



Invited Review / Davetli Derleme

Scale Development in Behavioral Sciences (Based on Exploratory Factor Analysis)

Davranış Bilimlerinde Ölçek Geliştirme (Açıklayıcı Faktör Analizine Dayalı)

İrfan Yurdabakan*¹, Sait Çüm²

ÖZET

Davranış bilimlerinde, herhangi bir özelliğin ölçülmesi söz konusu olduğunda iki temel yaklaşım öne çıkmaktadır. Bu yaklaşımlar, psikometrik ve izlenimci yaklaşımdır. Psikometrik yaklaşım, bilimsel yöntemlere dayanarak geçerliği ve güvenilirliği yüksek ölçme araçlarının geliştirilmesine önem verir, aynı zamanda yol gösterir. Bu makalede, davranış bilimlerinde ölçek geliştirme süreci, karşılaşılabilecek sorunlar, izlenmesi gereken adımlar ve kullanılabilecek istatistiksel yöntemler ayrıntılı olarak ele alınmış, ölçek boyutlarını keşfetmek amacıyla başvuru olan açıklayıcı faktör analizi (AFA) tekniğine yer verilmiştir. Makalenin, ölçek geliştirmek isteyen araştırmacılara yol göstereceği umulmaktadır.

Anahtar kelimeler: Ölçme, ölçek geliştirme, açıklayıcı faktör analizi

ABSTRACT

In the measurement of a trait in behavioral sciences two basic approaches stand out: Those are psychometric and impressionistic approaches. Psychometric approach places emphasis on development of measurement instruments based on scientific methods with high validity and reliability and is guiding at the same time. In this study, scale development process in behavioral sciences, problems encountered, steps to be taken and statistical methods to be used are discussed in detail and exploratory factor analysis (EFA) as a technique used in search of dimensions of a scale is mentioned. This study is anticipated to guide the researchers in scale development process.

Keywords: Measurement, scale development, exploratory factor analysis

Received / Geliş tarihi: 05.05.2017, Accepted / Kabul tarihi: 17.05.2017

¹Dokuz Eylül Üniversitesi, Buca Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı

²Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü Eğitimde Ölçme ve Değerlendirme Anabilim Dalı

*Address for Correspondence / Yazışma Adresi: İrfan Yurdabakan, Dokuz Eylül Üniversitesi, Buca Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, İzmir-TÜRKİYE

E-mail: irfan.ybakan@gmail.com

Yurdabakan İ, Çüm S. Davranış Bilimlerinde Ölçek Geliştirme (Açıklayıcı Faktör Analizine Dayalı). TJFMPC, 2017;11(2): 108-126.

DOI: 10.21763/tjfmpe.317880

GİRİŞ

Bilimsel disiplinlerin, tarihsel gelişim süreçleri boyunca ileri sürülen farklı kuramlar ölçme ve değerlendirme konusunda farklı yaklaşımların ortaya çıkmasına neden olsa da, bilimin gereği olarak bilimsel çalışmaların tümünde ölçme yapmak temel bir zorunluluktur. Davranış bilimlerinde, belirli amaçlarla (yönetmel, öğretime ilişkin, rehberlik, araştırma, tanılama vb.) herhangi bir özelliğin ölçülmesi söz konusu olduğunda, iki temel yaklaşım öne çıkmaktadır: **psikometrik** ve **izlenimci** yaklaşım.¹ Bu yaklaşımlar, yapılan değerlendirmelerde, sürece veya sonuca odaklanma, kişilerin incelemeye konu olan niteliklerinin derinlemesine veya yüzeysel incelenmesi, nesnellikten ödün verilip verilmemesi gibi konular üzerinde farklı görüşleri içlerinde barındırır. Psikometrik yaklaşımın, izlenimci yaklaşıma göre daha objektif ve karşılaştırılabilir bilgiler sağladığı ileri sürülmektedir.² Bu nedenle, farklı alanlarda yapılan araştırmalarda psikometrik yaklaşımın etkileri yaygın olarak gözlemlenmektedir. Psikometrik yaklaşım doğrultusunda geliştirilen ve araştırmalarda kullanılan ölçme araçlarının, diğer bilgi toplama yöntemlerine göre daha kolay uygulanabilmesi, objektif olarak puanlanabilmesi, geçerli ve güvenilir gözlemler yapmaya olanak sağlaması gibi nedenlerle bu yaklaşımın sıklıkla tercih edildiği söylenebilir.¹

Anastasi'ye³ göre, ölçekler genellikle bireyler arasındaki farklılıkları veya aynı bireylerin farklı durumlarda gösterdikleri tepkiler arasındaki farklılıkları ölçmek amacıyla kullanılmaktadır. McRaynolds'a⁴ göre ise ölçme, sadece bireyler arası farklılıklarla ilgilenmez, bunun yanı sıra birey ile ilgili olan farklı değişkenler arasındaki ilişkiler ve belirli bir süre içerisinde bireyde meydana gelen değişimler de psikolojik ölçmenin konusudur. Zeka düzeyinin tanımlanabilmesi amacıyla ortaya çıkan ihtiyaç, psikolojik ölçmenin geliştirilmesini gerekli kılan nedenlerden birisi olmuştur. Ölçeklerin zihinsel düzeyleri belirleme konusundaki işlevi, günümüze kadar önemli bir işlev olma özelliğini korumuştur. Kliniklerde, olağan dışı davranışlar üzerinde yapılan çalışmalar, eğitim ortamlarında ortaya çıkan sorunlar, endüstriye ve orduya personel seçimi gibi farklı alanlarda meydana gelen ihtiyaçlar psikolojik ölçme araçlarının geliştirilmesini gerekli kılmıştır.⁵

Günümüzde kliniklerde, objektif veya projektif kişilik ölçekleri, psikolojik rahatsızlıkları tanılayıcı ölçekler, zeka testleri ve beyin hasarlarını saptama süreçlerinde nöropsikolojik test bataryaları, psikologlar tarafından yaygın bir

şekilde kullanılmaktadır. Davranış bilimlerinde, ölçmenin tarihçesinde yer alan önemli olaylar tarihleri ile birlikte Tablo 1'de sunulmuştur.⁶

Tablo 1. Davranış bilimlerinde ölçmenin tarihçesinde dönüm noktaları

Tarih	
M.Ö. 1000	Çin'de kamu sektöründe testlerin kullanılması
1850-1900	Amerika Birleşik Devletleri'nde kamu sektöründe testlerin kullanılması
1900-1920	Bilişsel yeteneğin ölçülmesi amacıyla birey ve grup testlerinin geliştirilmesi, psikometrik teorinin geliştirilmesi
1920-1940	Faktör analizinin geliştirilmesi, projektif ölçme araçlarının geliştirilmesi ve standart kişilik envanterlerinin geliştirilmesi
1940-1960	Mesleki ilgi envanterlerinin geliştirilmesi, psikopatoloji üzerine ölçme girişimleri
1960-1980	Madde tepki kuramının geliştirilmesi, nöropsikolojik testlerin geliştirilmesi

Davranış Bilimlerinde Ölçme Süreçlerinde Karşılaşılan Bazı Sorunlar

Bilim dalları arasında birçok benzerlik olsa da, davranış bilimlerinde yapılan ölçmelerde fiziksel bilimlerde karşılaşılmayan bazı sorunlarla karşılaşılabilir. Bu sorunlardan bazıları, psikolojik kavramların karmaşıklığından kaynaklanmaktadır. Zeka, öz-saygı, kaygı, depresyon, vb. birçok psikolojik kavramın nasıl tanımlanacağı konusunda farklı görüşler bulunmaktadır. Psikolojik kavramların karmaşıklığının yanı sıra, yanıtlayıcıların kendilerini yansıtmayan tepkileri de bu değişkenlerin ölçülmesinde karşılaşılan sorunlardandır. Örneğin, bazı yanıtlayıcılar, ölçeceği yanıtladıkları sırada çok fazla gergin veya heyecanlı olabilirler ve bu durum normalden farklı yanıtların verilmesine yol açabilir. Bazı yanıtlayıcılar ise, var olan düşüncelerini yansıtmak yerine, diğer kişiler tarafından daha çok takdir edilen yanıtlar vererek, ölçmeyi yapan kişiye yanıltıcı bilgiler verebilirler. Bu duruma **sosyal**

beğenilirlik etkisi adı verilir. Örneğin, ırkçı düşüncelere sahip olan birisi gerçek düşüncelerini yansıtmak yerine, davranışlarını toplumun kendisine tepki göstermeyeceği şekilde değiştirebilir. Bazı yanıtlayıcılar ise, olduklarından daha kötü bir durumu yansıtmaya eğiliminde olabilirler. Örneğin, suç işleyen bir kişi cezasını hafifletmek umuduyla akli dengesinin yerinde olmadığı izlenimi vermeye çalışabilir.⁷ Ayrıca, yanıtlayıcılarda gözlemlenen davranışlar zamana ve koşullara bağlı olarak değişiklik gösterebilir. Örneğin, şartlı tahliye edilen şiddete eğilimli bir mahkûm, şartlı tahliye memuruna karşı bu eğilimini bastırarak kendini çok farklı göstermesine rağmen, arkadaşlarına ve ailesine karşı daha şiddet yanlısı davranışlar sergileyebilir. Başka bir deyişle, kişi yakın çevresine oldukça sıkıcı ve aciz bir insan olarak görülebilirken, iş arkadaşları arasında çekici ve aşırı derecede etkileyici birisi olarak görülebilir.⁸

Gözlemlenen davranışların, her zaman bireyin ölçülen özelliğinin derecesine ilişkin doğru bilgi vermeyeceğine dair, sosyolog La Piere tarafından 1934 yılında ABD’de yapılan bir araştırmayı Tavşancıl⁹ şöyle aktarmaktadır. “La Piere, genç bir Çinli karı-kocayla birlikte 66 otel ve motele ve 184 lokantaya gitmiştir, her gittikleri yerde kabul edilmişler ve kendilerine servis yapılmıştır. Yalnızca üçüncü sınıf bir motele kabul edilmemişlerdir. Bu ilk deneyden altı ay sonra La Piere, bütün gittikleri otel, motel ve lokantaya mektup yazarak Çinli müşteri kabul edip etmeyeceklerini sormuş ve mektupların yarısına cevap almıştır. Cevap yazanların %92’si Çinlilerin kabul edilemeyeceğini bildirmiştir. Bu araştırmayla, daha önce gösterilen olumlu davranış ile daha sonra belirtilen olumsuz tutum arasında bir çelişki ortaya konmuştur.” Kişilerin içinde bulunduğu koşulların ve farklı durumların, kişinin sergilediği davranışların oluşumunda önemli bir etkisinin olduğu unutulmamalıdır. Psikolojik yapılar üzerinde değerlendirme yapacak olan kişinin, değerlendirilenlerin içinde bulunduğu ortam ve koşullar, yaşadığı toplum veya çevre, mesleki deneyimler ve baskı altında olup olmama gibi faktörleri göz önünde bulundurması gerekmektedir.⁸

Yanıtlayıcıdan kaynaklanan bir diğer sorun ise, yanıtlayıcıların rastgele yanıtlar vermesi durumudur. Bu sorun, yanıtlayıcıların yeterince motive olmadığı ölçme ortamlarında veya yanıtlayıcıların karşılaştıkları görevleri yerine getirmekte yetersiz kaldıkları durumlarda ortaya çıkabilir. Kişilerin, ölçeceği yanıtlamak konusunda isteksiz olduğu durumlarda da rastgele yanıtlama sorunuyla karşılaşılabilir. Örneğin, disiplin suçu işlediği gerekçesiyle rehberlik servisine

yönlendirilen bir öğrenci, rehberlik uzmanı ile iş birliği yapmak istemeyebilir ve bu durumda kendisine verilen ölççekleri istekli olarak yanıtlamayabilir.⁶ Bu sorun ile karşılaşmamak için, ölççekler gönüllü yanıtlayıcılara uygulanmalıdır. Yanıtlayıcılar yeterince motive edilmeli ve gerekli durumlarda çeşitli özendiriciler ile yanıtlama istekleri artırılmalıdır.

Davranış bilimlerinde, veri toplayan kişiden kaynaklı olan sorunlar da ortaya çıkabilmektedir. Araştırmacılar, veri toplarken yanlı davranabilirler ve araştırma verilerini kendi beklentileri doğrultusunda düzenleyebilirler. Örneğin, öğrencilerin bilişsel becerilerini geliştirmesi beklenen bir eğitimden sonra araştırmacı, öğrencilere bir zeka testi uygulayarak verilen eğitimin işe yarayıp yaramadığını öğrenmek istiyor olabilir. Böyle bir durumda, verilen eğitimin yararlı olacağı konusunda beklentileri olan bir araştırmacı, veriler üzerinde bilinçli olarak çarpıtma yapabilir. Araştırmacılar, bilinçsiz olarak da benzer hatalar yapabilirler. Araştırma verilerinde, bilinçli veya bilinçsiz olarak yapılan bu tür hataları tespit etmek oldukça zordur. Davranış bilimlerindeki ölçme süreçlerinde karşılaşılan bir diğer sorun ise, kullanılan ölçme araçlarının araştırmacı tarafından psikometrik nitelikleri dikkate alınmadan seçilmiş olmasıdır. Psikometrik nitelikleri zayıf olan ölççeklerle elde edilen veriler, araştırmacıya yeterince doğru bilgi sunmaz. Bu yüzden, davranış bilimlerinde üzerinde çalışmalar yapan araştırmacıların, ölççek geliştirme veya yaptıkları araştırma için doğru ölççekleri seçme konusunda bilgi sahibi olmaları gerekmektedir.⁷

Bir ölççeğin bilim çevreleri tarafından kabul görebilmesi için, aracın belirli standartlara uygun olarak geliştirilmesi gerekmektedir. Gelişigüzel üretilmiş maddeler, bilimsel araştırmalar için uygun değildir.¹⁰ Ölçek geliştirmek isteyen araştırmacıların, hem ölçmeye çalıştıkları değişkenin yapısı ve değişkenin kavramsal-kuramsal çerçevesine ilişkin yeterli alan bilgisine sahip olması, hem de söz konusu değişkenin nasıl ölçüleceği ve ölçme bilimi alanında yeterli düzeyde bilgisinin olması gerekmektedir.^{8,11} Burada sözü edilen yapı, psikolojik özelliğin varlığını açıklayan bir kavramdır. Yapılar görülemez, duyulamaz veya koklanamazlar fakat varlıkları gözlenen davranışlar yoluyla anlaşılabilir. İngilizcede, binlerce psikolojik özellik terimi bulunmaktadır. Bu terimler arasında zekâ, özel yetenek, bilişsel stil, ilgi, tutum, tercihler, psiko-patoloji, kişilik gibi kavramlar yer almaktadır. Psikolojik bir özelliği ölçmek isteyen araştırmacı, önce ölçmeyi amaçladığı özelliği dikkatlice tanımlamalıdır. Örneğin, saldırganlık

üzerine çalışma yapılmak istenilsin, günlük hayatta birçok farklı ortamda saldırgan kelimesini duymak mümkündür. Saldırgan kavramı, saldırgan satış temsilcisi, saldırgan katil, saldırgan patron, saldırgan çocuk, saldırgan futbolcu gibi kullanıldığı farklı yerlerde farklı anlamlar taşıyabilir. Eğer bu yapı ile ilgili bir ölçek geliştirilecek olursa elde edilecek puanlar, test geliştiricinin özelliği nasıl tanımladığına bağlı olarak yorumlanacaktır. Eğer yapı net olarak tanımlanmamışsa, hem yapının ölçülmesinde hem de elde edilecek puanların yorumlanmasında büyük hatalar yapılabilir⁸. Ölçülecek yapıyı bilmeden ve ölçme sürecinde uyulması gereken teknik ayrıntıları dikkate almadan, geliştirilip veri toplanan araştırmaların sonuçları hatalı olmakta ve bilgi kirliliğine yol açmakta ayrıca, harcanan emek ve zaman boşa gitmektedir. Çünkü bir değişkenin ölçüleri ile ilgili herhangi bir kuşku varsa, kuşkulu ölçülere dayalı olarak incelenen ilişki ve fark testlerinden elde edilen sonuç ve bunlara dayalı olarak yapılan yorumlar da kuşkulu olacaktır⁵. Ayrıca geliştirilme sürecinde teknik hataların yapıldığı ölçekler, daha sonra başka araştırmacılar tarafından psikometrik özellikleri incelenmeksizin kullanılabilir. Bu ölçekler kanalıyla elde edilen verilere dayanarak, bireyler hakkında önemli kararlar verilebilmekte ve bu kararlar sonucunda bireylerin yaşamlarını olumsuz yönde etkileyen durumlar ortaya çıkabilmektedir. Bu makalede, ölçek geliştirme adımlarının detaylı şekilde ele alınması ve sürecin hatasız ilerletilebilmesi bakımından önemli noktaların vurgulanması amaçlanmıştır.

ÖLÇEK GELİŞTİRME

Araştırmacılar, incelemek istedikleri yapı ya da yapılara ilişkin ölçme yapmak istediklerinde geçerli ve güvenilir bir ölçme yapabilmeye olanak tanıyacak bir ölçme aracına ihtiyaç duyarlar. Bu durumda, eğer kendi kültürlerine uygun ve kendi dillerinde geliştirilmiş psikometrik nitelikleri güçlü bir ölçme aracı yoksa ortaya iki seçenek çıkar. Bunlar, yeni bir ölçek geliştirmek ya da başka dilde geliştirilmiş bir ölçeği hedef kültüre uyarlamaktır. Uyarlama çalışmaları daha kolay gibi görünmekle birlikte, bir ölçeği geliştirirken izlenen basamakların ve yürütülen işlemlerin birçoğu, hatta neredeyse hemen hepsi ölçek uyarlaması için de gerekli ve zorunludur. Hatta bazı durumlarda, geliştirilme sürecinde yapılan çalışmalardan daha fazlasının yapılması gerekli olabilmektedir. Uyarlama aşamaları, ölçeklerdeki ifadelerin eş anlama gelip gelmediğinin denetlenmesinden başlayarak ölçeğin içyapısında ve psikometrik özelliklerinde değişmeler olup olmadığının ortaya çıkarılmasına kadar giden bir

dizi anlamsal, kültürel ve istatistiksel işlemi içermektedir¹². Buna karşın, kültürlerarası karşılaştırma yapmak gibi bir amacın olması durumunda ölçek uyarlamak araştırma için daha uygun olabilir. Diğer yandan, uyarlama çalışmalarında amaç kültürlerarası karşılaştırma yapmak değilse yeni bir ölçek geliştirmek, uyarlama sırasında karşılaşılabilecek birçok sorunu ortadan kaldırması bakımından daha uygun olacaktır¹³. Türkiye’de psikoloji ve eğitim bilimleri dergilerinde, 2005-2013 yılları arasında yayımlanan 121 makale üzerinde yaptıkları araştırmada¹⁴, uyarlama çalışmalarının büyük çoğunluğunun kültürlerarası karşılaştırma amacı taşımadığı belirlenmiştir. Bu tür bir amacın olmamasına karşın, ölçek uyarlamayı daha cazip kılan etmenlerden biri olarak araştırmacıların madde yazımı konusundaki çekinceleri gösterilebilir. Ölçek maddesi oluşturmanın, uzmanlık gerektiren bir iş olduğu düşünüldüğünde haklı bir çekince olarak değerlendirilebilecek bu sorunun ölçme ve değerlendirme uzmanları ile alan uzmanlarının ekip çalışmasıyla aşılabileceği unutulmamalıdır. Ölçek geliştirmek, bireysel uğraşının ötesinde iş bölümü gerektiren daha kapsamlı bir organizasyondur. Dünya üzerinde yayımlanmış birçok ölçek bulunmaktadır. Sadece İngilizce yayımlanmış 5000’den fazla ölçek vardır, fakat bunlardan çok azı sıklıkla kullanılmaktadır¹⁰. Bilim çevrelerinde kabul gören ve araştırmalarda yaygın olarak kullanılan bir ölçeği uyarlamak pek çok bakımdan avantaj sağlayabilir ve hedef kültürdeki çalışmalara bilimsel katkı sunabilir. Ancak, yeterince kabul görmemiş ve kullanımı yaygınlaşmamış bir ölçeği, üstelik hedef kültüre yeterince uyum göstermemesine rağmen çeşitli yapısal müdahalelerle uyarlamaya çalışmak doğru ölçümler elde edebilmek bakımından ciddi sorunları doğurabilir. Bunun yerine belirli standartlara ve gerekli adımlara uygun şekilde uzman bir ekiple birlikte ölçek geliştirmenin daha doğru olduğu iddia edilebilir.

Ölçek Geliştirme Adımları

Ölçek geliştirme süreci, birbirini takip eden çeşitli adımlar altında özetlenebilir. Turgut¹⁵, ölçek ve test geliştirme sürecini 16 adımda ele almıştır. Bu adımlar aşağıda öncelik sırasına göre verilmiştir.

- Ölçeğin amacının belirlenmesi
- Ölçeğin kapsamının belirlenmesi
 - *Hangi davranışların ölçüleceğinin kararlaştırılması*
 - *Bu davranışların ölçülmesinde hangi malzemeden ya da içerikten yararlanılacağına belirlenmesi*

- *Davranış-içerik kesişimlerinin bağlı ağırlıklarının neler olacağıının belirlenmesi*
- Belirtkelerin, maddelerle örneklenme yönteminin belirlenmesi
- Kullanılacak madde türünün ve bu türlere göre madde sayısının belirlenmesi
- Seçilen türde veya türlerde maddelerin üretilmesi
- Üretilen maddelerin incelenerek geliştirilmesi
- Ölçeğin deneme formunun düzenlenmesi (konu birliği sağlama, kolaydan zora sıralama vb.)
- Puanlama yönteminin belirlenmesi
- Ölçeğin deneme formunun, yönergelerinin yazılması, baskıya hazırlanması ve çoğaltılması
- Deneme uygulaması örnekleminin belirlenmesi
- Deneme uygulamasının planlanması (yer, zaman, uygulayıcılar vb.)
- Deneme uygulamasının öngörülen koşullarda gerçekleştirilmesi
- Test ve madde analizlerinin yapılması
- Analiz sonuçlarının yorumlanması ve bunlar ışığında madde seçimi
- Seçilen ve geliştirilen maddelerden ölçeğin oluşturulması
 - *Sürenin ve uygulama koşullarının belirlenmesi*
 - *Yönergelere son şeklinin verilmesi*
 - *Seçilen maddelerin ölçek düzenine sokulması*
 - *Son şekliyle ölçeğin ortalama, standart sapma gibi istatistiklerinin ve güvenilirlik, geçerlik gibi teknik özelliklerinin tahmin edilmesi.*
- Ölçeğin çoğaltılması
- Eğer ölçek standardize edilecekse, yeni bir uygulamanın planlanması gerçekleştirilmesi ve normların çıkarılması
- Ölçeğin el kitabının hazırlanması

Crocker ve Algina⁵, ölçek geliştirme sürecini 10 adımda ele almışlardır. Bu adımlar aşağıda öncelik sırasına göre verilmiştir.

- Ölçek puanlarının hangi amaçla kullanılacağıının belirlenmesi
- Psikolojik yapıyı temsil eden davranışların tanımlanması
- Belirlenen davranışlar doğrultusunda belirtke tablosunun hazırlanması
- Madde havuzunun oluşturulması

- Maddelerin gözden geçirilmesi ve gerekli düzenlemelerin yapılması
- Küçük bir grup üzerinde, ön deneme uygulamasının yapılması ve gerekli düzenlemelerin yapılması
- Ölçeğin hedef kitleyi temsil eden daha büyük bir örneklem üzerinde uygulanması
- Elde edilen veriler üzerinde istatistiksel analizlerin yapılması ve uygun olmayan maddelerin ölçekten çıkartılması
- Ölçeğin geçerlik ve güvenilirlik çalışmalarının yapılması
- Standardizasyon çalışmalarının yapılması

Rust ve Golombok¹⁶, ölçek geliştirme sürecini 9 adımda ele almışlardır. Bu adımlar aşağıda sırasıyla verilmiştir.

- Ölçeğin amacının belirlenmesi
- Taslak hazırlanması
 - *Ölçeğin kapsamının belirlenmesi*
- Maddelerin oluşturulması
- Ölçeğin düzenlenmesi
 - *Yanıtlayıcı bilgi formunun oluşturulması*
 - *Yönergelerin yazılması*
- Deneme uygulamasının gerçekleştirilmesi
- Madde analizlerinin yapılması
- Güvenirlik çalışmalarının yapılması
- Geçerlik çalışmalarının yapılması
- Standardizasyon çalışmalarının yapılması

Murphy ve Davidshofer⁶'e göre ölçek geliştirme süreci, üç ayrı ana başlıkta ele alınmalıdır. Bu başlıklar, **ölçeğin oluşturulması**, **ölçeğin standardizasyonu** ve ölçeğin geliştirilmesinden belli bir zaman sonra, güncelliğini kaybetmesi sonucu yeniden gözden geçirilmesi sürecini kapsayan, **ölçeğin güncellenmesi** şeklinde adlandırılmaktadır. Aşağıda bu ana başlıklar ve alt başlıkları özetlenerek sunulmuştur.

- Ölçeğin oluşturulması
 - *Ölçülmesi amaçlanan değişkenin davranışsal göstergelerinin belirlenmesi*
 - *Madde türünün belirlenmesi*
 - *Madde üretimi*
 - *Ölçeğin deneme uygulamasının yapılması ve madde seçimi*
- Ölçeğin standardizasyonu
 - *Standardizasyon örnekleminin belirlenmesi*
 - *Ölçeğin normlarının oluşturulması*

- Geçerlik ve güvenilirlik çalışmalarının yapılması
- Ölçeğin el kitabının hazırlanması
- Ölçeğin güncellenmesi
 - Ölçeğin izlenmesi ve gerektiğinde yeniden gözden geçirilmesi

Tezbaşaran¹⁷, tutum ölçeği geliştirme adımlarını 3 ana başlık altında ele almıştır. Bu başlıklar sırasıyla aşağıdaki gibidir:

- Ölçülecek tutumun (özeliğin) tanımlanması
 - Tutum kapsamının belirlenmesi
 - Kapsama uygun gözlenebilir işaretçilerin belirlenmesi
- Ölçeğin deneme formunun düzenlenmesi ve deneme uygulamasının gerçekleştirilmesi
 - Ölçek materyalinin hazırlanması
 - Yönergelerin hazırlanması
 - Maddelerin ölçek içindeki düzeninin belirlenmesi
 - Ön incelemenin yapılması
 - Deneme uygulamasının gerçekleştirilmesi
- Deneme uygulamasından elde edilen verilerin analizinin yapılması
 - Maddelere verilen cevapların puanlanması
 - Bireylerin ölçekten aldığı ham puanların hesaplanması
 - Ham puan dağılımının özelliklerinin belirlenmesi
 - Madde puanları dağılımının özelliklerinin belirlenmesi
 - Madde analizlerinin yapılması

Cohen ve Swerdlik⁸, ölçek geliştirme sürecini 5 temel adımda ele almış ve bu adımları şöyle ifade etmişlerdir:

- Ölçülecek değişkenin kavramsal-kuramsal çerçevesinin belirlenmesi
- Ölçeğin yapılandırılması
 - Ölçekleme tekniğine karar verilmesi
 - Madde türüne karar verilmesi
 - Madde üretimi
- Ölçeğin uygulanması
- Madde analizlerinin yapılması
- Gerekli düzenlemelerin yapılması

Coaley¹⁸, ölçek geliştirme sürecini 8 adımda incelemiştir. Bu adımlar aşağıda öncelik sırasına göre verilmiştir.

- Ölçeğin hangi amaçla geliştirileceğine karar verilmesi
- Ölçülecek değişkenin tanımlanması
- Ölçek planının oluşturulması
- Maddelerin üretilmesi
- Ölçeğin uygulanması
- Maddelerin seçimi
 - Madde analizi
 - Ölçek analizi
 - Faktör analizi
- Güvenirlik ve geçerlik çalışmalarının yapılması
- Standardizasyon çalışmalarının yapılması

Son olarak, Erkuş¹¹, ölçek geliştirme sürecini 12 başlık altında ele almış ve sıralamayı şöyle yapmıştır:

- Ölçeğin hangi amaçla geliştirileceğine karar verilmesi
- Neyin ölçüleceğine karar verilmesi ve tanımlanması
 - Ölçülecek değişkenin kavramsal-kuramsal çerçevesinin çizilmesi ve tanımının yapılması
- Kavramsal tanımın davranışsal göstergelerinin bulunması
- Ölçek geliştirme tekniğine karar verilmesi
- Uyarıcıların ve uygun tepki kategorilerinin üretilmesi
- Maddelerin gözden geçirilmesi ve düzeltilmesi
- Ölçeğin açıklama ve yönergelerinin yazılması ile biçimsel yapısının incelenmesi
- Küçük bir grup üzerinde ölçeğin ön deneme uygulamasının gerçekleştirilmesi
- Deneme uygulamasının gerçekleştirilmesi
- Ölçek ve madde analizi için hazırlık
- Ölçek toplam ham puanları ile madde puanlarının betimsel istatistiklerinin incelenmesi
- Eğer gerekiyorsa tekrar deneme uygulamalarına başvurulması

Ölçek Geliştirme Süreci: Öneriler ve Uyarılar

İlgili alanyazında, ölçek geliştirme sürecinde izlenmesi gereken adımların ve her bir adımda yapılması gereken işlemlerin birbiriyle benzerlik gösterdiği görülmektedir. Ölçek geliştirme süreci, aşağıda detaylı olarak ele alınmış ve bu alanda çalışacak olan araştırmacılara çeşitli öneriler getirilmiştir.

1. Ölçeğin Hangi Amaçla Geliştirileceğine Karar Verilmesi

Ölçek geliştirmenin ilk adımında, araştırmacının kendisine soracağı soru “Geliştireceğim ölçek ne işe yarayacak?” sorusudur. Bu soru, açık ve net olarak yanıtlanmadıkça, ölçekten elde edilen bilgiler araştırmacının elde etmek istediği bilgiler olmayacaktır.¹⁶ Ölçeğin geliştirilme amacı;

a) ölçülmesi hedeflenen özellik

b) bireyler hakkında verilecek karar şeklinde almak gerekir. Bireyler hakkında verilecek kararlar, kurumsal olarak seçme, yerleştirme, tanı gibi amaçları; istatistiksel olarak da sınıflama ve sıralama kararlarını içerir.¹¹

Bu aşamada araştırmacının bazı sorulara yanıt vermesi gerekmektedir.⁸

Bu sorular:

Ölçek hangi yapıyı ölçmek için geliştirilecektir?

Ölçek geliştirici, bu sorunun yanıtı olarak ölçmek istediği psikolojik yapıyı nasıl tanımladığını ve bu geliştirilecek olan testin benzer yapıyı ölçen, önceden geliştirilmiş diğer ölçekler ile benzer ve farklı yönlerinin ne olacağını belirlemelidir.

Ölçek hangi kullanım amacıyla geliştirilecektir?

Ölçek geliştiricinin, ölçeğin hangi amaçla kullanılacağını belirlemesi gerekir. Ayrıca, geliştirilecek olan testin benzer amaçla daha önce geliştirilmiş olan ölçekler ile benzer ve farklı yönlerinin ne olacağını belirlenmelidir.

Yeni bir ölçeğe gereksinim var mıdır?

Ölçek geliştiriciden, geliştirilecek olan ölçeğin daha öncekilerden üstün yönlerini belirlemesi beklenmektedir. Eğer yeni ölçeğin, aynı yapıyı ölçen mevcut ölçeklerden daha üstün yönleri olmayacaksa, geliştirilmesinin de anlamı olmayacaktır.

Geliştirilen ölçeği kimler uygulayabilecekler?

Geliştirilen ölçek, amacına uygun olarak kimler tarafından doğru bir şekilde uygulanabilir sorusunun yanıtı ölçek geliştirici tarafından verilmelidir.

Geliştirilen ölçek kimlere uygulanacak?

Geliştirilen ölçeğin hedef kitlesinin, yaş, eğitim düzeyi, meslek, sosyo-ekonomik düzey, medeni durum gibi değişkenler açısından kimlere yönelik geliştirildiğinin belirlenmesi gerekmektedir.

Ölçekten elde edilen puanlar nasıl yorumlanacaktır?

Ölçeği yanıtlayan bireyin ölçekten elde ettiği puanın, ölçeği yanıtlayan diğer bireylerin ölçekten elde ettikleri sonuçlara göre mi, yoksa daha önceden belirlenen mutlak ölçütlere göre mi, yorumlanacağını ölçeği geliştiren kişi tarafından belirlenmesi gerekmektedir.

2. Ölçülecek Değişkenin Kavramsal-Kuramsal Çerçevesinin Çizilmesi ve İşevuruk Tanımının Yapılması

Bu adımda, ölçeği geliştirenlerin ölçmek istedikleri psikolojik yapı ile ilgili alanyazın taraması yapması ve yapının kavramsal çerçevesini net olarak ortaya koyması beklenmektedir.⁸ Bir ölçme işlemi, ölçülecek özelliğin belirlenmesi ile başlar, tanımlanamayan özellikler ölçülemez. Ölçülmesi amaçlanan değişkenin kuramsal tanımından yola çıkılarak değişkenin, gözlenebilir ve ölçülebilir bir biçimde işevuruk olarak tanımlanması gerekmektedir.¹⁷

Araştırmacı, değişkenin kuramsal çerçevesini oluştururken yapıyı belirginleştirmek için bir veya birden fazla yöntem kullanabilir.⁵ Bu yöntemlerden bazıları aşağıda özetlenmiştir.

İçerik analizi: Bu yöntemde araştırmacı, hedef kitleyi temsil eden bir gruba ölçülmesi amaçlanan yapıyla ilgili açık uçlu sorular sorabilir.

Yapılmış olan araştırmaların incelenmesi: Araştırmacı, üzerinde çalıştığı psikolojik yapı ile ilgili daha önceden yapılan araştırmalarda yapılan tanımları gözden geçirebilir.

Doğrudan gözlemele: Araştırmacı, üzerinde çalıştığı psikolojik yapıyı doğrudan bireyler üzerinde gözlemleyebilir. Örneğin mesleki stres değişkeni üzerinde çalışan bir araştırmacı, doğrudan çalışanları iş saatleri içerisinde gözlemleyebilir.

Uzman görüşü: Araştırmacı, ilgili psikolojik yapı konusunda uzman olan kişilerin görüşlerine başvurabilir.

3. Kavramsal Olarak Tanımlanmış Değişkenin Davranışsal Göstergelerinin Bulunması

Bu adımda, ölçek geliştirenlerin ölçmeyi amaçladıkları değişkenin davranışsal göstergelerini belirlemeleri beklenmektedir. Göstergeler, ortama ve hedef kitleye bağlı olarak şekillenen, tanımlanan kavramın gerçek yaşamdaki somut karşılıkları, yapının temsilcisi

davranışlardır.¹¹ Bir örnek olarak, Cordes ve Dougherty¹⁹ aile yaşamında bozulma, fiziksel ve psikolojik sorunlar yaşama, sigara içme, uyuşturucu kullanma ve aşırı düzeyde alkol tüketme gibi göstergeleri “tükenmişlik” kavramı ile ilişkilendirmiştir. Kavram, “öğretmen tükenmişliği” olarak sınırlandırıldığında ise sık hastalanma, devamsızlık, uygunsuz davranışlar ve düşük öğretmenlik performansı gibi göstergeler de ortaya çıkmaktadır.²⁰ Belirlenen göstergelerin, her birinin bir ölçek maddesine dönüştürüleceği göz önünde bulundurulduğunda bu adımın, ölçülmesi amaçlanan değişkenin miktarının belirlenebilmesi bakımından son derece önemli olduğu ifade edilebilir. Göstergeleri kesin olarak bilinmeyen değişkenlerin ölçülebilmesi de mümkün değildir. Bu tür bir değişkenin ölçülmesi söz konusu olduğunda, öncelikle değişkenin göstergelerinin bulunması amacıyla çeşitli bilimsel araştırmaların yapılması ve hatta araştırmalardan elde edilen sonuçların tartışmaya açılması gerekebilir. Dolayısıyla araştırmacılara, henüz deneysel çalışmalara konu olmamış ve somut davranışlarla bağlantısı kurulmamış olan soyut kavramları ölçme girişiminde bulunmaktan uzak durmaları önerilebilir.

4. Ölçek Geliştirme Tekniğine Karar Verilmesi

Bu adımda, ölçek geliştiricinin, ölçek geliştirme tekniğine karar vermesi beklenmektedir. Ölçek geliştirme teknikleri arasında, Bogardus’un Toplumsal Uzaklık Ölçeği, Thurstone Eşit Görünümlü Aralıklar Ölçeği, Likert Tipi Ölçekleme Tekniği, Osgood Anlamsal Farklılıklar Ölçeği gibi teknikler ön plana çıkmaktadır. Ölçek geliştirme tekniklerinden biri diğerinden daha üstündür veya en iyi ölçek geliştirme tekniği şudur, şeklinde bir belirlemede bulunulamaz. Ölçek geliştirici, ölçek geliştirme tekniğine karar verirken tekniğin ölçülmesi amaçlanan psikolojik yapıya uygunluğunu araştırmalıdır.⁸ Ölçek geliştirme tekniğine karar verilirken, geliştirme kolaylığı da ön plana çıkabilir. Örneğin, bir tutum ölçeği geliştirileceğinde zaman-emek vb. açısından Thurstone tipi ölçek geliştirme yerine, Likert tipi ölçek geliştirme tercih edilebilir.¹¹

5. Uyarıcıların (Maddelerin) ve Uygun Tepki Kategorilerinin Üretilmesi

Burada uyarıcılardan kasıt, ölçek maddeleridir. Madde türleri şöyle sınıflandırılabilir; iki seçeneqli maddeler, çoktan seçmeli maddeler, dereceleme türü maddeler, sıralama gerektiren türdeki maddeler, eşleştirmeli maddeler, açık uçlu maddeler, senaryo tipi maddeler. Her madde, türünün zayıf veya güçlü yönleri madde yazarı tarafından bilinmeli ve ölçülecek değişkenin

doğasına en uygun madde türüne karar verilmelidir. Örneğin, iki seçeneqli maddeler, yanıtlayıcının “doğru” veya “yanlış”, “evet” veya “hayır” gibi iki seçenek arasından birisini seçmesini gerektiren maddelerdir. Bu tür maddeler, genellikle bilgi temelli testlerde ve bazen de kişilik ölçeklerinde kullanılmaktadır. İki seçeneqli maddeler, klinik tanı ölçekleri veya ilgi ve tutum ölçekleri için uygun değildir. Çünkü ilgi ve tutum gibi, birçok psikolojik değişken için verilecek olan yanıtlar kesin doğru veya kesin yanlış olarak nitelendirilemez. Buna rağmen zaman zaman bu tür testlerde de bu madde türünün kullanımına rastlanılmaktadır.¹⁶ Madde türü ne olursa olsun, maddeler hedef kitle tarafından rahatlıkla anlaşılabilir düzeyde, gereksiz sözcüklerden arınık olabildiğince kısa ve yalın olmalıdır. Maddelerin, farklı yaşantı geçmişine sahip ya da farklı kültürel etki altındaki bireyler tarafından aynı şekilde anlaşılması gerekmektedir. Bu bakımdan, algılayışta farklılık yaratabilecek ifadelerden kaçınılmalıdır. Her bir madde, ölçülen özelliğin yalnızca bir tek göstergesini içerecek şekilde yazılmalı birden fazla yargı içermemelidir.

Deneme uygulaması için, geniş bir madde havuzu oluşturulması, geliştirilecek olan ölçeğin kapsam geçerliği bakımından önem taşımaktadır. Deneme uygulaması sonrası, yapılacak olan elemelerden sonra geriye kalan maddeler, eğer ölçülmesi amaçlanan yapıyı yeterince temsil etmiyorsa, ölçek geliştiricinin yazılmış olan maddeler üzerinde düzenlemeler yapması, ya da yeniden madde yazması gerekecektir. Bu aşamada, ölçek geliştirici, konu uzmanlarından veya konu ile ilgili kişisel deneyimi olanlardan yardım talep edebilir.⁸

6. Ölçeğin Açıklama ve Yönergesinin Yazılması

Ölçeğin, yönergeleri açık ve net olmalıdır. Yönergelerde, yanıtlayıcılara, ölçeğin hangi amaçla oluşturulduğu, ölçeği nasıl yanıtlayacakları, seçenekleri nasıl işaretleyecekleri, yanıtlama süreleri gibi bilgiler, yanıtlayıcıların anlayabileceği şekilde anlaşılır ifadelerle verilmelidir. Yanıtlayıcıların, ölçeğin amacından haberdar olması durumunda, yanıtlama davranışlarının değişiklik göstereceği düşünülmüyorsa amaç gizlenebilir. Çoğu zaman, katılımcıların gönüllülük esasına dayalı olarak ölçeği yanıtladığı düşünülürse, yönergenin yazımında emir kalıpları içeren soğuk bir dil yerine, motive edici samimi bir dil tercih edilmelidir. Yönerge, hedef kitleye uygun bir hitap ile başlamalı ve sonunda mutlaka teşekkür ile bitirilmelidir. Yönergede, yanıtlayıcılardan isim istenip istenmediği belirtilmelidir. Yanıtlayıcılardan isim alınmayacaksa ve eğer

ölçek aynı gruba tekrar uygulanacaksa, unutmayacakları bir rumuz yazmaları istenebilir.

7. Ön Deneme Uygulamasının Gerçekleştirilmesi

Ön deneme uygulamasının amacı, veri toplamak ve toplanan verileri işlemek değildir. Bu adımda asıl amaç, maddelerin okunurluğu, anlaşılabilirliği ve farklı kişilerce aynı şekilde algılanıp algılanmadığı gibi konularda, araştırmacının gözden kaçırdığı ya da öngöremediği sorunları tespit etmektir. Ayrıca, uygulama sayesinde ortalama yanıtlama süresi hesaplanarak daha sonraki uygulamalarda karşılaşılabilecek sorunların önüne geçilebilir. Ön deneme uygulaması, hedef kitleyi temsil eden oldukça küçük bir grup üzerinde yürütülür. Gruptaki kişi sayısı, hedef kitledeki çeşitlilik oranında değişiklik gösterebilir. Araştırmacı, mutlaka uygulamanın gerçekleştirildiği ortamda bulunmalı ve yanıtlayıcılarla iletişim halinde olmalıdır. Uygulama tamamlandığında, yanıtlayıcılarla maddeler üzerinde konuşulmalı ve eleştiriler not edilmelidir. Deneme uygulamasına geçmeden önce gerekli değişiklikler yapılmalı, sorunların çözüldüğü konusunda tereddüt edilmesi halinde yapılan değişiklikler tekrar yanıtlayıcılarla tartışılmalıdır.

8. Deneme Uygulamasının Gerçekleştirilmesi

Deneme uygulaması, olanaklar çerçevesinde ulaşılabilen maksimum sayıda birey üzerinde gerçekleştirilmelidir.¹⁶ Deneme uygulaması örnekleminde yer alan birey sayısının faktör analizi ve güvenilirlik kestirimleri için de yeter sayıda olması önem taşımaktadır. Ayrıca örneklem, ölçeğin hedef kitlesi içerisinde yer alan tüm alt grupları kapsayacak çeşitlilikte olmalıdır. Ölçeğin deneme uygulaması sırasında ölçeği yanıtlayan tüm bireyler eşit şartlar altında katılım göstermelidir. Açıklamalar ölçeği yanıtlayan bireylerin hepsine aynı şekilde yapılmalı, herkese eşit süre verilmelidir. Ayrıca ölçek, uygulamaya katılan her bireye fiziki koşulları benzer ortamlarda verilmelidir.⁸

10. İstatistiksel Teknikler ile Ölçeğin Son Şeklinin Verilmesi

Bu adımda ölçek geliştiren kişinin, madde ve test analizlerini yapması ayrıca faktör analizi ile ölçeğin faktör yapısını incelemesi beklenmektedir. Maddenin, bir gruptan elde edilen verilere göre, sayısal özellikleri olarak tanımlanan madde analizinin temel amacı, ölçekte yer alan ve psikometrik özellikleri zayıf olan maddeleri çıkartmak veya revizyondan geçirmektir. Bu adımda, madde güçlük indeksi, madde standart

sapması ve varyansı, madde ayırıcılık indeksi ve madde güvenilirlik katsayısı gibi istatistiksel ölçüler elde edilmektedir. Bir maddeyi değerlendirirken bu ölçülerin ayrı olarak değil, birlikte değerlendirilmesi gerekmektedir.² Fakat tipik davranış ölçeklerinde (ilgi, tutum, kişilik vb.) madde güçlüğünden söz etmek doğru değildir. Bu tip ölçeklerde, ilgili maddenin ölçülen özelliğin miktarı bakımından bireyleri ayırt edip edemediğinin incelenmesi gerekmektedir. Başka bir ifadeyle başarı ve yetenek testlerinde maddenin doğru yanıtlanma olasılığına ilişkin bilgi veren madde güçlük indeksi cevapların doğru ya da yanlış şeklinde nitelendirilmediği tipik davranış ölçeklerinde hesaplanmaz. Fakat, madde ayırt edicilik indeksi, tüm ölçek tiplerinde madde seçiminde kullanılan bir istatistiktir. Eğer ölçek, Madde Tepki Kuramı'na dayalı olarak geliştiriliyorsa madde seçiminde, belirlenen modele özgü madde parametreleri ve madde bilgi fonksiyonları kullanılabilir. Fakat bu makalede, Madde Tepki Kuramı detaylı olarak ele alınmayacaktır. Sözü edilen madde analizi çıktılarının yanı sıra, ölçeğin geliştirilme amacı ne olursa olsun nihai formda yer alacak maddelerin seçiminde ve ölçeğin yapısının oluşturulmasında kullanılan bir diğer istatistiksel teknik faktör analizidir. Günümüzde, bilgisayar teknolojilerinin ve istatistik yazılımlarının gelişmesiyle birlikte, araştırmacılar için kolay uygulanabilir hale gelen faktör analizi, ölçek geliştirme sürecinde kesinlikle başvurulması gereken bir teknik olması bakımından çalışmamızda detaylı bir şekilde ele alınmıştır.

Faktör Analizi

Faktör analiziyle ilgili teknikleri, Açıklayıcı Faktör Analizi (AFA) ve Doğrulayıcı Faktör Analizi (DFA) olmak üzere, iki ana başlık altında incelemek mümkündür. DFA, ölçülen özelliğe ilişkin kuramsal olarak öngörülen, ya da daha önceden açıklanmış olan bir yapının gözlem sonuçlarına uyum sağlayıp sağlamadığını test etmek için kullanılır.²¹ Bu bakımdan, genellikle DFA araştırmacının ölçeğin yapısına ilişkin önsel bilgi sahibi olduğu ölçek uyarılama çalışmalarında kullanılırken, AFA ise yapıyı açıklamaya yönelik olarak ölçek geliştirme çalışmalarında kullanılır. Bu bölümde detaylı olarak açıklayıcı faktör analizi üzerinde durulmuştur.

Ölçek geliştirme sürecinde açıklayıcı faktör analizi, üç temel amaca hizmet eder. Bunlardan ilki, çok sayıda değişkeni (maddeyi) faktör adı altında daha az sayıda değişkene indirgemektir. İkincisi, ölçülen özelliğin altında yatan gizil boyutları ortaya çıkarmaktır. Üçüncüsü ise, ölçeğin yapısını oluşturmak ve ölçümün yapısal

geçerliğine kanıt elde etmektir.^{22,23} Buna göre, faktör analizinin ölçek geliştirme sürecindeki asıl rolü değişkenler (maddeler) arasındaki ilişki katsayıları matrisinden birbirleri ile ilişkili olan maddeleri bir araya getirerek kendi içerisinde tutarlı boyutlar (faktörler) oluşturmak, ölçülen yapıyı az sayıda değişkenle ve mümkün olan en yüksek varyansla açıklamaktır.

Williams, Onsman ve Brown²⁴, açıklayıcı faktör analizi sürecini beş temel adımda incelemiştir. Bunlar:

- 1) Veri setinin faktör analizine uygunluğunun incelenmesi,
- 2) faktör çıkarma (faktörleştirme) yönteminin belirlenmesi,
- 3) faktör sayısının belirlenmesinde kullanılacak ölçütün belirlenmesi,
- 4) faktör döndürme yönteminin seçilmesi
- 5) yorumlama ve adlandırma adımlarıdır. Bu adımlar sırasıyla şöyle ele alınabilir.

1. Veri Setinin Faktör Analizine Uygunluğunun İncelenmesi

Örneklem Büyüklüğü: Örneklemdeki kişi sayısı, faktör analizi için oldukça önemli olsa da, alanyazında sayı konusunda çok sayıda farklı görüşlerin olduğu bilinmektedir. Tabachnick ve Fidell'e²⁴, göre örneklem en az 300 kişiden oluşmalıdır. Hair ve diğerleri²⁵ ise, kişi sayısının 100'ün üzerinde olması gerektiğini öne sürmüştür. Comrey ve Lee²⁶ 100 kişinin az, 200 kişinin idare eder sayıda, 300 kişinin uygun, 500 kişinin çok uygun ve 1000'den fazla kişinin de mükemmel sayıda olacağı görüşündedir. Bunların yanı sıra, alanyazında, örneklemdeki kişi sayısının madde sayısına oranı şeklinde bir minimum oran öneren araştırmacıların olduğu görülmektedir. Kimi araştırmacılar, bu oranın en az 5:1 olması gerektiğini ileri sürmektedirler.^{27,28,29,30} Yaygın olarak benimsenen bir diğer görüş de bu oranın en az 10:1 olması gerektiği yönündedir.^{31,32,33} Bu makalede, araştırmacılara ulaşabildikleri kadar geniş bir örneklemden veri toplamaları önerilmektedir. Örneklem, büyüdükçe yapılacak analizin gücünün artacağı ve hataların azalacağı söylenebilir.

Kayıp Değerler ve Uç Değerler: Her ne kadar araştırmacılar eksiksiz veri setleri elde etme çabasında olsalar da, araştırmalar kapsamında üzerinde çalışılmak istenilen veriler bazı nedenlerle istenildiği gibi eksiksiz bir şekilde toplanamayabilir. Özellikle, büyük gruplar üzerinde yürütülen çalışmalarda eksiksiz veri setlerinin elde edilmesi neredeyse olanaksızdır.³⁴ Kayıp verilerin varlığında araştırmacılar, (1) veriye yeni gözlemlerin eklenmesi, (2) kayıp verili

gözlemlerin veri setinden çıkartılması, (3) kayıp verilere ilişkin kestirimlerin yapılması ve elde edilen yaklaşık değerlerin kayıp veriler yerine kullanılması yöntemlerinden birini kullanarak olası sorunlara yönelik önlem alabilirler. Veriye yeni gözlemler ekleme sürecinin zaman ve emek maliyeti ortaya çıkaracağı göz önünde bulundurulmalıdır. Eksik verili gözlemlerin, veri setinden çıkartılması ise gözlem sayısında ciddi bir azalmaya yol açabilir ve yeterli sayıda oluşturulmuş bir örneklem yetersiz sayıdaki bir örnekleme dönüşebilir. Bu durum, yapılacak olan istatistiksel analizlerin gücünün azalmasına neden olacaktır.^{35,36} Üstelik kayıp veri içeren gözlemlerin, veri setinden çıkartılabilmesi için kayıp verilerin tamamen rastlantısal olarak dağılıyor olması gerekmektedir. Kayıpların, analize dâhil edilen başka değişkenlerle ilişkili olduğu durumlarda, yapılacak olan silme işlemi önemli bir yanlılığa yol açabilir.^{25,37,38} Bu bağlamda değerlendirildiğinde, kayıp veriler yerine, yaklaşık değer atama yöntemleri, araştırmacıların hem zamandan ve emekten tasarruf edecekleri, hem de topladıkları verileri koruyabilmelerini sağlayacak bir yol olarak ortaya çıkmaktadır. Yaklaşık değer atama yöntemleri konusunda araştırmacılara, en çok olabilirlik kestirimine dayalı Beklenti Maksimizasyonu Algoritması (Expectation Maximization Algorithm), çoklu atamaya dayalı yöntemler olarak adlandırılan Eğilim Puanları Eşleştirme (Propensity Score Matching), Markov Zincirleri Monte Carlo (Markov Chain Monte Carlo) gibi yöntemler ve son yıllarda kullanımı artmaya başlayan Bayesci Veri Atama (Bayesian Imputation) Stokastik Regresyon Ataması (Stochastic Regression Imputation), Mahalanobis Uzaklığı Ataması (Mahalanobis Distance Imputation), K-Ortalama Kümeleme Ataması (K-Means Clustering Imputation) gibi farklı gelişmiş yöntemleri önerilebilir. Kayıp değerlerin yanı sıra, faktör analizi sürecinde sorun oluşturabilecek olan bir diğer durum, veri setinde uç (aşırı) değerlerin yer almasıdır. Ölçek geliştirme sürecinde uç değer araştırmacıdan kaynaklı olarak verinin yanlış girilmesi sonucu (Araştırmacının aslında 5 olan bir değeri 55 olarak girmesi gibi...) ortaya çıkabilir. Analiz öncesinde bu tip durumlar tespit edilmeli ve düzeltilmelidir.

Korelasyon Matrisinin Faktörleşebilirliği: Faktör analizine başlamadan önce, korelasyon matrisinin faktörleşebilirliği test edilmelidir. Faktörleşebilme, değişkenler (ölçek maddeleri) arası ilişkilerin belirli bir düzeyin üzerinde olmasına bağlıdır. İlişkilerin gücü, Bartlett Küresellik Testi (Bartlett's Test of Sphericity) ve Kaiser-Meyer-Olkin Örneklem Yeterliliği Ölçütü

(Kaiser-Meyer-Olkin Test of Sampling Adequacy) ile değerlendirilebilir. Bartlett Küresellik Testi, gözlenen korelasyon matrisi (maddeler arası korelasyon matrisi) ile birim matris arasındaki farkın anlamlılığını sınar. Bir başka ifadeyle, matrisin faktörleşebilmesi için maddeler arası korelasyonların, sıfırdan istatistiksel olarak anlamlı derecede farklı olması gerekmektedir. Bu bakımdan, veriyi analize uygun kabul edebilmek için matrisler arası farkın anlamlı çıkması beklenir.³⁹

Keiser-Meyer-Olkin Örneklem Yeterliliği Ölçütü ise, maddeler arası ortak varyansa (shared variance) ilişkin bir ölçüttür. Bu anlamda KMO, yaygın kanının aksine uygulama örnekleminin büyüklüğünün yeterliliğine ilişkin bir test değil, analize tabi tutulan değişkenlerin (maddelerin) faktörleşebilirliğine ilişkin bir ölçüttür. Elde edilen KMO değerine ilişkin değerlendirme aşağıdaki tabloya dayalı olarak yapılabilir.⁴⁰

2. Faktör Çıkarma (Faktörleştirme) Yönteminin Belirlenmesi

Faktörleştirme yönteminin belirlenmesi söz konusu olduğunda, iki ana yaklaşım ortaya çıkmaktadır. Bunlardan biri Temel Bileşenler Analizi (Principal Components Analysis), diğeri

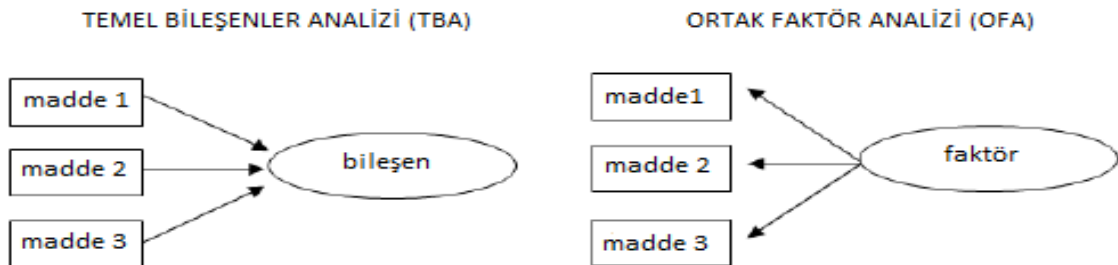
ise Ortak Faktör Analizidir (Common Factor Analysis).

Ortak Faktör Analizi bir yöntemler ailesi olup, Temel Eksenler Faktörleştirme (Principal Axis Factoring), En Çok Olabilirlik Yöntemi (Maximum Likelihood), Ağırlıklandırılmamış En Küçük Kareler Yöntemi (Unweighted Least Squares), Genelleştirilmiş En Küçük Kareler Yöntemi (Generalised Least Squares), Alfa Faktörleştirme (Alpha Factoring) ve İmaj Faktörleştirme (Image Factoring) gibi yöntemleri içermektedir. Temel Bileşenler Analizi (TBA) ile Ortak Faktör Analizi (OFA) arasında hem teorik hem de istatistiksel olarak farklılıklar vardır. TBA, veriyi temel bileşene katkı sağlayan bir grup doğrusal değişken olarak ayrıştırmaktadır. Bu bakımdan ayırma analiziyle benzerlik göstermektedir.²³

TBA'da temel amaç çok sayıda değişkeni, veriyi açıklayan daha az sayıda değişkene indirgemektir. OFA'da ise temel amaç, doğrudan gözlemlenemeyen fakat gözlemlenebilir değişkenler tarafından yansıtılan arka plandaki gizil yapıyı ortaya çıkarmaktır.^{41,42,43} Söz konusu tartışmaya ilişkin görsel anlatım Şekil 1'de sunulmuştur

Tablo 2. KMO değerlerine ilişkin değerlendirme ölçütleri

KMO değeri	Matrisin Faktörleşebilirlik Düzeyi
0.90 - 1.00	Mükemmel
0.80 - 0.89	İyi
0.70 - 0.79	Orta Düzey
0.60 - 0.69	Kötü
0.50 - 0.59	Çok Kötü
0.00 - 0.49	Faktörleşemez



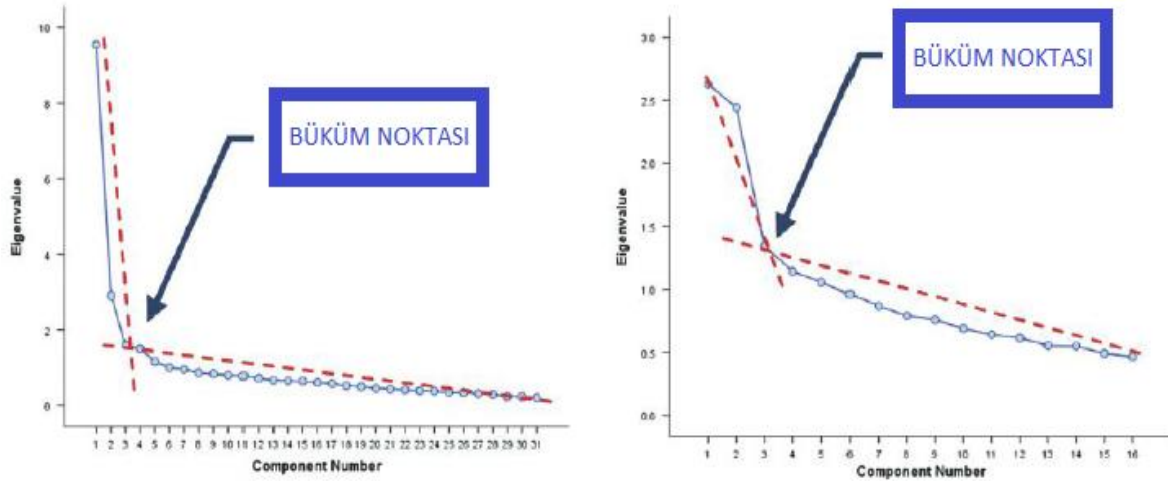
Şekil 1. TBA ve OFA'da ilişkilerin yönü *Kaynak: Beavers ve diğerleri (2013).*

Araştırmacının amacı, toplam varyans üzerinden değişken indirgemek olduğunda TBA bir seçenek olarak değerlendirilebilir. Ancak amaç, biricik varyans ve hata varyansını ayırıştırarak değişkenlerce belli bir oranda açıklanan bir gizil yapıyı ortaya çıkarmaksa OFA tercih edilmelidir. Temel Bileşenler Analizinde hata varyansı ayırıştırılmaz, oysa hata varyansı yapının gözlenen değişkenlerle açıklanamayan kısmıdır ve özellikle psikoloji araştırmalarında önemli yer tutmaktadır.⁴⁴ Yapılan tartışmalar doğrultusunda, ölçek geliştiren araştırmacıların amacı ile OFA'nın kullanım amacının birbiriyle örtüştüğü ifade edilebilir. Ayrıca TBA'da bileşen yüklerinin olduğundan yüksek (şişkin) çıkması ve koşullar değiştiğinde kararsızlık göstermesi de farklı bir sorun olarak gündeme gelmektedir.^{45,45} Söz konusu nedenlerle bu makalede ölçek geliştiren araştırmacılara istatistik programlarının çoğunda önceden tanımlanmış (default) olarak gelen Temel Bileşenler Analizi yerine Ortak Faktör Analizi yöntemlerini kullanmaları önerilmektedir. OFA yöntemlerinden hangisinin seçilmesi gerektiği konusunda ise alanyazında fikir birliğinin olmadığı görülmektedir. Fabrigar, Wegener, MacCallum ve Strahan,⁴⁶ Temel Eksenler Faktörleştirilmesi (TEF) ve En Çok Olabilirlik Yöntemi'nin (EÇÖY) en sık kullanılan yöntemler olduğunu belirtmiş, veriler normal dağıldığında EÇÖY'un, normal dağılmadığında ise bu varsayıma ihtiyaç duymayan TEF'in kullanılması gerektiğini öne

sürmüşlerdir. Coughlin,⁴⁷ farklı koşullar (örneklem büyüklüğü, değişken sayısı, faktör sayısı, ikili puanlanan madde sayısı) altında yürüttüğü simülasyon çalışmasında Ağırlıklandırılmamış En Küçük Kareler Yöntemi'nin, TEF ve EÇÖY yöntemlerinden daha az yanlışlık ve hata içerdiği sonucuna ulaşmış ve kullanılmasını önermiştir. Erkuş'a¹¹ göre ise faktörleştirme sürecinde birden çok yöneme başvurulmalı ve "a priori" kurama uygun en iyi yapı bulununcaya kadar denemeye devam edilmelidir.

3. Faktör Sayısının Belirlenmesinde Kullanılacak Ölçütün Belirlenmesi

Faktör sayısının belirlenmesi, analizin en önemli adımlarından biridir. Olduğundan fazla ya da, olduğundan az sayıda faktör tanımlamak ölçeğin yapısıyla ilgili ciddi sorunlara yol açabilir. Alanyazında faktör sayısının belirlenmesi için, önerilen birden fazla yöntem ve ölçüt olduğu görülmektedir. Bunlardan, öz-değeri (eigenvalue) 1'den büyük olan faktörlerle devam edilmesi gerektiğini ileri süren "Kaiser Ölçütü" ve öz-değerlere ilişkin bir yamaç grafiği sunan "Scree Plot" yöntemi araştırmacılar tarafından sıklıkla kullanılmaktadır. Scree Plot yönteminde eğrinin büküm noktası faktör sayısını belirlemektedir. Yönteme ilişkin görsel Şekil 2'de sunulmuştur.



Şekil 2. "Scree Plot" yöntemiyle faktör sayısının belirlenmesi. Kaynak: Field, A.(2009).

Faktör sayısı belirleme üzerine yürütülen çok sayıda çalışmada, Kaiser Ölçütü ve Scree Plot yöntemleri ile çoğu kez hatalı sonuçlar elde edildiği ve söz konusu yöntemlerin olduğundan

fazla sayıda faktör (overfactor) belirleme eğilimi gösterdikleri sonucuna ulaşılmıştır.^{48,49,50} İstatistik programlarının çoğunda bulunmadığı için, araştırmacılar tarafından yaygın şekilde kullanılmayan bir diğer faktör sayısı belirleme yöntemi ise Paralel Analiz (Parallel Analysis)

yöntemidir. Paralel analiz yöntemi, iki yöntemden daha doğru sonuçlar verdiği gerekçesiyle pek çok araştırmacı tarafından önerilmekte, dahası faktör sayısı belirlemede en iyi yöntem olarak nitelendirilmektedir.^{49,47,51,52,53,54,55,23} Bu yöntemlerin yanı sıra, açıklanan varyans yüzdesi de faktör sayısına karar vermek için kullanılan bir başka ölçüttür. Belirlenen sayıdaki faktörün ölçülen özelliği belirli bir oranın üzerinde açıklaması istenmektedir. Davranış bilimlerinde açıklanan toplam varyansın, %50-%60 civarlarında olması kabul edilebilir bir orandır.^{26,40,41} Alanyazında bu adım için sıkça getirilen bir öneri ise faktör sayısı belirlemede, birkaç yöntemin eşzamanlı olarak yürütülmesi ve elde edilen sonuçları karşılaştırdıktan sonra bir karara varılması yönündedir.²⁴ Bu öneriye destek vermekle birlikte bu makalede, eğer mümkünse, pek çok araştırmacı tarafından diğer yöntemlere üstünlüğü ortaya koyulmuş olması nedeniyle “Paralel Analiz” yönteminin kullanımı önerilmektedir.

Diğer yandan, yapılan tartışmalarla birlikte, ölçek geliştirme sırasında boyut belirleme işleminin yalnızca istatistiksel sonuçlara bakılarak mekanik kararlar alınan bir süreç olmadığı söylenebilir. Faktör sayısına karar verme aşamasında belirlenen faktörlerin kuramsal olarak anlaşılabılır ve yorumlanabilir olması son derece önemlidir.³¹ Bu bakımdan, araştırmacının ya da araştırmacıların bilgi ve deneyimi her zaman sürece dahil olmak durumundadır. Buradan hareketle araştırmacılara, elde ettikleri istatistiksel verilerle birlikte açıklanması amaçlanan yapıya ilişkin önsel bilgilerini de işin içine katarak bir karara varmaları önerilebilir.

4. Faktör Döndürme Yönteminin Seçilmesi

En uygun yapıyı keşfetmek amacıyla, faktör sayısına karar vermede başvurulabilecek bir diğer yol döndürme (rotation) yöntemini kullanmaktır. Hangi maddenin, hangi faktöre gideceğini belirginleştirmede döndürme işlemi öncel yapıya uygun sonuçların elde edilmesini sağlayabilir⁴¹ (Beavers ve diğerleri, 2013). Doğrudan çözümlenmede, bazı maddeler birden fazla faktörle yüksek ilişki gösterebilir ve bu durum maddelerin yeri açısından bir belirsizlik oluşturabilir. Bu nedenle, döndürme işlemi yapılmaksızın hangi maddenin hangi faktör altında yer aldığını belirlemek oldukça zordur. Döndürme işlemi, bir maddenin yükünü en yakın olduğu faktör için artırıp, diğer faktörler için ise azalttığından ulaşılan sonuçları yorumlamak kolaylaşmaktadır.^{25,40,37} Faktör döndürme yöntemleri, faktörlerin birbirleriyle ilişkisiz olduğu varsayımına dayanan dik döndürme

(orthogonal rotation) ve faktörlerin birbirleriyle ilişkili olduğu varsayımına dayanan eğik döndürme (oblique rotation) yöntemleri olarak iki ana başlık altında incelenebilir. En sık kullanılan dik döndürme teknikleri “varimax”, “quartimax” ve “equamax”; eğik döndürme teknikleri ise “direct oblimin” ve “promax”tır. Dik döndürme yöntemleri (özellikle varimax), araştırmacılar tarafından en çok tercih edilen yöntemler olsa da^{47,44} davranış bilimleri araştırmaları için eğik döndürme yöntemlerinin çoğu zaman daha uygun olduğu öne sürülmektedir.⁴¹ Bunun nedeni, psikolojik boyutların arasında korelasyon olacağı beklentisidir. Daha da ötesi bir psikolojik boyutun diğer psikolojik boyutlarla ilişki göstermeyeceğine dair bir varsayımın bulunması gerçekçi olmayacaktır.²³

Tabachnick ve Fidell²⁵ faktörlerin öncelikle eğik döndürülmesi ve faktörler arası korelasyon matrisinin incelenmesi yönünde görüş belirtmişlerdir. Buna göre, korelasyon matrisinde 0,32 ve üzerinde bir katsayı değeri bulunmuş ise ve eğer dik döndürme için ciddi bir gerekçe yok ise eğik döndürme yöntemlerinin kullanılmasının daha uygun olacağı söylenebilir. Kim ve Mueller’e⁵⁶ göre ise, dik döndürme yöntemlerinin daha basit yorumlanabilir sonuçlar üretmesi tercih edilmesi için yeterli bir gerekçedir ve döndürme yalnızca maddelerin faktörlerle ilişkisini yorumlamak için yapılıyorsa, bu gerekçe kabul edilebilir. Fakat seçilecek olan yöntemin faktör puanlarını etkileyeceği unutulmamalıdır. Thompson’a⁵⁵ göre, daha sonraki istatistiksel analizlerde faktör puanlarının kullanılması planlanıyorsa ve faktörler arasında kayda değer oranda ilişki varsa eğik döndürme yöntemleri kullanılmalıdır. Faktörlerin ilişkili olduğu durumlarda dik döndürme yapmak bilgi kaybına neden olmakta, bu tip durumlarda eğik döndürme yöntemleri daha hatasız sonuçlar üretmektedir.⁴⁴ Diğer yandan, analize geçmeden önce faktörler arası ilişkiler bilinmediğinden en uygun sonucu bulmaya yardımcı olabilecek yöntemlerin sınanması gerektiğini iddia eden araştırmacılar da vardır.¹¹

5. Yorumlama ve Adlandırma

Bu aşama, maddelerin faktörlerle hem istatistiksel, hem de kuramsal olarak ilişkilendirilmesi ve yapıya son şekli verildikten sonra faktörlerin adlandırılması süreçlerini içermektedir. İstatistiksel olarak, herhangi bir faktöre yeterli düzeyde yük vermeyen (genellikle 0,30’un altında) ve ölçülen özelliğe sağladığı katkı bakımından vazgeçilmez olmadığı düşünülen maddelerin analizden çıkarılarak faktörleştirme işleminin tekrarlanması önerilmektedir.²⁴ Bazı maddelerde

birden çok faktöre 0,30'un üzerinde yük verebilir (crossloading item). Bu tip maddelerin analizden çıkarılması noktasında araştırmacıların daha dikkatli olmaları gerekebilir. Birden fazla faktöre anlamlı düzeyde yük veren bir maddenin, yapıyla ilişkili olmadığı öne sürülemez. Yapıyla çok boyutlu ilişki gösteren bir maddenin ölçekten çıkarılması tartışıldığında, kuramsal anlamlandırma sürecinin işletilmesi daha önemli hale gelmektedir. İlgili maddenin yüksek yük verdiği faktörlerden, herhangi biriyle kuramsal olarak daha fazla ilişkili olduğu yönünde bir karar alınabiliyorsa maddenin ölçekte kalması yönünde karar verilebilir. Tüm bu değerlendirmeler yapılırken belirlenen faktörlerde yer alacak olan madde sayısına da dikkat edilmelidir. Üç maddeden daha az sayıda maddeyle ilişkilendirilen faktörler zayıf ve kararsızdır. Arzu edilen ise beş ve daha fazla sayıda maddenin faktöre yüksek yükü (0,50 ve üzeri) bağlanmasıdır. Ancak, bu tip faktörler güçlü ve kararlı olarak nitelendirilebilir.⁴⁴ Söz konusu ölçütler doğrultusunda, analizden madde çıkarma işlemi, ölçülen yapıyı yansıtırma konusunda en güçsüz olan maddelerden başlanılarak birer, birer yapılmalı ve arzu edilen kararlılıkta bir yapıya ulaşana dek, daha önce alınan kararlar çerçevesinde (faktör sayısı, döndürme yöntemi gibi) analiz tekrarlanmalıdır.

Faktör Skorları

Faktör skoru basit anlamda, bireylerin maddelere verdikleri tepkilerin (maddeden alınan puan) faktör yükleri ile ağırlıklandırılmasıyla elde edilen puanlardır. Bu ağırlıklandırma için çeşitli yöntemler kullanılmaktadır (Regresyon, Bartlett, Anderson-Rubin yöntemleri). Her bir birey için, ölçeğin faktör sayısı kadar, faktör skoru hesaplanmaktadır (Örneğin, iki faktörlü bir yapıda her birey için iki ayrı faktör skoru elde edilir). Faktör skorlarını, bireylerin ölçülen özellik bakımından düzeyini (bireyin alt-test ya da test puanı) belirlemek için kullanılabilir gibi, elde edilen verilerle daha başka analizler yapmak için de (ağırlıklandırılmamış toplam puanlar yerine) kullanılabilir. Faktör skorlarının kullanımı, bazı analiz yöntemleri için gerekli olan varsayımların ihlal edilmesi durumunda karşılaşılabilecek sorunların ortadan kalkmasını da sağlayabilir.²³ Ağırlıklandırılmamış toplam puanların kullanılması, aynı maddelere farklı şekilde tepki vermiş bireylerin ölçekten aynı puanı alması gibi sorunlara yol açabilmektedir. Oysa, çoğu zaman ölçekten eşit puan alan bireylerin yanıt örüntüleri birbirinin aynı olmaz (Bu sorun Madde Tepki Kuramı'na dayalı geliştirilen ölçeklerle aşılabilmektedir, fakat Klasik Test Kuramı'na dayalı geliştirilen ölçekler için sorun olmaya devam etmektedir). Yanıt örüntüleri farklı, fakat

ölçek toplam puanları aynı olan bireylerin, ölçülen özellik bakımından aynı düzeyde olduğu söylenebilir mi? Her bir maddenin ölçülen özelliği farklı miktarlarda açıkladığı (faktör yükleri farklı) göz önünde bulundurulduğunda, puanların ağırlıklandırılmasının daha ayırt edici ölçmeler yapabilmek adına olumlu etkilerinin olacağı iddia edilebilir.

Ölçeğin Güvenirlik Analizlerinin Yapılması

Güvenirlik bir ölçme aracının duyarlı, tutarlı, objektif ölçme sonuçları verebilmesidir.⁵⁷ Güvenirlik bir bakıma, ölçme sonuçlarının tesadüfi hatalardan arınıklık derecesini ifade eden bir kavramdır.⁵⁸

Ölçümlerde, gözlenen toplam değişkenlik içindeki hata kaynaklı pay arttıkça, ölçeğin güvenirliliği düşmekte, bu pay azaldıkça, ölçeğin güvenirliliği artmaktadır. Hatasız, gerçek ölçümler bilinemediğinden, güvenirlilik dolaylı yoldan belirlenmeye, tahmin edilmeye çalışılır.⁹ Güvenirlilik kestiriminde kullanılan çeşitli yöntemler vardır. Bunlar; test-tekrar, eşdeğer formlar, test-yarı ve iç tutarlılık katsayısı saptama yöntemleridir.⁵⁷ Ölçek geliştirme çalışmalarında, güvenirlilik kestirimi söz konusu olduğunda, kolay ve tek uygulamaya dayandığından, en yaygın olarak iç tutarlılık katsayısı hesaplandığı ve bu hesaplama için genellikle Cronbach Alpha (α) formülünün kullanıldığı söylenebilir. Çüm ve Koç'un¹⁴, Türkiye'de geliştirilen ölçekler üzerine yaptığı araştırmada, incelenen çalışmaların yaklaşık %67'sinde güvenirlilik kestirimi olarak, yalnızca Cronbach α iç tutarlılık katsayısının rapor edildiği ve çalışmaların tamamında diğer yöntemlerle yapılan kestirimlerin yanında Cronbach α katsayısının da yer aldığı belirlenmiştir. Yang ve Green'e⁵⁹ göre, Cronbach α 'nın bu kadar yaygın kullanılmasında, katsayının yorumlanmasının kolay olması ve diğer kestirim yöntemlerine göre araştırmacının sübjektif kararlarına ihtiyaç duymaması etkili olmaktadır. Örneğin, test-yarı güvenirlilik kestirimi yöntemi kullanılmak istendiğinde araştırmacının ölçeği ikiye bölme konusunda nasıl bir yol izleyeceğiyle ilgili bir karar alması gerekmektedir. Benzer şekilde, test- tekrar test yöntemi kullanılmak istendiğinde araştırmacı, iki uygulama arasında ne kadar süre olması gerektiğine ilişkin bir karar almak zorundadır. Ayrıca, diğer yöntemlerin gerektiği ikinci bir uygulama yapma ya da paralel bir form oluşturma gibi zorluklar, araştırmacıları hesaplaması kolay olan iç tutarlılık kestirimine yönlendirmektedir. Oysa, alanyazında Cronbach α 'nın yanlış kullanıldığı ve yorumlandığına dair pek çok görüş öne sürülmüştür.^{60,61,62,63,60}

Cronbach α hesaplamak için, ölçeğin tek boyutlu olması varsayımı vardır, fakat α 'nın yüksek çıkması ölçeğin tek boyutlu olduğu anlamına gelmeyebilir. Buna ancak faktör analizi tekniğiyle karar verilebilir. Tan⁶³ KPSS verileri üzerine yaptığı çalışmada, Kaiser Ölçütü'ne göre 71 faktörden oluşan tüm testten hesaplanan KR-20 katsayısının (Cronbach α formülünden türetilir) oldukça yüksek bir değer (0,92) elde ettiği sonucuna ulaşmıştır. Böylesine bir durumda, ölçeğin iç tutarlılığının yüksek olduğu gibi yanlış bir yoruma gidilebilir, oysa gerçek öyle değildir. Ölçeğin birden fazla faktörden oluşması durumunda, her bir faktör için ayrı bir Cronbach α katsayısının hesaplanması gerekmektedir.^{33,64,65,66} Çok faktörlü yapılarda, tüm ölçek için bir tek katsayı elde edilmek isteniyorsa Cronbach, Schonemann ve McKie⁶⁷ tarafından önerilen "Tabakalı Cronbach α " (Stratified Alpha) katsayısı, Moiser⁶⁸ tarafından önerilen "Moiser'in Bileşik Güvenirlik Katsayısı" ve Wang ve Stanley⁶⁹ tarafından önerilen "Bileşik Güvenirlik Katsayısı" kullanılabilir.

Diğer taraftan, Cronbach α ve sözü edilen diğer iç tutarlılık katsayıları maddelerin paralel, eşdeğer (tau-equivalent) ya da eşbiçimli (essentially tau-equivalent) olduğu durumlarda, güvenirligi yansız olarak kestirmektedir. Başka bir ifadeyle, α katsayısının kullanılabilmesi için maddelerin faktör yüklerinin eşit olması varsayımı sağlanmalıdır.⁷⁰ Fakat ölçek geliştiren araştırmacıların, tüm maddelerin faktör yüklerinin eşit olduğu bir durumla karşılaşmaları pek olası değildir. Yurdugül⁷¹ (2006), faktör yüklerinin eşit olduğu durumlarda Cronbach α , Revelle β , Heise ve Bohrnstedt Ω gibi iç tutarlılık katsayılarının kullanılabileceğini fakat faktör yüklerinin eşit olmadığı durumda yalnızca McDonald'ın ω katsayısının kullanılması gerektiğini ifade etmiştir. McDonald'ın ω katsayısı, aynı zamanda "yapısal güvenirlik katsayısı" olarak da adlandırılmaktadır. Bu yöntemle güvenirlik kestirimi yapılırken, maddelerin faktör yükleri de hesaba katılmaktadır. Raykov ve Shrou⁷² McDonald'ın ω katsayısı formülünü çok faktörlü yapılarda tüm ölçek için güvenirlik kestirimi yapabilecek şekilde düzenlemişler ve yeni katsayıya "Konjenerik Ölçmeler için Bileşik Güvenirlik Katsayısı" adını vermişlerdir. McDonald'ın Omega'sı ve ondan türetilen güvenirlik katsayıları, genelde Doğrulamalı Faktör Analizi (DFA) ile elde edilen faktör yükleri üzerinden hesaplanmaktadır. Ölçeğin nihai formu, deneme uygulaması örnekleme gibi hedef kitleyi yansıtan farklı bir örnekleme uygulanarak DFA yapılabilir. Bu işlem sonrası ω iç tutarlılık katsayısı elde edilebileceği gibi ölçeğin yapısal kararlılığına ilişkin daha fazla

kanıt elde edilebilir. Bu makalede araştırmacılara, çok faktörlü yapılarda her alt test için ayrı ayrı iç tutarlılık kestirimi yapmaları ve bir de ölçeğin tümü için uygun bir yöntem seçerek güvenirlik kestirimi yapmaları önerilmektedir.

Ölçeğin Geçerlik Analizlerinin Yapılması

Bir ölçme aracının geçerliği, aracın ölçme amacına hizmet etme derecesi ve ölçümlere ölçülen özellik dışında başka değişkenleri karıştırmaması olarak ifade edilebilir⁵⁸. Başka bir deyişle bir ölçeğin geçerliği, onun, istenilen özelliği ölçme ve bu işi diğer özelliklerin etkilerine kapalı kalarak yani onların etkilerini ölçümlere yansıtmadan yapma derecesidir denilebilir⁹. Geleneksel olarak, geçerlik türleri, **kapsamla bağlantılı, ölçütle bağlantılı ve yapıyla bağlantılı** geçerlik olarak gruplanmaktadır. Bu kategoriler arasında kesin ayrımlar yapmak mümkün değildir. Örneğin ölçüt veya kapsamla bağlantılı geçerlik verileri, aynı zamanda yapı geçerliğiyle ilişkilidir. İdeal bir geçerlik saptama, bu üç geleneksel kategoriye de kapsayan bilgileri içermelidir.⁷³

Bu makalede, geçerliğe ilişkin incelemelerin son adımı verilmiş olması, geçerlik kanıtlarının ölçek geliştirme sürecinin sonunda elde edileceği anlamına gelmemektedir. Örnek olarak, ölçeğin madde örnekleminin ölçülmesi amaçlanan yapıyı ne derece yansıttığının bir göstergesi olan kapsam geçerliğinin sağlanmasına yönelik literatür taraması, hasta görüşmeleri, uzman görüşü gibi etkinlikler, madde üretiminin öncesinde yapılmaktadır. Deneme uygulaması sonrası yürütülen faktör analizi aşamasında da ölçeğin yapısal geçerliğine ilişkin kanıtlar elde edilmektedir. Araştırmacı, AFA'nın yanı sıra DFA'da yapacaksa ortalama açıklanan varyanslar (OAV) üzerinden ölçeğin yakınsama geçerliğini (convergent validity) ve ayırt edici geçerliğini (discriminant validity) inceleyerek yapısal geçerliğe ilişkin kanıt elde edebilir.⁷⁴ Diğer tüm aşamalar tamamlandıktan sonra nihai form elde edildikten sonra da, ölçeğin geçerliği kanıtlanmış bir dış ölçüt (başka bir ölçme aracı) karşılaştırılması ile ölçüt geçerliğine ilişkin kanıt elde edilebilir.

SONUÇ

Davranış bilimlerinde, veri toplama gereksinimi ortaya çıktığında, araştırmacılar için üç durum söz konusudur. İlki, başkaları tarafından geliştirilmiş ve aynı özellikleri yoklayan bir ölçme aracı bularak veri toplamak, ikincisi, aynı özellikleri yoklayan ve başka dil ve kültür için geliştirilmiş bir ölçeği Türkçeye uyarlamak, üçüncüsü ise bunlar mümkün değilse, yeni bir ölçme aracı geliştirmektir. Araştırmacılar için her üç durumda da önemli olan

şey, ölçme aracının hedef kitle için yapıyı güvenilir ve geçerli bir biçimde ölçüp ölçmediğinin kanıtlanmasıdır. Bu konuda, kanıt arama işlemi uzmanlık gerektiren bir dizi işlemi içerir ve bu işlemler her biri titiz ve dikkatli bir biçimde yerine getirilmesi gereken kapsamlı bir süreci vurgular. Araştırmacılar bu süreçte, başından sonuna kadar izlenecek her bir ayrı yöntem ve tekniğin, ölçeğin nihai formu üzerinde önemli farklılıklara yol açabileceğini unutmamalıdır. Örneğin, yalnızca faktör analizi aşamasında bile araştırmacının tercih edebileceği pek çok faktörleştirme, faktör sayısı belirleme ve döndürme tekniği vardır. Araştırmacıların, bu teknikler konusunda bilgili olması, kendi koşullarını en iyi karşılayan yönetime karar vermesi açısından önemlidir. Çünkü, farklı teknikler ölçeğin nihai formunun psikometrik özellikleri üzerinde farklı etkiler meydana getirebilir. Burada önemli olan, alınacak kararların ölçülen değişkenin özelliklerine ve elde edilen verilerin ortaya koyduğu ilişkilerin doğasına uygun olması ve hangi yöntem ve tekniklerin hangi durumlara uygun olduğunun araştırmacı tarafından bilinmesidir. Unutulmamalıdır ki, ölçek geliştirme işlemi aynı zamanda hem bilgi hem de deneyim gerektiren bir araştırma sürecidir. Araştırma, ilgili yapıyı en az hata ve en yüksek güvenilirlik ve geçerlik katsayısını elde edene kadar sürdürülmelidir. Bu makalede, ölçek geliştirmeye ilgili işlemler, büyük ölçüde Klasik Test Kuramı (gerçek puan kuramı-true score theory) ve açıklayıcı faktör analizine (AFA) dayalı olarak ele alınmıştır. Madde Tepki Kuramı'na (item response theory) ve Sato Test Kuramı'na (sato test theory) dayalı madde seçimi, güvenilirlik ve geçerlik analizleri ile ölçek geliştirme sürecinde doğrulayıcı faktör analizine (DFA) dayalı uygulamalar sonraki çalışma veya çalışmalarda ayrıca ele alınabilir. Bu makalenin ölçek geliştiricilere yol göstereceği umulmaktadır.

KAYNAKLAR

1. Cronbach, L. J. Essentials of psychological testing. New York: Harper Collins Publishers; 1960; 105-126.
2. Özgüven, İ. E. Psikolojik testler. Ankara: Pdem Yayınları. 2011; 46-52.
3. Anastasi, A. Psychological testing. New York: Macmillan Publishing Co., Inc. 1982; 34.
4. McReynolds, P. The motivational psychology of Jeremy Bentham: I. Background and general approach. *J. Hist. Behav. Sci.*, 1968; 4: 230-244.
5. Crocker, L. ve Algina, J. Introduction to classical and modern test theory. New York: Holt, Rinehart and Winston, Inc. 1986; 66-84.
6. Murphy K.R. ve Davidshofer C.O. Psychological testing: principles and applications. New Jersey: Pearson Education International; 2005; 30-260.
7. Furr M.R. ve Bacharach V.R. Psychometrics: an introduction. California: Sage Publications; 2008; 24-180.
8. Cohen R.J. ve Swerdlik M.E. Psychological testing and assessment. Boston: McGraw-Hill Companies; 2010; 120- 190.
9. Tavşancıl, E. Tutumların ölçülmesi ve SPSS ile veri analizi. Ankara: Nobel Yayıncılık; 2010; 34-85.
10. Edenborough R. Using psychometrics: a practical guide to testing and assessment. London: Kogan Page; 1999; 54-55.
11. Erkuş, A. Psikolojide ölçme ve ölçek geliştirme-1: Temel kavramlar ve işlemler. Ankara: Pegem Akademi; 2012a; 25-112.
12. Şahin, N. Psikoloji araştırmalarında ölçek kullanımı. *Türk Psikoloji Dergisi*, 1994; 9(33), 19-26.
13. Hambleton, R.K. ve Patsula, L. Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology*, 1999; 1(1), 1-30.
14. Çüm, S. ve Koç, N. Türkiye’de psikoloji ve eğitim bilimleri dergilerinde yayımlanan ölçek geliştirme ve uyarlama çalışmalarının incelenmesi. *Eğitim Bilimleri ve Uygulama*, 2013; 12(24), 115-135.
15. Turgut, F. Test geliştirme teknikleri: Ders notları. Ankara: Hacettepe Üniversitesi Yayınları. 1978. 15-17
16. Rust J. ve Golombok S. Modern Psychometrics: the science of psychological assessment. New York: Routledge. 1997; 46-81.
17. Tezbaşaran, A. Likert tipi ölçek hazırlama kılavuzu. Ankara: Türk Psikologlar Derneği Yayınları. 2008. 27-48
18. Coaley K. Psychological assessment and psychometrics. California: Sage Publications. 2010; 255-262.
19. Cordes, C.L. ve Dougherty, T.W. A review and an integration of research on job burnout. *Academy of Management Review*, 1993; 18(4), 621-656.
20. Huberman, A.M. ve Vandenberghe, R. Understanding and Preventing Teacher Burnout. Cambridge: Cambridge University Press. 1999; 62.
21. Brown, T.A. Confirmatory factor analysis for applied research. New York: The Guilford Press. 2015; 19-33.
22. Field, A. Discovering statistics using SPSS. London: Sage Publications Ltd. 2009; 627-681.
23. Williams, B., Onsmann, A. ve Brown, T. Exploratory factor analysis: A five step guide for novices. *Journal of Emergency Primary Health Care*, 2010; 8(3), 1-13.
24. Tabachnick, B. ve Fidell, L. Using multivariate statistics. New York: Herper Collins College Publishers. 1996; 764-792.
25. Hair, J., Anderson, R.E., Tatham, R.L. ve Black, W.C. Multivariate data analysis. 4th ed. New Jersey: Prentice-Hall Inc. 1995; 42-78.
26. Comrey, A.L. ve Lee, H.B. A first course in factor analysis. Hillsdale, NJ: Erlbaum. 1992; 22-24.
27. Gorsuch, R. L. Factor analysis (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum. 1983; 33-40.
28. Hatcher, L. A step-by-step approach to using the SAS. Cary, NC: SAS Institute, Inc. 1994; 210.
29. Bryant, F.B. ve Yarnold, P.R. Principal components analysis and exploratory and confirmatory factor analysis. Washington, DC: American Psychological Association. 1995; 270.
30. Suhr, D. Exploratory or Confirmatory Factor Analysis. SAS Users Group International Conference 2006; pp. 1 – 17, Cary: SAS Institute, Inc.
31. Everitt, S. Multivariate analysis: The need for data and other problems. *British Journal of Psychiatry*. 1975;126, 227-240.
32. Nunnally, J.C. Psychometric theory. New York: McGraw-Hill. 1978; 430-520.
33. Velicer, W.F. ve Fava, J.L. Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, 1998 . 3(2), 231-251.
34. Cool, A. L. A review of methods for dealing with missing data (rapor). Annual Meeting of the Southwest Educational Research Association. Dallas. 2000.
35. Roth, P. L. Missing data: A conceptual review for applied psychologists. *Personnel Psychology*, 1994;3(1), 537-560.

36. Alpar, R. Çok deęişkenli istatistiksel yöntemler. Ankara: Detay Yayıncılık. 2011; 286-301.
37. Schafer, J. L. Multiple imputation: a primer. *Statistical Methods on Medical Research*, 1999; 8(1), 3-15.
38. Osborne, J. W. Best practices in data cleaning. California: Sage Publication, Inc. 2013; 18-60.
39. Pett, M.A., Lackey, N.R. ve Sullivan, J.J. Making Sense of Factor Analysis: The use of factor analysis for instrument development in health care research. California: Sage Publications. 2003; 23-70.
40. Beavers, A.S., Lounsbury, J.W., Richards, J.K., Huck, S.W., Skolits, G.J. ve Esquivel, L. Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research and Evaluation*, 2013; 18(6), 1-13.
41. Byrne, B.M. Structural equation modeling with AMOS: Basic concepts, applications and programming. Mahwah, NJ: Lawrence Erlbaum Associates. 2001; 92-93.
42. Brown, J.D. Principal component analysis and exploratory factor analysis: Definitions, differences and choices. *JALT Testing and Evaluation Newsletter*, 2009; 13(1), 26-30.
43. Osborne, J.W. ve Costello, A.B. Best practices in exploratory factor analysis: Four recommendations for getting the most from you analysis. *Pan-Pacific Management Review*, 2009; 12(2), 131-146.
44. Widaman, K.F. Common factor analysis versus principal component analysis: Differential bias in representing model parameters. *Multivariate Behavioral Research*, 1993; 28(3), 263-311.
45. Snook, S.C. ve Gorsuch, R.L. Component analysis versus common factor analysis: A monte carlo study. *Psychological Bulletin*, 1989; 106, 148-154.
46. Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., ve Strahan, E. J. Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 1999; 4(3), 272-299.
47. Coughlin, K.B. An analysis of factor extraction strategies: A study of the relative strenghts of principal axis, ordinary least squares and maximum likelihood factor extraction methods in research contexts. Yayımlanmamış doktora tezi, University of South Florida, Tampa, FL. 2013.
48. Horn, J. L. A rationale and test for the number of factors in factor analysis. *Psychometrika*, 1965; 30(2), 179-85.
49. Linn, R. L. A Monte Carlo approach to the number of factors problem. *Psychometrika*, 1968; 33, 37-71.
50. Silverstein, A. B. Note on the parallel analysis criterion for determining the number of common factors or principal components. *Psychological Reports*, 1987; 61, 351-354.
51. Eaton, C. A., Velicer, W. F. ve Fava, J. L. Determining the number of components: An evaluation of parallel analysis and the minimum average partial correlation procedures. Unpublished manuscript. 1999.
52. Zwick, W. R. ve Velicer, W. F. Factors influencing five rules for determining the number of components to retain. *Psychological Bulletin*, 1986; 99, 432-442
53. Hayton, J.C., Allen, D.G. ve Scarpello, V. Factor retention decision in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, 2004; 7(2), 191-205.
54. Thompson, B. Exploratory and confirmatory factor analysis: understanding concepts and applications. Washington, DC: American Psychological Association. 2004; 44-47.
55. Henson, R.K. ve Roberts, J.K. Use of exploratory factor analysis in published research: common errors and some comment on improved practice. *Educational and Psychological Measurement*, 2006; 66, 393-416.
56. Kim, J.O. ve Mueller, C.W. Introduction to factor analysis: What it is and how do it. Beverly Hills, CA: Sage Publications. 1978; 54.
57. Yurdabakan, İ. Eğitimde Kullanılan Ölçme Araçlarının Nitelikleri. In S. Erkan & M. Gomleksiz (Eds.), *Eğitimde ölçme ve değerlendirme*. Ankara: Nobel Yayın Dağıtım. 2008; 38-66.
58. Turgut, F. ve Baykul, Y. Eğitimde ölçme ve değerlendirme. Ankara: Pegem Akademi Yayıncılık. 2012; 76-89.
59. Yang, Y. ve Green, S.B. Coefficient alpha a reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment*. 2011; 29(4), 377-392.
60. Schmitt, N. Uses and abuses of coefficient alpha. *Psychological Assessment*, 1996; 8, 350-353.
61. Osburn, H. G. Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods*, 2000; 5, 343-355.
62. Cronbach, L. J. My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement*, 2004; 64, 391-418.
63. Tan, Ş. Misuses of KR-20 and Cronbach's Alpha Reliability Coefficients. *Education and Science*, 2009; 34(152), 101-112.
64. Erkuş, A. Psikometri üzerine yazılar. Ankara: Türk Psikologlar Derneği Yayınları. 2003; 69.
65. Tavakol, M. ve Dennick, R. Making sense of cronbach's alpha. *International Journal of Medical Education*, 2011; 2, 53-55
66. Widhiarso, W. ve Ravand, H. Estimating reliability coefficient for multidimensional measures: A

-
- pedagogical illustration. *Review of Psychology*, 2014; 2, 111-121
67. Cronbach, L. J., Schonemann, P. ve McKie, D. Alpha coefficients for stratified-parallel tests. *Educational and Psychological Measurement*, 1965; 25, 291-312.
68. Moiser, C. On the reliability of a weighted composite. *Psychometrika*, 1943; 8(3), 161-168.
69. Wang, M.V. ve Stanley, J.C. Differential weighting: A review of methods and empirical studies. *Review of Educational Research*, 1970; 40, 663-705.
70. Novick, M.R. ve Lewis, C. Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. *Structural Equation Modeling*, 1967; 2, 255-273.
71. Yurdugül, H. The comparison of reliability coefficients in parallel, tau-equivalent and congeneric measurements. *Journal of Educational Sciences*, 2006; 39(1), 15-37.
72. Raykov, T. ve Shrout, P.E. Reliability of scales with general structure: Point and interval estimation using a structural equation modeling approach. *Structural Equation Modeling*, 2002; 9(2), 195-212.
73. Amerikan Eğitim Araştırmaları Birliği, Amerikan Psikoloji Birliği, Eğitim Ölçümleri Uluslar arası Konseyi. Eğitimde ve psikolojide ölçme standartları (S. Hovardaoğlu ve N. Sezgin, Çeviri). Ankara: Türk Psikologlar Derneği ve ÖSYM yayını. 1997; 62-70.
74. Yurdugül, H. ve Alsancak Sırakaya, D. Çevrimiçi öğrenme hazır bulunuşluluk ölçeği: Geçerlik ve güvenirlik çalışması. *Eğitim ve Bilim*, 2013;169(38), 391-406.