

Türkiye Adresli Eğitim Araştırmaları Dergilerinde Yayınlanmış Faktör Analizi Sonuçlarının Tekrarlanabilirliği Üzerine Bir Ön Çalışma*

Burak AYDIN**

Mehmet KAPLAN***

Hakan ATILGAN****

Sungur GÜREL*****

Öz

Akademik çalışmalarda istatistiksel çıkarımlarla erişilen sonuçları tekrar edebilmek; evrenin, örnekleme yönteminin ve çözümleme süreçlerinin şeffaf olarak tanımlanmaması durumunda mümkün değildir. Bu çalışmada Türkiye’de eğitim araştırmalarında saygın olarak kabul edilen dört dergide 2010-2017 yılları arasında yayınlanmış makalelerde yer alan faktör analizleri, sonuçlarının tekrarlanabilirliği açısından incelenmiştir. Toplamda 275 makalenin betimsel içerik analizi yapılmıştır. Çalışmaların %77.8’inde evren tanımının açıkça yapılmadığı, %50.9’unda örnekleme yönteminin belirtilmediği veya uygun örneklem kullanıldığı, %76’sında kayıp verilerden bahsedilmediği görülmüştür. Model-veri uyumu açısından çalışmaların yaklaşık yarısının eleştirilebilir olduğu değerlendirilmiştir. Araştırmalarda verilerin yapısının (sürekli/kategorik) dikkate alınmadığı da tespit edilememiştir. İncelenen çalışmalarda raporlanan sonuçları tekrar edilebilmenin büyük oranda mümkün olmayacağı değerlendirilmiştir.

Anahtar Kelimeler: Tekrarlanabilirlik, faktör analizi, betimsel içerik analizi

GİRİŞ

Açık Bilim İşbirliği ekibi üç farklı saygın psikoloji dergisinin yayınlarını incelemiş, seçtikleri toplam 100 deneysel veya korelasyonel çalışmada raporlanmış sonuçların tekrarlanabilirliğini araştırmış ve sonuçların çoğunluğunun tekrar edilemediğini belirtmiştir (Open Science Collaboration, 2015). Raporun sonuçları hem olumsuz eleştirilere (ör. Gilbert, King, Pettigrew, & Wilson, 2016) hem de destekleyici görüşlere (ör. Anderson ve ark., 2016) konu olmuştur. Gilbert ve ark. (2016) Açık Bilim İşbirliği tarafından tamamlanan çalışmada örneklemlerin hatalı ve küçük seçildiğini, sonuçların yanlış yorumlandığını ve dolayısıyla tekrarlanabilirliğin aslında daha yüksek olduğunu ileri sürmüşlerdir. Anderson ve ark. (2016) ise Gilbert ve ark. tarafından kaleme alınan eleştirilerin istatistiksel yanılığlara dayandığını ve gereğinden fazla iyimser olduğunu savunmuşlardır. Açık Bilim İşbirliği tarafından raporlanan sonuçlar, “psikoloji araştırmalarında tekrarlanabilirlik krizi” tasviri ile geniş yankı uyandırmıştır. Örneğin yayınladıkları çalışmalarda kullanılan veri setlerinin ve çözümleme esnasında kullanılan yazılım betiklerinin okuyucular ile paylaşılmasını teşvik eden dergi politikaları (ör. Behavioral Research Methods) ve açık bilimi teşvik eden akademik iş birlikleri (ör. Moshontz ve ark., 2018) oluşmuştur. Sonuçları tekrarlanamayan çalışmalar, bu çalışmaların güvenilirliğini ve geçerliğini şüpheli hale getirmektedir. Bu çalışmalardaki düşük güvenilirliğin en önemli sebeplerinden biri hataların çokluğudur. Özellikle sosyal bilimler alanında yapılan araştırmalarda örneklem

* Preliminary results of this work were presented at the 5th Congress on Measurement and Evaluation in Education and Psychology, Antalya 2016.

** Adjunct Professor, Recep Tayyip Erdoğan University, Department of Education, Rize-Turkey, burak.r.aydin@gmail.com, ORCID ID: 0000-0003-4462-1784

***Assistant Professor, Artvin Çoruh University, Department of Education, Artvin-Turkey, mehmet.kaplan2@gmail.com, ORCID ID: 0000-0002-4175-3899

**** Associate Professor, Ege University, Department of Education, Izmir-Turkey, hakan.atilgan@ege.edu.tr, ORCID ID: 0000-0002-5562-3446

*****Assistant Professor, Siirt Üniversitesi, Department of Education, Siirt-Turkey, s.gurel@siirt.edu.tr, ORCID ID: 0000-0003-3425-858X

Bu makaleye atıfta bulunmak için:

Aydın, B., Kaplan, M., Atılğan, H., & Gürel, S. (2019). A preliminary study to evaluate the reproducibility of factor analysis results: The case of educational research journals in Turkey. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 10(1), 1-11. DOI: 10.21031/epod.482393

Received: 13.11.2018

Accepted: 13.02.2019

seçiminde ve verilerin çözümlenme sürecinde yapılan dikkatsizlikler önemli hata kaynaklarından ikisidir.

Örneklem seçimi, yapılan araştırmalarda tutarlı sonuçlar elde edebilmek için önemlidir. Çünkü evrenin temsil edilememesi çalışmanın dışsal geçerliğini zedelemektedir. Örneklem seçim yöntemleri çeşitli şekillerde detaylı olarak sınıflandırılabilir (ör. Balcı, 2000; Lavrakas, 2008; Levy & Lemeshow, 2013; Kish, 1965; Neuman, 2013), fakat temelde olasılıklı (probability) veya olasılıklı olmayan (non-probability) yöntemler olarak ikiye ayrılır. Olasılıklı yöntemlerde örneklemin kullanılacak çözümlenmeye bağlı olarak yeterince büyük olmaması tesadüfi hatanın artmasına sebep olur. Olasılıklı olmayan yöntemlerde ise örneklemin içerisinde olması gerektiği halde dışarıda bırakılanlar hata kaynağıdır, bu durum seçmede yanlılık (selection bias) olarak bilinmektedir. Her iki örneklem yönteminde de karşılaşılabilecek diğer hatalar yanıtlamama (non-response) ve yanıtlanma (response) kaynaklı olabilir. Yanıtlamama durumu, bir diğer ifade ile kayıp veri, kişilere göre değişebilir ve bu hatalı sonuçlara yol açabilir. Yanıtların samimi olmaması veya etki altında verilmesi durumu da hata kaynağıdır (Bryman, 2004).

Sosyal bilimlerde yapılan araştırmalarda sıkça karşılaşılan diğer hata kaynağı ise çözümlenme süreci ile ilgilidir ve örneklemden bağımsız değildir. Sterba (2009) olasılıklı ve olasılıklı olmayan örneklem türleri üzerinde istatistiksel çözümlenme yaklaşımlarını Neyman ve Fisher tarafından tanımlanan çerçeveleri temel alarak incelemiştir. Fisher tabanlı yaklaşımlarda, olasılıklı olmayan örneklem türlerinden doğru istatistiksel çıkarım yapabilmek için üç şart gereklidir: a) araştırmacı çözümlenmesinde bir model kullanır ve bu model ile gözlenen değerler ile gözlenememiş değerler (sınırsız evrenden) arasında bağlantı kurulur, b) gözlenen değişkenlerin dağılımı hakkında varsayımlar yapılır, c) koşulluluk prensibi sağlanır (Sterba, 2009). Koşulluluk prensibi örneklemin tabakalara veya kümelere dayalı olması halinde bu durumun dikkate alınmasını gerektirir. Buna ek olarak, orantısız (disproportionate) seçimin, sonuçları büyük ölçüde değiştireceği durumlarda bu prensip sağlanamaz. Bir diğer ifade ile olasılıklı olmayan yöntemlerle elde edilen örneklemin, olasılıklı olarak elde edilecek örneklem türlerinden tehlikeli şekilde uzak olması, koşulluluk prensibinin sağlanamaması demektir. Buradan anlaşılacağı gibi Fisher'in koşulluluk prensibi özünde, olasılıklı örneklemin mümkün olmadığı durumlarda, ona en yakın örnekleme taklit etmeye (mimic) çalışmaktır. Koşulluluk prensibinin sağlanamaması durumunda örneklem türlerinden elde edilen sonuçların, evrenden elde edilebilecek sonuçlara kıyasla nasıl değişeceği kestirilemez. Fisher'in aksine Neyman yaklaşımında ise olasılıklı örneklem şarttır (Sterba, 2009).

Örneklemin doğru seçilmiş olması ve çözümlenme sürecinin hatasız tamamlanması her ne kadar bir araştırmacının tekrarlanabilirliğini garanti altına almasa da çalışma sonuçlarının benzer örneklem türleri ile tekrarlanabilirliği için şarttır. Açık Bilim İşbirliği tarafından raporlanan tekrarlanabilirlik çalışmasının bir benzerinin, Türkiye merkezli yayın yapan dergilerde yer alan çalışmalar baz alınarak tamamlanması önemli bilgilere ulaşmayı sağlayabilir. Çalışmada bu yönde atılacak bir adım için ön çalışma niteliğinde betimsel içerik analizi yapılması hedeflenmiştir. İçerik konusu olarak araştırma ekibinin ilgi alanına giren ve eğitim araştırmalarında önemli bir yeri olan faktör analizleri seçilmiştir.

Eğitim Araştırmalarında Faktör Analizi

Eğitim alanında tamamlanmış bir çalışmanın sonuçları, kullanılan ölçme aracından elde edilen puanlar üzerine inşa edilmiş ise, araştırma sonuçlarının doğruluğu ölçekten gelen puanların geçerliğine dayalıdır. Kullanılan ölçme aracının geçerliği bir bütün olarak ele alınmalıdır, fakat geçerliğe yönelik toplanacak kanıtlar çeşitli alt başlıklara ayrılabilir ve bu alt başlıklardan bir tanesi yapı geçerliği olarak bilinir. Davranışların ölçülmesi çerçevesinde, doğrudan görünür olmayan fakat bilginin ve bilimsel aklın ortak ürünü olarak ortaya çıkarılmış psikolojik bir yapının iki önemli özelliği vardır, (i) ölçülebilirliktir ve (ii) hem gerçek dünyaya ait kriterlerle hem de teorik bir sistemin parçası olarak tanımlanmış diğer yapılar ile ilişkili olmalıdır (Crocker & Algina, 1986; Lord & Novick, 1968). Yapının ölçülebilirliğine, bir diğer ifade ile ölçme aracının yapı geçerliğinin sağlandığına yönelik kanıt sunabilecek analitik araçlardan biri faktör analizidir. Diğer geçerlik alt başlıkları ve yapı

geçerliğine kanıt sunabilecek diğer araçlar birçok kaynakta detaylı olarak ele alınmıştır (Atılğan, Kan, & Aydın, 2017; Crocker & Algina, 1986; Nunnally & Bernstein, 1994).

Faktör analizlerinin eğitim araştırmalarında kullanımı iki çatı altında toplanabilir. Bunlardan birincisi bir ölçeğin geliştirilmesi ve uyarlanması aşamasıdır. Bu aşamada açımlayıcı faktör analizleri (AFA) ve doğrulayıcı faktör analizleri (DFA) kullanılabilir. İkinci durum ise daha önce geliştirilmiş ve geçerliğine yönelik kanıtlar sunulmuş bir ölçme aracının, yeni bir araştırmada veri toplamak amacıyla kullanıldığında, beklendiği şekilde çalışıp çalışmadığının belirlenmesidir ve bu aşamada ise DFA kullanılabilir. Ülkemizdeki eğitim araştırmalarında faktör analizi kullanılabilecek durumların yaygınlığı çeşitli içerik analizlerinde tespit edilmiştir. Örneğin Göktaş ve ark. (2012) 2005-2009 yılları arasında toplam 19 farklı dergide yayınlanmış 2111 makaleyi incelemiş ve bu çalışmalardan 1794 tanesinin veri toplama aracını belirttiklerini tespit etmişlerdir. Bu 1794 araçtan 590'ı (%22,3) ölçek, 343'ü (%13) ise başarı testleridir. Karadağ (2011) 2003-2007 yılları arasında eğitim bilimleri alanında tamamlanmış, görüşme formu, gözlem formu, ölçek veya başarı testi kullanan toplam 324 doktora tezinden 211'ine ulaşmış ve çoğunluğunun ölçek veya başarı testi kullandığını belirtmiştir. Benzer sonuçlara Yılmaz ve Altınkurt (2012), Sözbilir, Güler ve Çiltaş (2012), Selçuk, Palancı, Kandemir ve Dünder (2014), Kozinoğlu ve Senemoğlu (2016) ve Gökmen ve ark. (2017) tarafından da ulaşılmıştır. Yüksek etki faktörüne sahip uluslararası dergilerin içerik analizini yapan Yalçın, Yavuz ve Dibek (2016), inceledikleri 789 makalede kullanılan ölçme araçlarının yaklaşık yarısının ölçek ve başarı testi olduğunu rapor etmiştir. Ülkemizdeki ölçek geliştirme ve uyarlama çalışmalarının yaygınlığı ise akademik çalışmalara ilham olmuştur. Örneğin Öztürk, Eroğlu ve Kelecioğlu (2015) 2005-2014 yılları arasında, 10 farklı dergide yer alan 108 uyarlama çalışması tespit etmiştir ve bu çalışmaların hemen hemen hepsinde faktör analizi kullanılmıştır. Benzer bir durum Gül ve Sözbilir (2015) tarafından da tespit edilmiştir. Bu noktada faktör analizlerinin ölçek geliştirme ve uyarlama sürecinin tüm yükünü tek başına taşıyamayacağı ve bu süreçlerin zahmetli ve uzun bir süreç olduğu hatırlanmalıdır (Acar, 2014; Çüm & Koç, 2013; Erkuş, 2016; Güvendir & Özkan, 2015; Öztürk ve ark., 2015; Yurdugül & Bayrak, 2012). Örneğin Amerika Birleşik Devletleri'nde bir ölçeğin geliştirilmesi ve geçerliğine kanıt toplanmasına yönelik projelerin 3-4 seneye yayıldıkları ve milyon dolarlık bütçeler kullandıkları görülebilir. Faktör analizlerinin tarihsel gelişimi, temel kavramları, teknik detayları ve yanlış kullanımları birçok kaynakta detaylı olarak ele alınmıştır (Büyüköztürk, 2002; Erkuş, 2016; Kline, 2015; Prudon, 2015; Worthington & Whittaker, 2006; Wright, 2017).

Faktör analizine yer verilen çalışmalarda hata olasılığını azaltmada örneklem seçiminin önemini somut bir şekilde ortaya koyan çalışmalardan bir tanesi Simon (1979) tarafından tamamlanmıştır. Yazar sadece üniversite öğrencilerinden oluşan örneklem üzerinden geçerliği savunulmuş bir tutum ölçeğinin, üniversite öğrencisi olmayan bireylerde farklı şekilde çalışabileceğine dikkat çekmek istemiştir. İlk örneklemine sadece bir üniversiteden çalışmaya davet ettiği 188 öğrenciden, ikinci örneklemine ise ulusal bazda faaliyet gösteren bir vakfın yardımı ile 188 farklı bireyden oluşturmuştur. Farklı iki örneklem üzerinde tamamen aynı faktör analizi tekniklerini kullanan yazar farklı faktör yapılarına ulaşmıştır. Bu noktada, faktör analizlerinde örneklemin bir ülkeyi, coğrafyayı veya toplumu temsil etmesi değil, ölçülecek davranışları temsil edebilmesi gerektiği hatırlanmalıdır. Fakat ölçülecek davranışların bireylerden bağımsız olmadığı, çalışma evreninin tanımlanabilir olduğu, faktör analizlerinin istatistiksel modeller olduğu ve bir ölçeğin yapı geçerliğini sınamak için kullanılacağı kabul edilirse; her ne kadar olasılıklı olmayan yöntemlerden bazıları (ör. amaçlı örneklem, kota örnekleme) çok dikkatli şekilde kullanıldığında savunulabilir olsa dahi, örneklemin olasılıklı yöntemlerle belirlenmesi, koşulluluk prensibi düşünüldüğünde daha avantajlıdır. AFA'da örneklem seçiminin önemini somut bir şekilde ortaya koyan diğer bir çalışma Gaskin, Orellana, Bowe ve Lambert (2017) tarafından tamamlanmıştır. Yazarlar, Dünya Sağlık Örgütü tarafından kullanılan ve bireylerin genel anlamda sağlıklı olup olmadığını belirlemek üzere tasarlanmış bir ölçeğin yapı geçerliğini çalışmışlardır. Altı farklı ülkeden, 50 yaş üstü 31251 bireye ait verilerin çalışma evrenini oluşturduğu varsayılan çalışmada yazarlar, iki farklı örnekleme yöntemi kullanmışlardır. Birinci yaklaşımda 1000 farklı örneklem, toplam sağlık puanları referans alınarak ve 31251 bireye ait puanların çarpık dağılımını yansıtmak üzere basit seçkisiz (simple random sampling) olarak seçilmiştir. İkinci yaklaşımda 1000 farklı örneklem, bireylerin sağlık puanlarının normal dağılım

göstereceği şekilde tabakalı seçkisiz (stratified random sampling) olarak seçilmiştir. Seçilen örneklem üzerinde açıklayıcı faktör analizleri yapılmıştır. Basit seçkisiz yöntemle genellikle tek faktörlü yapı, diğer yöntemle ise genellikle iki faktörlü yapı bulunmuştur. Yazarlar, tabakalı seçkisiz örnekleme yöntemiyle elde edilen yapıyı daha savunulabilir bulmuşlardır. Bu sonuçlar örneklemin olasılığa dayalı yöntemler kullanarak seçildiğinde dahi farklı faktör yapılarını destekleyebileceğini göstermektedir. Ayrıca, bu sonuçlar, örneklem seçiminde evren hakkındaki mevcut bilgiyi (prior) kullanmanın önemini de vurgulamaktadır (Smith, 1983). Örneklem perspektifinden tekrarlanabilirliği zorlaştıracak etkenlerden bir diğeri ise uygun örnekleme (convenience) kullanmaktır. Uygun örnekleme yöntemi zaman ve paradan tasarruf karşılığında bulunacak sonuçların doğruluğundan taviz verebilir (Balcı, 2015). Uygun örnekleme yöntemi ile ulaşılan örneklemin kendinden büyük herhangi bir evreni temsil edebilme ihtimali genellikle çok düşüktür. Uygun örnekleme yöntemiyle elde edilecek sonuçların geçerliği yüksek derecede kaygı içerir ve bu durum birçok akademik çalışmaya konu olmuştur (Bornstein, Jager, & Putnick, 2013; Delice, 2010; Landers & Behrend, 2015; Peterson & Merunka, 2014; Tyrer & Heyman, 2016).

Evreni temsil edebilecek bir örnekleme yöntemi belirlendikten sonra, tekrarlanabilirlik için önemli olan diğer bir husus örneklem büyüklüğüdür. Faktör analizlerinde yanlı olmayan tahminlere ulaşabilmek için örneklemin yeterince büyük olması gerekir. Buradaki yeterince büyüklük faktör yapısının karmaşıklığına, faktör yüklerinin büyüklüğüne, kayıp verilere göre değişebilir ve kendi çalışmalarında gerekli örneklem büyüklüğünü belirlemek isteyen araştırmacılar Monte Carlo simülasyon çalışmalarını kullanabilirler (Wolf, Harrington, Clark, & Miller, 2013). Sonuç olarak faktör analizinde kullanılan verilerin hangi evrenden, hangi örnekleme tekniği ile hangi örneklem genişliğinde çekileceği önemli bir konudur ve ölçme aracının psikometrik niteliklerinin doğruluğunu etkiler. Nitekim faktör analiziyle elde edilen faktörler kullanılan örneklemden etkilenir (Kline, 2015). Bununla birlikte faktör analizi evren yerine örneklemden bilgi ya da varyansı yeniden üretme girişimidir ve örneklem evrenin makul bir temsilcisi ise örneklemden elde edilen faktörler evren faktörleri ile eşleşme eğiliminde olurlar (Thompson, 2004).

Teknik açıdan ele alındığında faktör analizleri bir boyut azaltma (dimension reduction) işlemidir. Bir ölçeği oluşturan n farklı soruya verilen yanıtlar, $n \times n$ boyutunda bir kovaryans matrisi oluşturur ve faktör analizleri bu matrisi daha az sayıda değişken kullanarak üretebilmenin yolunu arar (Crocker & Algina, 1986). Diğer bir ifade ile n farklı değişkenin sahip olduğu varyans daha az sayıda değişkenle, yani faktörle, temsil edilmeye çalışılır. Bu boyut azaltma işlemi, soruların sayısı, faktörlerin sorular ve birbiri arasındaki ilişkisi, kayıp verilerin nasıl ele alındığı, tahminleme yönteminin özellikleri gibi etkenlere bağlı olarak oldukça karmaşık bir hâl alabilir. Faktör analizlerinin bütün teknik parçalarını ele alan birçok kaynak mevcuttur (ör. Büyüköztürk, 2002; Crocker & Algina, 1986; Kline, 2015; Prudon, 2015; Thompson, 2004), AFA veya DFA sonucunda ortaya konulmuş bir yapının benzer örneklem ile tekrarlanabilirliği noktasında kayıp verilerin nasıl ele alındığı (Akbaş & Tavşancıl, 2015; Çüm & Gelbal, 2015; Kürşad & Nartgün, 2015), kullanılan yazılımın özellikleri (Hox, 1995), tahminleme yöntemi (Beauducel & Herzberg, 2006), verilerin sıralı veya sürekli değişken olarak ele alınması (Rhemtulla, Brosseau-Liard, & Savalei, 2012, Yang-Wallentin, Jöreskog, & Luo, 2010) ve model-veri uyum göstergeleri önemli bilgiler taşır (Prudon, 2015).

Sonuç olarak, akademik çalışmalarında rapor ettikleri faktör analizi sonuçlarının tekrarlanabilirliğini artırmak isteyen araştırmacılar örneklem seçimi ve çözümleme sürecini detaylı olarak açıklamalıdır. Tekrarlanabilirlik çalışmalarının yapılabilmesi için incelenecek araştırmaların hangi koşul ve kabuller ile tamamlandığının bilinmesi gerekir. Eğer çalışmaların koşul ve kabulleri bilinmiyorsa tekrarlanabilirlik çalışması yapılamaz ve dolayısıyla sonuçların tekrar edilebileceği söylenemez. Bu çalışmanın amacı, Türkiye merkezli dergilerde yayınlanmış faktör analizi sonuçlarını tekrarlanabilirlik açısından incelemek isteyen bir araştırma ekibinin, çalışmaya başlamadan önce zorluklarla karşılaşmış karşılaşılmayacağını belirlemektir. Belirlenen amaç doğrultusunda, Türkiye’de eğitim araştırmalarını kapsam olarak belirleyen ve saygın olarak kabul edilen dergilerde yapılan çalışmalar incelenmiş ve şu sorulara yanıt aranmıştır;

1. Çalışmalarda evren tanımı açık ve anlaşılır olarak yapılmış mıdır?
2. Hangi örnekleme yöntemleri kullanılmıştır?

3. Örneklem büyüklüğü, soru ve faktör sayıları nelerdir?
4. Kayıp veriler nasıl ele alınmıştır?
5. Hangi yazılımlar kullanılmıştır?
6. Kullanılan veri türü (kategorik/sürekli) ve tahminleyiciler (estimator) nelerdir?
7. Rapor edilen uyum bilgileri nelerdir?

Belirlenen bu yedi soru ile tekrarlanabilirlik, örneklem seçimi ve çözümleme süreci olarak iki boyutta incelenmiştir.

YÖNTEM

Çalışma kapsamı eğitim bilimleri alanında olan ve Türkiye’den yayınlanan, uluslararası indekslerce taranan dört dergi ile sınırlandırılmıştır. Bu dergiler; Eurasian Journal of Educational Research (EJER), Kuram ve Uygulamada Eğitim Bilimleri (KUYEB), Hacettepe Üniversitesi Eğitim Fakültesi (HÜEF) ile Eğitim ve Bilim (EB) dergileridir. Bu dergilerde yapılan bütün çalışmaların küçük bir araştırma ekibi ile incelenmesi mümkün olmadığından, kapsamın sınırları iki ana hat ile çizilmiştir. İlk sınırlama yayın tarihleridir ve nispeten güncel verilere ulaşmak hedefi ile 2010 Ocak- 2017 Aralık arasında yayınlanmış makaleler incelenmiştir. İkinci sınırlama ise konu seçimidir, araştırma ekibinin ilgi alanına giren ve eğitim araştırmalarında önemli bir yeri olan faktör analizleri odak olarak belirlenmiştir.

Çalışma kapsamına alınan dört dergide, belirlenen tarih aralığında yayınlanmış makalalarda faktör analizi içerenleri tespit etmek üzere, *geliştirme, uyarlama, faktör analizi, yapısal eşitlik modellemesi, geçerlik, güvenilirlik, doğrulayıcı, açımlayıcı, DFA, AFA* veya *Cronbach* anahtar kelimelerine veya bu terimlerin İngilizcelerine makale özetlerinde yer veren 341 yayın veri tabanlarından indirilmiştir. Her bir dergi seçkisiz olarak belirlenen bir yazar tarafından ön incelemeye alınmış, 341 makaleden 275’inin DFA, AFA veya Temel Bileşenler Analizi (TBA, İng. Principal Component Analysis) kullandığı tespit edilmiştir. Çalışma kapsamına alınan 275 makale, yine her bir dergi seçkisiz olarak belirlenen bir yazar tarafından, betimleyici araştırma (descriptive research) çerçevesinde incelenmiştir. Betimleyici araştırma nicel veri analizi yöntemlerinden biridir ve genellikle frekans, ortalama, ortanca, varyans gibi temel istatistiklerin raporlanmasını içerir (Gall, Gall, & Borg, 2003; Stapleton & Leite, 2005).

Verilerin Toplanması

Araştırma sorularını yanıtlamak üzere 275 makalede yer alan veriler, araştırmacılar tarafından geliştirilen ve gözden geçirilen bir form aracılığıyla toplanmıştır. Öncelikle her bir makalenin künyesi, yayınlanma yılı, yayımlandığı dergi ve genel amacı kaydedilmiştir. Faktör analizi kullanan makalelerin genel amacı ölçek geliştirme, ölçek uyarlama veya diğer olarak kodlanmıştır, diğer kategorisinde t-test, varyans analizi, regresyon, yapısal eşitlik modelleri, yol analizleri gibi nicel veri analizi teknikleri yer almıştır. Sonrasında tekrarlanabilirliğin ilk boyutu olarak ele alınan örneklem özellikleri belirlenmeye çalışılmış ve makalelerde sırasıyla, evren tanımının yapılıp yapılmadığı, örneklemin içeriği ve örnekleme yönteminin olasılığa dayalı olup olmadığı tespit edilmeye çalışılmıştır. Kodlamanın kolaylığı açısından örneklemin içeriği, öğrenci, öğretmen veya öğretmen adayı, akademisyen, yönetici ve diğer kategorilerine ayrılmıştır. Aynı zamanda araştırmacıların örnekleme yöntemini tanımlamaları durumunda bu bilgi kaydedilmiştir. Son aşamada, tekrarlanabilirliğin ikinci boyutu olarak ele alınan çözümleme sürecine yönelik veriler belirlenmeye çalışılmıştır ve makalelerde sırasıyla, kayıp verilerden bahsedilip bahsedilmediği, kullanılan çözümleme tekniği, aynı örneklem üzerinde AFA ve DFA yapılıp yapılmadığı, örneklemin büyüklüğü, madde sayısı, faktör sayısı, verilerin türü, analizde kullanılan yazılım ve model-veri uyum bilgileri aranmıştır. Çözümleme tekniği iki kategori olarak kodlanmış, açımlayıcı teknik kabul edilen TBA (İng. PCA), AFA olarak kodlanmıştır (Bryant & Yarnold, 1995). İncelenen makale içerisinde birden fazla ölçek kullanıldıysa madde sayısı, faktör sayısı, verilerin türü ve uyum bilgileri aynı makale künyesinde farklı bir satıra kaydedilmiştir. Ayrıca aynı ölçek için birden fazla model test edildiye, sadece nihai modelin bilgileri dikkate alınmıştır. Eklenen satırlarla birlikte veri seti toplamda 448 satırdan oluşmuştur. Model-veri

uyum bilgileri olarak, Ki-karenin serbestlik derecesine oranı, yaklaşık hataların ortalama karekökü (the root mean square error of approximation, aka, RMSEA), standardize edilmiş hataların ortalama karekökü (standardized root mean square residual, aka, SRMR), karşılaştırma uyum indeksi (comparative fit index, aka, CFI), Tucker-Lewis indeksi (TLI veya NNFI), normatif uyum indeksi (NFI), uyum iyiliği indeksi (goodness of fit index, aka, GFI), düzeltilmiş GFI (AGFI), artımlı uyum indeksi (incremental fit index, aka, IFI) ve görelî uyum indeksi (relative fit index, aka, RFI) bilgileri toplanmıştır.

BULGULAR

İncelenen 275 makalenin 2010-2017 yıllarına dağılımı sırasıyla, 35 (%12.7), 32 (%11.6), 35 (%12.7), 46 (%16.7), 53 (%19.3), 28 (%10.2), 18 (%6.5) ve 28 (%10.2); EB, EJER, HÜEF ve KUYEB dergilerine dağılımı sırasıyla, 94 (%34.2), 56 (%20.4), 40 (%14.5) ve 85 (%30.9); geliştirme, uyarlama ve diğer kategorilerine dağılımı sırasıyla, 108 (%39.3), 99 (%36.0) ve 68 (%24.7) olarak bulunmuştur. Tablo 1 incelenen 275 makalenin yıl, dergi ve genel amaç değişkenlerine göre dağılımını gösterir.

Tablo 1. İncelenen Makalelerin Yıl, Dergi ve Genel Amaç Değişkenlerine Göre Dağılımını

	EB			EJER			HÜEF			KUYEB			Top.
	GEL.	UYA.	D.	GEL.	UYA.	D.	GEL.	UYA.	D.	GEL.	UYA.	D.	
2010	3	5	5	1	3	1	4	4	0	6	3	0	35
2011	3	8	3	1	0	1	3	1	0	6	5	1	32
2012	4	3	2	2	3	0	3	6	0	9	1	2	35
2013	5	11	2	2	3	1	3	5	0	6	6	2	46
2014	12	12	4	5	0	5	3	1	0	5	3	3	53
2015	1	0	2	1	2	7	4	1	2	2	0	6	28
2016	0	2	2	2	2	4	0	0	0	2	4	0	18
2017	1	2	2	5	2	3	0	0	0	4	1	8	28
Top.	29	43	22	19	15	22	20	18	2	40	23	22	275

Not: GEL.= Geliştirme, UYA= Uyarlama, D= Diğer, Top.= Toplam

Evren Tanımı ve Örneklem Seçimi

Yapılacak olan bir araştırma öncesinde çalışmanın gerçekleştirileceği evrenin açık bir şekilde tanımlanması araştırma sonrası elde edilecek olan sonuçların geçerliğini sağlamada önemli noktalardan biridir. Ancak yapılan bu çalışmada incelenen 275 makaleden sadece 61'inde (%22.2) yazarların açık ve net bir evren tanımı yaptığı tespit edilebilmiştir. Bu durum bu çalışmalardan elde edilen sonuçların benzer evrenler için de geçerli olabilmesini şüpheli hale getirmektedir. Tablo 2 açık ve net bir evren tanımının yapıldığı çalışmaları yıl ve genel amaç bazında gösterir. Ölçek geliştirme ve uyarlama çalışmalarında açık ve net bir evren tanımı yaklaşık olarak her 5 çalışmanın 1'inde, diğerlerinde ise yaklaşık olarak her 3 çalışmanın 1'inde yer almıştır. Çalışmalarda yaygın bir şekilde tanımlanmamış olan evrenin aksine incelenen makalelerin 227'sinde (%82.6) örneklem seçiminin olasılığa dayalı olup olmadığı tespit edilebilmiştir. Daha spesifik olarak örneklem seçimi belirtilen bu 227 çalışmanın 169'unda olasılıklı olmayan, 58'inde ise olasılığa dayalı olan tekniklerin kullanıldığı gözlemlenmiştir. Örneklemenin olasılıklı olmadığı görülen 169 çalışmanın 112'sinde tekniğin net olarak ifade edildiği ve bunların 92'si uygun veya gönüllü, 11'i amaçsal, 5'i tabakalı, 2'si maksimum çeşitlilik, 1'i kartopu ve 1'i tipik durum örnekleme olduğu tespit edilmiştir. Genel olarak bakıldığında, 275 çalışmanın 48'inde (%17.4) örnekleme tekniği belirlenememiş, belirlenenler arasında 92'sinin (%33.5) uygun veya gönüllü (convenience) örnekleme kullandığı görülmüştür. Şekil 1 uygun veya gönüllü örnekleme kullanıldığını belirtilen çalışmaların yıllara ve genel amaca göre oranını gösterir. Toplam oranlara bakıldığında uygun veya gönüllü örnekleme, geliştirme, uyarlama ve diğer amaçlar için sırasıyla çalışmaların %31, %49 ve %29'unda kullanılmıştır. Ayrıca, bütün makalelerde örneklemin içeriği tanımlanmıştır. Daha spesifik olarak, çalışmaların 205'i (%74.6) sadece öğrencileri, 40'ı (%14.6) öğretmen veya öğretmen adaylarını, 8'i (%2.9) sadece akademisyenleri, 4'ü (%1.4)

sadece yöneticileri ve geri kalan 18'i (%6.4) bu gruplardan en az ikisini veya diğer bireyleri içermiştir (ör. veli ve yetişkinler gibi).

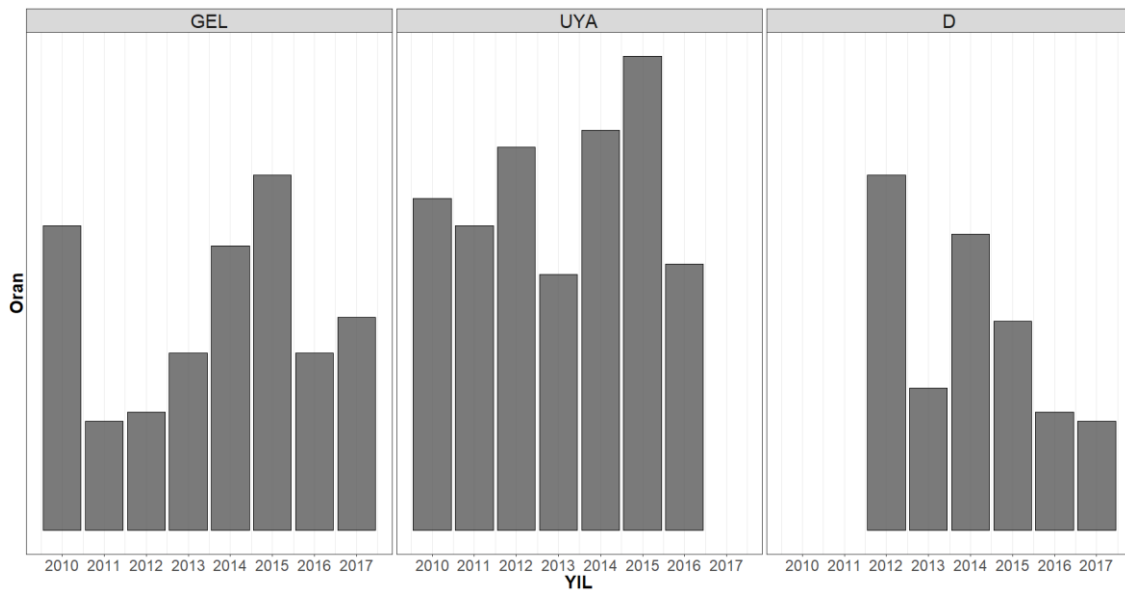
Örneklem Büyüklüğü, Madde ve Faktör Sayıları ve Kullanılan Veri Türü

Yapılan bu çalışmada, araştırmalarda kullanılan örneklem büyüklüğü, madde sayısı, faktör sayısı ve verilerin türü taranan 275 makalede 448 farklı analiz için ayrı ayrı raporlanmıştır. Gözlemlenen örneklem büyüklüklerinin, ölçeklerde kullanılan madde sayılarının ve elde edilen faktör sayılarının ortanca değerleri sırasıyla, 398, 25 ve 3'tür. Ayrıca, madde başına düşen örneklem büyüklüğü 14.8, faktör başına düşen madde sayısı ise 7 ortanca ile dağılmıştır. Tablo 3 448 analizde raporlanan örneklem büyüklüğü, faktör başına düşen madde ve madde başına düşen örneklem ortancalarının yıllara ve amaca göre dağılımını gösterir. Bütün yılların ortalamasına bakıldığında, faktör başına düşen madde sayısı ortancalarının amaca göre benzer olduğu, örneklem büyüklüğü ortancalarının ölçek geliştirme (381) ve uyarlama çalışmalarında (400) benzer fakat diğer amaçlı çalışmalardan (459) nispeten daha düşük olduğu ve madde başına düşen örneklem ortancalarının ölçek geliştirme çalışmalarında (12.2) uyarlama (16.6) ve diğer amaçlara (17.7) nispeten daha düşük olduğu görülmüştür.

Tablo 2. Amaç ve Yıllara Göre İncelenen 275 Makalede Evren ve Kayıp Veri Bilgisinin Verilme Oranları

Yıl	Evren bilgisi oranı			Kayıp veri bilgisi oranı		
	GEL	UYA	D	GEL	UYA	D
2010	21	33	17	29	13	33
2011	15	21	40	31	21	40
2012	28	8	50	22	23	25
2013	25	12	40	13	20	0
2014	12	6	25	20	13	25
2015	0	67	41	25	67	29
2016	0	38	50	75	25	17
2017	40	20	8	10	40	46
Bütün yıllar	19	19	31	23	21	29
Genel	22			24		

Not: GEL.= Geliştirme, UYA= Uyarlama, D= Diğer



Şekil 1. Uygun veya Gönüllü Örneklem Oranının Yıllara ve Amaca Göre Dağılımı.

Buna ek olarak, yapılan arařtırmalarda kullanılan veri türü incelendiğinde, taranan 275 makalenin 228'inde (%82.9) kategorik (ör. Likert), 7'sinde (%2.5) ikili (binary), 4'ünde (%1.5) sürekli deęişkenlerin kullanıldığı gözlemlenmiş ancak 36'sında (%13.1) hangi veri türünün kullanıldığı yazarlar tarafından belirtilmemiştir. Makalelerde yer alan analizler incelendiğinde, kullanılan veri türü 318 analizde belirtilmiştir, tespit edilen türlerin 304'ü kategoriktir, en çok tercih edilen seçenek sayısı beştir (209), üçlü, dörtlü, altılı, yedili, dokuzlu ve onlu seçenekler ise sırasıyla 11, 36, 16, 24, 5 ve 3 çalışmada tespit edilmiştir.

Table 3. 448 Analizde Raporlanan Örneklem Büyüklüğü, Faktör Başına Düşen Madde ve Madde Başına Düşen Örneklem Ortancalarının Yıllara ve Amaca Göre Dağılımı

Yıl	Örneklem büyüklüğü			Faktör başına düşen madde			Madde başına düşen örneklem		
	GEL	UYA	D	GEL	UYA	D	GEL	UYA	D
2010	464	358	367	8.9	8.3	8.5	12.2	21.3	10.0
2011	461	341	214	8.5	7.6	4.0	12.4	12.8	16.6
2012	336	529	258	6.1	7.0	12.5	10.8	13.6	6.0
2013	388	407	605	6.5	6.0	4.0	10.7	21.9	97.9
2014	317	436	256	6.0	5.0	7.0	12.9	25.6	14.9
2015	384	357	657	10.7	6.3	6.1	13.0	9.4	49.2
2016	330	462	556	4.3	5.3	7.0	12.3	15.5	20.4
2017	303	270	719	7.3	5.7	6.6	11.0	16.4	27.8
Bütün	381	400	459	7.3	6.5	7.0	12.2	16.6	17.7
Genel	398			7				14.8	

Not: GEL.= Geliştirme, UYA= Uyarlama, D= Diğer

Kayıp Veri ve Çözümleme Süreci

İncelenen 275 makalenin sadece 66'sında (%24) kayıp verilerin nasıl ele alındığı tespit edilebilmiştir. Bu 66 çalışmanın 62'sinde sıralı-silme (listwise) ve 3'ünde atama (imputation) yöntemi kullanmış ve 1 çalışmada ise hiç kayıp veri olmadığı belirtilmiştir. Tablo 2 kayıp veri bilgisi içeren makalenin oranlarını yıl ve amaç bazında gösterir. Açık ve net bir evren tanımı içeren makalelerin oranlarına benzer olarak, bütün yıllar göz önüne alındığında, ölçek geliştirme ve uyarlama makalelerinde kayıp veri bilgisinin diğer amaçlı makaleler kıyasla daha az raporlandığı görülmüştür, oranlar sırasıyla %23, %21 ve %29'dur. Buna ek olarak, çalışmaların 84'ünde (%30.6) sadece DFA, 57'sinde (%20.7) sadece AFA ve 134'ünde (%48.7) hem AFA hem de DFA kullanıldığı tespit edilmiştir. Hem AFA hem de DFA kullanan 134 makalenin 90'ı, çözümlemeleri aynı örneklem veya aynı örneklemin ikiye bölünmesi ile tamamlanmıştır. Ayrıca incelenen 275 makalenin 183'ünde çözümlemede kullanılan yazılım bilgisi tespit edilebilmiştir. Bu çalışmalar incelendiğinde SPSS ve Lisrel beraber, Lisrel, SPSS, AMOS, SPSS ve AMOS beraber, Mplus, ve EQS sırasıyla 71, 49, 29, 19, 12, 2 ve 1 çalışmada tercih edilen yazılımlar olarak belirtilmiştir. Burada dikkat çeken nokta bu çalışmalar için kullanılan yazılımların tamamının ticari yani ücretli oluşudur.

Model-veri Uyum Bilgileri

Ki-karenin serbestlik derecesine oranı 183 çalışmada 2.66 ortalama [1.01,9.45]; RMSEA 245 çalışmada 0.06 ortalama [0,0.44]; SRMR 131 çalışmada 0.05 ortalama [0.004, 0.11]; CFI 233 çalışmada 0.96 ortalama [0.70,1]; TLI 143 çalışmada 0.96 ortalama [0.69, 1]; NFI 146 çalışmada 0.95 ortalama [0.64,1]; GFI 197 çalışmada 0.92 ortalama [0.47,1]; AGFI 157 çalışmada 0.90 ortalama [0.07,1]; IFI 54 çalışmada 0.95 ortalama [0.81,1]; RFI 41 çalışmada 0.96 ortalama [0.62,1] ile dağılım göstermiştir. Kullanılan tahminleyici sadece 39 analizde tespit edilebilmiştir, bunların 30'u en çok olabilirlik (İng. Maximum likelihood), 7'si dirençli en çok olabilirlik (İng. Robust maximum likelihood) ve 2'si en küçük kareler yöntemidir.

TARTIŞMA ve SONUÇ

Psikoloji araştırmalarında tekrarlanabilirlik krizinden (Open Science Collaboration, 2015) yola çıkarak bu çalışmada, Türkiye’de eğitim araştırmalarını kapsam olarak belirleyen ve saygın olarak kabul edilen dört dergide 2010-2017 arasında yayınlanmış ve faktör analizi kullandığı belirlenmiş 275 makale incelenmiştir. Tekrarlanabilirlik, örneklem seçimi ve çözümleme süreci olarak iki boyutta ele alınmış, belirlenen 275 makalede toplam 448 analizden veri toplanmıştır.

Faktör analizlerinin %75.3 oranında ölçek geliştirme ve uyarlama çalışmalarında, %24.7 oranında ise diğer nicel veri analizi yöntemleri kullanan çalışmalarda raporlandığı görülmüştür. İncelenen çalışmaların %77.8’sinde açık ve net bir evren tanımı bulunamamıştır. Örneklem yönteminin belirlenemediği veya uygun/gönüllü örneklem olarak tespit edildiği makale sayısı 140’tır (%50.9). Çalışmaların %76’sında kayıp verilerin nasıl ele alındığı tespit edilememiştir, tespit edilenlerin ise hemen hemen hepsinde en basit ve eski yöntemlerden biri olan sıralı silme kullanıldığı görülmüştür. 275 çalışmanın 90’ında aynı örneklemle hem AFA hem de DFA raporlanmıştır. Açık ve anlaşılır bir evren tanımının, genellenebilir sonuçlar üretebilecek bir örneklem tekniğinin önemi düşünüldüğünde bu tespitler tekrarlanabilirlik adına kaygı vericidir. Ayrıca aynı veri üzerinde hem AFA hem de DFA yaparak raporlanan sonuçların başka çalışmalarda tekrarlanabilirliği de tartışmaya açık bir konudur (Erkuş, 2016; Van Prooijen & Van Der Kloot, 2001). Sosyal bilimlerin genelinde olduğu gibi (Schafer, 1997; Schlomer, Bauman, & Card, 2010), faktör analizlerinde de kayıp verilerin nasıl ele alındığı önemli bir araştırma konusudur (Allison, 2003; Çüm, & Gelbal, 2015). İncelenen çalışmalarda kayıp verilerin nasıl ele alındığının açıkça belirtilmemesi de tekrarlanabilirlik kaygısını artırmıştır. Türkiye’de eğitim araştırmalarında kayıp verilere ve kayıp veri tekniklerine gereken hassasiyetin gösterilmediği Demir ve Parlak (2012) tarafından da saptanmıştır. Çüm ve Gelbal (2015) kayıp veri tekniklerinin yanlış kullanılması durumunda sonuçların araştırmacıları yanıltabileceğini belirtmiştir, bu durum sonuçların tekrarlanabilirliği ile doğrudan ilgilidir. İncelenen her dört çalışmanın üçünde kayıp verilerden bahsedilmemesinin sebebi eğer toplanan verilerde hiç kayıp verinin olmaması ise yine bu durum endişe vericidir, katılımcıların boş yanıt bırakmamaya zorlanması veya boş yanıt bırakma seçeneğinin olmaması sakıncalıdır (Ray, 1990; Xiao, Liu & Li 2017).

Faktör analizlerinde, sonuçların tekrarlanabilirliği açısından önemli diğer konu yeterli örneklem büyüklüğünün sağlanabilmesidir (Wolf, Harrington, Clark, & Miller, 2013). Yeterli büyüklük ise modelin karmaşıklığına ve faktör yüklerinin büyüklüğüne göre değişir. Monte Carlo benzetimleri yeterli büyüklüğü tespit etmede kullanılabilir güçlü tekniklerdir, fakat alan yazında, katılımcı sayısının madde sayısına oranına yönelik tavsiyeler de mevcuttur, örneğin 1’e 20 ve 1’e 10 (akt. Hogarty, Hines, Kromrey, Ferron, & Mumford, 2005). İncelenen çalışmalarda bu oranın ortanca değeri yaklaşık 15 bulunmuştur ve genel olarak örneklem büyüklüğüne gereken önemin verildiği değerlendirilmiştir. Faktör başına düşen madde sayısının çalışmaların yarısında 7’den az olduğu görülmüştür. Teoride, modelde birden fazla faktör varsa, iki madde ile bir faktör tanımlanabilir, fakat faktör başına düşen madde sayısının en az 3,4 veya 5 olması tavsiye edilir (Kline, 2015). Madde sayısının artırılması yapının daha güçlü olarak tanımlanabilmesini, dolayısıyla tekrarlanabilirliğin artmasını sağlayabilir. Genel olarak bakıldığında, faktör başına düşen madde sayısına gereken önemin vermediği değerlendirilmiştir.

Faktör analizlerinde kullanılan model-veri uyum bilgileri sonuçların tekrarlanabilirliğine ipucu teşkil eder. Uyum değerleri, açıklanamayan varyans kaynakları olduğuna veya model doğru kurulamadığında düşük çıkar ve bu durum tekrarlanabilirlik için risk oluşturur. Model-veri uyumu bilgilerinde kesme değerlerin ne olması gerektiği birçok çalışmanın konusudur (Kline, 2015; Marsh, Balla, & McDonald 1988; O’Boyle & Williams 2011; Prudon, 2015), RMSEA <0.06, SRMR <0.08, CFI, TLI(NNFI), NFI, GFI, AGFI için >0.95 kesme değerlerinin iyi uyumu işaret ettiği varsayımı kabul edildiğinde, incelenen çalışmaların yaklaşık yarısının bu kriterleri sağlamakta zorlandığı değerlendirilmiştir. Ki-karenin serbestlik derecesine oranı ise değerlendirmede dikkate alınmamıştır, bu uyum bilgisinin kullanılmaması gerekmektedir (Kline, 2015). Ayrıca kullanılan tahminleyicinin analizlerin çoğunda belirtilmemiş olması verilerin özelliklerinin çözümleme sürecinde dikkate alınıp alınmadığının tespit edilmesini engellemiştir. Normal dağılım varsayımı karşılanmadığında kategorik

(ör. Likert) değişkenlerin sürekli değişken olarak kabul edilmesi tekrarlanabilirliği zedeleyeceğinden (Li, 2016), bu durum diğer bir endişe kaynağıdır.

Sonuç olarak, Türkiye merkezli eğitim araştırmaları dergilerinde yayınlanmış faktör analizi sonuçlarını tekrarlanabilirlik açısından incelemek isteyecek bir araştırma ekibinin zorluklarla karşılaşacağı, dolayısıyla genel olarak sonuçları tekrar edemeyeceği değerlendirilmiştir. Tekrarlanabilirlik ölçek geliştirme ve uyarılma çalışmalarında diğer amaçlı makalelere bakıldığında daha problemleri görünmektedir çünkü bu çalışmalarda açık ve net bir evren tanımının ve kayıp veri bilgisinin daha az bulunduğu, örneklemelerin nispeten daha küçük olduğu ve uygun veya gönüllü örneklemenin daha sık kullanıldığı görülmüştür.

Bu ön çalışmanın sınırlılıkları vardır. Öncelikle kapsam oldukça geniş tutulmuştur, ileride yapılacak araştırmalar daha dar kapsamlı ve dolayısıyla daha detaylı bilgilere ulaşabilmek amacıyla yola çıkabilir, örneğin sadece AFA ve rotasyon seçenekleri (ör. Kline, 2015; Osborne, 2015; Saracli, 2011), DFA ve model modifikasyonları (ör. Asparouhov & Muthen, 2009; Mueller & Hancock, 2008) tekrarlanabilirlik çerçevesinden incelenebilir. İkinci sınırlılık model-veri uyumlarının en azından örneklem büyüklüğü, tahminleyici ve kullanılan faktör modelinin özelliklerinden etkilenebildiğidir. Fakat bu ön çalışmada model-veri uyumu tekrarlanabilirlik için asıl kanıt olarak kullanılmamıştır, model-veri uyumsuzluğu tekrarlanabilirliği zedeleyebilecek bir ipucu olarak ele alınmıştır. Üçüncü sınırlılık ise bu ön çalışma sonunda, uygulama yapacakların tekrarlanabilirliği artırmak amacıyla kullanabileceği bir kontrol listesi oluşturulmamıştır. Fakat tekrarlanabilirliğin açık ve net bir evren tanımı olmadan, örneklemenin tam olarak nasıl tamamlandığı ve kayıp verilerin nasıl ele alındığı belirtilmeden ve çözümlenme sürecinin ayrıntıları raporlanmadan mümkün olmayacağına dikkat çekilmek istenmiştir. Bu uyarıları dikkate alacak araştırmacılar, faktör analizlerinde dikkat edilmesi gereken diğer konular için en azından Büyüköztürk (2002), Erkuş (2016), Kline (2015), Öztürk, Eroğlu and Kelecioğlu (2015), Prudon (2015), Worthington and Whittaker (2006) ve Wright (2017) tarafından kaleme alınan çalışmalardan faydalanabilir. Mümkün olan her durumda araştırmacıların kullandıkları veri setlerini ve çözümlenme betiklerini erişime açmaları, tekrarlanabilirlik çalışmalarını kolaylaştıracaktır. Bu ön çalışmada incelenen 275 makalenin listesi ve toplam 448 analizin bilgilerini içeren veri seti ek dosya olarak paylaşılmıştır.

KAYNAKLAR

- Acar, T. (2014). Ölçek geliştirmede geçerlik kanıtları: Çapraz geçerlik, sınıflama ve sıralama geçerliği uygulaması. *Kuram ve Uygulamada Eğitim Bilimleri*, 14(2), 1-11. DOI: 10.12738/estp.2014.3.2107
- Akbaş, U., & Tavşancıl, E. (2015). Farklı Örneklem büyüklüklerinde ve kayıp veri örüntülerinde ölçeklerin psikometrik özelliklerinin kayıp veri baş etme teknikleri ile incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(1) 38-57.
- Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology*, 112, 545-557.
- Anderson, C. J., Bahník, Š., Barnett-Cowan, M., Bosco, F. A., Chandler, J., Chartier, C. R., ... & Della, N. P. (2016). Response to Comment on "Estimating the reproducibility of psychological science". *Science (New York, NY)*, 351(6277), 1037-1037.
- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural equation modeling: a multidisciplinary journal*, 16(3), 397-438.
- Atılğan, H., Kan, A., & Aydın, B. (2017). *Eğitimde Ölçme ve Değerlendirme*. Ankara: Anı Yayıncılık.
- Balcı, A. (2015) *Sosyal Bilimlerde Araştırma: Yöntem, Teknik ve İlkeler*. Ankara: Pegem Yayınları.
- Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling*, 13(2), 186-203. DOI: 10.1207/s15328007sem1302_2
- Bornstein, M. H., Jager, J., & Putnick, D. L. (2013). Sampling in developmental science: Situations, shortcomings, solutions, and standards. *Developmental Review*, 33(4), 357-370. DOI: 10.1016/j.dr.2013.08.003
- Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99-136). Washington, DC, US: American Psychological Association.
- Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Kuram ve Uygulamada Eğitim Yönetimi*, 32, 470-483.

- Crocker, L. M., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart, Winston.
- Çüm, S., & Gelbal, S. (2015). Kayıp veriler yerine yaklaşık değer atamada kullanılan farklı yöntemlerin model veri uyumu Üzerindeki etkisi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 1(35), 87-111.
- Çüm, S., & Koç, N. (2013). Türkiye’de psikoloji ve eğitim bilimleri dergilerinde yayımlanan ölçek geliştirme ve uyarlama çalışmalarının incelenmesi. *Journal of Educational Sciences & Practices*, 12(24), 115-135.
- Delice, A. (2010). Nicel araştırmalarda örneklem sorunu. *Kuram ve Uygulamada Eğitim Bilimleri*, 10(4), 1969-2018.
- Demir, E., & Parlak, B. (2012). Türkiye’de eğitim araştırmalarında kayıp veri sorunu. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 3(1), 230-241.
- Erkuş, A. (2016). Ölçek geliştirme ve uyarlama çalışmalarındaki sorunlar ile yazım ve değerlendirilmesi. *Pegem Atf İndeksi*, 1211-1224.
- Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction*. White Plains, NY: Longman USA.
- Gaskin, C. J., Orellana, L., Bowe, S. J., & Lambert, S. D. (2017). Why sample selection matters in exploratory factor analysis: Implications for the 12-item world health organization disability assessment schedule 2.0. *BMC Medical Research Methodology*, 17(1), 40.
- Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on estimating the reproducibility of psychological science. *Science*, 351(6277), 1037-1037.
- Gökmen, Ö. F., Uysal, M., Yaşar, H., Kirksekiz, A., Güvendi, G. M., & Horzum, M. B. (2017). Türkiye’de 2005-2014 yılları arasında yayınlanan uzaktan eğitim tezlerindeki yöntemsel eğilimler: Bir İçerik analizi. *Eğitim ve Bilim*, 42(189), 1-25.
- Göktaş, Y., Küçük, S., Aydemir, M., Telli, E., Arpacık, Ö., Yıldırım, G., & Reisoğlu, I. (2012). Türkiye’de eğitim teknolojileri araştırmalarındaki eğilimler: 2000-2009 dönemi makalelerinin içerik analizi. *Kuram ve Uygulamada Eğitim Bilimleri Dergisi*, 12(1), 177-199.
- Gül, Ş., & Sözbilir, M. (2015). Fen ve matematik eğitimi alanında gerçekleştirilen Ölçek geliştirme araştırmalarına yönelik tematik içerik analizi. *Eğitim ve Bilim*, 40(178), 85-102.
- Güvendir, M. A., & Özkan, Y. Ö. (2015). Türkiye’deki eğitim alanında yayımlanan bilimsel dergilerde ölçek geliştirme ve uyarlama konulu makalelerin incelenmesi. *Elektronik Sosyal Bilimler Dergisi*, 14(52), 23-33.
- Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M., & Mumford, K. R. (2005). The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination. *Educational and Psychological Measurement*, 65(2), 202-226. DOI: 10.1177/0013164404267287
- Hox, J. J. (1995). Amos, EQS, and Lisrel for windows: A comparative review. *Structural Equation Modeling: A Multidisciplinary Journal*, 2(1), 79-91.
- Karadağ, E. (2011). Eğitim bilimleri doktora tezlerinde kullanılan Ölçme araçlar: Nitelik düzeyleri ve analitik hata tipleri. *Kuram ve Uygulamada Eğitim Bilimleri*, 11(1), 311-334.
- Kish, L. (1965). *Survey sampling*. Oxford, England: John Wiley & Sons.
- Kline, R. B. (2015). *Principles and practice of structural equation modeling*. New York, NY, US: Guilford publications.
- Kozikoğlu, I., & Senemoğlu, N. (2016). Eğitim programları ve Öğretim alanında yapılan doktora tezlerinin içerik analizi (2009-2014). *Eğitim ve Bilim*, 40(182), 29-41.
- Kürşad, M. Ş., & Nartgün, Z. (2015). Kayıp veri sorununun çözümünde kullanılan farklı yöntemlerin Ölçeklerin geçerlik ve güvenilirliği bağlamında karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(2), 254-267. DOI: 10.21031/epod.95917
- Landers, R. N., & Behrend, T. S. (2015). An inconvenient truth: Arbitrary distinctions between organizational, mechanical turk, and other convenience samples. *Industrial and Organizational Psychology*, 8(2), 142-164.
- Lavrakas, P. J. (2008). *Encyclopedia of survey research methods*. Thousand Oaks: Sage Publications.
- Levy, P. S., & Lemeshow, S. (2013). *Sampling of populations: Methods and applications*. John Wiley & Sons.
- Lewis-Beck, M. S., Bryman, A. , & Liao, T. F. (2004). *The Sage Encyclopedia of social science research methods*. Thousand Oaks: Sage.
- Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48(3), 936-949. DOI: 10.3758/s13428-015-0619-7.
- Lord, F. M., and Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, Mass.: Addison-Wesley

- Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, *103*(3), 391-410. DOI: 10.1037/0033-2909.103.3.391
- Moshontz, H., Campbell, L., Ebersole, C. R., Iljerman, H., Urry, H. L., Forscher, P. S., ... & Castille, C. M. (2018). The Psychological Science Accelerator: Advancing psychology through a distributed collaborative network. *Advances in Methods and Practices in Psychological Science*, *1*(4), 501-515. DOI: 10.1177/2515245918797607
- Mueller, R. O., & Hancock, G. R. (2008). Best practices in structural equation modeling. In J. W. Osborne (Ed.), *Best practices in quantitative methods* (pp. 488-508). Thousand Oaks: Sage Publications Inc.
- Neuman, W. L. (2013). *Social research methods: Qualitative and quantitative approaches*. UK: Pearson education.
- Nunnally, J., & Bernstein, I. (1994). *Psychometric Theory (3rd Edition)*. New York: McGraw-Hill, Inc.
- O'Boyle, E. H., Jr., & Williams, L. J. (2011). Decomposing model fit: Measurement vs. theory in organizational research using latent variables. *Journal of Applied Psychology*, *96*(1), 1-12.
- Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). doi: 10.1126/science.aac4716
- Osborne, J. W. (2015). What is rotating in exploratory factor analysis. *Practical assessment, research & evaluation*, *20*(2), 1-7.
- Öztürk, N. B., Eroğlu, M. G., & Kelecioğlu, H. (2015). Eğitim alanında yapılan ölçek uyarlama makalelerinin incelenmesi. *Eğitim ve Bilim*, *40*(178), 123-137.
- Peterson, R. A., & Merunka, D. R. (2014). Convenience samples of college students and research reproducibility. *Journal of Business Research*, *67*(5), 1035-1041. DOI: 10.1016/j.jbusres.2013.08.010
- Prudon, P. (2015). Confirmatory factor analysis as a tool in research using questionnaires: A critique. *Comprehensive Psychology*, *4*, 1-19. DOI: 10.2466/03.CP.4.10
- Ray, J. J. (1990). Acquiescence and problems with forced-choice scales. *The Journal of Social Psychology*, *130*(3), 397-399. DOI: 10.1080/00224545.1990.9924595
- Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, *17*(3), 354-373. DOI: 10.1037/a0029315
- Saracli, S. (2011). Faktör analizinde yer alan döndürme metotlarının karşılaştırmalı incelenmesi üzerine bir uygulama. *Düzce Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi*, *1*(3), 22-26.
- Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. Chapman and Hall: CRC.
- Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology*, *57*(1), 1-10.
- Selçuk, Z., Palancı, M., Kandemir, M., & Dündar, H. (2014). Eğitim ve bilim dergisinde yayınlanan araştırmaların eğilimleri: İçerik analizi. *Eğitim ve Bilim*, *39*(173), 428-449.
- Simon, A. (1979). Effects of selective sampling on a factor analysis. *The Journal of General Psychology*, *101*(2), 259-264. DOI: 10.1080/00221309.1979.9920079
- Smith, T. (1983). On the validity of inferences from non-random sample. *Journal of the Royal Statistical Society. Series A (General)*, *146*(4), 394-403. DOI: 10.2307/2981454
- Sözbilir, M., Güler, G., & Çiltaş, A. (2012). Türkiye'de matematik eğitimi araştırmaları: Bir içerik analizi Çalışması. *Kuram ve Uygulamada Eğitim Bilimleri*, *12*(1), 565-580.
- Stapleton, L. M., & Leite, W. L. (2005). Teacher's corner: A review of syllabi for a sample of structural equation modeling courses. *Structural Equation Modeling*, *12*(4), 642-664.
- Sterba, S. K. (2009). Alternative model-based and design-based frameworks for inference from samples to populations: From polarization to integration. *Multivariate Behavioral Research*, *44*(6), 711-740.
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.
- Tyrer, S., & Heyman, B. (2016). Sampling in epidemiological research: Issues, hazards and pitfalls. *BJPsych Bulletin*, *40*(2), 57-60.
- Van Prooijen, J. W., & Van Der Kloot, W. A. (2001). Confirmatory analysis of exploratively obtained factor structures. *Educational and Psychological Measurement*, *61*(5), 777-792. DOI: 10.1177/00131640121971518
- Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, *73*(6), 913-934.
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, *34*(6), 806-838.

- Wright, A. G. (2017). The current state and future of factor analysis in personality disorder research. *Personality Disorders: Theory, Research, and Treatment*, 8(1), 14-25.
- Xiao, Y., Liu, H., & Li, H. (2017). Integration of the Forced-Choice Questionnaire and the Likert Scale: A Simulation Study. *Frontiers in Psychology*, 8, 806. DOI: 10.3389/fpsyg.2017.00806
- Yalçın, S., Yavuz, H. C., & Dibek, M. I. (2016). En yüksek etki faktörüne sahip eğitim dergilerindeki makalelerin İçerik analizi. *Eğitim ve Bilim*, 40(182), 1-28.
- Yang-Wallentin, F., Jöreskog, K. G., & Luo, H. (2010). Confirmatory factor analysis of ordinal variables with misspecified models. *Structural Equation Modeling*, 17(3), 392-423.
- Yılmaz, K., & Altinkurt, Y. (2012). An examination of articles published on preschool education in Turkey. *Educational Sciences: Theory and Practice*, 12(4), 3227-3241.
- Yurdugül, H., & Bayrak, F. (2012). Ölçek geliştirme Çalışmalarında kapsam geçerlik Ölçüleri: Kapsam geçerlik indeksi ve kappa istatistiğinin karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, Özel Sayı*, 2, 264-271.