

FAKTÖR ÇIKARMA YÖNTEMLERİNİN PARALEL ANALİZ SONUÇLARINA ETKİSİ

THE EFFECT OF FACTOR EXTRACTION METHODS ON THE PARALLEL ANALYSIS RESULTS

Abdullah Faruk KILIÇ¹, İbrahim UYSAL²

ÖZ: Açımlayıcı faktör analizinden en önemli kararlardan biri faktör sayısını belirlemektir. Bunun için birçok yöntem geliştirilmiş olsa da paralel analiz, hala en çok önerilen ve kullanılan yöntemler arasında yer almaktadır. Sıklıkla kullanılan birçok modifikasyonun yapılmasına ve analizle ilgili araştırmaların yoğunlaşmasına neden olmuştur. Bu araştırmada açımlayıcı faktör analizinde kullanılan faktör çıkarma yöntemlerinin paralel analiz sonuçlarına etkisinin incelenmesi amaçlanmıştır ve Monte Carlo simülasyon çalışması gerçekleştirilmiştir. İki kategorili veri setleriyle gerçekleştirilen simülasyon çalışmasında ortalama faktör yükü, madde sayısı, ölçme modeli, örneklem büyüklüğü ve kullanılan korelasyon matrisi koşulları manipüle edilmiştir. Paralel analizde uygulanan en küçük kalıntı, temel bileşenler, temel eksenler, en çok olabilirlik, ağırlıklandırılmamış en küçük kareler, en küçük kare ve optimal paralel analizde uygulanan en küçük rank faktör çıkarma yöntemleri karşılaştırılmıştır. Araştırma sonucunda tetrakorik korelasyon matrisiyle gerçekleştirilen optimal paralel analiz yönteminin uygulandığı en küçük rank yönteminin en iyi sonucu verdiği gözlemlenmiştir. Bununla birlikte ortalama faktör yükü .70 olan koşullarda Pearson korelasyon matrisiyle gerçekleştirilen analizlerde tüm yöntemler yeterli performans gösterirken tetrakorik korelasyon matrisinin kullanılmasıyla paralel analiz uygulanan temel bileşenler ve optimal paralel analiz uygulanan en küçük rank yöntemleri hariç diğer yöntemlerin çok sayıda faktör çıkardığı söylenebilir. Araştırma bulgularına göre tetrakorik (polikorik) korelasyon matrisiyle en küçük rank yöntemiyle optimal paralel analizin kullanılması önerilmektedir.

Anahtar sözcükler: Açımlayıcı faktör analizi, optimal paralel analiz, faktör sayısını belirleme, iki kategorili veri

ABSTRACT: One of the most important issues regarding exploratory factor analysis is determining the number of factors. Although many methods have been developed for this, parallel analysis is still among the most widely recommended and used. Its frequent use has led to many modifications and intensive method studies. In this study, the aim was to investigate the effect of factor extraction methods used in exploratory factor analysis on the results of parallel analysis. In a Monte Carlo simulation study carried out with binary data sets, the average factor loading, number of items, measurement model, sample size and correlation matrix conditions were manipulated. The minimum residual principal component, principal axis, maximum likelihood unweighted least squares and minimum chi-square used in the parallel analysis minimum rank factor extraction method used in optimal parallel analysis were compared. As a result of the research, it was observed that the minimum rank method with the optimal parallel analysis performed with tetrachoric correlation matrix gave the best results. In this analysis performed with the Pearson correlation matrix at an average factor loading of .70, all methods showed sufficient performance. It can be said that other methods except for the principal components and minimum rank used in optimal parallel analysis, have overfactoring results in this condition where the correlation matrix is tetrachoric. According to the research findings, use of optimal parallel analysis with the minimum rank method with a tetrachoric (polychoric) correlation matrix can be recommended.

Keywords: Exploratory factor analysis, optimal parallel analysis, parallel analysis, factor retention, binary data

Bu makaleye atf vermek için:

Kılıç, A. F. ve Uysal, İ. (2021). Faktör Çıkarma Yöntemlerinin Paralel Analiz Sonuçlarına Etkisi, *Trakya Eğitim Dergisi*, 11(2), 926-942

Cite this article as:

Kilic, A. F. ve Uysal, I. (2021). The effect of factor extraction methods on the parallel analysis results, *Trakya Journal of Education*, 11(2), 926-942

¹ Dr. Arş. Gör., Adıyaman Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, abduhfarukkilic@gmail.com, ORCID: 0000-0003-3129-1763

² Dr. Arş. Gör., Bolu Abant İzzet Baysal Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, ibrahimuysal06@gmail.com, ORCID: 0000-0002-6767-0362

EXTENDED ABSTRACT

Introduction

Factor analysis is one of the methods employed to obtain fewer variables by using relationships between the variables. Deciding on the number of factors is one of the most important issues in exploratory factor analysis, which is frequently used in social and educational science to examine constructs. Many methods have been developed to decide on the number of factors in factor analysis, including the K1 rule proposed by Kaiser (1960), the scree plot developed by Cattell (1966), the minimum average partial (MAP) test developed by Velicer (1976) and parallel analysis developed by Horn (1965).

In parallel analysis as developed by Horn (1965), random data sets are produced in line with the number of variables and observations in the data set whose dimensions are being investigated. The eigenvalues obtained from the correlation matrices produced are compared with the eigenvalues obtained from the actual data. When comparing eigenvalues, the threshold point of 95% eigenvalue is often used (Timmerman & Lorenzo-Seva, 2011).

In traditional parallel analysis, the number of variables and the sample size in the real data set are taken into account. Timmerman and Lorenzo-Seva (2011) suggested that random data should be produced using eigenvalue sampling distribution with a permutation of the raw data with the minimum rank factor (minrank) extraction method. It was reported that parallel analysis gave more accurate results, especially in the analysis of data sets using polychoric correlations. They called this method “optimal parallel analysis”.

In this study, parallel analysis results obtained with factor extraction methods (maximum likelihood, minimum chi-square, minimum residual, principal component, principal axis, unweighted least squares) and optimal parallel analysis results with the minimum rank factor extraction method were compared. The research question was: “How do factor extraction methods affect the results of parallel analysis?” Accordingly, the following subproblems were created.

As a result of the parallel analysis performed with minimum residual, maximum likelihood, principal components, principal axes, unweighted least squares, minimum chi-square and optimal parallel analysis performed with minimum rank, which estimation methods are more accurate:

- 1) the percentage of correct estimates results?
- 2) the bias values?

Method

A Monte Carlo simulation study was carried out to examine the effect of factor extraction methods on parallel analysis results. In Monte Carlo simulation studies, sampling data suitable for the determined distribution are produced (Bandalos & Leite, 2013). In this study, binary data sets with normal distribution were produced.

In the research, test length (20 and 30 items), mean factor loading (.40 and .70), measurement model (single factor, two factors [interfactor correlation = .30]), sample size (200, 500 and 1000) and correlation matrix (Pearson and tetrachoric) simulation conditions were manipulated. Accordingly, $2 \times 2 \times 2 \times 3 \times 2 = 48$ simulation conditions were studied and 100 replications used for each condition. Since the focus of the research is achievement tests, data sets were made binary (1-0). For this, firstly data sets were generated as continuous and followed multivariate normal distribution. Then the data sets were categorized (1-0). When categorizing data sets, {0} was used as the threshold point.

The percentage of correct estimates (PCE) and bias values were used to evaluate the research results. The PCE was calculated by comparing the suggested number of factors and the real number of factors. In other words, the percentage of the data set in which the suggested number of factors and the real number of factors were equal was considered as the PCE.

Bias value was calculated by subtracting the real number of factors from the suggested number of factors. Accordingly (average suggested number of factors - real number of factors), the real number of factors equation was expressed as bias (Bandalos & Leite, 2013). Since the bias value is considered acceptable in the literature if it is less than .10 (Flora & Curran, 2004; Forero, Maydeu-Olivares & Gallardo-Pujol, 2009; Moshagen & Musch, 2014; Rhemtulla, Brosseau-Liard & Savalei, 2012), in the current study, bias values of less than .10 were considered acceptable.

In the research, simulation data were produced with the lavaan (Rosseel, 2012) package in R software (R Core Team, 2018). EFA.MRFA (Navarro-Gonzalez & Lorenzo-Seva, 2020) has been used in the analysis of the produced data sets with the optimal parallel analysis method and the psych (Revelle, 2018) package has been used in the analysis of the produced data sets with the parallel analysis method. The data of the

correlation matrices, while performing parallel analysis in other methods (minimum residual, maximum likelihood, principal components, principal axes, unweighted least squares, minimum chi-square) except for the minimum rank with optimal parallel analysis, are produced in accordance with the multivariate normal distribution. However, obtaining sample data from permutation affects parallel analysis results (Timmerman & Lorenzo-Seva, 2011). Therefore, other factor extraction methods (minimum residual, maximum likelihood, principal components, principal axes, unweighted least squares, minimum chi-square) are compared with the minimum rank with optimal parallel analysis.

Findings

It was observed that all factor extraction methods except for maximum likelihood (ML) performed with the Pearson correlation matrix had sufficient PCE (> 90%) in all conditions where the average factor loading was .70. Factor extraction methods except for principal component analysis (PCA) generally had a PCE value of less than 90% in conditions where the average factor loading was .40 and the correlation matrix used was Pearson. When the Pearson correlation matrix was used, the PCA had sufficient PCE value in all other conditions except for one. Only PCA (except for one condition) had a sufficient PCE value in the simulation conditions where the measurement model was two-dimensional, the average factor loading was .40 and the correlation matrix used was Pearson. In addition, it can be said that all other factor extraction methods except for ML had acceptable bias (< .10) in all conditions where the average factor loading was .70. In conditions where the average factor loading was .40, and the correlation matrix used was Pearson, all other factor extraction methods, except for PCA, generally had a bias value greater than .10. PCA has acceptable bias in all of these conditions. In conditions where the test length was 20 items, the measurement model was unidimensional, the average factor loading was .40 and the sample size was 200, the bias value of other methods, except for the minimum rank with optimal parallel analysis and ML with parallel analysis, were acceptable. However, in conditions where the measurement model was two-dimensional, the sample size was 1000 and the number of items was 20, the bias value of the PCA with parallel analysis and minimum rank with optimal parallel analysis methods were acceptable. For the same conditions, but where the number of items increased to 30, the bias of the PCA and unweighted least squares (ULS) methods was acceptable (< .10).

When analysed with a tetrachoric correlation matrix, it can be said that the minimum rank method with optimal parallel analysis had sufficient PCE in all simulation conditions. On the other hand, it was observed that principal component analysis (PCA) with parallel analysis was superior in terms of PCE compared to other methods. However, in contrast to the analyses using the Pearson correlation matrix, in the analysis performed with the tetrachoric correlation matrix, the PCA did not have sufficient PCE in most conditions where the average factor loading was .40. Except for PCA with parallel analysis and minimum rank with optimal parallel analysis, other factor extraction methods did not have sufficient PCE (< 90%) in the analysis performed with the tetrachoric correlation matrix. When the minimum rank with optimal parallel analysis results were examined, it was observed that two-dimensional 20-item tests with an average factor loading of .40 had a PCE of less than 90% with sample sizes of 200. In other words, minimum rank with optimal parallel analysis had sufficient PCE in all other conditions except for one in the analysis performed with tetrachoric correlation. In addition, it can be said that the minimum rank method with optimal parallel analysis had acceptable bias under all simulation conditions. In the analysis performed with the tetrachoric correlation matrix, there were conditions in which other methods except PCA had very high bias. The PCA method had acceptable bias in all conditions with an average factor loading of .70 in analyses performed with the tetrachoric correlation matrix. In the analysis performed with the Pearson correlation matrix, the y-axis showing the bias value was in the range -.20 to .60, while in the analysis performed with the tetrachoric correlation matrix, the bias of the y-axis was in the range 0 to 15.

Discussion and Conclusion

This research was conducted to investigate the effect of factor extraction methods on parallel analysis results. As a result of the study, it was observed that in the analyses using the Pearson correlation matrix, all factor extraction methods, except for principal component analysis (PCA), failed to accurately predict the number of factors under conditions where the average factor loading was low (.40). Similarly, in conditions where the average factor loading was low, all methods except for PCA estimated the number of factors to be larger than the real number of factors. It can be said that the sample size did not have much of an effect on the results of parallel analysis in the analyses performed with the Pearson correlation matrix. This finding is similar to that in the research performed by Weng and Cheng (2005).

In the analysis performed with the tetrachoric correlation matrix, it can be said that the minimum rank method with optimal parallel analysis estimated the number of factors more accurately than the other methods. All of the methods except for PCA with parallel analysis and minimum rank with optimal parallel analysis estimated the number of factors to be much higher the real number of factors in conditions where the correlation matrix used was tetrachoric. This situation is similar to the study conducted by Timmerman and Lorenzo-Seva (2011). The minimum rank with optimal parallel analysis performed with the polychoric correlation matrix showed sufficient performance for most of the simulation conditions. However, there may be convergence problems in small samples due to difficulties in calculating the polychoric correlation. In this study, simple smoothing in the psych packet (Revelle, 2018) was used when it was necessary to calculate the small sample size for tetrachoric correlation matrix. The sample size did not have a big effect on the performance of factor extraction methods in the analyses performed with the tetrachoric correlation matrix. Increasing the sample size did not affect the PCE values.

This research is limited to simulation conditions. However, according to the research findings, the use of minimum rank with optimal parallel analysis can be recommended in analysis performed with the tetrachoric correlation matrix in binary data sets. Minimum rank with optimal parallel analysis can also be performed with the Factor 10.10 (Lorenzo-Seva & Ferrando, 2019) package program. Researchers who have difficulty in using R software can be recommended to use this free package program. If the data set is small and the tetrachoric correlation matrix cannot be calculated, then analysis can be performed using the Pearson correlation matrix employing the principal component analysis method with parallel analysis. In this study, 95% can be recommended as the threshold point to decide on the factor number, since the value corresponding to 95% is used instead of the average of the eigenvalues.

The number of factors and interfactor correlation conditions in this study was limited. In future research, more comprehensive simulation conditions can be studied. In addition, the effect of the situation on the parallel analysis results, such as changing the number of categories of the observed variables in the data sets and the distribution of the variables, can be examined.

GİRİŞ

Faktör analizi, çok sayıdaki değişken arasındaki ilişkilerden yararlanılarak daha az sayıda değişken elde etme yöntemlerinden biridir. Hem yapı geçerliğine yönelik kanıt toplama sürecinde hem de değişkenlerin oluşturduğu yapıyı incelemek için sosyal bilimler ve eğitim bilimlerinde sıklıkla kullanılan faktör analizinde en önemli konulardan biri de faktör sayısına karar verebilmektir.

Faktör analizinde faktör sayısına karar verebilmek için birçok yöntem geliştirilmiştir. Örneğin Kaiser (1960) tarafından önerilen K1 kuralı, özdeğeri 1'den büyük olan değişken sayısını önemli faktör sayısı olarak değerlendirmektedir. Cattell (1966) ise geliştirdiği yamaç grafiği yöntemiyle özdeğerlerde meydana gelen ani düşüşleri incelemiş ve grafikte ani düşüş sayısını faktör sayısı olarak tanımlamıştır. Velicer (1976) tarafından geliştirilen kısmi ortalamaların en küçüğü (Minimum Average Partial [MAP]) testinde ise temel bileşenler analizi ile ilk faktör çıkarıldıktan sonra köşegen dışında kalan elemanların kareleri hesaplanır. Böylece kısmi korelasyon matrisi oluşmuş olur. İkinci aşamada oluşan korelasyon matrisi ile yine temel bileşenler analizi gerçekleştirilerek faktör çıkarılır. Köşegen dışında kalan elemanların kareleri alınarak işleme devam edilir. İşlemler, orijinal korelasyon matrisindeki elemanların karelerinin ortalamalarının en küçük kısmi ortalamalardan daha küçük olduğunda durdurulur. Faktör sayısı ise işlemlerin durdurulduğu adım olarak elde edilir (O'Connor, 2000).

Horn (1965) tarafından geliştirilen paralel analizde ise boyutluluğu araştırılan veri setindeki değişken ve gözlem sayısı kadar rassal veri seti üretilmektedir. Örneğin boyutluluğu araştırılan veri setinde 10 değişken ve 233 gözlem var ise rassal olarak üretilen korelasyon matrisleri de 10 değişken ve 233 gözlem üzerinden elde edilmektedir. Üretilen korelasyon matrislerinden elde edilen özdeğerler gerçek veriden elde edilen özdeğerler ile kıyaslanmaktadır. Bu kıyaslamada eşik değeri için Horn (1965) rassal olarak üretilen matrislerden elde edilen özdeğerlerin ortalamasını almayı önermiştir. Ancak güncel alanyazında ortalama değer faktör sayısını olduğundan daha yüksek kestirmesi nedeniyle daha çok %95 ya da %99'luk özdeğer ile kıyaslama gerçekleştirilmektedir (Buja ve Eyuboglu, 1992; Glorfeld, 1995; Timmerman ve Lorenzo-Seva, 2011).

Horn (1965) tarafından geliştirilen paralel analiz, temel bileşenler analiziyle gerçekleştirilmektedir. Buna göre rassal dağılımdan elde edilen özdeğerler ile ampirik veriden elde edilen özdeğerler kıyaslanır. Boyut sayısı da bu kıyaslama ile elde edilir. Buna göre ampirik olarak elde edilen özdeğerlerin rassal olarak elde edilen özdeğerlerden büyük olduğu sayı önerilen faktör sayısını verir (Timmerman ve

Lorenzo-Seva, 2011). Örneğin rassal veriden elde edilen özdeğerler ampirik olarak elde edilen özdeğerlerle karşılaştırıldığında üçüncü özdeğerden sonra rassal olarak elde edilen özdeğer, ampirik olarak elde edilenden daha büyükse önerilen faktör sayısı üç olacaktır.

Geleneksel paralel analizde gerçek veri setindeki değişken sayısı ve örneklem büyüklüğü dikkate alınarak korelasyon matrisleri üretilmekteyken Timmerman ve Lorenzo-Seva (2011) en küçük rank faktör çıkarma yöntemiyle birlikte permütasyona dayalı örnekleme yöntemi kullanılarak rassal verilerin üretilmesini önermiştir. Bu şekilde özellikle polikorik korelasyonların kullanıldığı veri setlerinde gerçekleştirilen analizlerde paralel analizin daha uygun sonuçlar verdiği raporlanmış olup bu yöntemle optimal paralel analiz adını vermişlerdir. Bu çalışmada faktör çıkarma yöntemleriyle (en küçük kalıntı, temel bileşenler, temel eksenler, en çok olabilirlik, ağırlıklandırılmamış en küçük kareler, en küçük ki-kare) gerçekleştirilen paralel analiz sonuçlarıyla birlikte en küçük rank yöntemiyle gerçekleştirilen optimal paralel analiz sonuçları incelenmiştir.

Temel bileşenler analizinin sürekli veri setleri için geliştirilmesi, faktör çıkarma yönteminin ve veri tipinin paralel analiz sonuçları üzerindeki etkisinin incelenmesini önemli hale getirmiştir. Bu nedenle Timmerman ve Lorenzo-Seva (2011) tarafından paralel analizin orijinal hali olan temel bileşenler yöntemi, temel eksenler ve en küçük rank faktör çıkarma yöntemleri kullanılarak paralel analiz sonuçları karşılaştırılmıştır. Araştırma sonucunda temel eksenler faktör çıkarma yönteminin aşırı faktör çıkardığı (over-factoring) raporlanmıştır. Ayrıca Horn (1965) tarafından önerilen paralel analiz ile permütasyona dayalı örnekleme yönteminin kullanıldığı en küçük rank faktör çıkarma yönteminin (optimal paralel analiz) benzer sonuçlar verdiği belirtilmiştir. Bu çalışmanın yanında Crawford ve diğerleri (2010) ve Steger (2006) de faktör çıkarma yöntemlerinin paralel analiz sonuçlarına etkisini incelemiştir.

Crawford ve diğerleri (2010) tarafından sürekli veri setleriyle gerçekleştirilen çalışmada, temel bileşenler (TBA) ve temel eksenler faktör çıkarma yöntemleri kullanılarak paralel analiz sonuçları karşılaştırılmıştır. Bu yöntemlerle birlikte eşik noktası olarak paralel analizde üretilen korelasyon matrislerinden elde edilen özdeğerlerin ortalamasının ve %95'lik değerlerinin kullanılması durumlarının da karşılaştırılması sağlanmıştır. Araştırma sonucunda TBA'nın tek boyutlu veya iki boyutlu fakat boyutlar arasındaki korelasyonun düşük olduğu modellerde temel eksenler kadar ya da ondan daha iyi sonuçlar verdiği belirtilmiştir. Özdeğerlerin ortalaması veya %95'lik değerinden elde edilen sonuçların faktör başına düşen değişkenle ilgili olduğunu vurgulamıştır.

Steger (2006) ise çok kategorili (1-5, 1-10) gerçek uygulama verisiyle gerçekleştirdiği çalışmada, TBA ile temel eksenler faktör çıkarma yöntemlerini K1 kuralı (Kaiser, 1960), yamaç grafiği ve geleneksel paralel analiz sonuçları üzerinde karşılaştırmıştır. Araştırma sonucunda TBA ve temel eksenler yöntemleriyle gerçekleştirilen paralel analizin, yamaç grafiği yönteminden daha fazla faktör çıkardığı ve yamaç grafiğinden elde edilen sonucun daha uygun olduğu raporlanmıştır.

Diğer taraftan alanyazın incelendiğinde, paralel analizin açılımlayıcı faktör analizi (AFA) çalışmalarında sıklıkla kullanıldığı ve diğer faktör sayısını belirleme yöntemleriyle karşılaştırıldığında daha iyi performans sergilediği gözlenmiştir (Dinno, 2009; Glorfeld, 1995; Zwick ve Velicer, 1986). Örneğin Goretzko, Pham ve Bühner (2019) tarafından gerçekleştirilen çalışmada 2007-2017 yılları arasında AFA gerçekleştirilen çalışmalar incelenmiş ve çalışmaların %42.1'inde paralel analizin kullanıldığı raporlanmıştır. Bu nedenle paralel analizin farklı faktör çıkarma yöntemlerinde nasıl sonuçlar verdiğinin incelenmesi de önemli hale gelmektedir.

Alanyazında paralel analizle ilgili gerçekleştirilen çalışmalar incelendiğinde; genellikle diğer faktör sayısını belirleme yöntemleriyle karşılaştırıldığı söylenebilir (Çokluk ve Koçak, 2016; Green, Xu ve Thompson, 2018; Keith, Caemmerer ve Reynolds, 2016; Koçak, Çokluk ve Kayri, 2016; Weng ve Cheng, 2005). Bunun yanında paralel analize yeni yaklaşımların getirildiği çalışmalar da mevcuttur (Glorfeld, 1995; Liu ve Rijmen, 2008; Preacher, Zhang, Kim ve Mels, 2013; Ruscio ve Roche, 2012; Xia, Green, Xu ve Thompson, 2018). Ancak faktör çıkarma yöntemlerinin karşılaştırıldığı çalışmaların genellikle temel eksenler ve temel bileşenler yöntemleriyle gerçekleştirildiği gözlemlendiğinden (Crawford ve diğerleri, 2010; Steger, 2006; Timmerman ve Lorenzo-Seva, 2011) bu çalışmada diğer faktör çıkarma yöntemlerinden elde edilen sonuçlar da incelenmiştir. Araştırma bu yönüyle alanyazındaki diğer araştırmalardan farklılaşmaktadır. Bu çalışmada en küçük kalıntı, en çok olabilirlik, temel bileşenler, temel eksenler, ağırlıklandırılmamış en küçük kareler, en küçük ki-kare yöntemleriyle gerçekleştirilen paralel analiz ile en küçük rank yöntemiyle gerçekleştirilen optimal paralel analiz sonuçları karşılaştırılmıştır. Bu nedenle bu çalışmanın alanyazındaki bu boşluğu kısmen de olsa doldurabileceği ve araştırmacılara önerilerde bulunabileceğinden önemli olduğu düşünülmektedir.

Araştırmanın problem cümlesi; "Faktör çıkarma yöntemleri paralel analiz sonuçlarını nasıl etkilemektedir?" şeklinde oluşturulmuş olup buna göre şu alt problemlere yanıt aranmıştır:

Gerçekleştirilen paralel analiz sonucunda en küçük kalıntı, en çok olabilirlik, temel bileşenler, temel eksenler, ağırlıklandırılmamış en küçük kareler, en küçük ki-kare ve optimal paralel analiz sonucunda en küçük rank yöntemlerinden elde edilen;

- 1) doğru kestirim yüzdesi sonuçları nasıldır?
- 2) yanlılık değerleri nasıldır?

YÖNTEM

Faktör çıkarma yöntemlerinin paralel analiz sonuçlarına etkisinin incelenmesi amacıyla Monte Carlo simülasyon çalışması gerçekleştirilmiştir. Monte Carlo simülasyon çalışmalarında belli dağılımlara uygun olacak şekilde örneklem verisi üretilmektedir (Bandalos ve Leite, 2013). Bu çalışmada da normal dağılım gösteren iki kategorili veri setleri üretilmiştir.

Faktör Çıkarma Yöntemleri

Açımlayıcı faktör analizi için geliştirilmiş birçok faktör çıkarma yöntemi bulunmaktadır. Temel eksenler (principle axis), temel bileşenler (principle components), ağırlıklandırılmamış en küçük kareler (unweighted least squares), en çok olabilirlik (maximum likelihood), alfa faktörleştirme, en küçük kalıntı (minimum residual), en küçük ki-kare (minimum chi-square) ve en küçük rank (minimum rank) bunlardan bazılarıdır. Bu çalışmada temel eksenler, en küçük kalıntı, en çok olabilirlik, ağırlıklandırılmamış en küçük kareler ve en küçük ki-kare faktör çıkarma yöntemleri paralel analizde en küçük rank yöntemi ise paralel analizde kullanılmıştır.

Temel eksenler faktör çıkarma yönteminde ortaklık (communalities) katsayıları, korelasyon matrisinin ana köşegeninde yer almaktadır. İteratif bir süreçte gerçekleşen temel eksenler faktör çıkarma yönteminde ortaklık katsayılarının değişimi, seçilen kriterden daha küçük olduğunda elde edilen sonuç açımlayıcı faktör analizinin sonucunda elde edilen değerleri oluşturur (Thompson, 2004). Analizin amacı en çok varyansı açıklayan dik faktörler elde edebilmektir (Tabachnik ve Fidell, 2012).

En küçük kalıntı (minimum residual) faktör çıkarma yönteminde ise ana köşegendeki değerler görmezden gelinir. Ana köşegendeki değerlere odaklanmak yerine köşegen dışında kalan korelasyon değerlerinin kalıntılarının karelerini en küçük yapacak model bulunmaya çalışılır (Gorsuch, 1974). En küçük kalıntı yönteminin avantajı, matrislerin tersini hesaplamayı gerektirmemesidir. Böylece tersi alınamayan matrislerle de işlem yapılabilir (Revelle, 2009).

ten Berge ve Kiers (1991) tarafından geliştirilen en küçük rank yönteminde ise faktörler verilen herhangi bir kovaryans matrisinin negatif olmayan tekil varyanslarının oluşturduğu köşegen matrisi kullanılarak çıkarılır. Bu faktör çıkarma yönteminde amaç, verilen az sayıda faktör için açıklanmayan ortak varyansı en küçük yapmaktır (Sočan ve ten Berge, 2003).

En çok olabilirlik faktör çıkarma yönteminin diğer yöntemlere göre üstün yönü uyum indekslerinin de hesaplanabilmesine imkân vermesidir. Ancak veri setinin analiz varsayımlarını (çok değişkenli normallik, uç değerler vb.) karşılaması gerekmektedir (Fabrigar, Wegener, MacCallum ve Strahan, 1999). Ancak varsayımlar karşılanmadığı durumda bu yöntemin hatalı sonuçlar verebileceği raporlanmıştır (Curran, West ve Finch, 1996). En çok olabilirlik faktör çıkarma yönteminde ortaklıklar (communalities) ki-kare uyum iyiliği testini en küçük yapacak değerlere göre hesaplanmaktadır (Revelle, 2018).

Ağırlıklandırılmamış en küçük kareler faktör çıkarma yöntemi en küçük kalıntı yöntemine çok benzer şekilde çalışmaktadır. İki yöntemde de uyum fonksiyonunun birinci düzeyde türevler kullanılmaktadır. Ancak ağırlıklandırılmamış en küçük kareler kestirim yönteminde orijinal korelasyon matrisinin tersine ihtiyaç duyulurken en küçük kalıntı yönteminde buna ihtiyaç duyulmamaktadır (Revelle, 2009). Bu nedenle de yakınsama problemi daha az ortaya çıkmaktadır.

En küçük ki-kare yöntemi ise en küçük kalıntı yönteminin bir çeşidi olup sadece kalıntıları en küçük yapmak yerine örnekleme ağırlıklandırılan kalıntıları en küçük yapmaya çalışmaktadır. Bu yöntemin eksik verinin, devasa tamamen rassal (Massively Missing Completely at Random [MMCAR]) olduğu durumda kullanılması için geliştirildiği belirtilmektedir (Revelle, 2018).

Simülasyon Koşulları

Araştırmada test uzunluğu (20 ve 30 madde), ortalama faktör yükü (.40 ve .70), ölçme modeli (tek faktörlü, iki faktörlü [boyutlar arası korelasyon = .30]), örneklem büyüklüğü (200, 500 ve 1000) ve kullanılan korelasyon matrisi (Pearson ve tetrakorik) koşulları manipüle edilmiştir. Buna göre 2x2x2x3x2 = 48 simülasyon koşulunda çalışılmış olup her bir koşul için 100 replikasyon yapılmıştır. Araştırmanın odak noktası başarı testleri olduğu için ikili yapıda (1-0) veri setleri kullanılmıştır. Bunun için öncelikle

çok deęişkenli normal daęılım gösteren sürekli veri setleri üretilmiştir. Üretilen veri setleri daha sonra kategorik hale (1-0) getirilmiştir. Veri setleri kategorik hale getirilirken {0} eşik noktası olarak kullanılmıştır.

Araştırmada test uzunluğu, 20 ve 30 madde olarak belirlenmiştir. Başarı testleri göz önüne alındığında Liselere Geçiş Sınavında (LGS) 20 maddelik formların da bulunduğu gözlenmiştir (MEB, 2020). Bu nedenle 20 madde koşulu araştırmada yer almaktadır. Madde sayısının 30 olması ise test uzunluğunun artmasının analiz sonuçlarına etkisinin incelenmesini sağlamaya yöneliktir.

Howard (2016) maddelerin faktör yüklerinin en az .40 olması gerektiğini belirtmiştir. Bu nedenle ortalama faktör yükü koşuluna .40 eklenmiştir. Ayrıca ortalama faktör yükünün artmasının araştırma sonuçları üzerindeki etkisinin incelenmesi amacıyla .70 ortalama faktör yükü koşulu da araştırmaya eklenmiştir. Araştırmada tüm maddelerin faktör yüklerinin eşit olması yerine farklılaşması, gerçek durumlara daha yakın olması nedeniyle tercih edilmiştir. Buna göre maddelerin faktör yükleri ek-A'da sunulmuştur.

Ölçme modeli tek boyutlu ve iki boyutlu tanımlanmıştır. İki boyutlu yapılarda boyutlar arasındaki korelasyon .30 olarak tanımlanmıştır. Bu değer araştırmalarda sıklıkla kullanılmakta olduğundan (Cho, Li ve Bandalos, 2009; Li, 2016a, 2016b) sonuçların diğer çalışmalarla karşılaştırılabilmesi için seçilmiştir. Ayrıca gerçek veri setiyle yürütülen çalışmalarda faktörler arasındaki korelasyonların genellikle .20-.40 aralığında bulunduğu belirtilmektedir (Li, 2016b).

Örneklem büyüklüğü koşullarından 200, faktör analizinde genellikle en küçük örneklem büyüklüğü olarak önerildiğinden (Gorsuch, 1974) seçilmiştir. Diğer örneklem büyüklükleri ise örneklem artışının analiz sonuçlarına etkisinin incelenmesi amacıyla eklenmiştir. Diğer taraftan 200 (küçük), 500 (orta) ve 1000 (büyük) örneklem büyüklükleri faktör analizi alanyazınında sıklıkla kullanılmaktadır (Beauducel ve Herzberg, 2006; Li, 2016a; West, Finch ve Curran, 1995).

Araştırmada aynı veri seti için iki farklı korelasyon matrisi kullanılmıştır. Horn (1965) tarafından geliştirilen paralel analiz yönteminde sürekli veri seti ve Pearson korelasyon matrisi kullanılmaktadır. Ancak kullanılan korelasyon matrisi analiz sonuçlarını etkiyebilmektedir. Diğer taraftan Weng ve Cheng (2005) gerçekleştirdiği çalışmada Phi ve tetrakorik korelasyon matrislerini kullanmış ve phi korelasyon matrisi çözümlerinin daha uygun olduğunu belirtmiştir. Bu nedenle araştırmada Pearson ve tetrakorik korelasyon matrisi kullanılmıştır. Paralel analiz, faktör sayısına karar verme yöntemi olduğu için hangi korelasyon matrisiyle daha doğru kestirim elde edildiği incelenerek daha doğru sonuçlar veren korelasyon matrisi kullanılabilir. Faktör analizinde veri setinin yapısına uygun korelasyon matrisi kullanılması gerekirken paralel analizde doğru faktör sayısının elde edildiği korelasyon matrisinin kullanılmasında bir problem yoktur.

Deęerlendirme Kriterleri

Araştırma sonuçlarının deęerlendirilmesi için doğru tahmin yüzdesi (DTY) ve yanlılık deęerleri kullanılmıştır. DTY, önerilen boyut sayısı ile gerçek boyut sayısının karşılaştırılması ile elde edilmektedir. Diğer bir deyişle önerilen boyut sayısı ile gerçek boyut sayısının birbirine eşit olduğu veri seti yüzdesi doğru tahmin yüzdesi olarak ele alınmıştır.

Yanlılık deęeri ise önerilen boyut sayısından gerçek boyut sayısı çıkarılarak elde edilmektedir. Buna göre (ortalama önerilen boyut sayısı - gerçek boyut sayısı) / gerçek boyut sayısı ifadesi yanlılık olarak ifade edilmektedir (Bandalos ve Leite, 2013). Yanlılık deęeri alanyazında .10'dan küçük olduğu durumda kabul edilebilir olarak deęerlendirildiğinden (Flora ve Curran, 2004; Forero, Maydeu-Olivares ve Gallardo-Pujol, 2009; Moshagen ve Musch, 2014; Rhemtulla, Brosseau-Liard ve Savalei, 2012) bu çalışmada da .10'dan küçük yanlılık deęerleri kabul edilebilir olarak ele alınmıştır.

Simülasyon Verilerinin Üretilmesi ve Veri Analizi

Araştırmada simülasyon verileri R yazılımında (R Core Team, 2018) bulunan lavaan (Rosseel, 2012) paketi ile üretilmiştir. Üretilen veri setlerinin minimum rank ile optimal paralel analiz yöntemiyle analizinde EFA.MRFA (Navarro-Gonzalez ve Lorenzo-Seva, 2020) diğer faktör çıkarma yöntemleri (en küçük kalıntı, en çok olabilirlik, temel bileşenler, temel eksenler, ağırlıklandırılmamış en küçük kareler, en küçük ki-kare) ile paralel analizde ise psych (Revelle, 2018) paketi kullanılmıştır. Optimal paralel analiz gerçekleştirilen en küçük rank yöntemi hariç diğer yöntemlerde (en küçük kalıntı, en çok olabilirlik, temel bileşenler, temel eksenler, ağırlıklandırılmamış en küçük kareler, en küçük ki-kare) paralel analiz yapılırken üretilen korelasyon matrislerinin verileri çok deęişkenli normal daęılıma uygun olacak şekilde üretilmektedir. Ancak örneklem verisinin permütasyonundan elde edilmesi paralel analiz

sonuçlarını etkilemektedir (Timmerman ve Lorenzo-Seva, 2011). Bu nedenle en küçük rank yönteminin kullanıldığı optimal paralel analizle birlikte paralel analiz yapılan diğer faktör çıkarma yöntemleri (en küçük kalıntı, en çok olabilirlik, temel bileşenler, temel eksenler, ağırlıklandırılmamış en küçük kareler, en küçük ki-kare) karşılaştırılmıştır. Ayrıca paralel analizde %95’lik özdeğer ile kıyaslanmanın daha sık kullanıldığı belirtildiğinden (Timmerman ve Lorenzo-Seva, 2011) %95’lik özdeğer eşik noktası olarak kullanılmıştır.

BULGULAR

Bu bölümde araştırma problemlerindeki sırayla bulgulara yer verilmiştir.

Yöntemlerden Elde Edilen Doğru Tahmin Yüzdesi

Faktör çıkarma yöntemlerine göre paralel analiz sonucunda elde edilen doğru tahmin yüzdeleri Tablo 1’de, bu değerlerle oluşturulan grafik ise şekil 1’de sunulmuştur.

Tablo 1.

Faktör çıkarma yöntemlerine göre elde edilen doğru tahmin yüzdeleri

Örnekleme Büyüklüğü	Ortalama Faktör Yüklüğü	Faktör Çıkarma Yöntemi	Kullanılan Korelasyon Matrisi							
			Pearson				Tetrakorik			
			Boyutluluk							
			Tek Boyutlu		İki Boyutlu		Tek Boyutlu		İki Boyutlu	
			Madde Sayısı							
			20	30	20	30	20	30	20	30
200	0.40	EK-Kalıntı	96	89	63	70	0	0	0	0
		TE	93	89	65	73	0	0	0	0
		EÇO	85	87	30	74	22	0	0	0
		UEKK	93	90	68	73	0	0	0	0
		EK-KiKare	93	89	68	72	0	0	0	0
		TBA	100	100	86	94	22	3	6	0
		OPA	83	75	61	54	97	99	83	96
	0.70	EK-Kalıntı	96	100	99	100	67	51	21	13
		TE	97	100	99	100	62	47	21	11
		EÇO	100	100	99	100	91	64	25	16
		UEKK	95	100	98	100	64	48	22	14
		EK-KiKare	96	100	99	100	66	46	20	13
		TBA	100	100	100	100	100	100	100	100
		OPA	100	100	100	100	100	100	100	100
500	0.40	EK-Kalıntı	85	94	71	65	0	0	0	0
		TE	86	95	75	67	0	0	0	0
		EÇO	81	89	65	64	30	0	0	0
		UEKK	84	94	76	68	0	0	0	0
		EK-KiKare	84	92	78	66	0	0	0	0
		TBA	99	100	97	97	56	31	35	8
		OPA	90	82	86	66	100	100	100	100

Tablo 1. devamı

Faktör çıkarma yöntemlerine göre elde edilen doğru tahmin yüzdeleri

Örnekleme Büyüklüğü	Ortalama Faktör Yüklü	Faktör Çıkarma Yöntemi	Kullanılan Korelasyon Matrisi									
			Pearson				Tetrakorik					
			Boyutluluk									
			Tek Boyutlu		İki Boyutlu		Tek Boyutlu		İki Boyutlu			
			Madde Sayısı									
20		30		20		30		20		30		
500	0.70	0.70	EK-Kalıntı	96	100	100	100	55	56	16	11	
			TE	99	100	100	100	58	55	19	13	
			EÇO	85	93	100	100	73	43	23	12	
			UEKK	96	100	100	100	54	54	16	13	
			EK-KiKare	96	100	100	100	57	51	23	13	
			TBA	100	100	100	100	100	100	100	100	
			OPA	100	100	100	100	100	100	100	100	
1000	0.40	0.40	EK-Kalıntı	79	90	72	81	0	0	0	0	
			TE	81	89	67	82	0	0	0	0	
			EÇO	81	87	62	82	35	0	0	0	
			UEKK	81	91	73	85	2	0	0	0	
			EK-KiKare	81	90	69	82	1	0	0	0	
			TBA	100	100	100	100	90	76	79	53	
	0.70	0.70	0.70	OPA	69	54	89	47	100	100	100	99
				EK-Kalıntı	96	100	100	100	76	51	22	15
				TE	97	100	100	100	72	57	22	15
				EÇO	88	93	100	100	71	41	23	16
				UEKK	95	100	100	100	70	56	21	10
				EK-KiKare	96	100	100	100	75	53	25	14
				TBA	100	100	100	100	100	100	100	100
OPA	100	99	100	100	100	100	100	100				

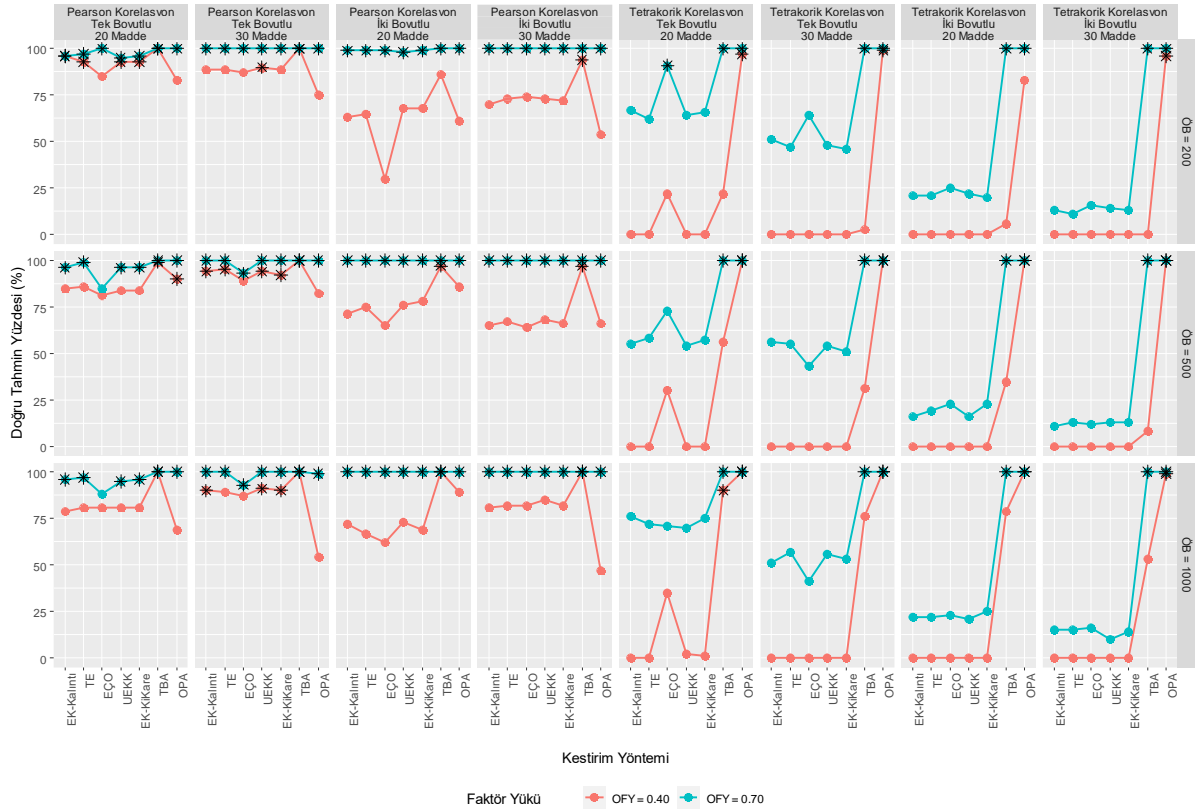
EKRrank: En küçük rank, EKKalıntı: En küçük kalıntı, TE: Temel Eksenler, EÇO: En çok olabilirlik, UEKK: Ağırlıklandırılmamış en küçük kareler, EK-KiKare: En küçük ki kare, OPA: Optimal paralel analiz, OFY: Ortalama faktör yükü.

Pearson korelasyon matrisiyle gerçekleştirilen paralel analizler Tablo 1 ve Şekil 1'de incelendiğinde, en çok olabilirlik (EÇO) hariç diğer faktör çıkarma yöntemlerinin ortalama faktör yükünün .70 olduğu tüm koşullarda yeterli DTY'ye sahip olduğu (DTY > %90) gözlenmiştir. EÇO ise ortalama faktör yükü .70 olan 20 maddelik tek boyutlu testlerde 500 ve 1000 örneklem büyüklüğünde %90'dan küçük DTY'ye sahiptir.

Ortalama faktör yükünün .40 olduğu koşullarda, temel bileşenler analizi hariç diğer faktör çıkarma yöntemleri genellikle %90'dan daha küçük DTY'ye sahiptir. Pearson korelasyon matrisi kullanıldığında temel bileşenler analizi (TBA) bir koşul hariç diğer tüm koşullarda yeterli düzeyde DTY'ye sahiptir. Ayrıca 20 maddelik tek boyutlu testlerde 200 örneklem büyüklüğünde optimal paralel analiz ile en küçük rank ve paralel analiz ile en çok olabilirlik hariç diğer yöntemlerin DTY'si yeterli düzeydedir. Aynı koşullar altında madde sayısının 30'a çıkması en küçük kalıntı, temel eksenler ve en küçük ki-kare yöntemlerinin DTY'sini düşürmüştür. Pearson korelasyon matrisi kullanılan analizler için ortalama faktör yükünün .40 olduğu iki boyutlu testlerde sadece TBA (bir koşul hariç) yeterli düzeyde DTY'ye sahiptir.

Tetrakorik korelasyon matrisiyle gerçekleştirilen analizler incelendiğinde optimal paralel analiz ile en küçük rank yönteminin tüm simülasyon koşullarında yeterli düzeyde DTY'ye sahip olduğu söylenebilir. Diğer taraftan temel bileşenler analizinin de diğer yöntemlere göre DTY açısından daha üstün olduğu gözlenmiştir. Ancak Pearson korelasyon matrisinin kullanıldığı analizlerin aksine tetrakorik

korelasyon matrisiyle gerçekleştirilen analizlerde TBA ortalama faktör yükünün .40 olduğu çoğu koşulda yeterli düzeyde DTY'ye sahip değildir. Paralel analiz ile TBA ve optimal paralel analiz ile en küçük rank yöntemleri hariç paralel analiz gerçekleştirilen diğer faktör çıkarma yöntemleri ise tetrakorik korelasyon matrisi ile gerçekleştirilen analizlerde yeterli düzeyde DTY'ye sahip değildir (< %90). En çok olabilirlik yöntemi ortalama faktör yükü .70 olan 20 maddelik tek boyutlu testlerde 200 örneklem büyüklüğünde %90'dan büyük DTY'ye sahiptir. Paralel analiz ile en çok olabilirlik yöntemi sadece bir simülasyon koşulunda yeterli düzeyde DTY'ye sahipken optimal paralel analiz ile en küçük rank ve paralel analiz ile temel bileşenler analizi yöntemleri hariç diğer tüm yöntemler %90'dan küçük DTY'ye sahiptir.



Şekil 1. Simülasyon koşullarından elde edilen DTY değerleri

EKRank: En küçük rank, EKKalıntı: En küçük kalıntı, TE: Temel Eksenler, EÇO: En çok olabilirlik, UEKK: Ağırlıklandırılmamış en küçük kareler, EK-KiKare: En küçük ki kare, OPA: Optimal paralel analiz, OFY: Ortalama faktör yükü, *:DTY>%90

Optimal paralel analiz ile en küçük rank yöntemine ait sonuçlar incelendiğinde ise sadece ortalama faktör yükü .40 olan iki boyutlu 20 maddelik testlerin 200 örneklem büyüklüğü koşulunda %90'dan küçük DTY'ye sahip olduğu gözlenmiştir. Diğer bir deyişle en küçük rank yöntemi ile optimal paralel analiz, tetrakorik korelasyonla gerçekleştirilen analizlerde bir koşul hariç diğer tüm koşullarda yeterli düzeyde DTY'ye sahiptir.

Yöntemlerden Elde Edilen Yanlılık Değerleri

Faktör çıkarma yöntemlerine göre paralel analiz sonucunda elde edilen yanlılık sonuçları Tablo 2'de, bu değerlerle oluşturulan grafik ise Şekil 2'de sunulmuştur.

Tablo 2.

Faktör çıkarma yöntemlerine göre elde edilen yanlılık değerleri

Örnekleme Büyüklüğü	Ortalama Faktör Yüklü	Faktör Çıkarma Yöntemi	Kullanılan Korelasyon Matrisi							
			Pearson				Tetrakorik			
			Boyutluluk							
			Tek Boyutlu		İki Boyutlu		Tek Boyutlu		İki Boyutlu	
			Madde Sayısı							
			20	30	20	30	20	30	20	30
200	0.40	EK-Kalıntı	.05	.12	.11	.21	6.85	11.18	3.13	5.23
		TE	.08	.12	.12	.16	7.09	11.08	3.13	5.26
		EÇO	.17	.13	-.24	.15	5.37	11.21	3.02	5.25
		UEKK	.08	.10	.07	.17	6.99	11.22	3.13	5.22
		EK-KiKare	.08	.12	.10	.17	6.89	11.23	3.12	5.21
		TBA	.00	.00	-.00	.03	1.66	3.78	1.20	2.25
		OPA	.19	.26	.21	.33	.03	.01	.00	.01
	0.70	EK-Kalıntı	.04	.00	.00	.00	.54	.98	.94	1.29
		TE	.03	.00	.00	.00	.68	1.12	.94	1.45
		EÇO	.00	.00	.00	.00	.18	.89	.84	1.26
		UEKK	.05	.00	.01	.00	.64	1.02	.93	1.37
		EK-KiKare	.04	.00	.00	.00	.70	1.16	.95	1.34
		TBA	.00	.00	.00	.00	.00	.00	.00	.00
		OPA	.00	.00	.00	.00	.00	.00	.00	.00
500	0.40	EK-Kalıntı	.16	.06	.20	.21	7.47	11.70	3.27	5.34
		TE	.14	.05	.14	.20	7.30	11.54	3.25	5.26
		EÇO	.21	.11	.00	.22	5.26	11.57	3.14	5.30
		UEKK	.17	.06	.17	.21	7.43	11.64	3.24	5.25
		EK-KiKare	.20	.09	.14	.21	7.57	11.45	3.21	5.29
		TBA	.01	.00	.01	.01	.53	1.38	.45	1.08
		OPA	.10	.20	.07	.19	.00	.00	.00	.00
	0.70	EK-Kalıntı	.04	.00	.00	.00	.83	.80	.95	1.38
		TE	.01	.00	.00	.00	.81	.94	.97	1.30
		EÇO	.15	.07	.00	.00	.49	1.07	.88	1.29
		UEKK	.04	.00	.00	.00	.82	.92	.97	1.32
		EK-KiKare	.04	.00	.00	.00	.77	.97	.88	1.32
		TBA	.00	.00	.00	.00	.00	.00	.00	.00
		OPA	.00	.00	.00	.00	.00	.00	.00	.00

Tablo 2. devamı

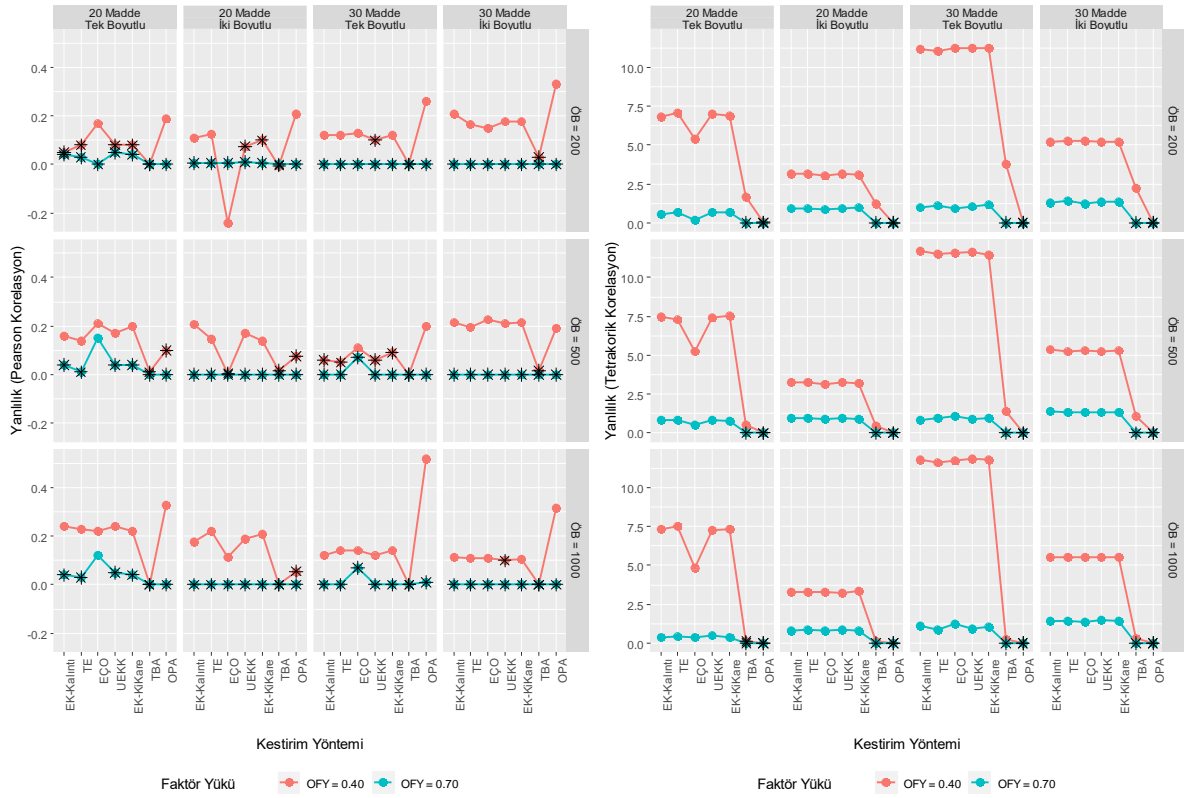
Faktör çıkarma yöntemlerine göre elde edilen yanlılık değerleri

Örnekleme Büyüklüğü	Ortalama Faktör Yüklüğü	Faktör Çıkarma Yöntemi	Kullanılan Korelasyon Matrisi							
			Pearson				Tetrakorik			
			Boyutluluk							
			Tek Boyutlu		İki Boyutlu		Tek Boyutlu		İki Boyutlu	
			Madde Sayısı							
20	30	20	30	20	30	20	30			
1000	0.40	EK-Kalıntı	.24	.12	.17	.12	7.32	11.79	3.27	5.51
		TE	.23	.14	.22	.11	7.54	11.64	3.31	5.52
		EÇO	.22	.14	.12	.11	4.84	11.71	3.29	5.51
		UEKK	.24	.12	.19	.10	7.29	11.88	3.25	5.54
		EK-KiKare	.22	.14	.21	.10	7.33	11.83	3.33	5.55
		TBA	.00	.00	.00	.00	.10	.26	.10	.30
		OPA	.33	.52	.05	.32	.00	.00	.00	.00
	0.70	EK-Kalıntı	.04	.00	.00	.00	.35	1.09	.80	1.42
		TE	.03	.00	.00	.00	.41	.87	.84	1.40
		EÇO	.12	.07	.00	.00	.38	1.23	.80	1.38
		UEKK	.05	.00	.00	.00	.47	.93	.86	1.51
		EK-KiKare	.04	.00	.00	.00	.36	1.02	.78	1.42
		TBA	.00	.00	.00	.00	.00	.00	.00	.00
		OPA	.00	.01	.00	.00	.00	.00	.00	.00

EKRank: En küçük rank, EKKalıntı: En küçük kalıntı, TE: Temel Eksenler, EÇO: En çok olabilirlik, UEKK: Ağırlıklandırılmamış en küçük kareler, EK-KiKare: En küçük ki kare, OPA: Optimal paralel analiz, OFY: Ortalama faktör yükü.

Pearson korelasyon matrisiyle gerçekleştirilen paralel analizler Tablo 2 ve Şekil 2'de incelendiğinde, en çok olabilirlik (EÇO) hariç diğer faktör çıkarma yöntemlerinin ortalama faktör yükünün .70 olduğu tüm koşullarda kabul edilebilir düzeyde yanlılığa sahip olduğu söylenebilir (<.10). EÇO ise 20 maddelik tek boyutlu testlerde 500 ve 1000 örneklem büyüklüğünde .10'dan büyük yanlılığa sahiptir.

Pearson korelasyon matrisiyle gerçekleştirilen ve ortalama faktör yükü .40 olan koşullarda, temel bileşenler analizi (TBA) hariç diğer faktör çıkarma yöntemleri genellikle .10'dan büyük yanlılığa sahiptir. TBA ise bu koşulların tümünde kabul edilebilir düzeyde yanlıdır. Ortalama faktör yükü .40 olan 20 maddelik tek boyutlu testlerde 200 örneklem büyüklüğünde optimal paralel analiz ile en küçük rank ve paralel analiz ile en çok olabilirlik hariç diğer yöntemlerin yanlılığı kabul edilebilir düzeydedir. Aynı koşullar altında madde sayısının 30'a çıkması en küçük kalıntı, temel eksenler ve en küçük ki-kare yöntemlerinin yanlılığını yükseltmiştir. Ortalama faktör yükünün .40 olduğu iki boyutlu testlerde yöntemlerin yanlılığı genellikle .10'dan büyüktür. Ancak örneklem büyüklüğünün 1000, madde sayısının 20 olduğu iki boyutlu testlerde paralel analiz ile temel bileşenler ve optimal paralel analiz ile en küçük rank yöntemlerinin yanlılığı kabul edilebilir düzeydeyken madde sayısının 30'a yükselmesiyle paralel analizle temel bileşenler ve ağırlıklandırılmamış en küçük kareler yöntemlerinin yanlılığı kabul edilebilir düzeyde kalmıştır (<.10).



Şekil 2. Simülasyon koşullarından elde edilen yanlılık değerleri

EKRrank: En küçük rank, EKKalıntı: En küçük kalıntı, TE: Temel Eksenler, EÇO: En çok olabilirlik, UEKK: Ağırlıklandırılmamış en küçük kareler, EK-KiKare: En küçük ki kare, OPA: Optimal paralel analiz, OFY: Ortalama faktör yükü, *: Yanlılık < 10

Tetrakorik korelasyon matrisiyle gerçekleştirilen analizler incelendiğinde optimal paralel analiz uygulanan en küçük rank yönteminin tüm simülasyon koşullarında kabul edilebilir düzeyde yanlılığa sahip olduğu gözlenmiştir. Tetrakorik korelasyon matrisiyle gerçekleştirilen analizlerde TBA hariç diğer yöntemlerin çok yüksek yanlılığa sahip olduğu koşullar bulunmaktadır. Örneğin en küçük kalıntı yönteminin yanlılık değeri ortalama faktör yükü .40 olan 30 maddelik tek boyutlu testlerde 10'dan daha büyüktür. Diğer bir deyişle bu yöntem faktör sayısını olduğundan çok daha fazla kestirmiştir. Benzer şekilde optimal paralel analiz yapılan en küçük rank yöntemi hariç diğer yöntemlerin yanlılıklarının da genel olarak yüksek olduğu söylenebilir. Ancak TBA yöntemi tetrakorik korelasyon matrisiyle gerçekleştirilen analizlerde ortalama faktör yükü .70 olan koşulların tümünde kabul edilebilir düzeyde yanlıyken ortalama faktör yükü .40 olan 1 koşul hariç diğer koşullarda yüksek düzeyde yanlıdır. Pearson korelasyon matrisiyle gerçekleştirilen analizlerde yanlılık değerini gösteren y-ekseninin ölçeği -.20 ile .60 aralığında yer alırken tetrakorik korelasyon matrisiyle gerçekleştirilen analizlerde yanlılık değerini gösteren y-ekseninin ölçeği 0-15 aralığındadır.

TARTIŞMA, SONUÇ ve ÖNERİLER

Faktör çıkarma yöntemlerinin paralel analiz sonuçlarına etkisinin iki kategorili veri tipinde incelendiği bu çalışma sonucunda Pearson korelasyon matrisinin kullanıldığı analizlerde ortalama faktör yükünün düşük olduğu (.40) koşullarda temel bileşenler hariç diğer yöntemlerin faktör sayısını yeterli düzeyde doğru kestirmediği gözlenmiştir. Benzer şekilde ortalama faktör yükünün düşük olduğu koşullarda temel bileşenler hariç diğer yöntemler, faktör sayısını genellikle olduğundan daha büyük kestirmiştir. Örneklem büyüklüğünün ise Pearson korelasyon matrisiyle gerçekleştirilen analizlerde paralel analiz sonuçları üzerinde çok etkili olmadığı söylenebilir. Bu sonuç Weng ve Cheng (2005) tarafından gerçekleştirilen çalışmayla da benzerlik göstermektedir. Weng ve Cheng (2005) kesme noktası olarak ortalama, %95 ve %99'luk özdeğerlerin kullanıldığı paralel analiz sonuçlarını karşılaştırmış ve faktör yükünün .45 olduğu durumda doğru tahmin yüzdesinin, faktör yükü .70 ve .90 olan durumlara göre daha düşük olduğunu belirtmiştir. Ayrıca Horn (1965) tarafından önerilen paralel analiz (temel bileşenler analizi ve özdeğerlerin ortalamasının kullanıldığı paralel analiz) Pearson korelasyon matrisiyle gerçekleştirilen analizlerde koşulların çok büyük kısmında yeterli performansı gösterdiği raporlanmıştır.

Tetrakorik korelasyon matrisiyle gerçekleştirilen analizlerde ise optimal paralel analiz ile en küçük rank yönteminin diğer yöntemlerle kıyaslandığında faktör sayısını çok daha fazla koşulda doğru kestirdiği söylenebilir. Paralel analiz ile temel bileşenler ve optimal paralel analiz ile en küçük rank yöntemleri hariç diğer kestirim yöntemleri kullanılarak tetrakorik korelasyon matrisiyle analizler gerçekleştirildiğinde faktör sayısının olduğundan çok daha büyük kestirildiği belirlenmiştir. Bu durum Timmerman ve Lorenzo-Seva (2011) tarafından gerçekleştirilen çalışmayla da benzerlik göstermektedir. Polikorik korelasyon matrisiyle en küçük rank yöntemi kullanılarak gerçekleştirilen optimal paralel analiz simülasyon koşullarının büyük kısmında yeterli performansı göstermiştir. Ancak polikorik korelasyonun hesaplanmasındaki güçlükler nedeniyle küçük örneklerde yakınsama problemleri olabilmektedir. Bu çalışmada küçük örneklem büyüklüğünde tetrakorik korelasyon matrisinin hesaplanmasında gerekli olduğu durumda psych paketinde (Revelle, 2018) bulunan basit yumuşatma (simple smoothing) kullanılmıştır. Örneklem büyüklüğü, tetrakorik korelasyon matrisiyle gerçekleştirilen analizlerde faktör çıkarma yöntemlerinin performansları üzerinde çok etkili değildir. Örneklem büyüklüğünün artması doğru tahmin yüzdesini etkilememiştir.

Bu araştırma, ele alınan simülasyon koşullarıyla sınırlıdır. Ancak araştırma bulgularına göre, iki kategorili veri setlerinde tetrakorik korelasyon matrisiyle gerçekleştirilen analizlerde en küçük rank yöntemi ile optimal paralel analizin kullanılması önerilebilir. Optimal paralel analiz Factor 10.10 (Lorenzo-Seva ve Ferrando, 2019) paket programı ile de gerçekleştirilebilmektedir. R yazılımının kullanımında güçlük yaşayan araştırmacılara ücretsiz olan bu paket programı kullanmaları önerilebilir. Eğer veri seti küçük ve tetrakorik korelasyon matrisi hesaplanamıyorsa bu durumda temel bileşenler faktör çıkarma yöntemiyle Pearson korelasyon matrisi kullanılarak paralel analiz gerçekleştirilebilir. Bu çalışmada faktör sayısına karar vermek için eşik noktası olarak özdeğerlerin ortalaması yerine %95'e karşılık gelen değer kullanıldığından araştırmacılara %95'in kullanılması önerilebilir.

Bu çalışmada boyut sayısı ve boyutlar arası korelasyon koşullarının sınırlı olması nedeniyle gelecek çalışmalarda daha fazla boyuta sahip yapılarda araştırma yapılabilir. Ayrıca veri setindeki değişkenlerin kategori sayısı, değişkenlerin çarpık dağıldığı durumda nasıl sonuçlar elde edileceği ilgi çekici problemler arasında yer alabilir.

KAYNAKÇA

- Bandalos, D. L., & Leite, W. (2013). Use of Monte Carlo studies in structural equation modeling research. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 625-667). Charlotte, NC: Information Age.
- 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: A Multidisciplinary Journal*, 13(2), 186-203. https://doi.org/10.1207/s15328007sem1302_2
- Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research*, 27(4), 509-540. https://doi.org/10.1207/s15327906mbr2704_2
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2), 245-276. https://doi.org/10.1207/s15327906mbr0102_10
- Cho, S.-J., Li, F., & Bandalos, D. L. (2009). Accuracy of the parallel analysis procedure with polychoric correlations. *Educational and Psychological Measurement*, 69(5), 748-759. <https://doi.org/10.1177/0013164409332229>
- Çokluk, Ö., & Koçak, D. (2016). Using Horn's parallel analysis method in exploratory factor analysis for determining the number of factors. *Educational Sciences: Theory & Practice*, 16(2), 537-552. <https://doi.org/10.12738/estp.2016.2.0328>
- Crawford, A. V., Green, S. B., Levy, R., Lo, W.-J., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Measurement*, 70(6), 885-901. <https://doi.org/10.1177/0013164410379332>
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16-29. <https://doi.org/10.1037/1082-989X.1.1.16>
- Dinno, A. (2009). Exploring the sensitivity of Horn's parallel analysis to the distributional form of random data. *Multivariate Behavioral Research*, 44(3), 362-388. <https://doi.org/10.1080/00273170902938969>
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299. <https://doi.org/10.1007/s10803-009-0816-2>
- Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory

- factor analysis with ordinal data. *Psychological Methods*, 9(4), 466–491. <https://doi.org/10.1037/1082-989X.9.4.466>
- Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor analysis with ordinal indicators: A monte carlo study comparing DWLS and ULS estimation. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(4), 625–641. <https://doi.org/10.1080/10705510903203573>
- Glorfeld, L. W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement*, 55(3), 377–393. <https://doi.org/10.1177/0013164495055003002>
- Goretzko, D., Pham, T. T. H., & Bühner, M. (2019). Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. *Current Psychology*. <https://doi.org/10.1007/s12144-019-00300-2>
- Gorsuch, R. L. (1974). *Factor analysis*. W. B. Saunders.
- Green, S., Xu, Y., & Thompson, M. S. (2018). Relative accuracy of two modified parallel analysis methods that use the proper reference distribution. *Educational and Psychological Measurement*, 78(4), 589–604. <https://doi.org/10.1177/0013164417718610>
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179–185. <https://doi.org/10.1007/BF02289447>
- Howard, M. C. (2016). A review of exploratory factor analysis decisions and overview of current practices: What we are doing and how can we improve? *International Journal of Human-Computer Interaction*, 32(1), 51–62. <https://doi.org/10.1080/10447318.2015.1087664>
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1), 141–151. <https://doi.org/10.1177/001316446002000116>
- Keith, T. Z., Caemmerer, J. M., & Reynolds, M. R. (2016). Comparison of methods for factor extraction for cognitive test-like data: Which overfactor, which underfactor? *Intelligence*, 54, 37–54. <https://doi.org/10.1016/j.intell.2015.11.003>
- Koçak, D., Çokluk, Ö. & Kayri, M. (2016). Faktör sayısının belirlenmesinde MAP testi, paralel analiz, K1 ve yamaç birikinti grafiği yöntemlerinin karşılaştırılması. *YYÜ Eğitim Fakültesi Dergisi*, 13(1), 330–359.
- Li, C.-H. (2016a). The performance of ML, DWLS, and ULS estimation with robust corrections in structural equation models with ordinal variables. *Psychological Methods*, 21(3), 369–387. <https://doi.org/10.1037/met0000093>
- Li, C.-H. (2016b). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48(3), 936–949. <https://doi.org/10.3758/s13428-015-0619-7>
- Liu, O. L., & Rijmen, F. (2008). A modified procedure for parallel analysis of ordered categorical data. *Behavior Research Methods*, 40(2), 556–562. <https://doi.org/10.3758/BRM.40.2.556>
- Lorenzo-Seva, U., & Ferrando, P. J. (2019). *Factor* (Version 10.10.01) [Computer software]. Universitat Rovira i Virgili.
- MEB. (2020). *Merkezi sınav başvuru ve uygulama klavuzu*. http://www.meb.gov.tr/meb_iys_dosyalar/2020_05/06105923_BasYvuru_ve_Uygulama_KYlavuzu_2020_GuYncel.pdf adresinden erişildi (Erişim Tarihi:02/06/2020).
- Moshagen, M., & Musch, J. (2014). Sample size requirements of the robust weighted least squares estimator. *Methodology*, 10(2), 60–70. <https://doi.org/10.1027/1614-2241/a000068>
- Navarro-Gonzalez, D., & Lorenzo-Seva, U. (2020). *EFA.MRFA: Dimensionality assessment using minimum rank factor analysis*. <https://cran.r-project.org/package=EFA.MRFA> adresinden erişildi (Erişim Tarihi:05/05/2020).
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, 32(3), 396–402. <https://doi.org/10.3758/BF03200807>
- Preacher, K. J., Zhang, G., Kim, C., & Mels, G. (2013). Choosing the optimal number of factors in exploratory factor analysis: A model selection perspective. *Multivariate Behavioral Research*, 48(1), 28–56. <https://doi.org/10.1080/00273171.2012.710386>
- R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.r-project.org/>. adresinden erişildi (Erişim Tarihi:04/04/2020).
- Revelle, W. (2009). *An introduction to psychometric theory with applications in R*. Springer. <http://www.personality-project.org/r/book/> adresinden erişildi (Erişim Tarihi:18/05/2020).
- Revelle, W. (2018). *psych: Procedures for psychological, psychometric, and personality research*. <https://cran.r-project.org/package=psych> adresinden erişildi (Erişim Tarihi:19/03/2020).
- 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. <https://doi.org/10.1037/a0029315>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1–36.
- Ruscio, J., & Roche, B. (2012). Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychological Assessment*, 24(2), 282–292. <https://doi.org/10.1037/a0025697>
- Sočan, G., & ten Berge, J. M. F. (2003). The determinants of the bias in minimum rank factor analysis (MRFA). In H. Yanai, A. Okada, K. Shigemasu, Y. Kano & J. J. Meulman (Eds.), *New Developments in Psychometrics* (pp. 95–102). Tokyo: Springer.
- Steger, M. F. (2006). An illustration of issues in factor extraction and identification of dimensionality in psychological assessment data. *Journal of Personality Assessment*, 86(3), 263–272. https://doi.org/10.1207/s15327752jpa8603_03
- Tabachnik, B. G., & Fidell, L. S. (2012). *Using multivariate statistics* (6th ed.). Boston: Pearson.
- ten Berge, J. M. F., & Kiers, H. A. L. (1991). A numerical approach to the approximate and the exact minimum rank of a covariance matrix. *Psychometrika*, 56(2), 309–315. <https://doi.org/10.1007/BF02294464>
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: APA.
- Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, 16(2), 209–220. <https://doi.org/10.1037/a0023353>
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41(3), 321–327. <https://doi.org/10.1007/BF02293557>
- Weng, L.-J., & Cheng, C.-P. (2005). Parallel analysis with unidimensional binary data. *Educational and Psychological Measurement*, 65(5), 697–716. <https://doi.org/10.1177/0013164404273941>
- West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with non-normal variables: Problems and remedies (pp. 56-75). In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications*. Thousand Oaks, CA: Sage.
- Xia, Y., Green, S. B., Xu, Y., & Thompson, M. S. (2019). Proportion of indicator common variance due to a factor as an effect size statistic in revised parallel analysis. *Educational and Psychological Measurement*, 79(1), 85-107. <https://doi.org/10.1177/0013164418754611>
- Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99(3), 432–442. <https://doi.org/10.1037/0033-2909.99.3.432>

Ek A: Veri üretiminde kullanılan madde-faktör yükleri

Tek Boyutlu Modeller					İki Boyutlu Modeller (Boyutlar arası korelasyon = .30)							
Madde No	Maddelerin Boyutlara Dağılımı				Maddelerin Boyutlara Dağılımı	Maddelerin Boyutlara Dağılımı			Maddelerin Boyutlara Dağılımı	Maddelerin Boyutlara Dağılımı		
	OFY = .40, 20 Madde	OFY = .70, 20 Madde	OFY = .40, 30 Madde	OFY = .70, 30 Madde		Madde No	OFY = .40, 20 Madde	OFY = .70, 20 Madde		Madde No	OFY = .40, 30 Madde	OFY = .70, 30 Madde
1	.464	.704	.396	.679	Birinci Boyut	1	.464	.704	Birinci Boyut	1	.396	.679
2	.404	.691	.399	.705		2	.404	.691		2	.399	.705
3	.417	.716	.363	.703		3	.417	.716		3	.363	.703
4	.412	.712	.430	.719		4	.412	.712		4	.430	.719
5	.446	.732	.374	.717		5	.446	.732		5	.374	.717
6	.429	.684	.415	.708		6	.429	.684		6	.415	.708
7	.404	.694	.421	.717		7	.404	.694		7	.421	.717
8	.387	.702	.403	.683		8	.387	.702		8	.403	.683
9	.407	.706	.394	.702		9	.407	.706		9	.394	.702
10	.389	.685	.438	.695		10	.389	.685		10	.438	.695
11	.437	.710	.367	.728	İkinci Boyut	11	.437	.710	İkinci Boyut	11	.367	.728
12	.399	.709	.426	.680		12	.399	.709		12	.426	.680
13	.39	.704	.406	.728		13	.39	.704		13	.406	.728
14	.394	.712	.424	.695		14	.394	.712		14	.424	.695
15	.371	.689	.422	.685		15	.371	.689		15	.422	.685
16	.388	.716	.394	.701		16	.388	.716		16	.394	.701
17	.428	.692	.41	.691		17	.428	.692		17	.410	.691
18	.42	.688	.471	.695		18	.420	.688		18	.471	.695
19	.381	.734	.382	.699		19	.381	.734		19	.382	.699
20	.382	.727	.414	.699		20	.382	.727		20	.414	.699
21		.407	.707	İkinci Boyut	21		.407	İkinci Boyut	21	.407	.707	
22		.448	.686		22		.448		22	.448	.686	
23		.44	.703		23		.44		23	.44	.703	
24		.387	.709		24		.387		24	.387	.709	
25		.434	.717		25		.434		25	.434	.717	
26		.351	.669		26		.351		26	.351	.669	
27		.396	.734		27		.396		27	.396	.734	
28		.418	.730		28		.418		28	.418	.730	
29		.395	.693		29		.395		29	.395	.693	
30		.425	.682		30		.425		30	.425	.682	