



## Açımlayıcı ve Doğrulayıcı Faktör Analizlerinin Kullanımı: Bir Doküman İncelemesi \*

İlhan Koyuncu <sup>1</sup>, Abdullah Faruk Kılıç <sup>2</sup>

### Öz

Bu çalışmanın amacı, Türkiye’de 2006-2016 yılları arasında yayınlanan ölçek geliştirme çalışmalarını açımlayıcı faktör analizi (AFA) ve doğrulayıcı faktör analizinin (DFA) kullanımı bakımından incelemektir. Bu kapsamda, yapılan çalışmaların yıllara ve kullanılan faktör analizi türüne göre nasıl bir dağılım gösterdiği, her iki analiz yöntemi için varsayımların sağlanma durumları ile açımlayıcı ve doğrulayıcı faktör analizi sonuçlarının makalelere göre nasıl bir dağılım gösterdiği incelenmiştir. Böylece, teorik olarak sıkça değinilen AFA ve DFA’nın pratikte nasıl uygulandığına yönelik önemli bulguların elde edilmesi amaçlanmıştır. Derleme çalışması niteliğinde olan bu araştırmada, 2006 ile 2016 yılları arasında Türkiye Bilimsel ve Teknik Araştırma Kurumu (TÜBİTAK) Dergi Park veri tabanında taranan dergilerde eğitim ve sosyal bilimler alanlarında yayınlan toplam 131 ölçek geliştirme çalışması incelenmiştir. Verilerin incelenmesinde frekans tabloları, yüzdeler, sütun, histogram ve çizgi grafikleri ve merkezi eğilim ölçüleri kullanılmıştır. Araştırmada incelenen 131 ölçek geliştirme çalışmasının %70’inden fazlası 2013 yılından sonra yapılmıştır. En fazla çalışma 2015 yılında, en az çalışma ise 2008 yılında yapılmıştır. Makalelerin 59’unda sadece AFA, dördünde sadece DFA ve 58’inde her iki analiz birlikte yapılmış, onunda ise her iki analiz de yapılmamış sadece madde analizlerine yer verilmiştir. AFA için kullanılan ortalama örneklem büyüklüğü 395 kişiyken DFA için 529 kişidir. Buna göre, faktör analizleri için örneklem büyüklüklerinin genel olarak yeterli olduğu görülmüştür. 36 çalışmada AFA ve DFA aynı örneklem üzerinde gerçekleştirilirken, 22’sinde farklı örneklem üzerinde analizler yapılmıştır. AFA yapılan 117 çalışmanın %65’inde kestirim yöntemi olarak Temel Bileşenler Analizi (TBA), %2’sinde Maximum Likelihood Estimation (MLE), %1’inde Paralel Analiz (PA) ve %3’ünde Paralel Axis Factoring (PAF) kullanılmış, %29’unda belirtilmemiştir. AFA analizinin varsayımları çalışmaların %93’ünde incelenmiştir. Faktör sayısına karar verirken çalışmaların %50’sinde yamaç grafiği göz önünde

### Anahtar Kelimeler


Açımlayıcı faktör analizi (AFA)  
Doğrulayıcı faktör analizi (DFA)  
Doküman incelemesi  
Ölçek geliştirme  
Geçerlik  
Yapı geçerliği


### Makale Hakkında

Gönderim Tarihi: 15.12.2017  
Kabul Tarihi: 12.02.2019  
Elektronik Yayın Tarihi: 22.04.2019

DOI: 10.15390/EB.2019.7665

\* Bu makale IV. Uluslararası Avrasya Eğitim Araştırmaları Kongresi’nde sunulan "Türkiye’de 2006-2016 Yılları Arasında Yayınlanan Ölçek Geliştirme Çalışmalarının Faktör Analizi Açısından İncelenmesi" başlıklı bildirinin genişletilmiş sürümüdür.

<sup>1</sup>  Adıyaman Üniversitesi, Eğitim Fakültesi, Eğitimde Ölçme ve Değerlendirme, Türkiye, [ilhankync@gmail.com](mailto:ilhankync@gmail.com)

<sup>2</sup>  Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Eğitimde Ölçme ve Değerlendirme, Türkiye, [abdullahfarukkilic@gmail.com](mailto:abdullahfarukkilic@gmail.com)

bulundurulmuştur. AFA yapılan araştırmaların %67'sinde Varimax döndürme yöntemi kullanılmıştır. Açıklanan ortalama toplam varyans oranı %53.05'tir. DFA yapılan 62 çalışmanın %87'sinde analizin varsayımları incelenmemiş, 51'inde yol diyagramları, 52'sinde faktör yükleri verilmiştir. Çalışmaların %90'ında parametre kestirim yöntemiyle ilgili herhangi bir bilgi verilmemiştir. Çalışmalarda model uyum/uyumsuzluk indeksleri olarak Parsimony Goodness of Fit Index (PGFI) dışında diğer tüm indekslerin kabul edilebilir düzeyde olduğu görülmüştür. Çalışmanın son bölümünde incelenen araştırmalardan elde edilen bulgular, ilgili alan-yazın ışığında değerlendirilerek, teori ve uygulamaya dönük bazı önerilerde bulunulmuştur.

## Giriş

Psikolojide ve davranış bilimlerinde belirli bir yapının göstergeleri kullanılarak bu yapı ile ilgili birtakım sayısal ölçümlerin elde edilmesinde sistematik bir süreç olan ölçek geliştirmeye başvurulur. Davidshofer ve Murphy (2005) ölçek geliştirmenin genel olarak ölçeğin oluşturulması, standardizasyonu ve güncellenmesi olmak üzere üç ana kısımdan oluşan bir süreç olduğunu ifade etmişlerdir. Ancak, bazı araştırmacılar ölçek geliştirmeyi, ölçeğin oluşturulması süreci ile sınırlarken, bazıları standardizasyonunu da sürece dâhil etmişlerdir. Crocker ve Algina (1986) ölçek geliştirme sürecinin sırasıyla ölçek puanlarının kullanım amacının belirlenmesi, davranışların tanımlanması, belirtke tablosunun hazırlanması, madde havuzunun oluşturulması, madde redaksiyonu, ön deneme uygulaması, deneme uygulaması, madde analizleri, geçerlik ve güvenilirlik çalışması ile standardizasyon adımlarından oluştuğunu belirtmişlerdir. Clark ve Watson (1995) ise ölçek geliştirmenin, hedef yapının net bir tanımını, dikkatli bir madde havuzu oluşturma sürecini, havuzdaki maddelerin temsil edici bir örneklem üzerinde test edilmesini ve maddeler arası korelasyon ve faktör analizi ile boyutluluk ve ayırıcılık geçerliklerini kapsayan bir süreç olduğunu ifade etmiştir. Erkuş (2012) tarafından önerilen ölçek geliştirme sürecinde ise Crocker ve Algina'nın (1986) adımlarına ek olarak ölçek geliştirme tekniğinin seçilmesi, açıklamalar ve yönergelerin yazılması ve eğer gerekiyorsa tekrardan deneme uygulamalarına başvurulması yer almaktadır. Tanımlamalar ya da sürecin aşamaları farklılıklar gösterse de ölçek geliştirmenin ortak öncelikli amacı, altta yatan psikolojik yapının geçerli bir ölçüsünü oluşturmaktır (Clark ve Watson, 1995).

Bir ölçme aracının geçerliği, test puanlarına dayanarak yapılan çıkarımların desteklenmesi amacıyla kanıtların toplanması sürecidir (Cronbach, 1984). American Psychological Association [APA], American Educational Research Association [AERA] ve National Counseling on Measurement in Education [NCME] (1974) tarafından yapılan genel kabul görmüş sınıflandırmada geçerlik kanıtları kapsam, yapı ve ölçüt temellidir. Kapsam geçerliği, maddelerin, konu ve davranış alanının yeterince temsil edici bir örnekleme olması ile ilgilidir (Cronbach ve Meehl, 1955). Kapsam geçerliği için kanıt sunmak amacıyla genellikle uzman görüşlerine başvurulur. Ölçüt geçerliği testin ölçmek isteneni ölçüyor olması ile ilgilidir (Cureton, 1951). Ölçüt temelli kanıtlar elde etmek amacıyla ölçeğin aynı yapıyı veya alt yapıları ölçen geçerliği ve güvenilirliği kanıtlanmış alan yazında yer alan diğer ölçeklerle uyum derecesine bakılır. Yapı geçerliği ise test puanlarının ölçülmek istenen yapıyı sadece bu yapının göstergeleri ile ölçüp ölçmediğini gösterir (Kline, 2016). Başka bir ifadeyle, yapı geçerliği, geliştirilen test veya ölçekle ölçülmek istenen yapının ölçülme derecesine karşılık gelir. Cronbach ve Meehl (1955), yapı geçerliğine ilişkin kanıtlar elde etmek amacıyla grup farklılıkları, korelasyon matrisleri ve faktör analizi, iç yapı çalışmaları, şartlarının değişiminin incelendiği çalışmalar ve süreç çalışmaları yöntemlerinin kullanılabileceğini ifade etmiştir.

Uygulamalı araştırmalarda, faktör analizi testlerin psikometrik değerlendirmesinde (yapı geçerliği için) en yaygın kullanılan yöntemdir (Brown, 2015). Kline'a (2016) göre faktör analizi, karmaşık verileri indirgemeyi ya da basitleştirmeyi hedefleyen ve bir korelasyon ya da kovaryans matrisini temel alan istatistikî tekniklerin birleşimidir. Bu nedenle, faktör analizi basit madde analizlerine göre bir

yapıyı en iyi açıklayan (varyans miktarı açısından) maddelerin ya da faktör yapılarının tespit edilmesinde son derece etkili bir yöntemdir. Temelleri 1900'lü yıllarda atılan (örn. Spearman, 1904) bu analiz yönteminde açımlayıcı ve doğrulayıcı olmak üzere iki teknik kullanılmaktadır. Faktör analizi teknikleri için önemli varsayımlar veri tipi ve kalitesinin uygun olması, değişkenlerin dağılımının bilinmesi ve yeterli örneklem büyüklüğüne ulaşılmasıdır (Floyd ve Widaman, 1995). Açımlayıcı faktör analizinde (AFA), araştırmacılar en uygun faktör sayısını belirlemek ve ölçülen değişkenlerin (maddelerin) çeşitli gizil boyutun makul göstergeleri olup olmadığını ortaya çıkarmayı amaçlar (Brown, 2015). Doğrulayıcı faktör analizi (Jöreskog, 1967, 1971) ise hipotetik bir faktör yapısının, ölçülen değişkenlerin gözlenen kovaryans yapısına uygunluğunu test eder (Floyd ve Widaman, 1995). Açımlayıcı ve doğrulayıcı faktör analizleri benzer bazı varsayımlar gerektirmenin yanında kendilerine özgü süreçler içeren çok değişkenli istatistiksel yöntemlerdir. Bu çalışmada, açımlayıcı ve doğrulayıcı faktör analizlerinin kullanımları ele alındığından, her bir yöntemin hangi süreçlerden oluştuğuna ve bu süreçlerin içeriklerine değinilmiştir. Bu amaçla, öncelikle açımlayıcı daha sonra ise doğrulayıcı analizle ilgili detaylı bilgilere yer verilmiştir.

Açımlayıcı faktör analizinin gerçekleştirilmesinde, örneklem büyüklüğünün yeterli olup olmaması, kullanılan faktör çıkartma yönteminin seçilmesi, Kaiser-Meyer-Olkin (KMO) ve Bartlett testlerinin incelenmesi, yamaç grafiği ve faktör yüklerinin incelenmesi, faktör döndürme yönteminin seçilmesi, faktörlerin adlandırılması ve toplam açıklanan varyans oranlarının rapor edilmesi başlıca işlemlerdir. AFA için önerilen örneklem büyüklüğü en az 300 kişi olmalı ve her bir değişken (madde) için gözlem sayısı 5 ile 10 kişi arasında olmalıdır (Comrey ve Lee, 1992; Tabachnick ve Fidell, 2012). Ayrıca, Comrey ve Lee (1992) örneklem büyüklüğü için 100'ün az; 200'ün orta; 300'ün iyi; 500'ün çok iyi; 1000 kişi ve daha fazlasının mükemmel olduğunu ifade etmişlerdir. Değişken başına en az 10 gözlem ya da faktör başına 30 gözlem şeklinde öneriler olmasına rağmen, daha büyük örneklem büyüklüğü veride daha az hata oluşmasını ve dolayısıyla AFA'nın daha iyi çalışmasını sağlar (Yong ve Pearce, 2013). Ancak, Guadagnoli ve Velicer (1988) faktör yüklerinin birçok madde için 0,80'den büyük olması durumunda 150 kişi ve üzeri bir örneklem yeterli olabileceğini öne sürmüşlerdir. AFA için örneklem büyüklüğünün yeterliliği için KMO, verilerin çok değişkenli normallik varsayımlarını sağlayıp sağlamadığını test etmek için Bartlett testi yapılır. Bu amaçla, KMO değerinin 0,50'den büyük, Bartlett testinin anlamlı olması beklenir (Hair, Anderson, Tatham ve Black, 1995; Tabachnick ve Fidell, 2012).

Analizlerde uygun bir faktörleştirme yönteminin seçilmesi, elde edilen faktör yapısının geçerliği açısından son derece önemlidir. Faktör çıkartma yöntemi olarak; Temel Bileşenler Analizi (TBA), Temel Faktörleştirme Analizi (TFA), En Çok Olabilirlik, İmaj Faktörleştirme, Alfa Faktörleştirme, En Küçük ve Genelleştirilmiş Kareler Yöntemi gibi çok çeşitli yöntemler olmasına rağmen en çok kullanılanları TBA ve TFA'dır (Henson ve Roberts, 2006; Tabachnick ve Fidell, 2012). TBA yöntemi ise TFA'ya göre daha çok kullanılmaktadır (Thompson, 2004). TFA yöntemi genellikle çok değişkenli normallik varsayımı sağlanmadığında tercih edilir (Costello ve Osborne, 2005).

Açımlayıcı analizde, her bir maddenin belirli bir faktördeki yükünü maksimum yapmak ve faktör yapısının daha iyi yorumlanmasını sağlamak amacıyla faktör döndürmesine başvurulur (Rummel, 1988; Yong ve Pearce, 2013). Faktör döndürme yöntemleri, genel olarak dik ve eğik döndürme olmak üzere ikiye ayrılır (Tabachnick ve Fidell, 2012). Eğik döndürme, genellikle faktörler birbirleriyle ilişkili olduğunda; dik döndürme ise bu ilişki daha az olduğunda kullanılır (Costello ve Osborne, 2005; Rummel, 1988). En çok kullanılan dik döndürme yöntemleri Quartimax ve Varimax yöntemleridir. Quartimax, birden fazla maddeye yüksek yük veren faktör sayısını; Varimax yöntemi ise birden fazla faktörde yüksek faktör yükü olan madde sayısını minimize etmeyi amaçlar (Gorsuch, 1983; Yong ve Pearce, 2013). En çok kullanılan eğik döndürme yöntemleri ise Direct Oblimin ve Promax'tır. İki yöntem de çıktıdaki yapıyı basitleştirmeyi amaçlamasına rağmen Promax büyük ve karmaşık veride daha etkilidir (Gorsuch, 1983).

Faktör sayısına karar verilmesi aşamasında genellikle birden fazla kriter göz önünde bulundurulur (Hair vd., 1995). Bu yöntemlerden bazıları Kaiser ölçütü (özdeğerler  $\geq 1$ ), yamaç grafiği testi, açıklanan toplam varyans miktarı ve paralel analizdir. Faktör çıkartma işlemi, toplam açıklanan

varyans oranı %95'e ulaştığında durdurulur. Sosyal bilimlerde toplam açıklanan varyans oranı genellikle %50-60'a kadar düşebilmektedir. Yamaç grafiği testi için kişi sayısı en az 200 olmalıdır. Yamaç grafiği testi (Cattell, 1978) için kırılma noktası tespit edilmesi gerekmektedir. Bunun için eğrinin her bir uç noktasından yatay ve dikey doğrular çizilerek kesme noktası belirlenir (Yong ve Pearce, 2013). Costello ve Osborne (2005) bir maddenin birden fazla faktördeki yükleri 0,32'den büyük ise bir çapraz yükleme olduğunu ifade eder. Bu durumda, madde doğasından dolayı bu şekilde ise aynen bırakılması, değilse modelden çıkarılması veya faktör yükleri arasında belirli bir fark olması (genellikle 0,1) gibi kriterler göz önünde bulundurulabilir (Yong ve Pearce, 2013). Teorik olarak, her bir faktörde en az iki ya da üç madde yer almalıdır (Henson ve Roberts, 2006). Açımlayıcı faktör analizinin son aşamasında faktör adlandırması yer alır. Bu aşamada herhangi kesin bir kural yer almamakla birlikte faktörü en iyi temsil eden adlandırmanın yapılması önemlidir (Yong ve Pearce, 2013).

Doğrulamalı faktör analizi için ise örneklem büyüklüğünün yeterli olup olmaması, varsayımların kontrol edilmesi, parametre kestirim yönteminin seçilmesi, yol diyagramı,  $t$  değerleri, faktör yükleri,  $\chi^2$  (Kay-kare),  $\chi^2$ /Serbestlik Derecesi (sd) ve uyum indeksleri değerleri incelenir. Sık raporlanan başlıca uyum indeksleri;

- Yakınsama hatalarının kareleri ortalamalarının karekökü (Root Mean Square Error of Approximation [RMSEA]),
- Artıkların kareleri ortalamasının karekökü (Root Mean Square Residual [RMR]),
- Standartlaştırılmış artıkların kareleri ortalamasının karekökü (Standardized Root Mean Square Residual [SRMR]),
- Normalleştirilmiş uyum indeksi (Normed Fit Index [NFI]),
- Normalleştirilmemiş uyum indeksi Non-normed Fit Index [NNFI]) ya da Tucker Lewis indeksi (Tucker Lewis Index [TLI])
- Karşılaştırılmalı uyum indeksi (Comparative Fit Index [CFI]),
- Uyum iyiliği indeksi (Goodness of Fit Index [GFI]),
- Düzeltilmiş uyum iyiliği indeksi (Adjusted Goodness of Fit Index [AGFI]),
- Fazlalık uyum indeksi (Incremental Fit Index [IFI]),
- Göreli uyum indeksi (Relative Fit Index [RFI]),
- Sıkı normalleştirilmiş uyum indeksi (Parsimonious Normed Fit Index [PNFI]),
- Sıkı normalleştirilmiş uyum iyiliği indeksi (Parsimonious Goodness of Fit Index [PGFI]) değerleridir.

Kline (2016) DFA'nın yapıldığı araştırmalarda en azından RMSEA ve %90 güven aralığı,  $\chi^2$ , serbestlik derecesi ve manidarlık değeri, CFI ve SRMR değerlerinin raporlanmasını önermektedir. Cabrera-Nguyen'a (2010) göre ise model doğru tanımlanmalı, sabit ve değişken parametreleri belirlenmeli, birden fazla uyum indeksi kullanılmalı, faktör yükleri manidarlık değerleri ile birlikte verilmeli, yapılan modifikasyonların nedenleri açıklanmalı ve uygun parametre kestirim yöntemi seçilmelidir. Brown (2015) ise bir DFA çalışmasında rapor edilmesi gerekenleri model tanımlama, verinin analize hazırlanması, model kestirimleri, model değerlendirme ve sonuçlar başlıkları altında açıklamış ve örnek bir raporlamaya yer vermiştir.

DFA yapmak için sağlanması gereken varsayımlar, kullanılan parametre kestirim yöntemine göre farklılık göstermektedir. En Çok Olabilirlik (EÇO) ve En Küçük Kareler (EKK) yöntemleri verinin yapısına göre sıklıkla tercih edilen yöntemlerdir. EÇO yönteminin değişkenlerin sürekli olması, örneklemin geniş olması ve değişkenlerin dağılımlarının çok değişkenli normal olması gibi varsayımları vardır (Brown ve Moore, 2012). Bu varsayımlar sağlanmadığında oluşan model hataları sonucunda anlamlılık değerleri ve model uyum iyiliği indeksleri hatalı kestirilecektir. Değişkenler kategorik olduğunda ise EKK yöntemleri kullanılmaktadır. Küçük ve orta büyüklükteki örneklem ile normalliğin sağlanmadığı durumlar için güçlü (robust) EÇO (Bentler, 1995) ya da ağırlıklandırılmış EKK (Muthén, du Toit ve Spisic, 1997) yöntemleri sıklıkla tercih edilmektedir. EÇO, örneklem büyüklüğüne karşı oldukça hassas olmasına rağmen sıklıkla kullanılmaktadır (Brown ve Moore, 2012).

Bu nedenle, model hatalarının önüne geçmek açısından mümkün olduğunca büyük örneklem seçilmesi gerekmektedir. DFA için yeterli örneklem büyüklüğü farklı koşullara (parametre kestirim yöntemi, madde sayısı, faktör sayısı) göre değiştiğinden örneklem büyüklüğü için genel kabul görmüş bir sayı bulunmamakla birlikte (MacCallum, Widaman, Preacher ve Hong, 2001; Wolf, Harrington, Clark ve Miller, 2013), 300 ve üzeri bir örneklem olması beklenir (Worthington ve Whittaker, 2006). Bu noktada, AFA yapılan örneklemin aynısını DFA’da kullanmak ile ilgili tartışmalar da söz konusudur. Worthington ve Whittaker (2006) aynı örneklem üzerinde AFA ve DFA yapıldığında verinin yapısının deneysel olarak ortaya konacağını ileri sürmüşlerdir. Ancak, Fabrigar, Wegener, MacCallum ve Strahan (1999), örneklemin yeterince büyük olduğu durumda rassal olarak ikiye bölüp birinde AFA diğerinde ise DFA yapılmasını önermişlerdir. Yeterli örneklem büyüklüğüne ulaşıldığında her bir analiz aynı evrenden rastgele seçilen farklı örneklem üzerinde yapılabilir. Örneklemin yeterli olması varsayımı dışında, DFA yapmadan önce verideki eksik ve uç değerler belirlenmeli, tek ve çok değişkenli normallik incelenmeli ve çoklu doğrusal bağlantılık kontrol edilmelidir (Ullman, 2012). Böylece hem veri analizlere uygun hale getirilebilecek, hem de veri yapısına uygun bir parametre kestirim yöntemi seçilebilecektir.

DFA yapılan araştırmalarda faktör yüklerinin manidarlık değerleriyle birlikte verilmesi ve ölçme modellerinin grafiksel olarak gösterilmesi önerilmektedir (Cabrera-Nguyen, 2010). Birçok analiz programı (MPlus, EQS, LISREL, vb.) analiz çıktısında bu değerlere ve grafiklere yer vermektedir. Çıktılarda yer alan ve araştırmalarda sıklıkla raporlanan model değerlendirme ölçütü  $\chi^2$  istatistiğidir. Bu istatistiğin manidar çıkması, modelin veriye iyi uyum göstermediği anlamına gelir. Birçok araştırmada rutin olarak raporlanmasına rağmen,  $\chi^2$  değeri, hesaplanan dağılımın  $\chi^2$  dağılımı olmaması, örneklem büyük olduğunda manidar çıkması ve kestirilen ile beklenen modelin eşit olduğu varsayımına dayanmasından dolayı genellikle göz önünde bulundurulmaz ve ağırlıkla diğer uyum indekslerine göre model uyumuna karar verilir (Brown, 2015).  $\chi^2$  istatistiğinin bu yapısından dolayı araştırmalarda  $\chi^2/sd$  değerinin de model uyumu için iyi bir ölçüt olmadığı söylenebilir. Bazı araştırmacıların (Anderson ve Gerbing, 1984),  $\chi^2/sd$  değerinin 5’ten küçük olmasını model uyumu için kabul edilebilir bir ölçüt olarak görmesine rağmen, Kline (2016) bu değer model uyumu için bir ölçüt olarak kullanılmasının mantıksal ve istatistiksel bir temelini olmadığını belirtmiştir. Benzer şekilde, Wheaton (1987), bu değer model uyumunun değerlendirilmesinde kullanılmamasını önermiştir.

Model uyumunun değerlendirilmesinde kullanılan birçok indeks bulunmaktadır. Bu nedenle, hangi indekslerin raporlanması gerektiği konusunda çeşitli görüşler mevcuttur. Kline (2016) DFA çalışmalarında en azından RMSEA ve %90 güven aralığı,  $\chi^2$  değeri, CFI ve SRMR değerlerinin raporlanmasını, Cabrera-Nguyen (2010) ise birden fazla uyum indeksinin kullanılmasını önermiştir. Brown (2015) uyum indekslerinin mutlak uyum indeksleri ( $\chi^2$ , SRMR ve RMR), sıkı (parsimony) uyum indeksleri (RMSEA) ve karşılaştırmalı uyum indeksleri (CFI-IFI, TLI-NNFI) şeklinde üç gruba ayrıldığını ve raporlamada her gruptan en az bir indeks kullanılması önermiştir. Hu ve Bentler (1999) iki-indeks stratejisine göre NNFI (en az 0,96) ve SRMR (en fazla 0,09), RMSEA (en fazla 0,06) ve SRMR (en fazla 0,09), CFI en az 0,96 ve SRMR (en fazla 0,09) ikili ölçütlerin herhangi birisinin kullanılmasını önermektedir. GFI ve AGFI uyum indeksleri çalışmalarda sıklıkla kullanılmalarına rağmen, yapılan benzetim çalışmalarındaki (Hu ve Bentler, 1998; Marsh, Balla ve McDonald, 1988) zayıf performanslarından dolayı önerilmemektedir. Crowley ve Fan (1997) her bir uyum indeksinin model uyumunun ayrı bir yönüyle ilgili bilgi vermesinden dolayı mümkün olduğunca çok indeks raporlanmasını önermiştir. Genel olarak, uyum indeksleri, örneklem büyüklüğü, modelin karmaşıklık düzeyi, kestirim yöntemi (örn. EÇÖ ve EKK), veri tipi (sürekli veya kategorik), verinin normalliği, modelin yanlış tanımlanma biçimi ve miktarı gibi birçok faktörden farklı Grafiklerde etkilendiğinden dolayı kesin ölçütlerin ifade edilmesi uygun olmayacaktır (Brown, 2015). Bu nedenle, model uyumunun değerlendirilmesinde hangi indekslerin hangi şartlar altında kullanılmasının uygun olacağı ilgili alan-yazından (Hooper, Coughlan ve Mullen, 2008; Hu ve Bentler, 1999; Kline, 2016; MacCallum, Browne ve Sugawara, 1996; Marsh ve Hau, 1996; Steiger, 2007; Tabachnick ve Fidell, 2012; Wheaton, 1987) incelenebilir. Uyum indeksleri için kabul edilebilir değerler ilgili alan-yazından derlenerek Tablo 1’de verilmiştir.

**Tablo 1.** Uyum İndeksleri ve Kabul Edilebilir Kesme Noktaları

İstatistik	Yorum
$\chi^2/sd$	5'ten küçük olmalı (Anderson ve Gerbing, 1984); Kullanılması önerilmemekte (Kline, 2016; Wheaton,1987)
RMSEA	0.06'ya yakın veya az olmalı (Hu ve Bentler, 1999); 0.08'den az kabul edilebilir, 0.05'ten az iyi uyum, 0.1'den büyükse model ret (Browne ve Cudeck, 1993); 0.08-0.10 arası vasat uyum (MacCallum vd., 1996); 0.07'ye yakın veya az olmalı (Steiger, 2007)
RMR	Mümkün olduğunca düşük olmalı (Tabachnick ve Fidell, 2012); Ölçek düzeyine hassas olduğundan tercih edilmiyor (Hooper vd., 2008)
SRMR	0.08'e yakın veya az olmalı (Hu ve Bentler, 1999); 0.05'ten az iyi uyum (Byrne, 1998)
NFI	0.90'dan büyük iyi uyum (Bentler ve Bonnet, 1980); 0.95'ten büyük olmalı (Hu ve Bentler, 1999); 0.80'e kadar kabul edilebilir olduğuna dair görüşler var (Hooper vd.,2008)
NNFI-TLI	0.95'e yakın veya büyük olmalı, 0.90 ile 0.95 arası kabul edilebilir, 0.90'dan az model hatalı (Hu ve Bentler, 1998, 1999)
CFI	0.95'e yakın veya büyük iyi uyum, 0.90 ile 0.95 arası kabul edilebilir, 0.90'dan az model hatalı (Hu ve Bentler, 1998, 1999)
IFI	0.90 ve üzeri iyi uyum (Marsh ve Hau, 1996)
RFI	0.90 ve üzeri iyi uyum (Marsh ve Hau, 1996)
GFI	0.95'ten büyük olmalı (Miles, Shevlin, 1998); 0.90 ve üzeri iyi uyum (Hooper vd., 2008)
AGFI	0.90 ve üzeri iyi uyum (Hooper vd., 2008)
PNFI	Kesin bir sınır yok ama 0.50'ye kadar düşebilir (Hooper vd., 2008; Mulaik vd., 1989); 0.80'den büyük olmalı (Marsh ve Hau, 1996)
PGFI	Kesin bir sınır yok ama 0.50'ye kadar düşebilir (Hooper vd., 2008; Mulaik vd., 1989); 0.80'den büyük olmalı (Marsh ve Hau, 1996)

Tablo 1'de verilen uyum indeksi değerleri incelendiğinde, RMSEA değerinin 0,10'dan ve SRMR değerinin 0,8'den az olmaması önerilmektedir. Uyum iyiliği indekslerine bakıldığında, PNFI ve PGFI değerlerinin 0,50'den, NFI'nın 0,80'den ve diğer indeks değerlerinin 0,90'dan büyük olması tavsiye edilmiştir.

Bu çalışmada, ilgili alan yazında yer alan bazı araştırmaların faktör analizi süreçleri açısından incelenmesi amaçlanmıştır. Böylece, teorik olarak sıkça üzerinde durulan AFA ve DFA'nın pratikte nasıl uygulandığına yönelik önemli bulguların elde edilmesi amaçlanmıştır. Bu genel amaç bağlamında, daha önce ölçek geliştirme ve uyarılama (Acar-Güvendir ve Özer-Özkan, 2015; Bayık ve Gürbüz, 2016; Çüm ve Koç, 2013; Delice ve Ergene, 2015; Erkuş, 2016; Gül ve Sözbilir, 2015; Şahin ve Boztunç Öztürk, 2018; Öztürk, Eroğlu ve Kelecioğlu, 2015) ve faktör analizi süreçlerini inceleyen (Thompson ve Daniel, 1996; Worthington ve Whittaker, 2006) sınırlı sayıda araştırmaya rastlanmıştır. Bu araştırmalarda, Social Science Citation Index'te (SSCI) yer alan makaleler, araştırmaların bir örnekleme veya sadece analiz süreçlerinin nasıl gerçekleştirileceği ele alınmıştır. Ayrıca söz edilen bazı araştırmalarda sadece belirli alanlarda geliştirilen veya uyarılan ölçek geliştirme çalışmaları incelenmiştir. Bu çalışma ise TÜBİTAK ULAKBİM Ulusal Veri Tabanlarında 2006-2016 yılları arasında yayınlanan ölçek geliştirme çalışmalarının tamamını faktör analiz süreçleri açısından incelemesi açısından önem arz etmektedir.

Çalışma 2017 yılında toplanan veriden elde edildiğinden 2016 ve öncesindeki değişimin nasıl gerçekleştiğini görmek amaçlanmıştır. Türkiye’deki araştırmalar açısından en kapsayıcı veri tabanı olan TÜBİTAK ULAKBİM’de sosyal bilimler alanında son on yılda yer alan tüm ölçek geliştirme yayınlarının incelenmesi, AFA ve DFA’nın kullanımına dönük ileride yapılacak çalışmalara ışık tutması açısından önemli görülmektedir. Bu çalışma kapsamında ele alınan araştırma problemleri şu şekildedir: Türkiye’de 2006-2016 yılları arasında yayınlanan ölçek geliştirme çalışmalarında;

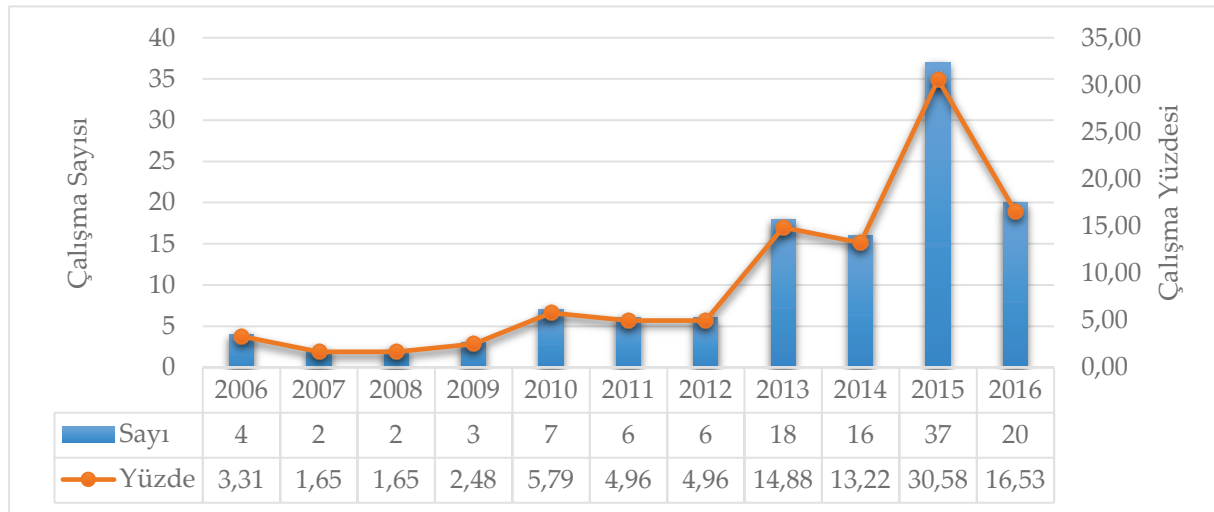
1. Açımlayıcı faktör analizinin kullanımına yönelik elde edilen bulgular nasıldır?
2. Doğrulayıcı faktör analizinin kullanımına yönelik elde edilen bulgular nasıldır?

### Yöntem

Bu çalışmada ölçek geliştirme makalelerinin açımlayıcı ve doğrulayıcı faktör analizleri açısından var olan özellikleri olduğu gibi raporlandığından betimsel bir araştırmadır. Betimsel araştırmalar, analiz birimine ait özelliklerin nedenlerinden çok var olduğu gibi aktarılmasını amaçlar (Fraenkel ve Wallen, 2011). Aynı zamanda bir derleme çalışması olan bu araştırmada doküman analizi yöntemi kullanılmıştır. Diğer araştırma yöntemlerinin tamamlayıcısı olarak kullanıldığı gibi tek başına bir teknik olarak da kullanılan doküman analizi, basılı veya elektronik materyallerin değerlendirilmesi ve gözden geçirilmesi için kullanılan sistematik bir prosedürdür (Bowen, 2009).

#### Çalışma Grubu

Bu araştırma kapsamında, 2006 ile 2016 yılları arasında TÜBİTAK Dergi Park veri tabanında taranan dergilerde eğitim ve sosyal bilimler alanlarında yayınlanan toplam 121 ölçek geliştirme çalışması incelenmiştir. Bu çalışmaların yıllara göre dağılımları Grafik 1’de verilmiştir.



Grafik 1. Çalışmaların Yıllara Göre Dağılımı

Grafik 1’e göre geliştirilen ölçek sayısı zaman içinde önemli bir artış göstermiştir. Araştırmada incelenen ölçek geliştirme çalışmalarının %70’inden fazlası 2013 yılından sonra yapılmıştır. En fazla çalışma 2015 yılında, en az çalışma ise 2008 yılında yapılmıştır (Grafik 1). Ölçek geliştirme çalışmalarının %97’sinde ( $N_{AFA}=117$ ) AFA ve %51’inde ( $N_{DFA}=62$ ) DFA yapılmıştır. Çalışmaların %48’inde ( $N_{iki}=58$ ) her iki analiz birlikte yapılmıştır. Hem AFA hem DFA’nın yapıldığı 58 çalışmanın, %62’sinde ( $N_{AYNI}=36$ ) AFA ve DFA aynı örneklem üzerinde gerçekleştirilirken, %38’inde ( $N_{FARKLI}=22$ ) farklı örneklem kullanılmıştır.

#### Prosedür

Araştırma kapsamında öncelikle hangi veri tabanını ya da tabanlarının incelenmesi üzerine bir inceleme yapılmıştır. Sadece belirli bir indekste yer alan çalışmaları incelemek yerine Türkiye’de gerçekleştirilen çalışmalar bakımından en kapsayıcı veri tabanı olan TÜBİTAK Ulakbim TR Dizinde yer alan Sosyal Bilimler alanındaki yayınlar incelenmiştir. Veri tabanında yapılan aramalarda “psikometrik”, “ölçek”, “faktör analizi”, “geçerlik” ve “geliştirme” anahtar kelimeleri kullanılmıştır. Arama sonuçlarında 2006-2016 yılları arasında yapılmış olan 131 çalışmadan 10’unda AFA ve/veya

DFA analizlerine yer verilmediği için inceleme dışında bırakılmıştır (EK 1). Daha sonra ise AFA ve DFA'nın gerçekleştirilmesi için karşılanması beklenen gerekliliklerin ve raporlanması beklenen olası çıktılarının neler olduğu üzerine alan-yazında yer alan birincil ve ikincil kaynaklardan (Brown, 2015; Cabrera-Nguyen, 2010; Kline, 2016; Tabachnick ve Fidell, 2012; Worthington ve Whittaker, 2006) detaylı bir inceleme yapılmıştır. Araştırmalar sonucunda AFA için örneklem büyüklüğü, faktör çıkartma yöntemi, Kaiser-Meyer-Olkin (KMO) ve Bartlett testleri, yamaç grafiği ve faktör yüklerinin incelenme durumları, faktör döndürme yöntemi, faktör adlandırması ve toplam açıklanan varyans miktarlarının raporlanma durumları incelenmiştir. DFA için ise örneklem büyüklüğü, varsayımların kontrol edilme durumu, parametre kestirim yöntemi, yol diyagramı, t değerleri, faktör yükleri ve  $\chi^2$  anlamlılık değerinin raporlanma durumları,  $\chi^2/sd$  ile RMSEA, RMR, SRMR, NFI, NNFI, CFI, GFI, AGFI, IFI, TLI, RFI ve PGFI incelenmiştir. Ayrıca, her iki analizin yapıldığı çalışmalarda doğrulayıcı faktör analizinin açıklayıcı faktör analizi ile aynı örneklemde yapıp yapılmadığının incelenmesinin de önemli olduğu görülmüştür. Çalışmaların incelenmesi için göz önünde bulundurulması gereken kriterler belirlendikten sonra Microsoft Office Excel 2016 yazılımına veri girişi yapılmıştır.

### *Veri Analizi*

Araştırma kapsamında kullanılan doküman analizi tekniğinin, birçok dökümana kolay ulaşabilme, bireyler veya kişilerle etkileşimin olmaması, geniş bir zaman dilimini inceleyebilme, geniş örneklem büyüklüğüne ulaşılabilmesi, bireyselliğin ve özgünlüğün olması, düşük maliyet ve nitelikli kaynaklara ulaşabilme gibi avantajları vardır. Döküman analizinin, bazı yanlışlık durumları, sadece çok bilenen veya tanınan eserlere ulaşabilme, dolaylı bilgi elde edildiğinden gerçek bilgi eksikliği, bazı durumlarda kaynaklara ulaşabilme zorluğu, örneklemin yanı olabilmesi ve kodlama zorluğu belli başlı eksiklikleri olarak sayılabilir (Bailey, 1982). Döküman incelemesi, dökümanlara ulaşma, dökümanların orijinalliğinin incelenmesi ve anlaşılması, verinin analiz edilmesi ve raporlanması aşamalarından oluşur. Bu çalışmada, on yıllık zaman diliminde, TR Dizinde yayınlanan eğitim ve sosyal bilimler alanlarındaki makalelere ulaşıldıktan sonra yayınlarda yer alan veri incelenerek uygun istatistiksel ve grafiksel yöntemlerle rapor edilmiştir.

Ölçek geliştirme çalışmalarında AFA ve DFA için raporlanan veri, Microsoft Office Excel 2016 programına girildikten sonra, araştırmaların yıllara göre dağılımına yönelik kümelenmiş sütun ve işaretçileri olan çizgi grafiklerine, farklı analiz türlerini kullanan araştırma sayıları ile analizleri aynı örneklemde gerçekleştiren araştırma sayıları ve yüzdelerine yer verilmiştir. AFA yapılan araştırmalar için;

- ✓ Ortalama örneklem büyüklüğü,
- ✓ Kullanılan faktör çıkartma yöntemi,
- ✓ Faktör sayıları ve açıklanan varyans oranları,
- ✓ Döndürme yöntemleri,
- ✓ Kaiser-Meyer-Olkin (KMO) ve Bartlett testleri,
- ✓ Yamaç grafiği ve faktör yüklerinin incelenme durumları ile
- ✓ Faktör adlandırmasının yapıp yapılmadığına yönelik frekanslara, yüzdeler, kümelenmiş sütun ve işaretçileri olan çizgi grafiklerine ve frekans tablolarına yer verilmiştir.

DFA yapılan araştırmalar için ise

- ✓ Ortalama örneklem büyüklüğü,
- ✓ Kullanılan parametre kestirim yöntemi,
- ✓  $\chi^2/sd$  ve ortalama uyum indeksi değerleri,
- ✓ Varsayımların kontrol edilme durumu,
- ✓ Yol diyagramı,
- ✓ t değerleri, faktör yükleri ile
- ✓  $\chi^2$  anlamlılık değerinin raporlanma durumlarına yönelik frekanslara, yüzdeler, kümelenmiş sütun ve işaretçileri olan çizgi grafiklerine ve frekans tablolarına yer verilmiştir.

Uyum indeksi değerlerinin yorumlanmasında Tablo 1'de verilen aralık ve eşik değerlerinden yararlanılmıştır.

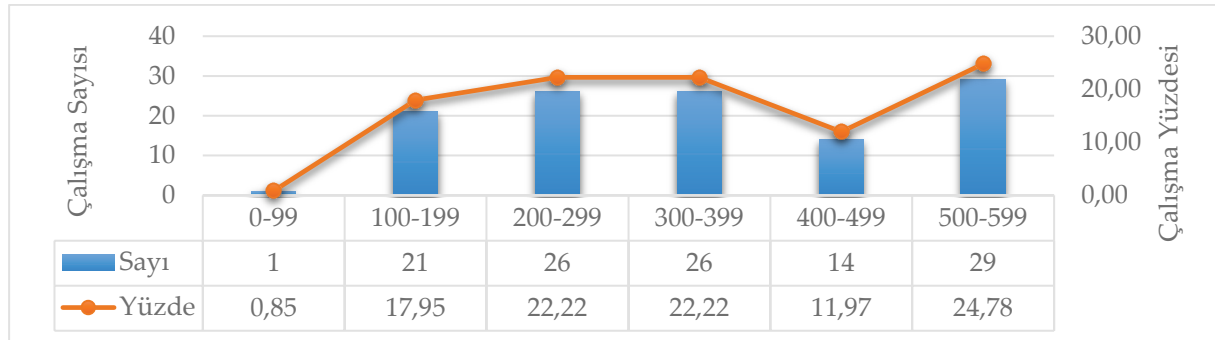


Doküman analizinin kodlama aşamasında güvenilirliğin sağlanması amacıyla 121 makale arasından 10 makale seçilerek iki araştırmacı tarafından incelenmiştir. Kodlayıcılar arası güvenilirliği sağlamak amacıyla Miles ve Huberman (1994) tarafından önerilen güvenilirlik formülü (Görüş birliği sayısı/(Ortak görüş sayısı + Farklı görüş sayısı)) formülü kullanılmış ve uzlaşma katsayısı 0,92 olarak hesaplanmıştır. Makalelerin incelenmesi sürecinde veriler arasında tutarlılık olması açısından araştırmacılar süreç boyunca gerekli yerlerde fikir alışverişinde bulunmuşlardır.

## Bulgular

### Açımlayıcı Faktör Analizinin Kullanımına Yönelik Bulgular

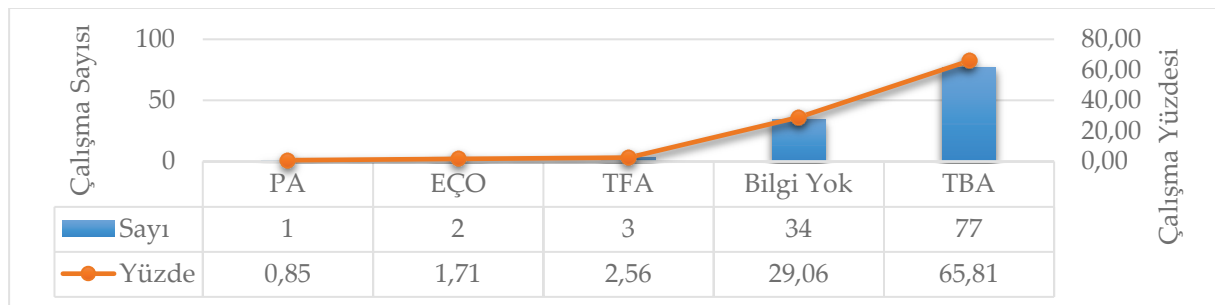
Araştırma kapsamında incelenen ve AFA yapılan çalışmaların örneklem büyüklüğüne göre dağılımı Grafik 2’de verilmiştir.



**Grafik 2.** AFA Yapılan Çalışma Sayılarının Örneklem Büyüklüğü Aralıklarına Göre Dağılımı

Grafik 2 incelendiğinde en çok çalışmanın 500’den büyük örneklem büyüklüğü ile yapıldığı görülmektedir ( $n=29$ , %24,78). Ortalama örneklem büyüklüğü ise tüm çalışmalar için yaklaşık 395 olarak gözlenmiştir. Örneklem büyüklüğü 300 ve üstü olan araştırma sayısı toplam araştırma sayısının yarısından fazladır ( $n=69$ , %58,97). Buna göre çalışmaların çoğunluğunun 300’den büyük örneklemde yapıldığı söylenebilir.

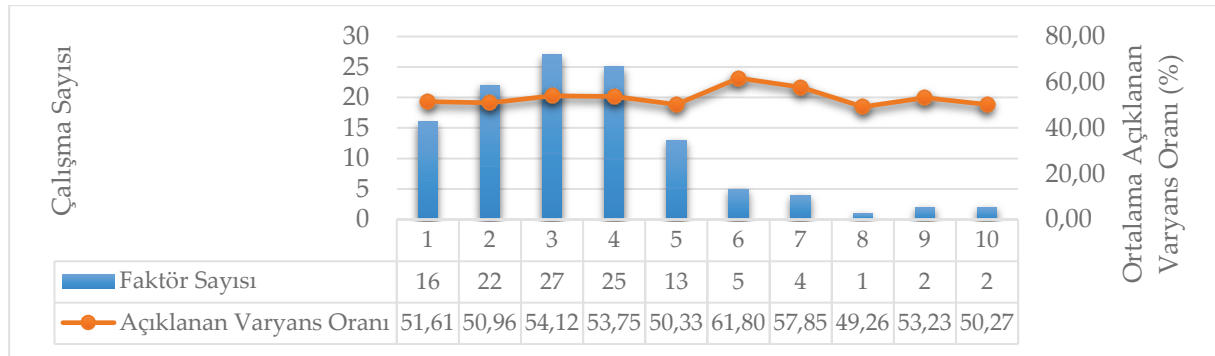
AFA yapılan çalışmaların kullanılan faktör çıkarma yöntemlerine göre dağılımı Grafik 3’te verilmiştir.



**Grafik 3.** Çalışma Sayılarının Kullanılan Faktör Çıkartma Yöntemine Göre Dağılımı

Grafik 3 incelendiğinde araştırmaların büyük çoğunluğunda (%65,81) TBA faktör çıkarma yönteminin kullanıldığı görülmektedir. Dikkat çeken bir diğer bulgu ise araştırmaların çeyreğinden fazlasında (%29,06) faktör çıkarma yöntemine dair bir bilginin yer almamasıdır. Araştırma kapsamındaki çalışmaların üçünün TFA ve iki tanesinin ise EÇÖ faktör çıkarma yöntemini kullandığı gözlenmiştir. Bir çalışmada ise faktör belirleme yöntemi olarak Paralel Analiz (PA) yöntemi kullanılmıştır.

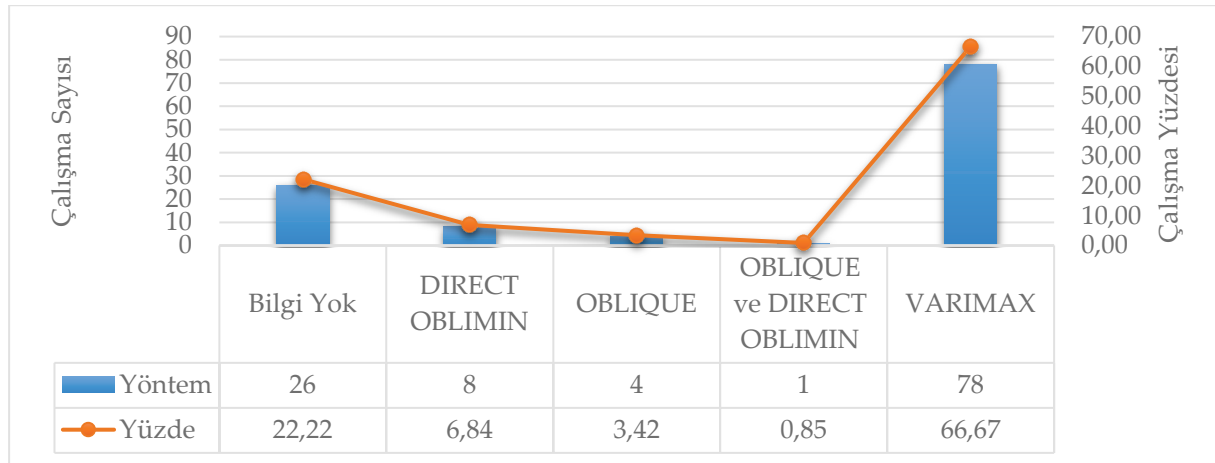
AFA yapılan çalışmaların elde edilen faktör sayısı ve açıklanan varyans oranı yüzdelere göre dağılımı Grafik 4'te verilmiştir.



**Grafik 4.** Çalışma Sayılarının Faktör Sayısı ve Ortalama Açıklanan Varyans Oranlarına Göre Dağılımı

Grafik 4 incelendiğinde araştırmaların %13,68'i tek boyutlu olarak elde edilmiştir ( $n=16$ ). Üç boyutlu çalışma sayısı ( $n=27$ ) en fazla iken, çalışmaların büyük çoğunluğu ( $n=103$ ) bir ile beş faktör arasındadır. Tek faktörlü yapılarda ortalama açıklanan varyans oranı %51,61 iken iki ve daha fazla faktörlü yapılarda %49 ile %62 arasında değişmektedir. Faktör sayısı arttıkça ortalama açıklanan varyans oranı önemli bir artış göstermemiştir.

AFA yapılan çalışmaların kullanılan faktör döndürme yöntemlerine göre dağılımı Grafik 5'te verilmiştir.



**Grafik 5.** Çalışma Sayılarının Kullanılan Faktör Döndürme Yöntemine Göre Dağılımı

Grafik 5 incelendiğinde araştırma kapsamındaki çalışmaların %66,67'sinde VARIMAX döndürme yöntemi kullanıldığı görülmektedir. Çalışmaların %22,22'inde ise döndürme yöntemine yönelik bir açıklamada bulunulmamıştır. Eğik döndürme yöntemlerinden OBLIQUE, dört çalışmada raporlanmış, sekiz çalışmada DIRECT OBLIMIN ve bir çalışmada her iki eğik döndürme yönteminin birlikte kullanıldığı belirtilmiştir.

Araştırmada incelenen ve AFA yapılan çalışmalarda KMO ve Bartlett değerlerinin, yamaç grafiğinin, faktör adlandırmasının ve faktör yüklerinin raporlanma durumuna yönelik olarak oluşturulan veriler Tablo 2'de verilmiştir.

**Tablo 2.** KMO ve Bartlett Değerlerinin, Yamaç Grafiğinin, Faktör Adlandırmasının ve Faktör Yüklerinin Raporlanma Durumları

	KMO Değeri		Bartlett Testi		Yamaç Grafiği		Faktör Adlandırması		Faktör Yükleri	
	n	%	n	%	n	%	n	%	n	%
Evet	110	94,02	109	93,16	58	49,60	93	91,18	112	95,73
Hayır	7	5,98	8	6,84	59	50,40	9	8,82	5	4,27
Toplam	117	100	117	100	117	100	102	100	117	100

Tablo 2 incelendiğinde AFA yapılan 117 çalışmanın %94,02'sinde KMO, %93,16'sında ise Bartlett istatistiklerinin raporlandığı gözlenmiştir. Yamaç grafiği, araştırmaların %49,60'ında raporlanırken, faktör yükleri %95,73 oranında çalışmalarda yer almıştır. Faktörlerin adlandırılması açısından incelendiğinde ise çalışmaların %91,18'inde faktörlerin adlandırıldığı, %8,82'sinde ise faktör adlandırılmasının yapılmadığı gözlenmiştir.

#### *Doğrulayıcı Faktör Analizinin Kullanımına Yönelik Bulgular*

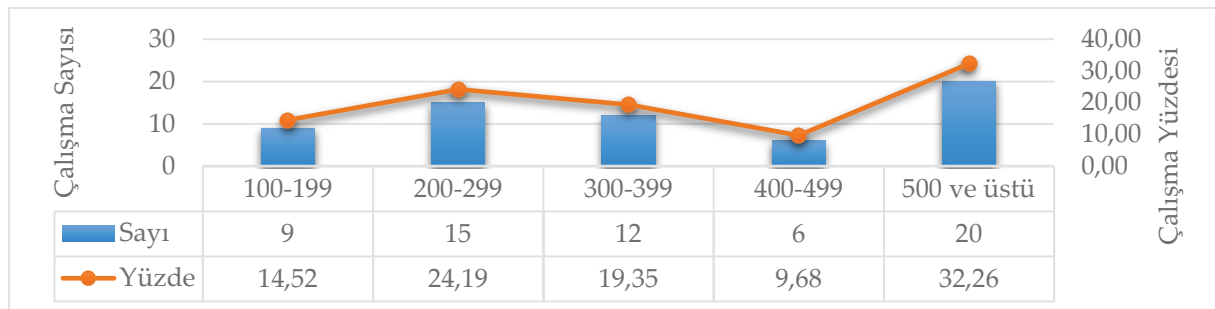
Araştırma kapsamındaki çalışmalardan DFA yapılan 62'sinin analiz varsayımlarının kontrol edilme durumlarının, ölçme modelinin Grafikselleştirilmesi, t değerlerinin, faktör yüklerinin ve  $\chi^2$  anlamlılık değerlerinin raporlanma durumları Tablo 3'te verilmiştir.

**Tablo 3.** Varsayımların Kontrolü, Ölçme Modelinin Grafikselleştirilmesi, t Değerleri, Faktör Yükleri ve  $\chi^2$  Anlamlılık Değerlerinin Raporlanma Durumları

	Varsayımların Kontrolü		$\chi^2$ Anlamlılık Değeri		Ölçme Modelinin Grafikselleştirilmesi		t-değerleri		Faktör Yükleri	
	n	%	n	%	n	%	n	%	n	%
Evet	8	12,90	46	74,19	51	82,26	16	25,81	52	83,87
Hayır	54	87,10	16	25,81	11	17,74	46	74,19	10	16,13
Toplam	62	100	62	100	62	100	62	100	62	100

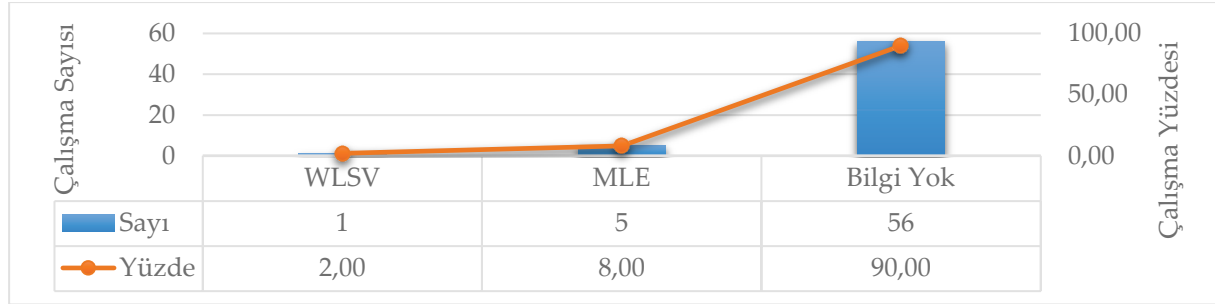
Tablo 3 incelendiğinde DFA yapılan 62 araştırmanın %87,10'unda analizin gerektirdiği varsayımların kontrol edilmediği gözlenmiştir. Eğer kontrol edildiyse de raporlanmamıştır. Araştırmaların %82,26'sında kurulan ölçme modelinin Grafikselleştirilmesi yer verilmiştir. Ölçme modelinin Grafikselleştirilmesi yer vermeyen 11 çalışmanın 3'ünde faktör yükleri raporlanırken 8'inde faktör yükleri raporlanmamıştır. Tüm araştırmaların %83,87'sinde faktör yüklerine yer verilmiştir. Araştırma kapsamındaki çalışmaların %25,81'inde t-değerleri raporlanırken, %74,19'unda  $\chi^2$  anlamlılık değerlerine yer verilmiştir.

DFA yapılan çalışmaların örneklem büyüklüğüne göre dağılımı Grafik 6'da verilmiştir.

**Grafik 6.** DFA Yapılan Çalışma Sayılarının Örneklem Büyüklüğü Aralıklarına Göre Dağılımı

Grafik 6 incelendiğinde DFA yapılan çalışmaların %38,71'inin 300'den az örnekleme yürütüldüğü gözlenmiştir. Örneklem büyüklüğüne göre DFA yürütülen araştırmalar incelendiğinde ortalama örneklem büyüklüğü 415 katılımcı olarak hesaplanmıştır. Çalışmaların %61,29'u 300 ve üzeri örnekleme yürütülmüş iken en çok çalışma ( $n=20$ , %32,26) 500 ve üzeri örnekleme yapılmıştır. Buna göre, küçük örneklemlerde yürütülen çalışma sayılarının nispeten az olduğu söylenebilir.

DFA yapılan çalışmaların kullanılan DFA kestirim yöntemlerine göre dağılımı Grafik 7'de verilmiştir.

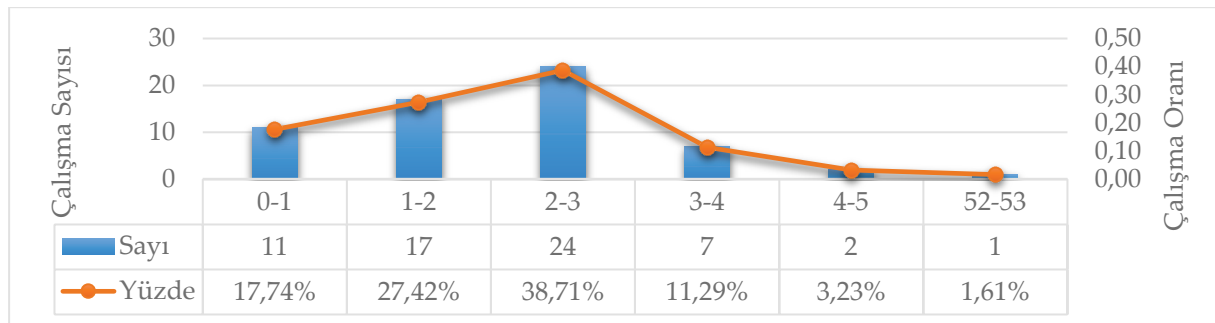


**Grafik 7.** Çalışma Sayılarının Kullanılan Parametre Kestirim Yöntemine Göre Dağılımı

Grafik 7 incelendiğinde araştırmaların %90'ında kullanılan kestirim yöntemi hakkında bilgi verilmemiştir. Araştırmaların %8'inde en çok olabilirlik yöntemi (ML), %2'sinde ise Ağırlıklandırılmış En Küçük Kareler (WLSV) yöntemi kullanılmıştır. Bunların dışında herhangi bir kestirim yöntemi raporlanmamıştır.

DFA yapılan çalışmalarda kullanılan uyum indeksi değerlerinin yeterliliğine ilişkin yapılan tüm yorumlar Tablo 1'de verilen kaynaklardaki kriterlere göre yapılmıştır.

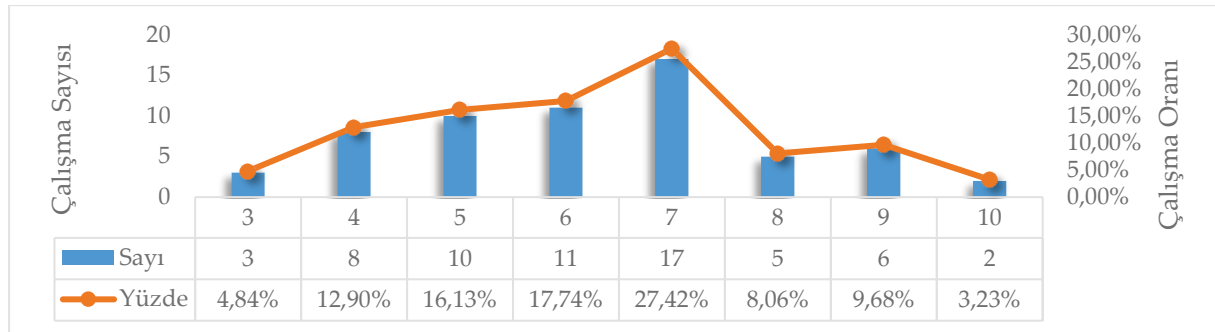
Model uyum istatistiklerinden  $\chi^2/sd$  oranına göre dağılımları Grafik 8'de verilmiştir.



**Grafik 8.** Çalışma Sayılarının  $\chi^2/sd$  Değer Aralıklarına Göre Dağılımı

Grafik 8'e göre DFA yapılan tüm araştırmalarda  $\chi^2/sd$  oranı raporlanmıştır. DFA yapılan bir araştırma hariç diğer tüm çalışmalarda  $\chi^2/sd$  oranının kabul edilebilir aralıkta olduğu söylenebilir. Bu çalışmaların çoğunda ( $n=24$ , %38,71) bu değer iki ile üç aralığındayken tüm çalışmalar bazında genel olarak üçün altındadır.  $\chi^2/sd$  oranının model uyumunun bir ölçüsü olarak kullanılması tartışmalı olduğundan diğer uyum indekslerinin çalışma başına yorumlanma durumları aşağıda raporlanmıştır.

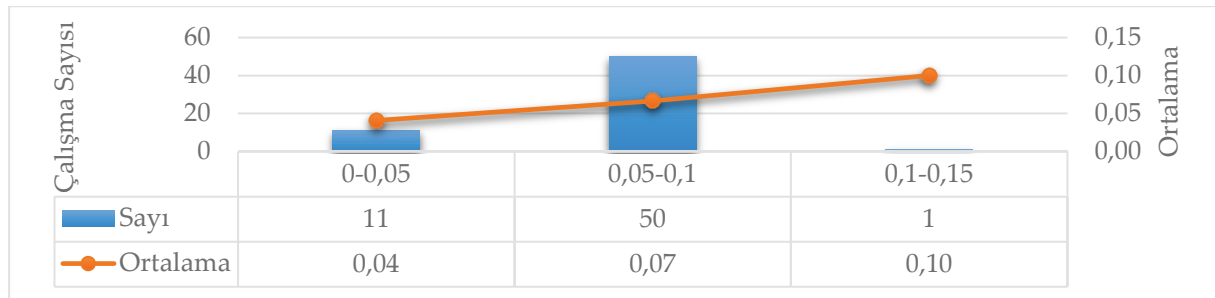
Araştırma kapsamında incelenen çalışmaların uyum istatistiklerinden kaçar tane kullandıklarına yönelik dağılımları Grafik 9'da sunulmuştur.



**Grafik 9.** Çalışma Başına Kullanılan Uyum İndeksi Sayıları

Grafik 9 incelendiğinde, DFA yapılan araştırmalarda, model veri uyumunu değerlendirmek amacıyla 3 ile 10 arasında indeksten yararlanılmış ve her bir çalışmada en az üç uyum indeksi kullanılmıştır. Araştırmaların %66,13'ünde ( $n=41$ ) altı veya daha fazla indekse bakılarak model uyumuna karar verilmiştir.

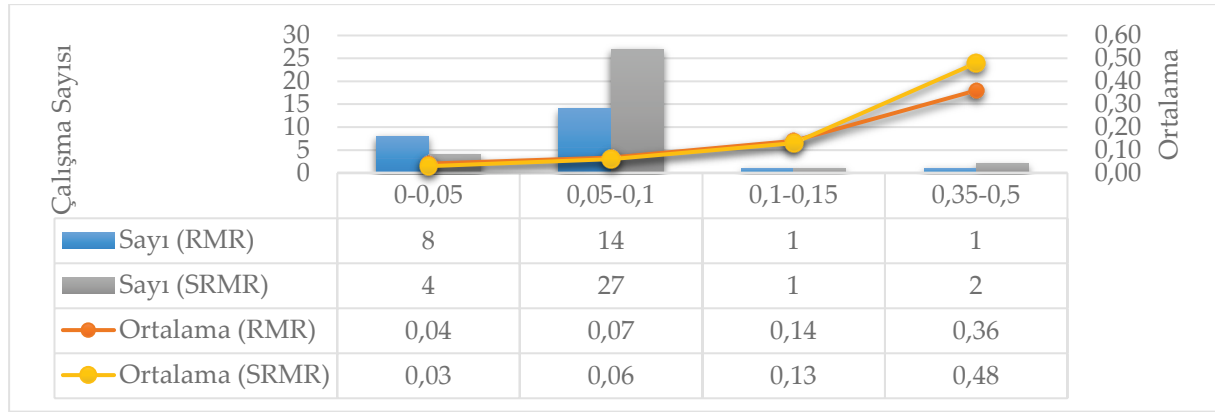
DFA yapılan çalışmaların uyum istatistiklerinden RMSEA değer aralıklarına göre dağılımları Grafik 10'da verilmiştir.



**Grafik 10.** Kritik RMSEA Değer Aralıklarına Düşen Çalışma Sayısı ve Ortalamaları

Grafik 10 incelendiğinde, RMSEA uyum istatistiğinin DFA yapan her çalışmada raporlandığı görülmektedir. Araştırmaların %17,74'ü mükemmel uyum aralığında yer alırken, %80,65'i kabul edilebilir uyum aralığındaki RMSEA değerlerine sahiptir. Araştırmaların küçük bir kısmı ise uyum aralıklarının dışında RMSEA değerine sahip olduğu söylenebilir. Buna göre araştırma kapsamındaki çalışmaların %98,39'unda RMSEA değerlerinin model veri uyumunu desteklediği yorumu yapılabilir.

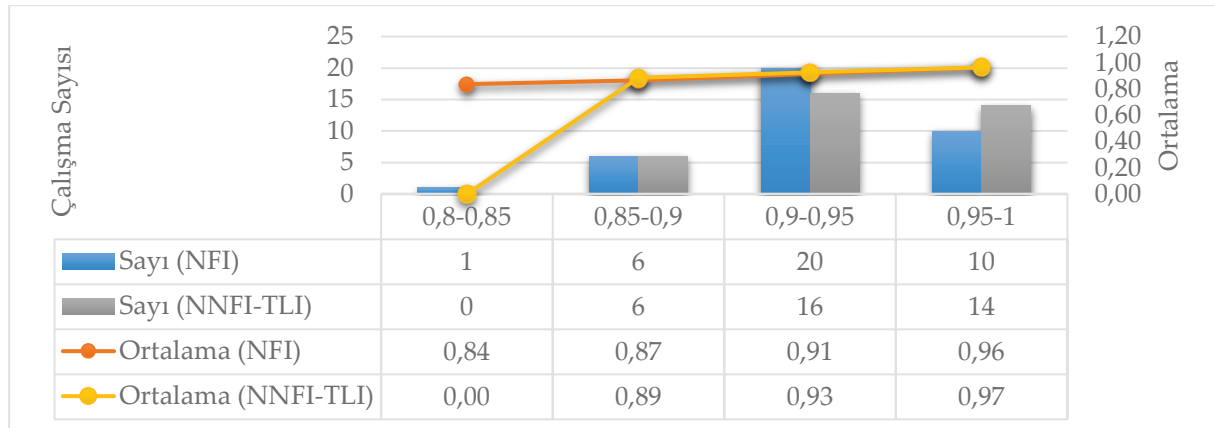
DFA yapılan çalışmaların RMR ve SRMR değer aralıklarına göre dağılımları Grafik 11'de verilmiştir.



**Grafik 11.** Kritik RMR ve SRMR Değer Aralıklarına Düşen Çalışma Sayısı ve Ortalamaları

Grafik 11'e göre, RMR uyum istatistiğinin DFA yapılan 62 çalışmanın 24'ünde (%38,71) raporlandığı görülmektedir. RMR istatistiğini raporlayan araştırmaların %33,33'ü mükemmel uyum aralığında yer alırken, %58,33'ü kabul edilebilir uyum aralığındaki RMR değerlerine sahiptir. RMR istatistiğini raporlayan araştırmaların %8,33'ünde bu değer uyum değerlerinin dışında kalmış olup model veri uyumunun sağlanamadığını göstermiştir. Buna göre araştırma kapsamındaki çalışmaların %91,67'sinde RMR değerlerinin model veri uyumunu desteklediği yorumu yapılabilir. SRMR uyum istatistiğinin ise DFA yapılan 62 çalışmanın 34'ünde (%54,84) raporlandığı görülmektedir. SRMR istatistiğini raporlayan araştırmaların %11,77'si mükemmel uyum aralığında yer alırken, %79,41'i kabul edilebilir uyum aralığındaki SRMR değerlerine sahiptir. SRMR istatistiğini raporlayan araştırmaların %8,82'sinde bu değer uyum değerlerinin dışında kalmış olup model veri uyumunun sağlanamadığını göstermiştir. Buna göre araştırma kapsamındaki çalışmaların %91,18'inde SRMR değerlerinin model veri uyumunu desteklediği yorumu yapılabilir.

DFA yapılan çalışmaların NFI ve NNFI-TLI değer aralıklarına göre dağılımları Grafik 12'de verilmiştir.

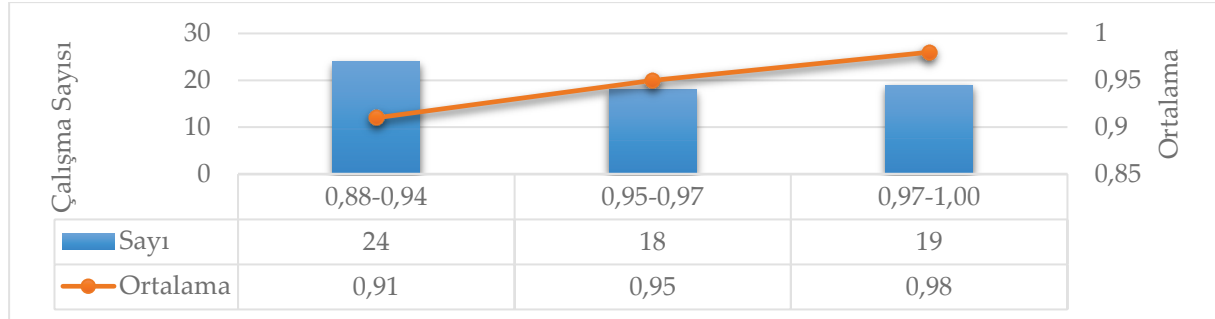


**Grafik 12.** Kritik NFI ve NNFI (TLI) Değer Aralıklarına Düşen Çalışma Sayısı ve Ortalamaları

Grafik 12 incelendiğinde, NFI uyum istatistiğini, DFA yapılan 62 çalışmanın 37'sinde (%59,68) raporlandığı görülmektedir. NFI istatistiğini raporlayan araştırmaların %27,03'ü mükemmel uyum aralığında yer alırken, %54,05'i kabul edilebilir uyum aralığındaki NFI değerlerine sahiptir. NFI istatistiğini raporlayan araştırmaların %18,92'sinde bu değer uyum değerlerinin dışında kalmış olup model veri uyumunun sağlanamadığını göstermiştir. Buna göre araştırma kapsamındaki çalışmaların %81,08'inde NFI değerlerinin model veri uyumunu desteklediği yorumu yapılabilir. NNFI uyum istatistiğinin ise DFA yapılan 62 çalışmanın 36'sında (%58,07) raporlandığı görülmektedir. NNFI istatistiğini raporlayan araştırmaların %38,89'u mükemmel uyum aralığında yer alırken, %44,44'ü kabul edilebilir uyum aralığındaki NNFI değerlerine sahiptir. NNFI istatistiğini raporlayan

araştırmaların %16,67'sinde bu değer uyum değerlerinin dışında kalmış olup model veri uyumunun sağlanamadığını göstermiştir. Buna göre araştırma kapsamındaki çalışmaların %83,33'ünde NNFI değerlerinin model veri uyumunu desteklediği yorumu yapılabilir.

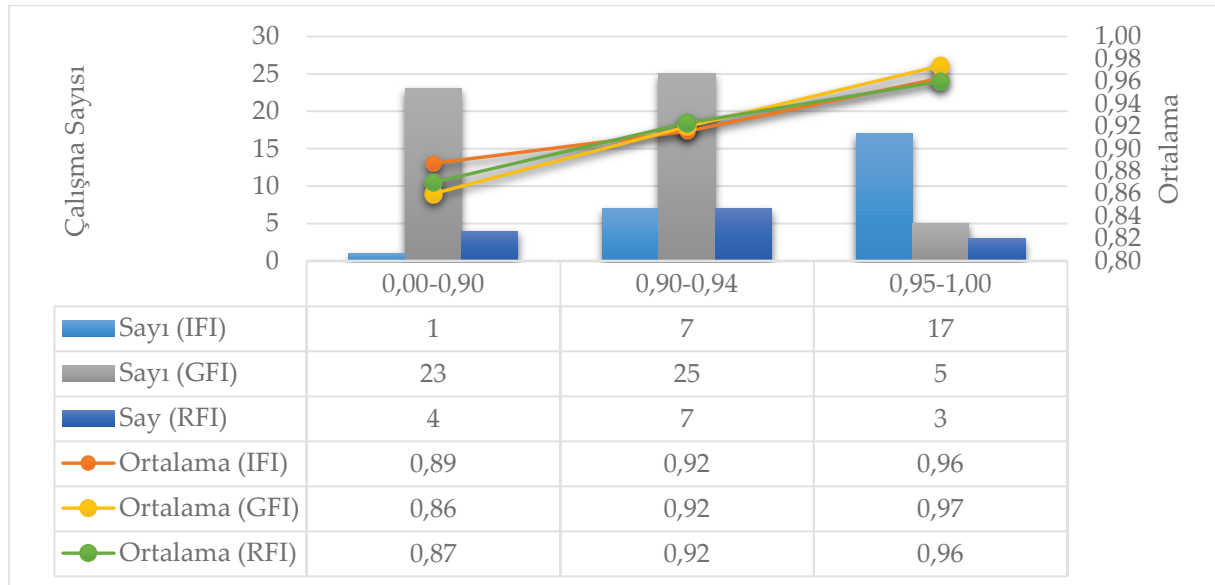
DFA yapılan çalışmaların CFI değer aralıklarına göre dağılımları Grafik 13'te verilmiştir.



**Grafik 13.** Kritik CFI Değer Aralıklarına Düşen Çalışma Sayısı ve Ortalamaları

Grafik 13'e göre, CFI uyum istatistiğinin DFA yapan 62 çalışmanın 61'inde (%98,39) raporlandığı görülmektedir. CFI istatistiğini raporlayan araştırmaların %31,15'i mükemmel uyum aralığında yer alırken, %29,51'i kabul edilebilir uyum aralığındaki CFI değerlerine sahiptir. CFI istatistiğini raporlayan araştırmaların %39,34'ünde bu değer uyum değerlerinin dışında kalmış olup model veri uyumunun sağlanamadığını göstermiştir. Buna göre araştırma kapsamındaki çalışmaların %60,66'sında CFI değerlerinin model veri uyumunu desteklediği yorumu yapılabilir.

DFA yapılan çalışmaların IFI, GFI ve RFI değer aralıklarına göre dağılımları Grafik 14'te verilmiştir.

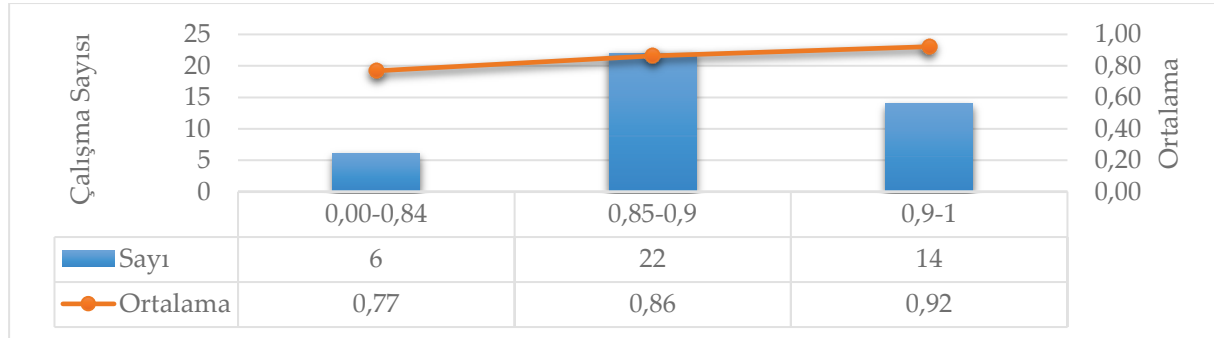


**Grafik 14.** Kritik IFI, GFI ve RFI Değer Aralıklarına Düşen Çalışma Sayısı ve Ortalamaları

Grafik 14 incelendiğinde, IFI uyum istatistiğinin DFA yapan 62 çalışmanın 25'inde (%40,32) raporlandığı görülmektedir. IFI istatistiğini raporlayan araştırmaların %68'i mükemmel uyum aralığında yer alırken, %28'i kabul edilebilir uyum aralığındaki IFI değerlerine sahiptir. IFI istatistiğini raporlayan araştırmaların %4'ünde bu değer uyum değerlerinin dışında kalmış olup model veri uyumunun sağlanamadığını göstermiştir. Buna göre araştırma kapsamındaki çalışmaların %96'sında IFI değerlerinin model veri uyumunu desteklediği yorumu yapılabilir. GFI uyum istatistiğinin DFA yapan 62 çalışmanın 53'ünde (%85,48) raporlandığı görülmektedir. GFI istatistiğini raporlayan araştırmaların %94,3'ü mükemmel uyum aralığında yer alırken, %47,17'si kabul edilebilir uyum

aralığındaki GFI değerlerine sahiptir. GFI istatistiğini raporlayan araştırmaların %43,40'ında bu değer uyum değerlerinin dışında kalmış olup model veri uyumunun sağlanmadığını göstermiştir. Buna göre araştırma kapsamındaki çalışmaların %56,60'ında GFI değerlerinin model veri uyumunu desteklediği yorumu yapılabilir. RFI uyum istatistiğinin ise DFA yapan 62 çalışmanın 14'ünde (%22,58) raporlandığı görülmektedir. RFI istatistiğini raporlayan araştırmaların %21,43'ü mükemmel uyum aralığında yer alırken, %50'si kabul edilebilir uyum aralığındaki RFI değerlerine sahiptir. RFI istatistiğini raporlayan araştırmaların %28,57'sinde bu değer uyum değerlerinin dışında kalmış olup model veri uyumunun sağlanmadığını göstermiştir. Buna göre araştırma kapsamındaki çalışmaların %71,43'ünde RFI değerlerinin model veri uyumunu desteklediği yorumu yapılabilir.

DFA yapılan çalışmaların AGFI değer aralıklarına göre dağılımları Grafik 15'te verilmiştir.



**Grafik 15.** Kritik AGFI Değer Aralıklarına Düşen Çalışma Sayısı ve Ortalamaları

Grafik 15 incelendiğinde, AGFI uyum istatistiğinin DFA yapılan 62 çalışmanın 42'sinde (%67,74) raporlandığı görülmektedir. AGFI istatistiğini raporlayan araştırmaların %33,33'ü mükemmel uyum aralığında yer alırken, %52,38'i kabul edilebilir uyum aralığındaki AGFI değerlerine sahiptir. AGFI istatistiğini raporlayan araştırmaların %14,29'unda bu değer uyum değerlerinin dışında kalmış olup model veri uyumunun sağlanmadığını göstermiştir. Buna göre araştırma kapsamındaki çalışmaların %61,91'inde AGFI değerlerinin model veri uyumunu desteklediği yorumu yapılabilir.

Araştırma kapsamında incelenen çalışmaların uyum istatistiklerinden PNFI'ya herhangi bir çalışmada yer verilmemiş, dört çalışmada (%6,45) ise PGFI raporlanmıştır. Raporlanan PGFI değerleri 0,65 ile 0,80 arasında değişmektedir.

### Tartışma, Sonuç ve Öneriler

Bu çalışmada, 2006-2016 yılları arasında Türkiye'de eğitim ve sosyal bilimler alanlarında yapılan ölçek geliştirme çalışmaları açılımcı ve doğrulayıcı faktör analizlerinin kullanımı açısından incelenmiştir. Araştırmada incelenen ölçek geliştirme çalışmalarının çoğunluğu 2013 yılından sonra yapılmıştır. Çalışmaların büyük çoğunluğunda AFA yapılırken, yaklaşık yarısında DFA yapılmış olması bazı araştırmalarda elde edilen ölçek faktör yapılarının hipotetik olarak doğrulanmaya ihtiyaç duyulmadığını göstermektedir. Faktör analizleri bir ölçeğin yapı geçerliği için kanıt oluşturur (Cronbach ve Meehl, 1955). Bu nedenle, faktör yapılarının DFA ile test edilmesi yapı geçerliğine dönük daha güçlü kanıtlar sağlayacaktır. Her iki analizin de yapıldığı araştırmaların yarısından çoğunda analizler aynı örneklem üzerinde yapılmıştır. Worthington ve Whitaker (2006), aynı örneklem üzerinde analiz gerçekleştirilmenin sorun teşkil etmeyeceğini, ancak Fabrigar ve diğerleri (1999) eğer örneklem yeterince büyükse rastgele ikiye ayırıp analizleri bu iki örneklemde gerçekleştirmeyi önermektedir.

AFA yapılan araştırmalarda, ortalama örneklem büyüklüğü 300'ün üzerindeyken araştırmaların yaklaşık %60'ı 300'den fazla kişiyle yürütülmüştür. Çalışmalarda yer alan madde sayıları ve faktör yükü büyüklüklerine göre ulaşılması beklenen örneklem büyüklüğü değişebilmektedir (Guadagnoli ve Velicer, 1988; Streiner, 1994; Yong ve Pearce, 2013). Ancak, AFA için genel olarak yeterli örneklem büyüklüğünün en az 300 kişi olması önerilirken (Comrey ve Lee, 1992; Tabachnick ve Fidell,



2012), Comrey ve Lee (1992), 300'den az örneklem büyüklerinin orta düzeyde veya yetersiz olduğunu belirtmişlerdir. Buna göre, incelenen araştırmaların yarısından fazlasında yeterli sayıda örneklem büyüklüğüne ulaşıldığı söylenebilir. Ancak çalışmaların yarısından azında ise orta büyüklükte veya kabul edilebilir sayıdan az örneklem büyüklüğüne ulaşıldığını göstermektedir. DFA'nın yapıldığı araştırmalarda da ortalama örneklem büyüklüğü 300'ün üzerinde iken araştırmaların yaklaşık %60'ı 300'den fazla kişiyle yürütülmüştür. DFA'da yeterli kabul edilebilecek örneklem büyüklüğü farklı koşullara göre değişebilmekle beraber (MacCallum vd., 2001; Wolf vd., 2013), genel olarak 300 ve üzeri örneklem olması beklenir (Worthington ve Whitaker, 2006). Buna göre, DFA yapılan araştırmalarında yarısından fazlasında yeterli örneklem büyüklüğüne ulaşılmasına rağmen azımsanmayacak sayıda çalışmada ise örneklemin küçük olduğu görülmüştür. Ölçek geliştirme çalışmalarının yeterli sayıda katılımcı ile yapılmaması elde edilecek faktör yapılarının ve bu ölçekten elde edilecek puanlara dayalı çıkarımların hatalı olmasına neden olacaktır. Bu nedenle, yapılacak çalışmalarda yeterli örneklem büyüklüğü saptanırken alan yazında yer alan ölçütlerden en az ikisinin göz önünde bulundurulması önerilmektedir (Çokluk, Şekercioğlu ve Büyüköztürk, 2012).

Araştırmaların büyük çoğunluğunda TBA faktör çıkarma yöntemi kullanılmasına rağmen bu yöntemin neden tercih edildiğine yönelik bir açıklamanın yer almaması, TFA, EÇO ve PA gibi yöntemlerin de verinin yapısına göre tercih edilebileceğini düşündürmüştür. Faktör analizinde kullanılan yöntemin varsayımlarının neler olduğu kullanılacak yöntem açısından belirleyici bir faktördür. Örneğin, TFA yöntemi çok değişkenli normallik sağlanmaması durumunda kullanılır (Costello ve Osborne, 2005). Benzer şekilde, araştırmaların büyük çoğunluğunda Varimax dik döndürme yöntemi kullanılmasına rağmen bu yöntemin neden tercih edildiğine yönelik bir açıklamanın yer almaması, diğer dik ve eğik döndürme yöntemlerin de verinin yapısına göre tercih edilebileceğini düşündürmüştür. Eğik döndürme, genellikle faktörler birbirleriyle ilişkili olduğunda; dik döndürme ise bu ilişki daha az olduğunda kullanılır (Costello ve Osborne, 2005; Rummel, 1988). Eğik ve dik döndürme yöntemleri de kendi içlerinde farklı koşullarda tercih edilmektedir (Gorsuch, 1983; Yong ve Pearce, 2013). Bu çalışmada incelenen araştırmaların çeyreğinden fazlasında, kullanılan faktör çıkartma ve döndürme yöntemine değinilmemesi ise ölçek yapılarının nasıl elde edildiğine ilişkin soru işaretlerine neden olmaktadır. Benzer şekilde, doğrulayıcı faktör analizi yapılan çalışmaların %87,10'unda kullanılan analiz yöntemleri ve dolayısıyla bu yöntemlere yönelik varsayımlar incelenmemiştir. Hâlbuki uygun analiz yönteminin seçilmesi doğrulayıcı faktör analizinin önemli adımlarından biridir (Brown, 2015; Cabrera ve Nguyen, 2010). Ayrıca, DFA yapmadan önce tek ve çok değişkenli normallik, çoklu bağlantılılık ve kayıp ile uç değerlerin incelenmesi önerilmektedir (Ullman, 2012). Sonuç olarak, açıklayıcı ve doğrulayıcı faktör analizlerinde kullanılacak faktör çıkartma, faktör döndürme ve parametre kestirim yöntemlerinin neden seçildiğinin açıkça ifade edilmesi, faktör yapılarının nasıl edildiği ve test edildiğine yönelik soru işaretlerini ortadan kaldıracaktır.

Her iki analiz yönteminde de incelenen çalışmaların %87'den fazlasında maddelerin faktör yükleri raporlanmıştır. Benzer şekilde, açıklayıcı analizde verinin analizlere uygunluğu (KMO ve Bartlet testleri) incelenmiş ve faktörler adlandırılmıştır. Ancak, çalışmaların sadece yarısında yamaç grafiğine yer verilmiştir. Faktör sayısına karar verirken yamaç grafiği göz önünde bulundurulmuş ölçütlerden biridir (Çokluk vd., 2012; Hair vd., 1995; Yong ve Pearce, 2013). Bu nedenle, faktör yapısı belirlenirken açıklayıcı analizin çıktılarında yer alan yamaç grafiğine yer verilmesi elde edilen faktör yapısı için kanıt oluşturacaktır. İncelenen araştırmaların %87'inden fazlası birden fazla faktörden oluşurken, elde edilen toplam açıklanan varyans oranları %50-60 arasındadır. Sosyal ve eğitim bilimleri alanları için genellikle toplam açıklanan varyans oranı %50-60 civarında olmaktadır (Yong ve Pearce, 2013). Ayrıca, tek faktörlü yapılar için toplam açıklanan varyans oranının en az %30 ve üzerinde olması önerilmektedir (Büyüköztürk, 2013). Buna göre elde edilen oranların genel olarak yeterli olduğu söylenebilir. Ancak faktör sayısı arttıkça toplam açıklanan varyans oranının artmaması faktör başına açıklanan varyans oranını düşüreceğinden, faktörlerin incelenen psikolojik yapıyı temsil edebilirliğine gölge düşürmektedir. Böyle durumlarda elde edilen faktör yapılarının mümkün olduğunca ölçülen yapı ile birlikte değerlendirilmesi ve tek başına göz önünde bulundurulmaması yararlı olacaktır.

Doğrulamalı faktör analizi yapılan araştırmalarda ölçme modelinin  $\chi^2$  anlamlılık değerlerine ve modelin grafiksel gösterimine %74 ve üzeri bir oranda yer verilmiş olması, modellerin incelenmesinde ve anlaşılmasında kolaylık sağlamaktadır. Ancak, incelenen araştırmaların %74'ünden fazlasında faktör yüklerinin  $t$  anlamlılık değerlerine yer verilmemiştir. DFA çalışmalarında faktör yüklerinin manidarlık değerleri de dahil model kestirimlerine yer verilmesi önerilmektedir (Brown, 2015; Cabrera ve Nguyen, 2010). Buna göre, DFA analizlerinde ölçme modelinin Grafiksel gösteriminde maddelerin anlamlılık değerlerine yer verilmesi elde edilen faktör yüklerinin değerlendirilmesi açısından önemli görülmektedir.

Araştırma kapsamında incelenen ve DFA yapılan tüm araştırmalarda  $\chi^2/sd$  değerlerinin raporlandığı ve hemen hepsinde beşten küçük olduğu görülmüştür. Bazı araştırmacılar (Anderson ve Gerbing, 1984)  $\chi^2/sd$  değerinin beşten küçük olması model uyumu için kabul edilebilir bir kriter olarak görürken, diğer araştırmacılar (Kline, 2016; Wheaton, 1987) bu değeri model uyumu için bir ölçüt olarak kullanılmasının mantıksal ve istatistiksel bir temelini olmadığını ifade etmişlerdir. Bu nedenle, incelenen araştırmalarda elde edilmiş olan değerler kabul edilebilir düzeyde olmasına karşın, raporlanması tartışma konusu olan bir indeksin tüm araştırmalarda yer alması üzerinde düşünülmesi gereken bir konudur. DFA çıktılarında model uyumu için kullanabilecek birçok indeks olmasına rağmen yeterli düzeyde mantıksal ve istatistiksel bir temeli olmayan bir değer model uyumu için bir ölçüt olarak görülmesi istenen bir durum değildir.

İncelenen tüm çalışmalarda model uyumunun değerlendirilmesinde RMSEA ve CFI uyum indeksi değerleri raporlanmıştır. Ayrıca, çalışmaların tümünde en az üç uyum indeksi ve büyük çoğunluğunda ise beşten fazla indeks kullanılmıştır. En çok kullanılan diğer indeksler sırasıyla GFI, AGFI, NFI, NNFI-TLI, SRMR, IFI, RMR, RFI ve PGFI'dir. DFA yapılan araştırmalarda, kaç uyum indeksinin kullanılması konusunda, bazı araştırmacılar belirli uyum indekslerinin ( $\chi^2$ , CFI-IFI, TLI-NNFI, SRMR-RMR, RMSEA) tercih edilmesini (Brown, 2015; Hu ve Bentler, 1999; Kline, 2016) önerirken, bazıları da mümkün olduğunca fazla indeks raporlanmasını (Cabrera ve Nguyen, 2010; Crowley ve Fan, 1997) önermiştir. Buna göre, DFA yapılan neredeyse tüm çalışmalarda belirtilen indekslerin yanında, model uyumunun değerlendirilmesinde yeterli sayıda indeksten yararlandığı söylenebilir. Ancak, GFI ve AGFI uyum indeksleri çalışmalarda diğer indekslere göre daha fazla sıklıkla kullanılmalarına rağmen, yapılan benzetim çalışmalarındaki (örn. Hu ve Bentler, 1998; Marsh vd., 1988) zayıf performanslarından dolayı önerilmemektedir. Bu nedenle, uyum indeksleri genel olarak modelin ve verinin yapısı ile kestirim yönteminden etkilendiğinden kesin ölçütler olmamakla birlikte (Brown, 2015), DFA çalışmalarında model uyum değerlendirilmesinde GFI ve AGFI yerine diğer indekslerden mümkün olduğunca fazla kullanılmalıdır. Özellikle CFI-IFI, TLI-NNFI, SRMR-RMR ikililerinin ve RMSEA değerlerinin tercih edilmesi yararlı olacaktır. Bunun yanında, hangi uyum indekslerinin kullanılmasına yönelik geliştirilen önerilerden (Brown, 2015; Hu ve Bentler, 1999; Kline, 2016) en az iki veya daha fazlasının göz önünde bulundurulması, model uyumunun doğru ve sağlıklı değerlendirilmesinde etkili bir yöntem olacaktır.

Model uyumu indekslerinin değerleri incelendiğinde, RMSEA, RMR, SRMR, NFI, NNFI-TLI, IFI ve PGFI'nun raporlandıkları araştırmalarda genel olarak kabul edilebilir düzeyde oldukları görülmüştür. CFI, GFI ve AGFI ve RFI değerlerinin raporlandığı araştırmaların yarısından çoğunda, bu değerler model uyumu için yine kabul edilebilir düzeydedir. İlgili alan yazından (Hooper vd., 2008; Hu ve Bentler, 1999; Kline, 2016; MacCallum vd., 1996; Marsh ve Hau, 1996; Steiger, 2007; Tabachnick ve Fidell, 2012; Wheaton, 1987) elde edilen eşik değerlere göre genel olarak DFA yapılan çalışmalarda uyum indekslerinin kabul edilebilir değerler aldığı söylenebilir.

İncelenen araştırmalardan elde edilen bulgular ve ilgili alan yazı doğrultusunda açıklayıcı ve doğrulamalı faktör analizlerinin kullanımlarıyla ilgili teoriye ve uygulamaya yönelik bazı öneriler geliştirilmiştir. Buna göre;

- Doğrulamalı ve açıklayıcı faktör analizlerinin varsayımları ve veri yapısının analizlere uygunluğu incelenmelidir. Bu amaçla, uç ve kayıp değerler kontrol edilmeli ve uygun şekilde

veri düzenlenmelidir. Tek ve çok değişkenli normallik ve çoklu doğrusal bağlantı ile tekillik gibi temel varsayımların sağlanıp sağlanmadığı kontrol edilmelidir.

- Verinin yapısına uygun yöntemler seçilmelidir. Örneklemin büyüklüğü, maddelerin sürekli veya kategorik (ikili, sıralı, vb.) olması ve temel varsayımların sağlanma durumlarına göre buna karar verilmelidir.
- Yeterli örneklem büyüklüğüne ulaşılmalıdır. Madde sayısına göre değişmekle birlikte her iki analiz yöntemi için genel olarak 300 ve üzeri bir örneklem yeterli olacaktır. 500 ve üzeri bir örneklem iyi düzeyde bir tercih olup, 200 ve altı bir örneklem zayıf veya yetersiz olacaktır.
- Analiz sonuçları mutlaka ilgili alan yazın ışığında yorumlanmalıdır.
- Açımlayıcı analizde:
  - KMO ve Bartlett testi sonuçları, faktör özdeğerleri ve toplam açıklanan varyans oranları, faktör yükleri, yamaç grafiği ve kullanılan faktör çıkartma ile döndürme yöntemleri yorumları ve gerekçeleri ile birlikte raporlanmalıdır.
  - Faktör sayısına karar verirken birden fazla yöntemden yararlanılmalıdır. Özdeğeri birden büyük faktörlerin seçilmesi, yamaç grafiği testi, paralel analiz, faktördeki kabul edilebilir madde sayısının üç ve üzeri sayıda olması gibi ölçütlerden yararlanılabilir.
- Doğrulayıcı analizde:
  - Model doğru tanımlanmalıdır.
  - Veri yapısı incelenerek analiz varsayımları kontrol edilmelidir.
  - Seçilen parametre kestirim yöntemleri kullanılma gerekçeleri ile birlikte verilmelidir.
  - Faktör yükleri, yol diyagramı,  $\chi^2$  değeri anlamlılık düzeyi ve serbestlik derecesi ile birlikte, RMSEA değeri, RMR-SRMR değerlerinden en az biri ve CFI ile NNFI değerleri raporlanmalıdır. Model değerlendirmede, bu minimum ölçütlerin yanında, verinin ve modelin yapısına göre diğer uyum indekslerinden de faydalanılabilir.
  - Ölçülmek istenen psikolojik yapıyı bozacak düzeyde çok modifikasyon yapılmamalı, eğer yapılacaksa nedenleriyle birlikte açıklanmalıdır.
- Aynı örneklem AFA ve DFA yapmak sorun oluşturmamakla birlikte (Worthington ve Whitaker, 2006) örneklemin yeterince büyük olduğu durumlarda ikiye bölünüp yarısında AFA yarısında DFA yapılması önerilmektedir (Fabrigar vd., 1999).

İncelenen araştırmalarda doğrulayıcı faktör analizinin yapıldığı çalışmalarda elde edilen ölçek yapılarının farklı gruplarda (cinsiyet, sosyoekonomik düzey, yaş grubu, vb.) da sağlanıp sağlanmadığına dönük bir kanıt sunulmamıştır. Bu nedenle, yapısal eşitlik modelleme ile farklı alt gruplarda elde edilen faktör yapıları hipotetik olarak test edilebilir. Böylece, ölçeğin yapı geçerliğine yönelik daha güçlü kanıtlar elde edilmiş olacaktır.

AFA ve DFA yapmadan önce analizlerin gerektirdiği varsayımların incelenmesi, analizlerin nasıl gerçekleştirileceğine ve çıktılarının nasıl yorumlanacağına yönelik gerekli bilgi ve donanımın sağlanması, daha geçerli ve güvenilir ölçme araçlarının geliştirilmesi açısından hayati öneme sahiptir. Araştırmacılar analizlerin gerçekleştirilmesi ve yorumlanması konusunda güvenilir kaynaklardan yararlanmalı, gerekirse ölçme değerlendirme ya da psikometri alanında uzman kişilerden destek almalıdırlar.

Gelecekte yapılacak çalışmalarda, eğitim, psikoloji ve davranış bilimlerinde dünya genelinde yapılan ve belirli indekslerde yayınlanan ölçek geliştirme çalışmaları incelenerek Türkiye'deki sonuçlarla karşılaştırılabilir.

## Kaynakça

- Acar-Güvendir, M. ve Özer-Özkan, Y. (2015). Türkiye'deki eğitim alanında yayımlanan bilimsel dergilerde ölçek geliştirme ve uyarlama konulu makalelerin incelenmesi. *Elektronik Sosyal Bilimler Dergisi*, 14(52), 23-33. doi: 10.17755/esosder.54872
- American Psychological Association, American Educational Research Association ve National Council on Measurement in Education. (1974). *Standards for educational and psychological tests and manuals*. Washington, DC: American Psychological Association.
- Anderson, J. C. ve Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, 49(2), 155-173.
- Bailey, K. D. (1982). *Methods of social research* (2. bs.). New York: The Free Press.
- Bayık, M. ve Gürbüz, S. (2016). Ölçek uyarlamada metodoloji sorunu: Yönetim ve örgüt alanında uyarlanan ölçekler üzerinden bir araştırma. *The Journal of Human and Work*, 3(1), 1-20.
- Bentler, P.M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software, Inc.
- Bentler, P. M. ve Bonnet, D. C. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588-606.
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal* 9(2), 27-40. doi: 10.3316/QRJ0902027.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2. bs.). New York: The Guilford Press.
- Brown, T. A. ve Moore, M. T. (2012). Confirmatory factor analysis. R. H. Hoyle (Ed.), *Handbook of structural equation modeling* içinde (s. 361-379). New York: Guilford Press.
- Browne, M. W. ve Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage Focus Editions*, 154, 136-136.
- Büyüköztürk, Ş. (2013). *Sosyal bilimler için veri analizi el kitabı: İstatistik, araştırma deseni, SPSS uygulamaları ve yorum* (18. bs.). Ankara: Pegem Akademi.
- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS and SIMPLIS: Basic concepts, applications and programming*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Cabrera-Nguyen, P. (2010). Author guidelines for reporting scale development and validation results in the Journal of the Society for Social Work and Research. *Journal of the Society for Social Work and Research*, 1(2), 99-103. doi: 10.5243/jsswr.2010.8
- Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences*. New York, NY: Plenum Press.
- Clark, L. A. ve Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309-319.
- Çokluk, O., Şekercioğlu, G. ve Büyüköztürk, Ş. (2012). *Sosyal bilimler için çok değişkenli SPSS ve LISREL uygulamaları*. Ankara: Pegem Yayıncılık.
- Comrey, A. L. ve Lee, H. B. (1992). *A first course in factor analysis* (2. bs.). Hillsdale, NJ: Lawrence Erlbaum.
- Costello, A. B. ve Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical assessment, research ve evaluation*, 10(7), 1-9.
- Crocker, L. ve Algina, J. (1986). *Introduction to classical and modern test theory*. Orlando: Holt, Rinehart and Winston.
- Cronbach, L. J. (1984). *Essentials of psychological testing*. New York: Harper.

- Cronbach, L. J. ve Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
- Crowley, S. L. ve Fan, X. (1997). Structural equation modeling: Basic Concepts and applications in personality assessment research. *Journal of Personality Assessment*, 68(3), 508-31.
- Cureton, E. E. (1951). Validity. E. F. Lindquist (Ed.), *Educational measurement* (1. bs.) içinde (s. 621-694). Washington, DC: American Council on Education.
- Çüm, S. ve Koç, N. (2013). Türkiye’de psikoloji ve eğitim bilimleri dergilerinde yayımlanan ölçek geliştirme ve uyarlama çalışmalarının incelenmesi. *Eğitim Bilimleri ve Uygulama*, 12(24), 115-135.
- Davidshofer, K. R. ve Murphy, C. O. (2005). *Psychological testing: principles and applications*. NJ: Pearson/Prentice-Upper Saddle River.
- Delice, A. ve Ergene, Ö. (2015). Ölçek geliştirme ve uyarlama çalışmalarının incelenmesi: Matematik eğitimi makaleleri örneği. *Karaelmas Eğitim Bilimleri Dergisi*, 3(1), 60-75.
- Erkuş, A. (2012). *Psikolojide ölçme ve ölçek geliştirme*. Ankara: Pegem Akademi Yayınları.
- Erkuş, A. (2016). Ölçek geliştirme ve uyarlama çalışmalarındaki sorunlar ile yazım ve değerlendirilmesi. *Pegem Atıf İndeksi*, 0, 1211-1224. doi: 10.14527/9786053183563.075
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. ve Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299. doi: 10.1037/1082-989X.4.3.272
- Floyd, F. J. ve Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7(3), 286-299.
- Fraenkel, J.R. ve Wallen, N.E. (2011). *How to design and evaluate research in education* (6. bs.). New York: McGraw-Hill, Inc.
- Gorsuch, R.L. (1983). *Factor analysis* (2. bs.). Hillside, NJ: Lawrence Erlbaum Associates.
- Guadagnoli, E. ve Velicer, W. F. (1988). Relation to sample size to the stability of component patterns. *Psychological bulletin*, 103(2), 265-275.
- Gül, Ş. ve Sözbilir, M. (2015). Fen ve matematik eğitimi alanında gerçekleştirilen ölçek geliştirme araştırmalarına yönelik tematik içerik analizi. *Eğitim ve Bilim*, 40(178), 85-102.
- Hair Jr, J. F., Anderson, R. E., Tatham, R. L., William, C. ve Black (1995). *Multivariate data analysis with readings*. New Jersey: Prentice Hall.
- Henson, R. K. ve Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, 66(3), 393-416.
- Hooper, D., Coughlan, J. ve Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), 53-60.
- Hu, L. T. ve Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424-453.
- Hu, L.T. ve Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Jöreskog, K. G. (1967). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika* 34(2), 183-202.
- Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, 36(2), 109-133.
- Kline, R. B. (2016). *Principle and practice of structural equation modelling* (4. bs.). New York, NY: The Guilford Press
- MacCallum, R. C., Browne, M. W. ve Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), 130-49.

- MacCallum, R. C., Widaman, K. F., Preacher, K. J. ve Hong, S. (2001). Sample size in factor analysis: The role of model error. *Multivariate Behavioral Research*, 36(4), 611-637. doi: 10.1207/S15327906MBR3604\_06
- Marsh, H. W., Balla, J. R. ve McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, 103, 391-410.
- Marsh, H. W. ve Hau, K. T. (1996). Assessing goodness of fit: Is parsimony always desirable? *The Journal of Experimental Education*, 64(4), 364-390.
- Miles, M. B. ve Huberman, A. M. (1994). *Qualitative data analysis*. Thousand Oaks, CA: Sage.
- Miles, J. ve Shevlin, M. (1998). Effects of sample size, model specification and factor loadings on the GFI in confirmatory factor analysis. *Personality and Individual Differences*, 25(1), 85-90.
- Mulaik, S. A., James, L. R., Van Alstine, J., Bennet, N., Lind, S. ve Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, 105(3), 430-45.
- Muthén, B., du Toit, S. H. C. ve Spisic, D. (1997). *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes* (Technical report). Los Angeles: University of California.
- Öztürk, N. B., Eroğlu, M. G. ve Kelecioğlu, H. (2015). Eğitim alanında yapılan ölçek uyarlama makalelerinin incelenmesi. *Eğitim ve Bilim*, 40(178), 123-137.
- Rummel, R. J. (1988). *Applied factor analysis*. Illinois, Evanston: Northwestern University Press.
- Spearman, C. (1904). General intelligence, objectively determined and measured. *The American Journal of Psychology*, 15(2), 201-292.
- Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences*, 42(5), 893-98.
- Streiner, D. L. (1994). Figuring out factors: the use and misuse of factor analysis. *Canadian Journal of Psychiatry*, 39(3), 135-140.
- Şahin, M. ve Boztunç Öztürk, N. (2018). Eğitim alanında ölçek geliştirme süreci: Bir içerik analizi çalışması. *Kastamonu Eğitim Dergisi*, 26(1), 191-199.
- Tabachnick, B. G. ve Fidell, L. S. (2012). *Using multivariate statistics* (6. bs.). Needham Heights, MA: Allyn ve Bacon.
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.
- Thompson, B. ve Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement*, 56(2), 197-208.
- Ullman, J. B. (2012). Structural equation modeling. B. G. Tabachnik ve L. S. Fidell (Ed.), *Using multivariate statistics* (6. bs.). Pearson.
- Wheaton, B. (1987). Assessment of fit in overidentified models with latent variables. *Sociological Methods ve Research*, 16(1), 118-154.
- Wolf, E. J., Harrington, K. M., Clark, S. L. ve Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *National Institutes of Health*, 76(6), 913-934. doi: 10.1177/0013164413495237.
- Worthington, R. L. ve Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34(6), 806-838. doi: 10.1177/0011000006288127
- Yong, A. G. ve 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.

### Ek 1. Araştırmada İncelenen Makale Listesi

- Ahmetoğlu, E., Mükerrer Ünal, A. ve Ergin, D. Y. (2016). Kaynaştırma uygulamalarının başarısını etkileyen etmenler ölçeğinin geliştirilmesi. *Trakya Üniversitesi Eğitim Fakültesi Dergisi*, 6(2), 167-175.
- Akaydın, Ş. ve Kurnaz, H. (2015). Lise öğrencilerine yönelik yazma tutum ölçeği: Geçerlik ve güvenilirlik çalışması. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 12(32), 246-261.
- Akbulut, N. ve Aslan, S. (2016). Örtük program ölçeği: Ölçek geliştirme çalışması. *Elektronik Sosyal Bilimler Dergisi*, 15(56), 169-176.
- Akın, A. (2006). 2x2 başarı yönelimleri ölçeği: Geçerlik ve güvenilirlik çalışması. *Sakarya Üniversitesi Eğitim Fakültesi Dergisi*, 12, 1-13.
- Akın, A. (2007). Öz-güven ölçeğinin geliştirilmesi ve psikometrik özellikleri. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 7(2), 167-176.
- Akın, A. (2010). İki boyutlu sosyal istenirlik ölçeğinin geliştirilmesi ve psikometrik özelliklerinin araştırılması. *Gazi Eğitim Fakültesi Dergisi*, 30(3), 771-784.
- Akın, U. ve Yıldırım, N. (2015). Öğretmenlerin eğitime inanma düzeyleri ölçeği (ÖEİDÖ): Geçerlik ve güvenilirlik çalışması. *e-International Journal of Educational Research*, 6(1), 70-83.
- Aksoy, V. ve Diken, İ. H. (2009). Rehber öğretmen özel eğitim öz yeterlik ölçeği: Geçerlik ve güvenilirlik çalışması. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Özel Eğitim Dergisi*, 10(1), 29-37.
- Altuntaş, S. ve Atli, A. (2015). Eş seçme stratejileri□ envanteri□ geliştirilmesi□. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 12(32), 123-135.
- Arıcak, O. T., Kınay, H. ve Tanrıkulu, T. (2012). Siber zorbalık ölçeğinin ilk psikometrik bulguları. *Hasan Ali Yücel Eğitim Fakültesi Dergisi*, 17, 101-114.
- Arslan, A. ve Kırık, A. M. (2013). Sosyal paylaşım ağlarında konum belirleme ölçeğinin geçerlik ve güvenilirlik çalışması. *Marmara Üniversitesi Öneri Dergisi*, 10(40), 223-231.
- Arslan, G. (2015). Psikolojik istismar ölçeği (PİÖ) geliştirme çalışması: Ergenler için psikometrik özelliklerinin incelenmesi. *Bartın Üniversitesi Eğitim Fakültesi Dergisi*, 4(2), 727. doi: 10.14686/buefad.v4i2.5000146983
- Arslan, S. ve Gelişli, Y. (2015). Algılanan öz-düzenleme ölçeğinin geliştirilmesi: Geçerlik ve güvenilirlik çalışması. *Sakarya University Journal of Education*, 5(3), 67-74. doi:10.19126/suje.91303
- Atıcı, M. ve Cabaroğlu, N. (2013). Davranış yönetiminde öğretmen yetkinlik algısı ölçeği. *Ç.Ü. Sosyal Bilimler Enstitüsü Dergisi*, 22(2), 223-238.
- Atli, A. ve Kaya, A. (2016). Lise öğrencilerinin mesleki değerleri. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 12(1), 314-330. doi: 10.17860/efd.78099
- Ayaz, M. F. (2016). Çokkültürlülük algı ölçeği□: Geçerli□k ve güveni□rli□k çalışması. *Elektronik Sosyal Bilimler Dergisi*, 15(57), 463-471. doi: 10.17755/esosder.11595
- Bakıoğlu, Ç. ve Kurtuldu, M. K. (2015). Pişano dersine yönelik tutum ölçeği geliştirme çalışması. *Alan Eğitimi Araştırmaları Dergisi (ALEG)*, 1(1), 33-39.
- Balcı Çelik, S., Batık, M. V., Özcan, K. ve Aksoy, S. (2015). Obez çocuklara yönelik tutum ölçeği: Geçerlik ve güvenilirlik çalışması. *e-International Journal of Educational Research*, 6(1), 40-53. <http://dergipark.ulakbim.gov.tr/ijer/article/view/1073000378> adresinden erişildi.
- Başer, E. H. ve Kılınc, E. (2015). Küresel sosyal sorumluluk ölçeği: Geçerlik ve güvenilirlik çalışması. *Sakarya University Journal of Education*, 5(3), 75-89. doi: 10.19126/suje.57328
- Bayoğlu, F. ve Atli, A. (2014). İnönü evlilik tutum ölçeği: Geçerlik ve güvenilirlik analizleri. *Ege Eğitim Dergisi*, 15(2), 397-415.
- Bütüner, S. Ö. ve Baki, A. (2011). Development of attitude scale toward the usage of history of mathematics. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 5(2), 278-311.
- Cavkaytar, A., Arduç, A. ve Aksoy, V. (2014). Aile gereksinimlerinin belirleme aracının geçerlik ve güvenilirliğinin güncellenmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Özel Eğitim Dergisi*, 15(2), 1-12.
- Çakıroğlu, Ü., Gököğlu, S. ve Çebi, A. (2015). Öğretmenlerin teknoloji entegrasyonlarına yönelik temel göstergeler: Bir ölçek geliştirme çalışması. *The Journal of Gazi Education Faculty*, 35(3), 507-522.

- Çankaya, C. ve İşçen, C. F. (2014). Fen bilgisi öğretmen adaylarına yönelik su tüketim davranış ölçeği: Geçerlik ve güvenilirlik çalışması. *e-Journal of New World Sciences Academy*, 9(3), 341-352. doi: 10.12739/NWSA.2014.9.3.1C0622
- Çelik, M. ve Yazgan İnanç, B. (2009). Evlilik doyumu ölçeği: Geçerlik ve güvenilirlik çalışmaları. *Ç.Ü. Sosyal Bilimler Enstitüsü Dergisi*, 18(2), 247-269.
- Çeliktürk, Z. ve Yamaç, A. (2015). İlkokul ve ortaokul öğrencileri için okuma kaygısı ölçeğinin geliştirilmesi: Geçerlik ve güvenilirlik çalışması. *Elementary Education Online*, 14(1), 97-107. doi: 10.17051/ieo.2015.03320
- Çengel, M. ve Türkoğlu, A. (2015). Öğrenciler tarafından algılanan sınıf iklimi ölçeği: Geçerlik ve güvenilirlik çalışması. *Eğitimde Kuram ve Uygulama*, 11(4), 1240-1257.
- Çetin, B., Bars, M. ve Bars, E. (2015). Matematik umutsuzluk ölçeği, güvenirlilik ve geçerlilik çalışması. *Elektronik Sosyal Bilimler Dergisi*, 14(53), 163-172.
- Çetin, Ş. (2006). Öğretmenlik mesleği tutum ölçeğinin geliştirilmesi (geçerlik ve güvenilirlik çalışması). *Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Dergisi*, (18), 28-37.
- Çıldır Pelitoğlu, F. ve Özgür, S. (2013). İlköğretim öğrencileri için sağlık tutum ölçeği geliştirilmesi. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 7(1), 32-56.
- Çögenli, M. Z. ve Asunakutlu, T. (2014). Akademisyenlere yönelik mobbing ölçeği geçerlik ve güvenilirlik çalışması. *Uşak Üniversitesi Sosyal Bilimler Dergisi*, (18), 92-105.
- Dağhan, G. ve Akkoyunlu, B. (2016). Çevrimiçi öğrenme ortamlarında kullanım sürekliliğini yordayabilecek yapılara ilişkin geçerlik ve güvenilirlik çalışması. *Ege Eğitim Dergisi*, 17(1), 198-224.
- \*Dalgıç, G. (2010). Öğretim elemanı performans değerlendirme envanterleri (ÖEPDE) geçerlik ve güvenilirlik çalışması. *Elektronik Sosyal Bilimler Dergisi*, 9(32), 92-105. doi: 10.17755/esosder.68526
- Dede, Y. ve Yaman, S. (2008). Fen öğrenmeye yönelik motivasyon ölçeği: Geçerlik ve güvenilirlik çalışması. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 2(1), 19-37.
- Demir, N., Kızılay, E. ve Bektaş, O. (2016). Development of an achievement test about solutions for 7th graders: A validity and reliability study. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 10(1), 209-237. doi: 10.17522/nefefmed.52947
- Demirci Seyrek, Ö., Ersanlı, K. ve Tunç, T. (2016). Duyarlık ölçeği. *Elektronik Sosyal Bilimler Dergisi*, 15(56), 284-294. doi: 10.17755/esosder.41322
- Deniş, H., Genç, H. ve Demirkaya, H. (2008). Milli parka yönelik tutum ölçeği geliştirilmesi. *GÜ, Gazi Eğitim Fakültesi Dergisi*, 28(2), 95-107.
- Dilekli, Y. ve Tezci, E. (2015). Öğretmenlerin düşünme becerilerinin öğretimine yönelik sınıf içi uygulamalar ölçeğinin geçerlik ve güvenilirlik çalışması. *e-Journal of New World Sciences Academy*, 10(4), 276-290. doi: 10.12739/NWSA.2015.10.4.1C0647
- Doğan, B. ve Çermik, H. (2016). Kitap okumaya yönelik tutum ölçeği: Geçerlik ve güvenilirlik çalışması. *e-Journal of New World Sciences Academy*, 11(3), 168-183. doi: 10.12739/NWSA.2016.11.3.1C0660
- Duruhan, K. ve Ersöz, Y. (2013). Ortaokulda demokrasi algıları ölçeği: Geçerlik ve güvenilirlik çalışması. *Journal Of History School*, 6(XVI), 657-672. doi: 10.14225/Joh355
- Ekici, E., Taşkın Ekici, F. ve Kara, İ. (2012). Öğretmenlere yönelik bilişim teknolojileri öz-yeterlik algısı ölçeğinin geçerlik ve güvenilirlik çalışması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 31, 53-65.
- Erdem, M. ve Arkün Kocadere, S. (2015). Yapılandırmacı öğrenme inanç ölçeğinin geliştirilmesi. *İlköğretim Online*, 14(2), 1260-1275. doi: 10.17051/ieo.2015.84249
- Eren, E., Yurtseven Avcı, Z. ve Seçkin Kapucu, M. (2014). Developing a scale for competencies and perceptions of necessity about using practical tools for content development. *Journal of Theory & Practice in Education*, 10(5), 1177-1189.
- Erol, M. ve Yıldırım, İ. (2016). The development of higher education life satisfaction scale. *Eğitimde Kuram ve Uygulama*, 12(1), 221-243.
- Ersanlı, K. ve Doğru Çabuker, N. (2015). Diğerkâmlık ölçeğinin psikometrik özellikleri. *Elektronik Sosyal Bilimler Dergisi*, 14(52), 43-53. <http://dergipark.ulakbim.gov.tr/esosder/article/view/5000079930> adresinden erişildi.



- Ersanlı, K. ve Uysal, E. (2015). Belirsizliğe karşı tutum ölçeğinin geliştirilmesi. *Elektronik Sosyal Bilimler Dergisi*, 14(54), 46-53.
- Eryılmaz, A. (2015). Amaçlar için mücadele ölçeği. *Uşak Üniversitesi Sosyal Bilimler Dergisi*, 8(4), 233-244.
- Eryılmaz, A. ve Dereli, E. (2011). Psychometric properties of time orientation during classroom disengagement scale. *Elementary Education Online*, 10(3), 1178-1190.
- Eryılmaz, A. ve Mammadov, M. (2016). Türk edebiyatı/Dil ve anlatım dersinde akış durumu ölçeğinin psikometrik özelliklerinin incelenmesi. *21. Yüzyılda Eğitim ve Toplum Eğitim Bilimleri ve Sosyal Araştırmalar Dergisi*, 5(13), 247-258.
- Eryılmaz, A. ve Mammadov, M. (2016). Developing the scale of factors increasing study motivation. *Journal Of European Education*, 6(2), 73-87. doi: 10.18656/jee.54739
- Gençtanırım, D. (2014). Riskli davranışlar ölçeği üniversite formu: Geçerlik ve güvenilirlik çalışmaları. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 5(1), 24-34. <http://dergipark.ulakbim.gov.tr/epod/article/view/1040000029> adresinden erişildi.
- Girgin, D. (2016). Çalgı eğitiminde algılanan aile desteği ölçeği geliştirme çalışması. *İlköğretim Online*, 15(3), 778-786. doi: 10.17051/io.2016.90950
- Gömlüksiz, M. N. ve Kan, A. Ü. (2012). Sosyal bilgiler dersi motivasyon ölçeğinin geçerlik ve güvenilirlik çalışması. *Fırat Üniversitesi Sosyal Bilimler Dergisi*, 22(2), 116-125.
- Güner, N., Akkurt Denizli, Z., Sezer, R. ve Bayraktar, A. (2015). 6.- 8. sınıf öğrencilerinin matematik dersi ve matematik ders kitapları hakkındaki görüşlerini belirlemeye yönelik ölçek geliştirme çalışması. *e-Journal of New World Sciences Academy*, 10(2), 54-65. doi: 10.12739/NWSA.2015.10.2.1C0632
- Gürbüz, H., Derman, M. ve Çakmak, M. (2013). Biyoçeşitlilik okuryazarlığı ölçeği: Geliştirme, geçerlik ve güvenilirliği. *Elektronik Eğitim Bilimleri Dergisi*, 2(3), 77-91.
- Güven, Y. ve Işık, B. (2006). Beş yaş çocukları için Marmara Sosyal Duygusal Uyum Ölçeğinin (MASDU-5 Yaş) geçerlik ve güvenilirlik çalışması. *M. Ü. Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi*, 23, 125-142.
- \*Güvenç, H. (2015). Öğretmen güdüsel desteği ölçeği geliştirme ve uyarılma çalışması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 15(1), 129-145.
- Harman, G. ve Alat, K. (2015). Fen bilgisi öğretmen adaylarının fen ve teknoloji dersinde model kullanımına yönelik tutum ölçeği geliştirilmesi. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, 17(1), 30. doi: 10.17556/jef.18249
- İlhan, A., Aslanargün, E. ve Shaukat, S. (2015). Developing teacher professionalism scale Validation and reliability study. *Eğitimde Kuram ve Uygulama*, 11(4), 1454-1474.
- Kan, A. (2007). Öğretmen adaylarının eğitim-öğretim özyeterkinliğine yönelik ölçek geliştirme ve eğitim-öğretim özyeterkinlikleri açısından değerlendirilmesi (Mersin Üniversitesi örneği). *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 3(1), 35-50.
- Kapıkıran, Ş. (2016). Ebeveyn akademik başarı baskısı ve desteği ölçeğinin psikometrik değerlendirmeleri ve yapısal geçerlik: Ortaokul ve lise öğrencileri. *Ege Eğitim Dergisi*, 17(1), 62-83.
- Kaplan, A., Öztürk, M. ve Doruk, M. (2014). Matematik uygulamaları dersine yönelik beklenti ölçeği: Geçerlik ve güvenilirlik çalışması. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, 16(2), 85-103. doi:10.17556/jef.19885
- Karaçanta, H. (2009). Öğretmen adayları için kamu personeli seçme sınavı kaygı ölçeğinin geliştirilmesi (geçerlik ve güvenilirlik çalışması). *Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Dergisi*, (25), 50-57. <http://esefdergi.gazi.edu.tr/makaleler/230.pdf> adresinden erişildi.
- Kararmak, Ö., Hacıfazlıoğlu, Ö. ve Şirin, S. (2014). Dengeli bütünleşme ve ayrışma ölçeğinin psikometrik özellikleri: Ergenlerde geçerlilik ve güvenilirlik çalışması. *Hasan Ali Yücel Eğitim Fakültesi Dergisi*, 11(22), 237-257.
- Karakuş, M. ve Çobanoğlu, N. (2013). İlkokullarda entelektüel sermayenin ölçülmesi ve bazı değişkenler açısından incelenmesi. *Uşak Üniversitesi Sosyal Bilimler Dergisi*, 6(4), 158-181.

- Karaman Kepenekçi, Y. ve Nayır, K. F. (2015). Örgütsel affedicilik ölçeğinin geliştirilmesi ve öğretmenlerin örgütsel affediciliğe ilişkin görüşleri. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 11(3), 922-934. doi: 10.17860/efd.93627
- Karaşar, B. ve Öğülmüş, S. (2016). Sosyal onay ihtiyacı ölçeği: Geçerlik ve güvenilirlik analizi. *Ege Eğitim Dergisi*, 17(1), 84-104.
- Karatekin, K. (2013). Öğretmen adayları için katı atık ve geri dönüşüme yönelik tutum ölçeğinin geliştirilmesi: Geçerlik ve güvenilirlik çalışması. *Uluslararası Avrasya Sosyal Bilimler Dergisi*, 4(10), 71-90.
- Kasapoğlu, F. (2015). Development of spiritual orientation scale: The study of validity and reliability. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 16(3), 51-68. doi:10.17679/ıuefd.16360640
- Kaya, A. ve Bozkur, B. (2015). Kadercilik eğilimi ölçeğinin geliştirilmesi: Geçerlik ve güvenilirlik çalışması. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 11(3), 935-946. doi: 10.17860/efd.55137
- Kaya, M. ve Doğan, U. (2014). Öğrenci sorumluluk: Ölçek geliştirme, güvenilirlik ve geçerlik çalışması. *Journal of European Education*, 4(1), 11-18.
- Keçeci, G., Alan, B. ve Kırbağ Zengin, F. (2016). Eğitsel bilgisayar oyunları destekli kodlama öğrenimine yönelik tutum ölçeği: Geçerlilik ve güvenilirlik çalışması. *e-Journal of New World Sciences Academy*, 11(3), 184-194. doi: 10.12739/NWSA.2016.11.3.1C0661
- Keskin, İ., Korkut, A. ve Can, S. (2016). Öğrenme güçlüğü yaşayan öğrencilere yönelik öğretmen yeterlilikleri ölçeğinin geliştirilmesi. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, 5(1), 133-155. doi: 10.17539/aej.43261
- Koç, F., Sak, R. ve Kayri, M. (2015). Okul öncesi eğitim programındaki etkinliklere yönelik öz-yeterlik ölçeğinin geçerlik ve güvenilirlik analizi. *İlköğretim Online*, 14(2), 1416-1427. doi: 10.17051/ıo.2015.50571
- Kolburan Geçer, A. ve Deveci Topal, A. (2015). E-derslere yönelik memnuniyet ölçeğinin geliştirilmesi: Geçerlik ve güvenilirlik çalışması. *Eğitimde Kuram ve Uygulama*, 11(4), 1272-1287.
- Korkut Owen, F. ve Bugay, A. (2014). İletişim becerileri ölçeğinin geliştirilmesi: Geçerlik ve güvenilirlik çalışması. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 10(2), 51-64. doi: 10.17860/efd.95021
- Kurtuldu, M. K. (2011). Piyano eğitiminde sosyal durum ölçeği geçerlik ve güvenilirlik çalışması. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 31, 85-96.
- Kurtuldu, M. K. (2010). Müzik öğretmeni adayı öz yeterlik ölçeği geçerlik ve güvenilirlik çalışması. *e-Journal of New World Sciences Academy*, 5(2), 178-189.
- Küçük, L. ve Miman, M. (2015). Çalışma yaşamında cinsiyetçilik ölçekleri geliştirme ve Toros Üniversitesinde bir uygulama. *Toros Üniversitesi İİSBF Sosyal Bilimler Dergisi*, 2(4), 117-137.
- Melanlıoğlu, D. (2013). Ortaokul öğrencileri için dinleme kaygısı ölçeğinin geçerlik ve güvenilirlik çalışması. *Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 6(11), 355-361.
- Oksal, A., Şenşekerci, E. ve Bilgin, A. (2006). Merkezi epistemolojik inançlar ölçeğinin geliştirilmesi: Geçerlik ve güvenilirlik. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi*, 19(2), 371-381.
- Onur, N. ve Nazik, M. H. (2014). Öğretmenler için bireysel finans alanında finansal tutum ölçeğinin geliştirilmesi (FTÖ): Geçerlik ve güvenilirlik çalışması. *e-Journal of New World Sciences Academy*, 9(4), 90-99. doi: 10.12739/NWSA.2014.9.4.2C0054
- Otrar, M. ve Kandemir, Y. (2016). Özel okullara yönelik bir tutum ölçeğinin geliştirilmesi: Geçerlik ve güvenilirlik çalışması. *International Journal of Social Sciences and Education Research*, 2(1), 284-297.
- \*Özcan, Z. Ç. ve Erkin, E. (2014). Matematik ödev alışkanlığı ölçeği: Geçerlik ve güvenilirlik çalışması. *Boğaziçi Üniversitesi Eğitim Dergisi*, 31(2), 27-47.
- Özer, A., Gençtanırım Kurt, D., Kızıldağ, S., Demirtaş Zorbaz, S., Arıcı Şahin, F., Acar, T. ve Ergene, T. (2016). Ergenler için yeterlik algısı ölçeğinin (YAÖ) geliştirilmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 7(1), 237-250. doi: 10.21031/epod.17205
- Özeren, E. ve Akpınar, B. (2016). İnternet kullanım alışkanlıkları ölçeği geçerlik ve güvenilirlik çalışması. *Elektronik Eğitim Bilimleri Dergisi*, 5(9), 20-29.
- Özlü, G., Özer Keskin, M. ve Gül, A. (2013). Çevre eğitimi öz-yeterlik ölçeği geliştirilmesi: Geçerlik ve güvenilirlik çalışması. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 33(2), 393-410.

- Özsoy-Güneş, Z., Derelioğlu, Y. ve Kırbaşlar, F. G. (2011). İşlemsel fizik ve kimya problemlerinde matematik kullanım ölçeği geliştirilmesi, geçerlik ve güvenilirliği çalışması. *Hasan Ali Yücel Eğitim Fakültesi Dergisi*, 16, 23-38.
- Perkmen, S. ve Tezci, E. (2015). Holland teorisinin ışığında meslek kişiliğinin ölçülmesi. *Necatibey Faculty of Education, Electronic Journal of Science and Mathematics Education*, 9(1), 184-204.
- Sarı, M. (2011). Lise yaşam kalitesi ölçeğinin geçerlik ve güvenilirlik çalışması. *Ç.Ü. Sosyal Bilimler Enstitüsü Dergisi*, 20(3), 253-266.
- Sarıçam, H. (2014). Akademik kontrol odağı ölçeği ergen formunun psikometrik özellikleri. *İlköğretim Online*, 13(4), 1135-1144. doi: 10.17051/io.2014.32429
- Seçer, İ. (2015). Zorbalıkla başa çıkma stratejileri ölçeğinin geliştirilmesi: Geçerlik ve güvenilirlik çalışması. *Atatürk Üniversitesi Kazım Karabekir Eğitim Fakültesi Dergisi*, 30, 85-105.
- Sımsıkı, H. ve Şendil, G. (2014). Baba katılım ölçeğinin (BAKÖ) geliştirilmesi. *Elektronik Sosyal Bilimler Dergisi*, 13(49), 104-123.
- Şahin, F. (2013). The scale for rating the behavioral characteristics of gifted and talented students: study of factor structure, reliability and validity. *Journal of Educational Science*, 38, 119-132. doi:10.15285/EBD.2013385570
- Şaşmaz Ören, F., Kıyıcı, G., Erdoğan, E. ve Sevinç, Ö. S. (2010). Çevre bilincine sahip öğretmen nitelikleri ölçeği: Geçerlik ve güvenilirlik çalışması. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 11(1), 133-152.
- Şengül, M. (2013). Ortaokul öğrencilerine yönelik yazma yeterliği ölçeği geliştirme çalışması. *Türkiye Sosyal Araştırmalar Dergisi*, 171(171), 81-94.
- Şimşek-Çetin, Ö. ve Alisinanoğlu, F. (2013). Okul öncesi dönemdeki çocukların yazı farkındalığını değerlendirme kontrol listesinin geçerlik ve güvenilirlik çalışması. *Karadeniz Sosyal Bilimler Dergisi*, 5(8), 15-27.
- Tatar, E., Tosun, C., Tüysüz, C., İlhan, N. ve Karakuyu, Y. (2015). Öğrencilerin fen bilimleri derslerindeki başarılarına etki eden faktörlerin belirlenmesine yönelik bir ölçeğin geliştirilmesi. *Uşak Üniversitesi Sosyal Bilimler Dergisi*, 8(4), 319-340.
- Tataroğlu, B. ve Erduran, A. (2010). Matematik dersinde akıllı tahtaya yönelik tutum ölçeğinin geliştirilmesi. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 1(3), 233-250. doi:10.16949/turcomat.65901
- Tok, M., Küçük, B. ve Kırmacı, Ö. (2015). Ortaokul kitap okuma alışkanlığı ölçeği: Geçerlik ve güvenilirlik çalışması. *Eğitimde Kuram ve Uygulama*, 11(2), 694-716.
- Tösten, R. ve Özgan, H. (2014). Pozitif psikolojik sermaye ölçeği: Geçerlik ve güvenilirlik çalışması. *Ekev Akademi Dergi Dsi D*, 18(59), 429-442.
- Türel, Y. K. (2011). Üniversite öğrencileri için akıllı tahta kullanımı değerlendirme ölçeği: Geçerlik ve güvenilirlik çalışması. *e-Journal of New World Sciences Academy*, 6(2), 1894-1903.
- Türkdoğan, T. ve Duru, E. (2012). Üniversite öğrencileri temel ihtiyaçlar ölçeğinin (ÜÖTİÖ) geliştirilmesi: Geçerlik ve güvenilirlik çalışması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 31, 81-91.
- Uğurlu, C. T. (2014). İnfomal iletişim ölçeği geçerlik ve güvenilirlik çalışması. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 15(3), 83-100. doi: 10.17679/iuefd.64723
- Uğurlu, C. T. ve Usta, H. G. (2016). Eğitimde denetim tutum ölçeği geçerlik ve güvenilirlik çalışması. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, 18(1), 137-159. doi: 10.17556/jef.28260
- Ünsal, S., Bağçeci, B., Meşe, N. N. ve Korkmaz, F. (2015). Eğitimin işlevlerini değerlendirme ölçeği: Geçerlik ve güvenilirlik çalışması. *Adıyaman Üniversitesi Eğitim Bilimleri Dergisi*, 5(2), 245-265. doi: 10.17984/adyuebd.26853
- Varış, Y. A. ve Cesur, D. (2012). Ortaöğretim düzeyi müzik dersine yönelik tutum ölçeği geliştirilmesi. *e-Journal of New World Sciences Academy*, 7(4), 361-374.
- Yalçınalp, S. ve Cabı, E. (2015). Eğitim teknolojileri kullanımı kaygı ölçeği (ETKKÖ): Geçerlik ve güvenilirlik çalışması. *İlköğretim Online*, 14(3), 1005-1016. doi: 10.17051/io.2015.50515

- Yalçınkaya, B. ve Eldemir, A. C. (2013). Bireysel çalgı dersine ilişkin tutum ölçeğinin geliştirilmesi. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 10(21), 29-36.
- Yaman, H. ve Dağtaş, A. (2013). Sözlüklere yönelik tutum ölçeği: Geçerlik ve güvenilirlik çalışması. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 10(24), 15-27.
- Yaman, H. ve Suroğlu-Sofu, M. (2013). Öğretmen adaylarına yönelik konuşma kaygısı ölçeğinin geliştirilmesi. *Türkiye Sosyal Araştırmalar Dergisi*, 173(173), 41-50. doi: 10.1080/10967490802095870
- Yaman, S. ve Tekin, S. (2010). Öğretmenler için hizmet-içi eğitime yönelik tutum ölçeği geliştirilmesi. *Bayburt Eğitim Fakültesi Dergisi*, 5(I-II), 76-88.
- Yaralı, D. (2015). Özel eğitim dersine ilişkin ölçek geliştirme çalışması: Geçerlik ve güvenilirlik. *Kafkas Üniversitesi, e - Kafkas Eğitim Araştırmaları Dergisi*, 2(2), 12-22.
- Yavuz, M., Balkan Kıyıcı, F. ve Atabek Yiğit, E. (2014). İlköğretim II. kademe öğrencileri için çevre okuryazarlığı ölçeği: Ölçek geliştirme ve güvenilirlik çalışması. *Sakarya University Journal of Education*, 4(3), 40-53.
- Yeşil, R. (2015). Sivil birey özalgı ölçeğinin geçerlilik ve güvenilirlik çalışması. *Journal of Educational Science*, 3(4), 104-118.
- Yeşilyurt, E. (2013). Yapılandırmacı öğrenme kuramına ilişkin bilişsel farkındalık ölçeği geliştirme çalışması: Bir ölçek revizyonu. *e-International Journal of Educational Research*, 8(2), 285-307. doi: 10.12739/NWSA.2013.8.2.1C058
- Yeşilyurt, E. (2012). Yapılandırmacı öğrenme konusunda öğretmen adaylarının yeterliği ölçeği: Geçerlik ve güvenilirlik çalışması. *Uluslararası Avrasya Sosyal Bilimler Dergisi*, 3(7), 29-45.
- Yıldırım, B. (2015). Fen bilimleri öğrenme kaygı ölçeği: Geçerlilik ve güvenilirlik çalışması. *Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi*, 3(1), 33. doi: 10.18506/anemon.71733

\*Bu araştırmalarda birden fazla ölçek geliştirildiği için iki çalışma olarak ele alınmıştır.