının 100 ve üzeri olması gerektiğini belirtmekte; Comrey ve arkadaşları ise 100 gözlem için zayıf, 200 gözlem için vasat, 300 gözlem için iyi, 500 gözlem için çok iyi, 1000 ve daha fazla sayıda gözlem için ise mükemmel ifadesini kullanmaktadır (Akt: Williams vd., 2010: 4). Ancak yaygın görüş gözlem sayısının, madde sayısının 10 katı olması gerektiği yönündedir (Shrestha, 2021: 4).

Verilerin faktör analizine uygunluğunu test etmenin bir diğer yolu da Kaiser Meyer Olkin örneklem ölçümü ve Bartlett küresellik testinden yararlanmaktır. Bu testler özellikle gözlem sayısı ve analize dâhil edilen değişken sayısının 1/5 oranını karşılamadığı durumlarda kullanılmaktadır. Kaiser Meyer Olkin testinin 0,5'ten büyük olması ve Bartlett küresellik testi sonucunun anlamlı olması ($p < 0,05$) verilerin faktör analizine uygunluğunu göstermektedir (Williams vd., 2010: 5). Kaiser (1974) verilerin faktör analizine uygunluğunun denetlenmesinde aşağıdaki tablodaki ölçütlerden yararlanılabileceğini belirtmektedir.

Tablo 3. KMO ölçütleri

KMO > ,90	Mükemmel
KMO > ,80	Çok iyi
KMO > ,70	Orta
KMO > ,60	Vasat
KMO > ,50	Kötü
KMO < ,50	Kabul edilemez

Kaiser, 1974: 35

Araştırmacıların sıkça başvurduğu diğer bir madde eleme yöntemi ise Kaiser'in (1960) özdeğer yöntemidir. Özdeğer kısaca faktör analizi işlemi sonucu elde edilen faktörlerde yer alan maddelerin faktör yüklerinin karesidir (Rummel, 1967: 467). Bu yöntemde analize dahil edilmesi gereken faktörlerin özdeğerlerinin 1'den büyük olması beklenmektedir. Bu yaklaşımın temel prensibi ise en önemli faktörlerin analize dahil edilmesidir (Floyd & Widaman, 1995: 291-292; Pituch & Stevens, 2016: 342).

Görüldüğü üzere faktör analizi için gerekli gözlem sayısına ilişkin çeşitli önermeler ileri sürülmekte, net bir gözlem sayısının yeterliliğinden söz edilememektedir. Faktör analizi için gerekli gözlem sayısına ilişkin bile bu kadar çeşitli yaklaşımın varlığı araştırmacıların kafasını daha çok karıştırmakta, sürekli olarak "Gözlem sayım acaba yeterli mi?" endişesini taşımalarına neden olmaktadır.

Faktör analizinde yararlanılabilecek temel bileşenler analizi yöntemi başta olmak üzere temel eksen yöntemi, görüntü yöntemi, en yüksek olasılıklar yöntemi, alfa yöntemi ve kanonik yöntem gibi çeşitli yöntemler tercih edilebilir. Ancak bu yöntemler arasında temel bileşenler analizi yöntemi en sık tercih edilen yöntem olup (Abdi & Williams, 2010: 433) bazı istatistik paket programlarında varsayılan analiz yöntemi olarak araştırmacılara sunulmaktadır.

Ayrıca açımlayıcı faktör analizi sürecinde araştırmacıların göz önünde bulundurması gereken birtakım kurallar vardır. Kısaca bu hususlara değinmek gerekirse temel bileşenler analizi

yönteminden yararlanarak madde indirgeme sırasında bazı maddelerin yüksek faktör yük değerlerine sahipken bazı maddelerin ise daha düşük faktör yük değerlerine sahip olduğu görülebilir. Söz konusu faktör yük değerleri, ölçekte yer alan ilgili maddenin içinde yer aldığı faktörle ilişkisini göstermektedir (Büyüköztürk, 2002). Yapılan literatür taramaları, faktör analizi sonucu ölçğe dahil edilmesi planlanan maddelerin faktör yük değerlerinin ne olması gerektiğinin araştırmacının kararına göre değişebileceğini işaret etmektedir. Diğer bir ifadeyle aynı madde havuzunda yer alan maddeler üzerinde faktör analizi gerçekleştiren iki farklı araştırmacı, faktör yük değerlerinin alt noktası olarak farklı eşikleri belirleyebilmektedir. Ancak genel olarak yapılan analizlerde 0,30'dan daha düşük faktör yüküne sahip maddelerin ölçekten çıkarıldığı görülmektedir (DiStefano, 2009: 3; Pituch & Stevens, 2016: 389).

Psikometrik ölçme araçlarının geliştirilmesi sürecinde göz önünde bulundurulması gereken bir diğer husus ise her bir faktörde yer alması gereken madde sayısıdır. Tek faktörden oluşan ölçeklerde yer alması gereken minimum madde sayısı 4 iken çok boyutlu ölçeklerde her bir faktörde yer alması gereken minimum madde sayısı 3 olarak ifade edilmektedir (Raubenheimer, 2004: 60). Ayrıca çok boyutlu ölçeklerde iki maddeden oluşan faktörler istisnai durumlar olarak nitelenmekte, bu tür durumlarda ise maddeler arasında yüksek düzeyde ilişki aranması gerektiği belirtilmektedir ($r > ,70$) (Raubenheimer, 2004: 60; Yong & Pearce, 2013: 80).

Araştırmacıların faktör analizi sürecinde sıklıkla karşılaştıkları bir diğer durum ise bir maddenin birden fazla faktörde yüksek faktör yük değeri almasıdır. Bu durum aynı maddenin birden fazla faktörle ilişkili olduğu anlamına gelip maddenin ölçekten çıkarılmasını gerektiren bir durumdur. Literatürde bir maddenin birden fazla faktörde yüksek faktör yükü alması durumunda faktör yük değerleri arasındaki farkın ne olması gerektiğine ilişkin kesin kurallar olmasa da faktör yük değerleri arasındaki fark en az 0,10 olması gerektiği görüşü benimsenmektedir (Bishop vd., 2013; Viapiana vd., 2016).

Likert Tipi Ölçeklere Yöneltilen Eleştiriler

Likert tipi ölçeklere yöneltilen eleştirilerden biri cevap seçeneklerinin eşit aralıklı olmadığını savunanlar tarafından yapıldığından (Chakrabartty, 2014: 32; Hodge & Gillespie, 2003; Jamieson, 2004; Pett, 1997) Likert tipi ölçeklerden yararlanarak çalışmalar yapan araştırmacılar, cevap seçeneklerinin diğer bir ifadeyle etiketleme seçeneklerinin belirlenmesinde çok dikkatli ve titiz davranmalıdırlar. Bu görüşün savunucuları bu tür ölçeklerden elde edilen veriler üzerinde parametrik testlerin yapılmasının mümkün olamayacağını ifade etseler de araştırmacılar, bu tür ölçeklerden elde edilen veriler üzerinde ölçekler iyi kurgulandığı ve cevap seçeneklerine doğru etiketleme yapıldığı takdirde parametrik testlerin yapılmasını mümkün olduğunu ifade etmektedirler (Batterton & Hale, 2017: 33; Norman, 2010; Pornel & Saldana, 2013: 12-13; Willits vd., 2016).

Likert tipi ölçeklere yöneltilen bir diğer eleştiri de tam bir orta noktası bulunmaması nedeni ile dördü, altılı ve sekizli gibi cevap seçeneklerinden oluşan ölçeklere yöneltilmektedir (Chakrabartty, 2014: 32; Krosnick & Fabrigar, 1997: 141). Öyle ki bazı araştırmacılar bu tür Likert ölçeklerin tam bir orta noktası bulunmadığını ifade etmektedirler (Chyung vd., 2017: 17). Bunun sonucunda cevap seçenekleri arasında orta noktası bulunmayan ölçekler kullanılarak yapılan araştırmalarda araştırma konusunda hiçbir fikri olmayan katılımcılar negatif ya da pozitif yönde görüş belirtmek durumunda kalmaktadır (Nadler vd., 2015: 74). Orta noktası bulunan ölçeklerde ise farklı sorunların görülmesi muhtemeldir. Örneğin katılımcıların yarısının pozitif uç yönde görüş belirttiği, diğer yarısının ise negatif uç yönde görüş belirttiği bir araştırmada katılımcı yanıtları negatif ve pozitif uçta toplandığından

katılımcı yanıtlarının aritmetik ortalaması nötr olarak ortaya çıkacaktır (Chakrabartty, 2014: 32).

Sonuç ve Tartışma

Likert tipinde geliştirilen ölçme araçlarının her ne kadar eksileri olsa da kolaylıkla kurgulanabilmesi, uyarlanabilmesi, ölçeklerden elde edilen verilerin istatistiki işlemlere olanak sağlaması, Likert tipi ölçeklerin veri toplamada sağladığı kolaylıklar ve güvenilir sonuçlar sunmasının bu tip ölçekleri son derece popüler veri toplama araçları haline getirdiği yadsınamaz bir sonuçtur (Li, 2013: 1609).

Ölçek geliştirme, sosyal bilimler alanında çalışan araştırmacılar için oldukça önemli bir süreç olup araştırmacı tarafından dikkatle izlenmesi gereken bir dizi süreçleri içermektedir. İlk olarak ölçek geliştirme aşamasında kapsamlı bir literatür taraması yapılarak ölçekte yer alması düşünülen maddeler için -ilgili literatür göz önünde bulundurularak- madde havuzu oluşturulmalıdır. Bu süreçte araştırmacının konuya ilişkin hem yerli hem de yabancı literatürü taraması, ölçme aracının kapsayıcılığını artıracaktır. Madde havuzunun oluşturulmasından sonra uzman görüşü alınıp uzman önerileri doğrultusunda gerekli düzenlemeler yapılarak pilot uygulamaya geçilmelidir. Pilot uygulama, araştırmacıya geliştirilen ölçek hakkında bilgiler sunan ve gerekli düzenlemeleri yapması için araştırmacıya yol gösteren bir kılavuz olarak düşünülebilir.

Pilot uygulama aşamasında ilk olarak verilerin faktör analizine uygunluğu kontrol edilmelidir. Bu bağlamda toplanan verilerin analize uygunluğunu belirlemede literatürde farklı ölçütler yer alsa da araştırmacıların bu amaçla sıklıkla Kaiser-Meyer-Olkin örneklem yeterliği ve Bartlett'in küresellik testini raporladığı dikkat çekmektedir. Pilot uygulama sürecinde verilerin analize uygunluğu belirlendikten sonra ölçekte birden fazla faktörde yüksek faktör yükü alan madde bulunup bulunmadığı, üçten az madde içeren faktör bulunup bulunmadığı gibi hususlar göz önünde bulundurularak faktör analizi gerçekleştirilerek ölçeğe nihai şekli verilmelidir.

Son olarak faktör analizinden sonra oluşan faktörlerin ve ölçek bütünüün güvenilirlik düzeyleri, benimsenen güvenilirlik analizi yöntemi ile analiz edilmelidir. Likert ölçeklerden yararlanarak çalışmalar yapan araştırmacılarının bu tür ölçeklere yöneltilen eleştirilere ilişkin bilgi sahibi olmasının da ölçek geliştirme sürecinde sergilenecek titizliğe katkı sunacağı düşünülmektedir.

Referanslar

- Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4), 433-459. <https://doi.org/10.1002/wics.101>
- Albaum G. (1997). The Likert scale revisited: An alternate version. *Market Research Society*, 39(2), 331-348.
- Batterton, K. A., & Hale, K. N. (2017). The Likert scale What it is and how to use it. *Phalanx*, 50(2), 32–39. <http://www.jstor.org/stable/26296382>
- Bishop, S. L., Hus, V., Duncan, A., Huerta, M., Gotham, K., Pickles, A., ...Lord, C. (2013). Subcategories of restricted and repetitive behaviors in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 43, 1287-1297. <https://doi.org/10.1007/s10803-012-1671-0>
- Bonett, D. G., & Wright, T. A. (2015). Cronbach's alpha reliability: Interval estimation, hypothesis testing, and sample size planning. *Journal of Organizational Behavior*, 36(1), 3-15. <https://doi.org/10.1002/job.1960>

- Bora Semiz, B. & Altunışık R. (2016). Pazarlama araştırmalarında Likert tipi ölçeklerin özelliklerinin cevaplama tarzları üzerindeki etkilerinin incelenmesi. *Bartın Üniversitesi İİBF Dergisi*, 7 (14), 577-598.
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27-40. <https://doi.org/10.3316/QRJ0902027>
- Bravo, G., & Potvin, L. (1991). Estimating the reliability of continuous measures with Cronbach's alpha or the intraclass correlation coefficient: Toward the integration of two traditions. *Journal of Clinical Epidemiology*, 44(4-5), 381-390. [https://doi.org/10.1016/0895-4356\(91\)90076-L](https://doi.org/10.1016/0895-4356(91)90076-L)
- Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Kuram ve Uygulamada Eğitim Yönetimi*, 32(32), 470-483.
- Chakrabartty, S. N. (2014). Scoring and analysis of Likert scale: Few approaches. *Journal of Knowledge Management and Information Technology*, 1(2), 31-44.
- Christmann, A., & Van Aelst, S. (2006). Robust estimation of Cronbach's alpha. *Journal of Multivariate Analysis*, 97(7), 1660-1674. <https://doi.org/10.1016/j.jmva.2005.05.012>
- Chyung, S. Y., Roberts, K., Swanson, I., & Hankinson, A. (2017). Evidence-based survey design: The use of a midpoint on the Likert scale. *Performance Improvement*, 56(10), 15-23. <https://doi.org/10.1002/pfi.21727>
- Cronbach, L. J., & Meehl, P., E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-352. <https://doi.org/10.1037/h0040957>
- Cudeck, R. (2000). Exploratory factor analysis. In *Handbook of applied multivariate statistics and mathematical modeling*, 265-296. <https://doi.org/10.1016/B978-012691360-6/50011-2>
- DeCoster, J. (1998). Overview of factor analysis. Retrieved from <http://www.stat-help.com/notes.html>
- DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation*, 14(20), 1-11. <https://doi.org/10.7275/da8t-4g52>
- Edmondson, D. (2005, May). Likert scales: A history. In *Proceedings of the Conference on Historical Analysis and Research in Marketing*, 12, pp. 127-133.
- Fabrigar, L. R., & Wegener, D. T. (2011). *Exploratory factor analysis*. Oxford University Press.
- Floyd, F., J., Widaman, K., F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7(3), 286-299. <https://doi.org/10.1037/1040-3590.7.3.286>
- Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education.
- Goodboy, A. K., & Martin, M. M. (2020). Omega over alpha for reliability estimation of unidimensional communication measures. *Annals of the International Communication Association*, 44(4), 422-439. <https://doi.org/10.1080/23808985.2020.1846135>

- Hodge, D. R., & Gillespie, D. (2003). Phrase completions: An alternative to Likert scales. *Social Work Research*, 27(1), 45–55. <https://doi.org/10.1093/swr/27.1.45>
- Hoşgörür, V. (1997). Bogardus, Guttman ve likert ölçekleri. *Ondokuz Mayıs University Journal of Education Faculty*, 10(1), 346-357.
- Huot, B. (1990). Reliability, validity, and holistic scoring: What we know and what we need to know. *College Composition and Communication*, 41(2), 201–213. <https://doi.org/10.2307/358160>
- Iacobucci, D., & Duhachek, A. (2003). Advancing alpha: Measuring reliability with confidence. *Journal of Consumer Psychology*, 13(4), 478–487. https://doi.org/10.1207/S15327663JCP1304_14
- Jamieson, S. (2004). Likert scales: How to (ab) use them?. *Medical Education*, 38(12), 1217-
<https://doi.org/10.1111/j.1365-2929.2004.02012.x>
- Jebb, A. T., Ng, V., & Tay, L. (2021). A review of key Likert scale development advances: 1995–2019. *Frontiers in Psychology*, 12, 637547. <https://doi.org/10.3389/fpsyg.2021.637547>
- Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. *British Journal of Applied Science & Technology*, 7(4), 396. <https://doi.org/10.9734/BJAST/2015/14975>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31–36. <https://doi.org/10.1007/BF02291575>
- Karakoç, F., Y. & Dönmez, L. (2014). Ölçek geliştirme çalışmalarında temel ilkeler. *Tıp Eğitimi Dünyası*, 13(40), 39-49.
- Kelley, T. L. (1927). Interpretation of educational measurements. New York: World Book Company.
- Krosnick, J. A. (2018). Questionnaire design. In: Vannette, D., Krosnick, J. (eds) The Palgrave Handbook of Survey Research. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-319-54395-6_53
- Krosnick, J. A. ve Fabrigar, L. R. (1997). Designing rating scales for effective measurement in surveys. In L. Lyberg, P. Biemer, M. Collins, L. Decker, E. DeLeeuw, C. Dippo, N. Schwarz, and D. Trewin (Eds.), Survey Measurement and Process Quality. New York: Wiley-Interscience
- Krosnick, J. A., & Presser, S. (2010). Question and questionnaire design. In J. D. Wright & P. V. Marsden (Eds.), Handbook of Survey Research (Second Edition). West Yorkshire, England: Emerald Group.
- Locke, E. A. (2012). Construct validity vs. concept validity. *Human Resource Management Review*, 22(2), 146-148. <https://doi.org/10.1016/j.hrmr.2011.11.008>
- Li, Q. (2013). A novel Likert scale based on fuzzy sets theory. *Expert Systems with Applications*, 40(5), 1609-1618. <https://doi.org/10.1016/j.eswa.2012.09.015>
- Nadler, J. T., Weston, R., & Voyles, E. C. (2015). Stuck in the middle: The use and interpretation of mid-points in items on questionnaires. *The Journal of General Psychology*, 142(2), 71-89. <https://doi.org/10.1080/00221309.2014.994590>

- Nemoto, T., & Beglar, D. (2014). Developing Likert-scale questionnaires. In N. Sonda & A. Krause (Eds.), *JALT2013 Conference Proceedings*. Tokyo: JALT
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Adv in Health Sci Educ*, 15, 625–632. <https://doi.org/10.1007/s10459-010-9222-y>
- Pett, M. A. (1997). *Nonparametric statistics for health care research: Statistics for small samples and unusual distribution*. London: SAGE Publication.
- Pituch, J. P., & Stevens, K. A. (2016). *Applied Multivariate Statistics for the Social Sciences*. Taylor & Francis.
- Pornel, J. B., & Saldaña, G. A. (2013). Four common misuses of the Likert scale. *Philippine Journal of Social Sciences and Humanities*, 18(2), 12-19.
- Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104, 1-15. [https://doi.org/10.1016/S0001-6918\(99\)00050-5](https://doi.org/10.1016/S0001-6918(99)00050-5)
- Rummel, R. J. (1967). Understanding factor analysis. *Journal of Conflict Resolution*, 11(4), 444-480. <https://doi.org/10.1177/002200276701100405>
- Raubenheimer, J. (2004). An item selection procedure to maximize scale reliability and validity. *SA Journal of Industrial Psychology*, 30(4), 59-64.
- Schmitt, N. W., & Klimoski, R. J. (1991). *Research methods in human resources management*. Cincinnati: South-Western Publishing.
- Shrestha, N. (2021). Factor analysis as a tool for survey analysis. *American Journal of Applied Mathematics and Statistics*, 9(1), 4-11. <https://doi.org/10.12691/ajams-9-1-2>
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika* 74, 107. <https://doi.org/10.1007/s11336-008-9101-0>
- Spiliotopoulou, G. (2009). Reliability reconsidered: Cronbach's alpha and paediatric assessment in occupational therapy. *Australian Occupational Therapy Journal*, 56(3), 150-155. <https://doi.org/10.1111/j.1440-1630.2009.00785.x>
- Symeonaki, M., Michalopoulou, C. & Kazani, A. (2015). A fuzzy set theory solution to combining Likert items into a single overall scale (or subscales). *Qual Quant* 49, 739–762. <https://doi.org/10.1007/s11135-014-0021-z>
- Thanasegaran, G. (2009). Reliability and validity issues in research. *Integration & Dissemination*, 4, 35-40.
- Warrens, M. J. (2015). Some relationships between Cronbach's alpha and the Spearman-Brown formula. *Journal of Classification*, 32(1), 127-137. <https://doi.org/10.1007/s00357-015-9168-0>
- Williams, B., Onsmann, A., & Brown, T. (2010). Exploratory factor analysis: A five step guide for novices. *Journal of Emergency Primary Health Care*, 8(3), 1–13. <https://doi.org/10.33151/ajp.8.3.93>
- Willits, F. K., Theodori, G. L., & Luloff, A. E. (2016). Another look at Likert scales. *Journal of Rural Social Sciences*, 31(3), 126-139. <https://egrove.olemiss.edu/jrsv/vol31/iss3/6/>

- Vagias, W. M. (2006). Likert-type scale response anchors. *Clemson International Institute for Tourism & Research Development, Department of Parks, Recreation and Tourism Management. Clemson University.*
- Vaske, J. J., Beaman, J., & Sponarski, C. C. (2017). Rethinking internal consistency in Cronbach's alpha. *Leisure Sciences, 39(2), 163-173.*
<https://doi.org/10.1080/01490400.2015.1127189>
- Viapiana, V.F., de Mendonça Filho, E.J., Fonseca, R.P., Giacomoni, C. H. & Stein, L. M. (2016). Development of the arithmetic subtest of the school achievement test-second edition. *Psicol. Refl. Crít., 29:39.* <https://doi.org/10.1186/s41155-016-0045-5>
- Voss, K. E., Stem, D. E., & Fotopoulos, S. (2000). A comment on the relationship between coefficient alpha and scale characteristics. *Marketing Letters, 11(2), 177-191.*
<https://doi.org/10.1023/A:1008146924781>
- Yong, A. G., & Pearce, S. (2013) A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology, 9(2), 79-94.*
<https://doi.org/10.20982/tqmp.09.2.p079>