EJAES

JOURNAL OF ADVANCED EDUCATION STUDIES
İleri Eğitim Çalışmaları Dergisi

5 (Special Issue): 288-312, 2023

PERFORMANCE OF FACTOR RETENTION METHODS IN SKEWED DISTRIBUTIONS

Gül GÜLER[1]
Abdullah Faruk KILIÇ[2]

Geliş Tarihi/Received: 09.09.2023 Elektronik Yayın / Online Published: 20.10.2023
DOI: 10.48166/ejaes.1357828

ABSTRACT

This research aims to evaluate the performance of dimensionality determination methods under various simulation conditions. Therefore, dimensionality determination methods were compared, including optimal parallel analysis, MAP, HULL, EGA (TMFG) estimation, EGA (glasso) estimation, and comparison data forest method. The type of distribution, sample size, number of items per factor, number of categories, and measurement model were specified as simulation conditions in the study. For each condition, 100 replications were conducted. A fully crossed simulation design was employed in the study. The results of this study, which examined the performance of factor determination methods under skewed distributions, indicated that the HULL method had the highest average considering the average accuracy values of all conditions. Meanwhile, the HULL method had the lowest Relative bias average. However, no method demonstrated adequate performance under all conditions. This study examined one-factor and two-factor structures with interfactor correlations of 0.00 and 0.30. Considering structures with more than two factors in education and psychology, future research could focus on working with data exhibiting skewed distributions involving more factors and items to compare the performance of methods.

Keywords: Factor retention; MAP; HULL; comparison factor forest; EGA

[1] Assist. Professor, Trakya University, Education Faculty, Department of Educational Science, Edirne, Türkiye, e-mail: gulyuce2010@gmail.com, ORCID: 0000-0001-8626-4901
[2] Assoc. Professor, Trakya University, Education Faculty, Department of Educational Science, Edirne, Türkiye, e-mail: abdullahfarukkilic@gmail.com, ORCID: 0000-0003-3129-1763

ÇARPIK DAĞILIMLARDA FAKTÖR SAYISI BELİRLEME YÖNTEMLERİNİN PERFORMANSLARININ İNCELENMESİ

ÖZET

Bu araştırmanın amacı faktör sayısı belirleme yöntemlerinin çeşitli simülasyon koşulları altında performanslarını değerlendirmektir. Bu amaç doğrultusunda boyutluluk belirleme yöntemlerinden optimal paralel analiz, MAP, HULL, EGA (TMFG) kestirimi, EGA (Glasso) kestirimi ve comparison data forest yöntemi karşılaştırılmıştır. Çalışmada simülasyon koşulları olarak dağılımın türü, örneklem büyüklüğü, faktör başına düşen madde sayısı, kategori sayısı ve ölçme modeli belirlenmiştir. Çalışmada her bir koşul için 100 replikasyon yapılmıştır. Çalışmada tamamen çaprazlanmış simülasyon deseni kullanılmıştır. Çarpık dağılımlarda faktör sayısı belirleme yöntemlerinin performanslarının incelendiği bu çalışma sonucunda tüm koşulların doğruluk değerlerinin ortalaması dikkate alındığında en yüksek ortalamaya HULL yönteminin sahip olduğu görülmüştür. Aynı zamanda en düşük göreli yanlılık ortalaması da HULL yöntemindedir. Ancak tüm koşullarda yeterli performansı gösteren bir yöntemin olmadığı söylenebilir. Diğer bir deyişle her koşulda doğru sonucu verecek bir yöntem bulunmamaktadır. Bu çalışmada tek faktörlü, faktörler arası korelasyonu 0.00 ve 0.30 olan iki faktörlü yapılar incelenmiştir. Eğitimde ve psikolojide ikiden fazla faktör sayısına sahip yapılar göz önünde bulundurulduğunda gelecekteki araştırmalarda çarpık dağılım gösteren verilerde daha fazla faktör ve madde sayısıyla çalışılarak yöntemlerin performansları karşılaştırılabilir.

Anahtar Kelimeler: Faktör sayısı belirleme; MAP; HULL; karşılaştırmalı faktör forest; EGA

1. INTRODUCTION

Latent traits attributed to individuals in education and psychology are considered constructs. Since these constructs cannot be directly observed, individuals' performance regarding the measured trait can be determined based on their responses to a measurement tool designed to assess the construct of interest. However, the validity of these performances should also be examined. Researchers often use Exploratory Factor Analysis (EFA) to examine the construct validity of measures (Cosemans et al., 2022; Finch, 2020; Haslbeck & Bork, 2022; Henson & Roberts, 2006; Svetina, 2011).

Deciding on the number of factors is one of the most essential steps in EFA (Cosemans et al., 2022; Finch, 2020; Reio & Shuck, 2015; Svetina, 2011; Zhang, 2007). In EFA, both overfactoring and underfactoring are problematic. When underfactoring occurs, variables are compressed into a smaller factor space, leading to loss of information, neglect of essential factors, and increased error loads (Cosemans et al., 2022). Overfactoring, on the other hand, can lead to the division of factors that are together or result in unimportant factors (Cosemans et al., 2022; Finch, 2020; Lee et al., 2023). Therefore, the criteria used in determining the number of dimensions become crucial.

In many studies that employ EFA, standard options in statistical software are more commonly preferred when determining the number of dimensions (e.g., Finch, 2020; Henson & Roberts, 2006; Montoya & Edwards, 2021; Schmitt & Sass, 2011). Goretzko et al. (2019) reported in their literature review that 55% of the studies they reviewed employed the Kaiser criterion (K1 rule), and 46% of them employed Cattell's Scree test (Cattell, 1966). However, using these methods alone to determine the number of dimensions has been criticized. For instance, in Cattell's Scree test method, eigenvalues are arranged from highest to lowest and connected by a line. However, this method is also criticized for being subjective (Ledasma & Mora, 2007). Considering the literature in Turkey, one could state that only Cattell's Scree test and the Kaiser criterion are still used to decide on the number of dimensions in scale development studies (Koyuncu & Kılıç, 2019). Making decisions solely based on methods like the scree plot, where researchers' subjective judgments play a role in determining the number of dimensions, may not yield accurate results (Ledasma & Mora, 2007).

Deciding on the dimensionality of a measurement instrument based on more than one method can also be problematic (Ledesma et al., 2015; Lee, 2023). Each method has strengths and weaknesses. Therefore, it becomes crucial to examine which method yields better results under what conditions of the data. In this case, the question of which methods to examine may arise. The Parallel Analysis (PA) method Horn (1965) suggested is widespread considering the factor retention methods. However, in addition to this method, there are also other methods such as Minimum Average Partial Correlation (MAP), HULL (Lorenzo-Seva et al., 2011), or, more recently, the Exploratory Graph Analysis (EGA) method, which has been used more frequently. With the widespread application of machine learning methods in various fields, some researchers have suggested using machine learning techniques as dimensionality determination methods (Goretzko & Ruscio, 2023).

When reviewing the literature on dimensionality determination methods, one may come across many studies working with categorical data (Goretzko & Bühner, 2020; Li et al., 2020; Svetina, 2011; Yang & Xia, 2015). Accordingly, the focus appears to be on studying the performance of dimensionality determination methods, specifically in dichotomous data. Some studies also compare various methods under different conditions in continuous data (Auerswald & Moshagen, 2019; Green et al., 2016). In this study, unlike other studies, we worked both on skewed data alone and manipulated the data to have 3 and 5 categories. Furthermore, the methods under investigation in terms of their performance may also differ from the literature. The dimensionality determination performance of EGA, which has been frequently used in recent years, is compared with the machine learning-based Comparison Data Forest method (Goretzko & Ruscio, 2023). The study investigated whether the machine learning method could solve skewed distributions. Therefore, this study may contribute to the literature in four aspects: i) examining which method performs better in skewed distributions, ii) examining the performance of machine learning methods in dimensionality determination, iii) examining whether machine learning methods can offer a solution for skewed distributions, and iv) examining the performance of the commonly preferred EGA in skewed and categorical datasets.

In this study, factor retention methods including Optimal Parallel Analysis (PA; Timmerman & Lorenzo-Seva, 2011), MAP (normal and revised), HULL, EGA with TMFG estimation [EGA(TMFG)], EGA with Glasso estimation [EGA(Glasso)], and the Comparison Data Forest method proposed by Goretzko and Ruscio (2023) were compared. The primary reason for preferring Optimal PA is that it accurately determines the number of dimensions even under challenging conditions (Golino et al., 2020; Nájera et al. 2021; Timmerman & Lorenzo-Seva, 2011). Conversely, EGA employs a network estimation method with a community detection algorithm that shows the number of dimensions and the distribution of items across relevant dimensions (Golino & Epskamp, 2017). Considering the literature, EGA is highly accurate in determining the number of dimensions (Lee, 2023). Furthermore, EGA is more resistant to differences in sample size, number of items, and correlations between dimensions (Golino & Epskamp, 2017).

Additionally, EGA is unaffected by researchers' a priori guidance (Lee, 2023), which is why it was preeffered in this study. The MAP method, which utilizes the partial correlation matrix and is based on principal component analysis, has been found to provide better results in determining the number of dimensions compared to other methods in a simulation study conducted by Kılıç and Uysal (2019). Therefore, the MAP method was also included in the study. On the other hand, the study included the HULL method because it suggests the number of dimensions based on fit indices. Finally, the study included the Comparison Factor Forest method (Goretzko & Ruscio, 2023), which utilizes machine learning methods of Random Forest and XGBoost algorithms and is based on the Comparison Data method (Ruscio & Roche, 2012). In this context, this research examines the performance of dimensionality determination methods under various simulation conditions. Within the research framework, the study compared the main effects of each condition and the interaction effects of conditions. Accordingly, answers were sought to the following questions:

(1) What are the accuracy values of dimensionality determination methods according to simulation conditions?

(2) What are the relative bias values of dimensionality determination methods according to simulation conditions?

2. METHOD

This study compares methods for determining dimensionality in a Monte Carlo simulation. In simulation studies, datasets generated based on desired characteristics (e.g., distribution, factor loadings, or number of items) are analyzed with the methods of interest, and the results are compared.

2.1. Simulation Conditions

Simulation factors such as distribution, sample size, number of items per factor, number of categories, and measurement model were determined in the study. In this study, 100 replications were conducted for each condition. Given that PA generates 500 datasets for each dataset, we preferred to conduct 100 replications considering the prolonged analysis time.

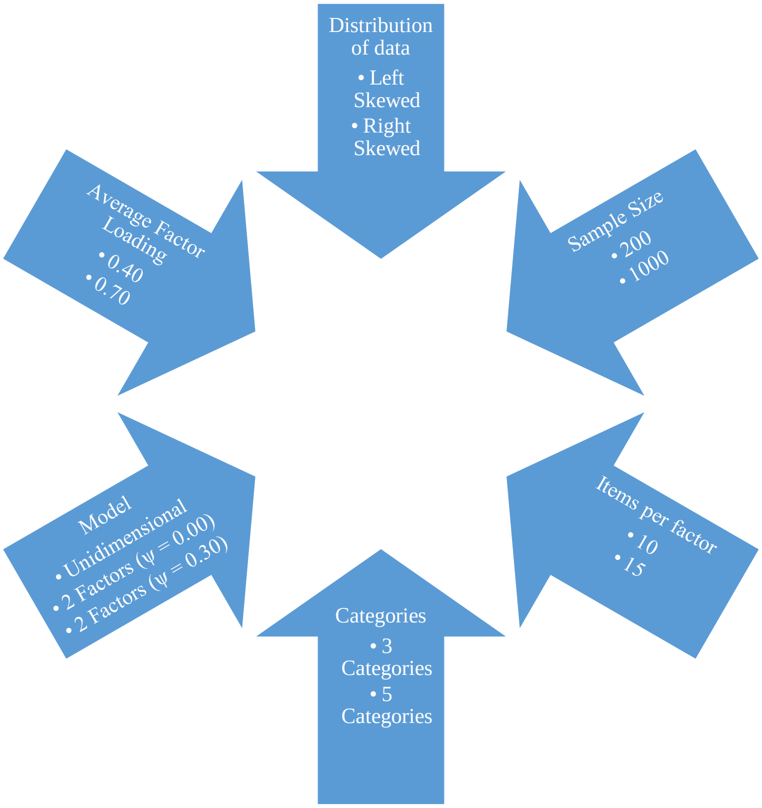

Figure 1. Simulation Conditions

We studied 2x2x2x2x3x2 = 96 simulation conditions (see Figure 1). Considering the distribution of the data, one of the simulation conditions, the datasets were skewed by a skewness coefficient of ±2.5. For this purpose, the dataset demonstrating a continuous normal distribution was generated, and then it was skewed using the cutoff points presented in Appendix 1. Generally, the skewness coefficient in real datasets falls within the range of ±2.00 (Garrido et al., 2011; Muthén & Kaplan, 1985). Therefore, a skewness coefficient of ±2.50 was chosen to examine extreme conditions.

Sample sizes of 200 and 1000 were determined as conditions. In simulation studies, sample sizes of 200 (small), 500 (medium), and 1000 (large) are commonly preferred (Beauducel & Herzberg, 2006; Li, 2016; West et al., 1995). In addition, Gorsuch (1974) proposed a minimum sample size of 200. Therefore, a sample size of 200 was included in this study. In addition, a sample size of 1000 was included as a simulation condition to examine the effect of increasing sample size on factor retention methods.

In their review study, Goretzko et al. (2021) reported that the number of items per factor in most studies (37.2%) was above 7. Therefore, this study determined the number of items per factor as 10 and 15. Since the study focuses on two-dimensional structures, when the number of items per factor is 15, the scale consists of 30 items. Therefore, thinking that longer scales would be less common, the number of items per factor was limited to 15.

In the condition related to the number of categories, there are 3 and 5 categories. As such, 5-point Likert-type items were included in the simulation condition, considering their common utilization (Lozano et al., 2008). Considering Likert-type scales, the number of categories would not be less than 3 in general. Dichotomous Likert-type scales may exist, but they are less common in practice. Therefore, the data was categorized into a minimum of 3 categories.

Under the model conditions analyzed, unidimensional conditions plus conditions for two factors with an inter-factor correlation of 0.00 and two factors with an inter-factor correlation of 0.30 were examined. The reason for examining unidimensional structures is to prevent artificial success, as methods would always have a 100% success rate when they suggest a unidimensional structure. In two-dimensional structures, the inter-factor correlation can influence the performance of methods. Therefore, data were generated with interfactor correlations of 0.00 and 0.30. One of the reasons for selecting an inter-factor correlation of 0.30 is that this value is more commonly found in practical studies (Li, 2016) and is also preferred in simulation studies (Cho et al., 2009; Curran et al., 1996; Flora & Curran, 2004; Foldnes & Grønneberg, 2017).

The average factor loading was manipulated as 0.40 and 0.70. Since the lowest recommended factor loading is generally 0.30 (Costello & Osborne, 2005) or 0.40 (Tabachnick & Fidell, 2019), an average factor loading condition of 0.40 was included. On the other hand, the condition of 0.70 was added as a simulation condition to examine structures with high factor loadings.

2.2. Evaluation Criteria

The accuracy was used as evaluation criteria. As evidenced in the literature, this statistic is used to compare the performance of methods (Goretzko & Bühner, 2022; Kılıç & Uysal, 2019). Accuracy is calculated as

$$Accuracy = \frac{Correct\ Estimates}{n_{rep}}.100 \qquad 1$$

Where correct estimates means cases where the number of factors was correctly identified by the respective method. The other evaluation criteria is relative bias (RB). RB is calculated as

$$RB = \frac{\hat{\zeta} - \zeta_{TRUE}}{\zeta_{TRUE}} \qquad 2$$

Where ζ_{TRUE} is true for the number of factors in simulation conditions (1 or 2). $\hat{\zeta}$ means the average of the number of factors estimates. |RB|> 0.10 indicates substantial bias (Flora & Curran, 2004; Forero et al., 2009). So we used the cut-off criteria as 0.10 for RB.

2.3. Data Analysis

We used the lavaan package (Rosseel, 2012) in the R software (R Core Team, 2022) to generate data. In addition, we used the EFA.MRFA package (Navarro-Gonzalez & Lorenzo-Seva, 2021) for Optimal PA and HULL methods, EFA.dimensions package (O'Connor, 2022) for MAP analysis, EGAnet package (Golino & Christensen, 2020) for EGA. For the Comparison Factor Forest method, we used the codes shared by Goretzko and Bühner (2022).

3. FINDINGS

This section presents the findings in the order of research problems.

3.1. Examination of Accuracy Values

Figure 2 shows the accuracy values obtained from the methods. Additionally, accuracy values are presented in Appendix 2 for researchers who wish to conduct a detailed examination. In addition, one-way ANOVA was conducted to identify the variables influencing accuracy values. The ANOVA results indicated that distribution of data [$F_{(1, 658)} = 0.03$, $p = 0.86$], model [$F_{(2,658)} = 0.93$, $p = 0.39$], and items per factor [$F_{(1,658)}$=0.72, $p = 0.40$] conditions differed in terms of accuracy values. Furthermore, the average factor loading [$F_{(1, 658)} = 96.80$, $p < 0.01$, $\eta^2 = 0.13$], number of categories in variables [$F_{(1, 658)}$=5.38, $p < 0.05$, $\eta^2 = 0.008$], sample size [$F_{(1,658)}$=73.39, $p<0.01$, η^2=0.10], and method [$F_{(1,658)}$=13.62, $p<0.01$, η^2=0.11] differed in terms of accuracy averages. The simulation condition that has the greatest effect on accuracy scores is average factor loadings. This is followed by factor retention method and sample size. Eta square values show that average factor loading has a significant effect on accuracy values, while sample size and method have a moderate effect.

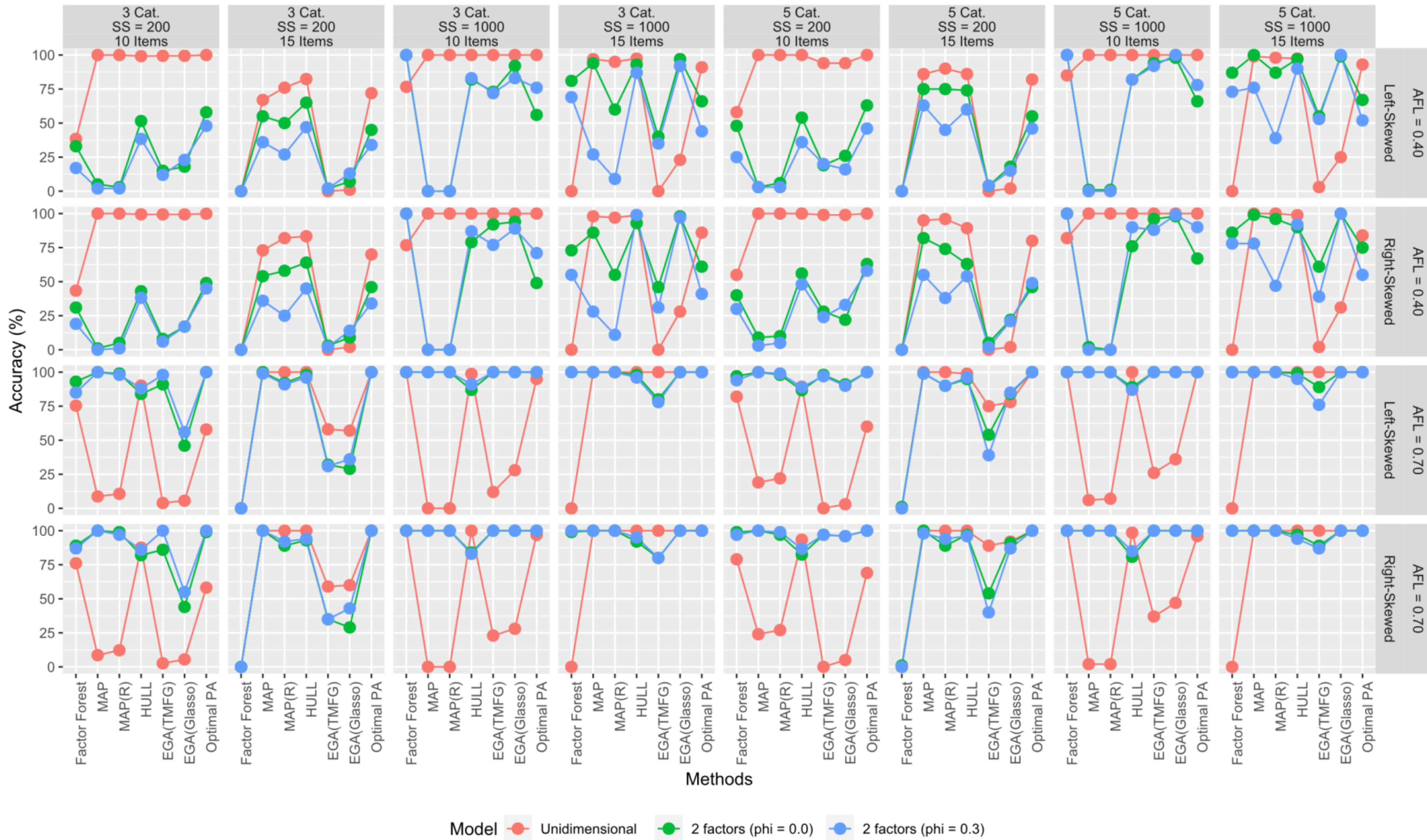

Figure 2. Accuracy Values of the Methods

Considering the factor determination methods in Figure 2, none of the methods exhibited adequate performance in all conditions. The average accuracy of the methods for all conditions was 54.32% for Factor Forest, 69.28% for MAP, 66.25% for MAP(R), 85.73% for HULL, 55.47% for EGA (TMFG), and 64.96% for EGA (Glasso). In all unidimensional conditions with an average factor loading of 0.40, except for the factor forest, other methods had sufficient accuracy (>90%). As the number of factors increased, the performance of methods also changed. Under the conditions of low average factor loading (0.40), two factors, and 10 items per factor, regardless of sample size, HULL, MAP, Revised MAP, and optimal PA did not demonstrate adequate performance. Under these conditions, Factor Forest had 100% accuracy when the sample size was 1000. Under the same conditions, but with a smaller sample size (n = 200), the Factor Forest method did not demonstrate adequate performance. Under the conditions of low average factor loading (0.40) two factors, and 10 items per factor, EGA methods had sufficient accuracy values in datasets consisting of 5-category variables and a sample size of 1000.

The study found that EGA (TMFG), MAP, Revised MAP, and Optimal PA methods demonstrated adequate performance in two-dimensional structures with 10 items and AFL of 0.70, regardless of the number of categories and sample size. EGA (Glasso) and Factor Forest demonstrated adequate performance in conditions other than those where the sample size was 200 and the number of categories was 3. However, under these conditions, the performance of the HULL method was lower than other methods. Only under one of these specified conditions (3 categories and a sample size of 1000), it had an accuracy rate of over 90%.

In unidimensional structures where the number of items per factor was 15 and AFL was 0.40, the performances of EGA methods and Factor Forest were quite low. Under these conditions, the accuracy value for Factor Forest was 0. However, the EGA methods had an accuracy value of around 25%. The HULL method, on the other hand, demonstrated adequate performance in 5-category data when the sample size was 1000. MAP and revised MAP methods exhibited adequate performance in both sample sizes when there were 5 categories. They also demonstrated adequate performance when there were 3 categories and the sample size was 1000.

Under all conditions with two-dimensional structures where the number of items per factor was 15 and AFL was 0.40 ($\psi = 0.00$ and $\psi = 0.30$), EGA (TMFG), Optimal PA, and Factor Forest did not demonstrate adequate performance. EGA (Glasso), on the other hand, demonstrated sufficient accuracy for the specified conditions when there were 5 categories and a sample size of 1000. The HULL method had sufficient accuracy with a sample size of 1000 for the specified conditions. In most of the specified conditions, MAP and revised MAP methods had accuracy rates of lower than 90%.

In unidimensional structures where the number of items per factor was 15 and AFL was 0.70, the accuracy value for Factor Forest was 0. In contrast, HULL, MAP, revised MAP, and Optimal PA methods demonstrated 100% accuracy. EGA methods showed sufficient accuracy under the sample size of 1000. However, EGA methods demonstrated inadequate performance under the specified conditions in small sample sizes.

In two-dimensional structures where the number of items per factor was 15 and AFL was 0.70 (ψ = 0.00 and ψ = 0.30), all conditions demonstrated that Optimal PA, Revised MAP, MAP, and HULL methods have sufficient accuracy. Under the specified conditions, Factor Forest exhibited adequate performance only with a sample size of 1000, while its performance was quite low with a sample size of 200. Put differently, under the conditions where the sample size was 1000, Factor Forest achieved at least 99% accuracy, while under the same conditions with a sample size reduced to 200, it achieved a maximum of 1% accuracy. EGA (TMFG) did not achieve sufficient accuracy under any of the specified conditions, while EGA (Glasso) achieved sufficient accuracy in all conditions where the sample size was 1000. The accuracy values of the methods did not show significant variations based on whether the data were right-skewed or left-skewed.

3.2. Examination of RB Values

RB values obtained from the methods can be seen in Figure 3. In addition, RB values are presented in Appendix 3 for researchers who wish to examine them in detail. One-way ANOVA was conducted to determine the simulation conditions influencing the RB values. ANOVA results indicated that the distribution of data [$F(1, 658) = 0.01$, $p = 0.92$] and number of categories in variables [$F(1, 658) = 0.23$, $p = 0.23$] did not differ in terms of RB values. In contrast, Model [$F(2, 658) = 31.91$, $p < 0.01$, $\eta2 = 0.09$], average factor loading [$F(1, 658) = 9.08$, $p < 0.01$, $\eta2 = 0.01$], items per factor [$F(1, 658) = 84.79$, $p < 0.05$, $\eta2 = 0.11$], sample size [$F(1, 658) = 28.20$, $p < 0.01$, $\eta2 = 0.04$], and method [F(1, 658) = 61.88, $p < 0.01$, $\eta2 = 0.36$] differed in terms of RB averages. The simulation condition that has the greatest effect on RB is factor retention method. This is followed by items per factor and sample size. Eta square values show that factor retention method has a significant effect on RB values, while items per factor has moderate and sample size has small effect. The RB averages of the methods for all conditions were 0.93 for Factor Forest, -0.25 for MAP, -0.23 for MAP(R), -0.01 for HULL, 0.46 for EGA (TMFG), and 0.42 for EGA (Glasso).

For two-factor structures with an AFL of 0.70, the MAP, revised MAP, HULL, and optimal PA methods had RB values within appropriate ranges ($|RB| < 0.10$) under all simulation conditions. Under the same conditions, the Factor Forest method did not demonstrate adequate performance when the sample size was 200 and the number of items was 15, overestimating the number of factors. EGA (Glasso) indicated adequate performance under the specified conditions with a sample size of 1000 but did not indicate adequate performance in most conditions with a sample size of 200. Similarly, EGA (TMFG) demonstrated inadequate performance in most conditions with a sample size of 200 while also performing inadequately in some conditions with a sample size of 1000.

Considering the unidimensional structures, one could express that the Factor Forest method tends to overestimate the number of factors in most conditions. The Factor Forest method exhibited an adequate RB value in unidimensional structures, where AFL was 0.70, the number of items was 10, and the sample size was 1000. The MAP and revised MAP methods generally had negative RB values under conditions where the number of items was low. The HULL method demonstrated adequate performance under most conditions for unidimensional structures. However, conditions where the sample size was 200 and AFL was 0.40 reduced the performance of the HULL method. The optimal PA method generally exhibited adequate RB values in unidimensional structures under most conditions. However, it made biased estimations under conditions where the sample size was small, AFL was low, and the number of items was 15. The EGA methods had adequate RB values under conditions with a sample size of 200, few items, and low AFL. They demonstrated adequate performance under the conditions where the sample size was 1000 and the AFL was 0.40, and also under conditions where the AFL was 0.70 and the number of items was 15.

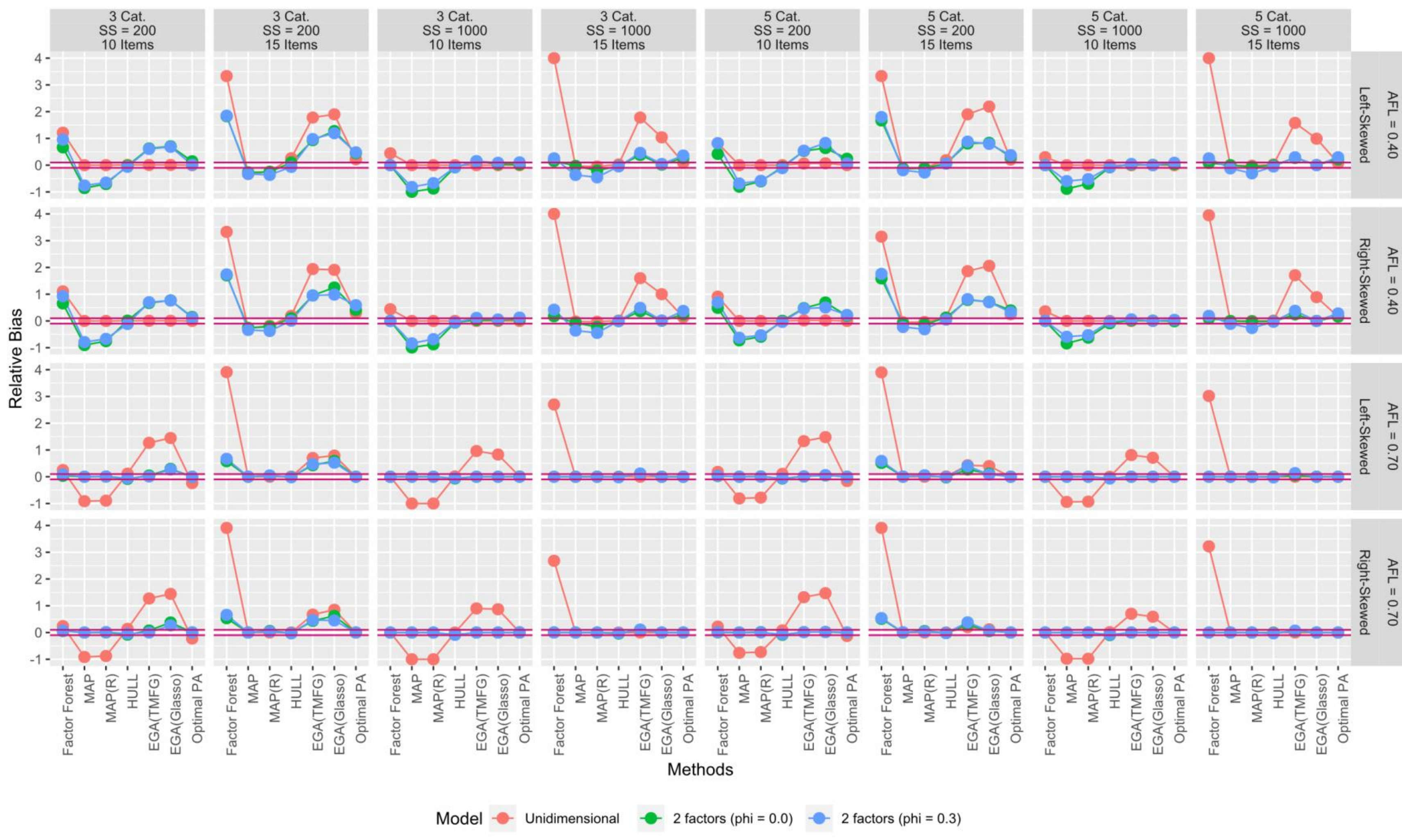

Figure 3. RB Values of the Methods

In two-dimensional structures where the AFL was 0.40, the HULL method had sufficient RB values under most conditions. The Factor Forest method did not perform adequately under any conditions when the sample size was 200. However, it showed better performance under conditions where the number of items was 10 rather than conditions where the sample size was 1000. The revised MAP method did not have sufficient RB values under nearly all specified conditions. MAP, on the other hand, had sufficient RB values under conditions where the sample size was 1000, the number of items was 15, and the interfactor correlation was 0. EGA (TMFG) demonstrated adequate performance under conditions where the sample size was 1000, the number of items was 10, and there was no correlation between dimensions. EGA (Glasso) had sufficient RB values under more conditions compared to EGA (TMFG). EGA (Glasso) had sufficient RB values under all conditions when the sample size was 1000 but did not have sufficient RB values under any of the conditions when the sample size was 200.

4. DISCUSSION AND CONCLUSION

This study examined the performance of factor retantion methods in skewed distributions. In other words, no method yields correct results under any condition. However, when examined in general, the Factor Forest method might be suitable for use in two-factor structures where there are a small number of items per factor and a high sample size. Similarly, in their study, Goretzko and Ruscio (2023) found that the Factor Forest method yielded more biased results in unidimensional structures compared to the comparison data (CD) method. Since the Factor Forest method yielded more inconsistent results than other methods in this research, using this method alone for determining the number of factors may increase the Type I error rate and reduce the test power. Therefore, the suggestion by Goretzko and Ruscio (2023) that this method can be used in conjunction with the CD method could be considered in future studies.

In this study, among the examined simulation conditions, the MAP method ($n_{condition}$ = 56) and the Optimal PA method ($n_{condition}$ = 55) demonstrated adequate performance in more conditions compared to other factor determination methods. These methods were followed by the Revised MAP ($n_{condition}$ = 53), HULL ($n_{condition}$= 51), EGA Glasso ($n_{condition}$= 48), Factor Forest ($n_{condition}$ = 33), and EGA (TMFG; n=31) methods, respectively. However, these results are valid considering all simulation conditions. Researchers can decide on the factor retention method by evaluating simulation conditions close to their own conditions. In this study, the skewness coefficients of the variables were determined to be ±2.5. Although the skewness coefficient was above the upper and lower limits estimated in real data, MAP and optimal PA, which demonstrated adequate performance under more than half of the conditions, were more robust to skewed distributions compared to other methods. This study's results support the literature findings that Optimal PA yields accurate results under challenging conditions (Golino et al., 2020; Nájera et al. 2021).

The EGA (TMFG) and EGA (Glasso) methods cannot be considered suitable factor determination methods for conditions with a sample size of 200. However, they may be suitable for use in larger samples (≥1000) and two-factor structures. The fact that EGA (TMFG) and EGA (Glasso) yield similar results when used for factor determination indicates consistency between these methods, and their combined use may increase the chances of accurately determining the number of factors.

5. RECOMMENDATIONS

The simulation conditions examined in this study are limited. This study examined unidimensional structures as well as two-dimensional structures with interfactor correlations of 0.00 and 0.30. Considering structures with more than two factors in education and psychology, future research could focus on working with data exhibiting skewed distributions involving a greater number of factors and items to compare the performance of different methods. This study skewed the variables by having 3 and 5 categories. In future studies, the researchers could compare the performance of methods using continuous datasets or datasets with dichotomous variables. On the other hand, practitioners could be advised to (i) avoid using the Factor Forest method alone and ii) consider the suggestions of the MAP, Optimal PA and HULL methods. However, it should be noted that these generalizations are limited to datasets exhibiting skewed distributions and simulation conditions.

REFERENCES

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

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

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10*(7), 27–29. https://doi.org/10.7275/jyj1-4868

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

Finch, W. H. (2020). Using Fit Statistic Differences to Determine the Optimal Number of Factors to Retain in an Exploratory Factor Analysis. *Educational and Psychological Measurement*, *80*(2), 217–241. https://doi.org/10.1177/0013164419865769.

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

Foldnes, N., & Grønneberg, S. (2017). The asymptotic covariance matrix and its use in simulation studies. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(6), 881–896. https://doi.org/10.1080/10705511.2017.1341320

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

Garrido, L. E., Abad, F. J., & Ponsoda, V. (2011). Performance of Velicer's minimum average partial factor retention method with categorical variables. *Educational and Psychological Measurement*, *71*(3), 551–570. https://doi.org/10.1177/0013164410389489

Golino, H. F., & Christensen, A. P. (2020). *EGAnet: Exploratory Graph Analysis – A framework for estimating the number of dimensions in multivariate data using network psychometrics*.

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. PLOS ONE, 12(6), 1–26. https://doi.org/10.1371/journal.pone.0174035.

Goretzko, D., & Bühner, M. (2020). One model to rule them all? Using machine learning algorithms to determine the number of factors in exploratory factor analysis. Psychological Methods, 25(6), 776–786.

Goretzko, D., & Bühner, M. (2022). Factor retention using machine learning with ordinal data. *Applied Psychological Measurement*, *46*(5), 406–421. https://doi.org/10.1177/01466216221089345

Goretzko, D., Pham, T. T. H., & Bühner, M. (2021). Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. *Current Psychology*, *40*(7), 3510–3521. https://doi.org/10.1007/s12144-019-00300-2

Goretzko, D., & Ruscio, J. (2023). The comparison data forest: A new comparison data approach to determine the number of factors in exploratory factor analysis. *Behavior Research Methods*. https://doi.org/10.3758/s13428-023-02122-4

Gorsuch, R. L. (1974). *Factor analysis*. W. B. Saunders.

Haslbeck, J. M. B., & van Bork, R. (2022). Estimating the number of factors in exploratory factor analysis via out-of-sample prediction errors. *Psychological methods*, 10.1037/met0000528. Advance online publication. https://doi.org/10.1037/met0000528.

Henson, R., & Roberts, J. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. Educational and Psychological Measurement, 66(3), 393–416.

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

Kılıç, A. F., & Uysal, İ. (2019). Comparison of factor retention methods on binary data: A simulation study. *Turkish Journal of Education*, *8*(3), 160–179. https://doi.org/10.19128/turje.518636

Koyuncu, İ., & Kılıç, A. F. (2019). The use of exploratory and confirmatory factor analyses: A document analysis. *Eğitim ve Bilim*, *44*(198), 361–388. https://doi.org/10.15390/EB.2019.7665

Kılıç, A. F., & Uysal, İ. (2021). Faktör çıkarma yöntemlerinin paralel analiz sonuçlarına etkisi. Trakya Eğitim Dergisi, 11(2), 926–942. https://doi.org/10.24315/tred.747075.

Montoya, A. K., & Edwards, M. C. (2021). The poor fit of model fit for selecting number of factors in exploratory factor analysis for scale evaluation. Educational and Psychological Measurement, 81(3), 413–440. https://doi.org/10.1177/0013164420942899.

Ledesma, R. D., Valero-Mora, P., & Macbeth, G. (2015). The Scree Test and the Number of Factors: a Dynamic Graphics Approach. *The Spanish journal of psychology*, *18*, E11. https://doi.org/10.1017/sjp.2015.13.

Lee, C., Park Y., & Cho B. (2023). Use of exploratory graph analysis in inspecting the dimensionality of the revised University of California Los Angeles (R-UCLA) Loneliness Scale Among Older Adults. Res Gerontol Nurs.16(1):15-20. https://doi.org/10.3928/19404921-20230104-03.

Li, C.-H. (2016). 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, Y., Wen, Z., Hau, K.-T., Yuan, K.-H., & Peng, Y. (2020). Effects of cross-loadings on determining the number of factors to retain. Structural Equation Modeling: A Multidisciplinary Journal, 27(6), 841–863. https://doi.org/10.1080/10705511.2020.1745075

Lorenzo-Seva, U., Timmerman, M. E., & Kiers, H. A. L. (2011). The Hull method for selecting the number of common factors. *Multivariate Behavioral Research*, *46*(2), 340–364. https://doi.org/10.1080/00273171.2011.56452

Lozano, L. M., García-Cueto, E., & Muñiz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology*, *4*(2), 73–79. https://doi.org/10.1027/1614-2241.4.2.73

Montoya, A. K., & Edwards, M. C. (2021). The Poor Fit of Model Fit for Selecting Number of Factors in Exploratory Factor Analysis for Scale Evaluation. *Educational and Psychological Measurement*, *81*(3), 413–440. https://doi.org/10.1177/0013164420942899

Muthén, B. O., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology*, *38*(2), 171–189. https://doi.org/10.1111/j.2044-8317.1985.tb00832.x

Navarro-Gonzalez, D., & Lorenzo-Seva, U. (2021). *EFA.MRFA: Dimensionality assessment using minimum rank factor analysis*. https://cran.r-project.org/package=EFA.MRFA

O'Connor, B. P. (2022). *EFA.dimensions: Exploratory factor analysis functions for assessing dimensionality*. https://cran.r-project.org/package=EFA.dimensions

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.r-project.org

Reio, T. G., & Shuck, B. (2015). Exploratory Factor Analysis: Implications for Theory, Research, and Practice. *Advances in Developing Human Resources*, *17*(1), 12–25. https://doi.org/10.1177/1523422314559804

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

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

Schmitt, T. A., & Sass, D. A. (2011). Rotation Criteria and Hypothesis Testing for Exploratory Factor Analysis: Implications for Factor Pattern Loadings and Interfactor Correlations. *Educational and Psychological Measurement*, *71*(1), 95–113. https://doi.org/10.1177/0013164410387348

Svetina, D. (2011). Assessing dimensionality in complex data structures: A performance comparison of DETECT and NOHARM procedures. Yayınlanmamış doktora tezi, Arizona State Üniversitesi, Arizona.

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.

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

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with non-normal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications*. Sage.

Yang, Y., & Xia, Y. (2015). On the number of factors to retain in exploratory factor analysis for ordered categorical data. *Behavior Research Methods*, *47*(3), 756–772. https://doi.org/10.3758/s13428-014-0499-2

Appendix-1.

The number of categories	Right Skewed (Skewness Coefficient = 2.5)	Left Skewed (Skewness Coefficient = -2.5)
3	$Y=\begin{cases}0, & y_i^*\leq 1\\ 1, & 1<y_i^*\leq 1.80\\ 2, & y_i^*>1.80\end{cases}$	$Y=\begin{cases}0, & y_i^*\leq -1.80\\ 1, & -1.80<y_i^*\leq -1\\ 2, & y_i^*>-1\end{cases}$
5	$Y=\begin{cases}0, & y_i^*\leq 0.75\\ 1, & 0.75<y_i^*\leq 1.28\\ 2, & 1.28<y_i^*\leq 1.645\\ 3, & 1.645<y_i^*\leq 2.05\\ 4, & y_i^*>2.05\end{cases}$	$Y=\begin{cases}0, & y_i^*\leq -2.25\\ 1, & -2.25<y_i^*\leq -1.80\\ 2, & -1.80<y_i^*\leq -1.30\\ 3, & -1.30<y_i^*\leq -0.8\\ 4, & y_i^*>-0.8\end{cases}$

Appendix-2. Accuracy Values of the Methods

Categories	Sample Size	Methods	10 Items												15 Items											
			Average Factor Loadings = 0.40						Average Factor Loadings = 0.70						Average Factor Loadings = 0.40						Average Factor Loadings = 0.70					
			Unidimensional		2 Factors Ψ = 0.00		2 Factors Ψ = 0.30		Unidimensional		2 Factors Ψ = 0.00		2 Factors Ψ = 0.30		Unidimensional		2 Factors Ψ = 0.00		2 Factors Ψ = 0.30		Unidimensional		2 Factors Ψ = 0.00		2 Factors Ψ = 0.30	
			Skewness of the Data																							
			LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS
3	200	M_1	38.4	43.4	33.0	31.0	17.0	19.0	75.4	76.1	93.0	89.0	85.0	87.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	200	M_2	100.0	100.0	5.0	1.0	2.0	0.0	8.7	8.6	100.0	100.0	100.0	100.0	67.0	73.0	55.0	54.0	36.0	36.0	100.0	100.0	100.0	100.0	99.0	100.0
3	200	M_3	100.0	100.0	3.0	5.0	2.0	1.0	10.6	12.2	99.0	99.0	98.0	97.0	76.0	82.0	50.0	58.0	27.0	25.0	100.0	100.0	92.0	89.0	91.0	92.0
3	200	M_4	99.2	99.4	51.5	43.0	38.4	38.1	90.1	87.6	84.0	82.0	88.0	86.0	82.4	83.3	65.0	64.0	47.0	45.0	100.0	100.0	98.0	92.9	96.0	94.0
3	200	M_5	99.4	99.4	15.0	8.0	12.0	6.0	3.9	2.6	91.0	86.0	98.0	100.0	0.0	0.0	2.0	3.0	2.0	2.0	58.0	59.0	32.0	35.0	31.0	35.0
3	200	M_6	99.4	99.4	18.0	17.0	23.0	17.0	5.6	5.5	46.0	44.0	56.0	55.0	1.0	2.0	7.0	9.0	13.0	14.0	57.0	60.0	29.0	29.0	36.0	43.0
3	200	M_7	100.0	99.9	58.0	49.0	48.0	45.0	57.9	58.2	100.0	99.0	100.0	100.0	72.0	70.0	45.0	46.0	34.0	34.0	100.0	100.0	100.0	100.0	100.0	100.0
3	1000	M_1	76.6	76.8	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	0.0	0.0	81.0	73.0	69.0	55.0	0.0	0.0	100.0	99.0	100.0	100.0
3	1000	M_2	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	100.0	100.0	97.0	98.0	94.0	86.0	27.0	28.0	100.0	100.0	100.0	100.0	100.0	100.0
3	1000	M_3	100.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0	100.0	100.0	95.0	97.0	60.0	55.0	9.0	11.0	100.0	100.0	100.0	100.0	100.0	100.0
3	1000	M_4	100.0	100.0	82.0	79.0	83.0	87.0	98.7	100.0	87.0	84.0	91.0	83.0	97.4	98.7	93.0	93.0	87.0	99.0	100.0	100.0	98.0	92.0	96.0	95.0
3	1000	M_5	100.0	100.0	73.0	92.0	72.0	77.0	12.0	23.0	100.0	100.0	100.0	100.0	0.0	0.0	40.0	46.0	35.0	31.0	100.0	100.0	80.0	80.0	78.0	80.0
3	1000	M_6	100.0	100.0	92.0	94.0	83.0	89.0	28.0	28.0	100.0	100.0	100.0	100.0	23.0	28.0	97.0	98.0	92.0	97.0	100.0	100.0	100.0	100.0	100.0	100.0
3	1000	M_7	100.0	100.0	56.0	49.0	76.0	71.0	95.0	97.0	100.0	100.0	100.0	100.0	91.0	86.0	66.0	61.0	44.0	41.0	100.0	100.0	100.0	100.0	100.0	100.0
5	200	M_1	58.0	55.0	48.0	40.0	25.0	30.0	82.0	79.0	97.0	99.0	94.0	97.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0
5	200	M_2	100.0	100.0	3.0	9.0	3.0	3.0	19.0	24.0	100.0	100.0	100.0	100.0	86.0	95.0	75.0	82.0	63.0	55.0	100.0	100.0	99.0	100.0	99.0	98.0
5	200	M_3	100.0	100.0	6.0	10.0	3.0	5.0	22.0	27.0	98.0	97.0	99.0	99.0	90.0	96.0	75.0	74.0	45.0	38.0	100.0	100.0	90.0	89.0	90.0	94.0
5	200	M_4	100.0	100.0	54.0	56.0	36.0	48.0	89.3	93.4	86.7	82.5	89.0	86.7	86.0	89.3	74.0	63.0	60.0	54.0	98.9	100.0	94.9	96.9	96.0	96.0
5	200	M_5	94.0	99.0	19.0	28.0	20.0	24.0	0.0	0.0	98.0	97.0	97.0	97.0	0.0	0.0	4.0	5.0	4.0	2.0	75.0	89.0	54.0	54.0	39.0	40.0
5	200	M_6	94.0	99.0	26.0	22.0	16.0	33.0	3.0	5.0	91.0	96.0	90.0	96.0	2.0	2.0	18.0	22.0	15.0	21.0	78.0	92.0	84.0	91.0	85.0	87.0
5	200	M_7	100.0	100.0	63.0	63.0	46.0	58.0	60.0	69.0	100.0	100.0	100.0	100.0	82.0	80.0	55.0	46.0	46.0	49.0	100.0	100.0	100.0	100.0	100.0	100.0
5	1000	M_1	85.0	82.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	0.0	0.0	87.0	86.0	73.0	78.0	0.0	0.0	100.0	100.0	100.0	100.0

Categories	Sample Size	Methods	10 Items												15 Items											
			Average Factor Loadings = 0.40						Average Factor Loadings = 0.70						Average Factor Loadings = 0.40						Average Factor Loadings = 0.70					
			Unidimensional		2 Factors				Unidimensional		2 Factors				Unidimensional		2 Factors				Unidimensional					
					Ψ = 0.00		Ψ = 0.30				Ψ = 0.00		Ψ = 0.30				Ψ = 0.00		Ψ = 0.30				Ψ = 0.00		Ψ = 0.30	
			Skewness of the Data																							
			LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS
5	1000	M_2	100.0	100.0	1.0	2.0	0.0	0.0	6.0	2.0	100.0	100.0	100.0	100.0	99.0	100.0	100.0	99.0	76.0	78.0	100.0	100.0	100.0	100.0	100.0	100.0
5	1000	M_3	100.0	100.0	1.0	0.0	0.0	0.0	7.0	2.0	100.0	100.0	100.0	100.0	98.0	100.0	87.0	96.0	39.0	47.0	100.0	100.0	100.0	100.0	100.0	100.0
5	1000	M_4	100.0	100.0	82.0	76.0	82.0	90.0	100.0	98.6	89.0	80.8	87.0	85.0	97.4	98.8	97.0	90.0	90.0	92.0	100.0	100.0	99.0	97.0	95.0	94.0
5	1000	M_5	100.0	100.0	94.0	96.0	92.0	88.0	26.0	37.0	100.0	100.0	100.0	100.0	3.0	2.0	55.0	61.0	53.0	39.0	100.0	100.0	89.0	89.0	76.0	87.0
5	1000	M_6	100.0	100.0	98.0	98.0	100.0	99.0	36.0	47.0	100.0	100.0	100.0	100.0	25.0	31.0	99.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
5	1000	M_7	100.0	100.0	66.0	67.0	78.0	90.0	100.0	96.0	100.0	100.0	100.0	100.0	93.0	84.0	67.0	75.0	52.0	55.0	100.0	100.0	100.0	100.0	100.0	100.0

M_1: Factor Forest, M_2: MAP, M_3: MAP(R), M_4: HULL, M_5: EGA(TMFG), M_6: EGA(Glasso), M_7: Optimal PA

Appendix-3. RB of the Methods

			10 Items												15 Items											
			Average Factor Loadings = 0.40						Average Factor Loadings = 0.70						Average Factor Loadings = 0.40						Average Factor Loadings = 0.70					
			Unidimensional		2 Factors				Unidimensional		2 Factors				Unidimensional		2 Factors				Unidimensional		2 Factors			
					Ψ = 0.00		Ψ = 0.30				Ψ = 0.00		Ψ = 0.30				Ψ = 0.00		Ψ = 0.30				Ψ = 0.00		Ψ = 0.30	
			Skewness of the Data																							
Categories	Sample Size	Methods	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS	LS	RS
3	200	M_1	1.2	1.1	0.7	0.7	1.0	0.9	0.2	0.2	0.0	0.1	0.1	0.1	3.3	3.3	1.8	1.7	1.8	1.7	3.9	3.9	0.6	0.5	0.7	0.7
3	200	M_2	0.0	0.0	-0.9	-0.9	-0.8	-0.8	-0.9	-0.9	0.0	0.0	0.0	0.0	-0.3	-0.3	-0.3	-0.3	-0.3	-0.3	0.0	0.0	0.0	0.0	0.0	0.0
3	200	M_3	0.0	0.0	-0.7	-0.8	-0.6	-0.7	-0.9	-0.9	0.0	0.0	0.0	0.0	-0.2	-0.2	-0.3	-0.2	-0.4	-0.4	0.0	0.0	0.0	0.1	0.0	0.0
3	200	M_4	0.0	0.0	-0.0	0.0	-0.1	-0.1	0.1	0.1	-0.1	-0.1	-0.0	-0.0	0.3	0.2	0.1	0.1	-0.1	0.0	0.0	0.0	-0.0	-0.0	-0.0	-0.0
3	200	M_5	0.0	0.0	0.6	0.7	0.6	0.7	1.3	1.3	0.0	0.1	0.0	0.0	1.8	1.9	0.9	1.0	1.0	0.9	0.7	0.7	0.4	0.4	0.5	0.5
3	200	M_6	0.0	0.0	0.7	0.8	0.7	0.8	1.4	1.4	0.3	0.4	0.3	0.3	1.9	1.9	1.3	1.2	1.2	1.0	0.8	0.8	0.6	0.6	0.5	0.4
3	200	M_7	0.0	0.0	0.1	0.1	0.0	0.1	-0.2	-0.2	0.0	0.0	0.0	0.0	0.2	0.3	0.4	0.4	0.5	0.6	0.0	0.0	0.0	0.0	0.0	0.0
3	1000	M_1	0.4	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	4.0	0.2	0.2	0.3	0.4	2.7	2.7	0.0	0.0	0.0	0.0
3	1000	M_2	0.0	0.0	-1.0	-1.0	-0.8	-0.8	-1.0	-1.0	0.0	0.0	0.0	0.0	-0.0	-0.0	-0.0	-0.1	-0.4	-0.4	0.0	0.0	0.0	0.0	0.0	0.0
3	1000	M_3	0.0	0.0	-0.9	-0.9	-0.7	-0.7	-1.0	-1.0	0.0	0.0	0.0	0.0	-0.0	-0.0	-0.2	-0.2	-0.5	-0.4	0.0	0.0	0.0	0.0	0.0	0.0
3	1000	M_4	0.0	0.0	-0.1	-0.1	-0.1	-0.1	0.0	0.0	-0.1	-0.1	-0.0	-0.1	0.0	0.0	-0.0	-0.0	-0.0	-0.0	0.0	0.0	-0.0	-0.0	-0.0	-0.0
3	1000	M_5	0.0	0.0	0.1	0.0	0.1	0.1	1.0	0.9	0.0	0.0	0.0	0.0	1.8	1.6	0.4	0.4	0.5	0.5	0.0	0.0	0.1	0.1	0.1	0.1
3	1000	M_6	0.0	0.0	0.0	0.0	0.1	0.1	0.8	0.9	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	1000	M_7	0.0	0.0	0.0	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.3	0.2	0.4	0.4	0.0	0.0	0.0	0.0	0.0	0.0
5	200	M_1	0.8	0.9	0.4	0.5	0.8	0.7	0.2	0.2	0.0	0.0	0.0	0.0	3.3	3.1	1.7	1.6	1.8	1.8	3.9	3.9	0.5	0.5	0.6	0.5
5	200	M_2	0.0	0.0	-0.8	-0.7	-0.7	-0.6	-0.8	-0.8	0.0	0.0	0.0	0.0	-0.1	-0.0	-0.1	-0.1	-0.2	-0.2	0.0	0.0	0.0	0.0	0.0	0.0
5	200	M_3	0.0	0.0	-0.6	-0.6	-0.6	-0.5	-0.8	-0.7	0.0	0.0	0.0	0.0	-0.1	-0.0	-0.1	-0.1	-0.3	-0.3	0.0	0.0	0.0	0.1	0.0	0.0
5	200	M_4	0.0	0.0	-0.1	0.0	-0.1	-0.0	0.1	0.1	-0.1	-0.1	-0.1	-0.1	0.2	0.1	0.1	0.1	0.1	0.0	0.0	0.0	-0.0	-0.0	-0.0	-0.0
5	200	M_5	0.1	0.0	0.5	0.5	0.5	0.5	1.3	1.3	0.0	0.0	0.0	0.0	1.9	1.9	0.8	0.8	0.9	0.8	0.4	0.2	0.3	0.3	0.4	0.4
5	200	M_6	0.1	0.0	0.6	0.7	0.8	0.5	1.5	1.5	0.0	0.0	0.1	0.0	2.2	2.1	0.8	0.7	0.8	0.7	0.4	0.1	0.1	0.0	0.1	0.1
5	200	M_7	0.0	0.0	0.2	0.2	0.1	0.2	-0.2	-0.1	0.0	0.0	0.0	0.0	0.2	0.2	0.3	0.4	0.4	0.3	0.0	0.0	0.0	0.0	0.0	0.0
5	1000	M_1	0.3	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	4.0	0.1	0.1	0.3	0.2	3.0	3.2	0.0	0.0	0.0	0.0
5	1000	M_2	0.0	0.0	-0.9	-0.8	-0.6	-0.6	-0.9	-1.0	0.0	0.0	0.0	0.0	-0.0	0.0	0.0	-0.0	-0.1	-0.1	0.0	0.0	0.0	0.0	0.0	0.0
5	1000	M_3	0.0	0.0	-0.7	-0.6	-0.5	-0.5	-0.9	-1.0	0.0	0.0	0.0	0.0	-0.0	0.0	-0.1	-0.0	-0.3	-0.3	0.0	0.0	0.0	0.0	0.0	0.0
5	1000	M_4	0.0	0.0	-0.1	-0.1	-0.1	-0.0	0.0	0.0	-0.1	-0.1	-0.1	-0.1	0.0	0.0	-0.0	-0.0	-0.0	-0.0	0.0	0.0	-0.0	-0.0	-0.0	-0.0
5	1000	M_5	0.0	0.0	0.0	0.0	0.0	0.1	0.8	0.7	0.0	0.0	0.0	0.0	1.6	1.7	0.3	0.2	0.3	0.4	0.0	0.0	0.1	0.1	0.1	0.1
5	1000	M_6	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.6	0.0	0.0	0.0	0.0	1.0	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	1000	M_7	0.0	0.0	0.0	-0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.2	0.2	0.3	0.3	0.0	0.0	0.0	0.0	0.0	0.0

M_1: Factor Forest, M_2: MAP, M_3: MAP(R), M_4: HULL, M_5: EGA(TMFG), M_6: EGA(Glasso), M_7: Optimal PA

GENİŞLETİLMİŞ TÜRKÇE ÖZET

ÇARPIK DAĞILIMLARDA FAKTÖR SAYISI BELİRLEME YÖNTEMLERİNİN PERFORMANSLARININ İNCELENMESİ

GİRİŞ

Açımlayıcı faktör analizinde (AFA), faktör sayısına karar vermek en önemli adımlardan biridir (Cosemans & et al., 2022; Finch, 2020; Reio & Shuck, 2015; Svetina, 2011; Zhang, 2007). AFA'da olması gerekenden az faktör çıkarmak değişkenleri daha küçük bir faktör uzayına sıkıştırır ve bu durum da bilgi kaybına, önemli faktörlerin ihmal edilmesine ve artan hata yüklerine neden olur (Cosemans vd., 2022). Olması gerekenden fazla faktör çıkarmak ise aslında bir arada olan faktörlerin bölünmesine veya önemsiz faktörlere neden olabilir (Cosemans & et al., 2022; Finch, 2020; Lee & et al, 2023). Bu nedenle boyut sayısına karar vermede kullanılacak kriterler de önemli hale gelmektedir.

Bir ölçme aracının boyutluluğuna karar verirken sadece bir yönteme göre karar vermek de problemli olabilir (Ledesma vd., 2015; Lee, 2023). Her bir yöntemin kendine ait üstün ve zayıf yanları bulunmaktadır. Bu nedenle hangi yöntemin verinin hangi koşulunda iyi sonuçlar verdiğinin incelenmesi önemli hale gelmektedir. Bu durumda da hangi yöntemlerin inceleneceği sorusu ortaya çıkabilir. Boyutluluk belirleme yöntemleri incelendiğinde genellikle Horn (1965) tarafından önerilen paralel analiz (PA) yönteminin popüler olduğu ancak bu yöntemin yanında Minimum Average Partial Correlation (MAP), HULL (Lorenzo-Seva vd., 2011) veya son zamanlarda daha sık kullanılmaya başlanan Açımlayıcı Grafikl Analizi (EGA) yöntemleri bulunmaktadır. Makine öğrenmesi yöntemlerinin birçok alanda uygulama bulmasıyla birlikte boyutluluk belirleme yöntemi olarak kullanımını öneren araştırmacılar da olmuştur. (Goretzko & Ruscio, 2023).

Bu çalışmada boyutluluk belirleme yöntemlerinden optimal paralel analiz (Timmerman & Lorenzo-Seva, 2011), MAP (normal and revize), HULL, EGA(TMFG kestirimiyle), EGA(Glasso kestirimiyle) ve Goretzko & Ruscio (2023) tarafından önerilen karşılaştırmalı faktör ormanı (comparison factor forest) yöntemi karşılaştırılmıştır. Bu doğrultuda bu araştırmanın amacı faktör sayısı belirleme yöntemlerinin çeşitli simülasyon koşulları altında performanslarını değerlendirmektir. Araştırma çerçevesinde her bir koşulun temel etkisinin yanı sıra koşulların etkileşiminin etkisini karşılaştırılmıştır. Bu amaç doğrultusunda şu sorulara yanıt aranmıştır?

- Simülasyon koşullarına göre faktör sayısı belirleme yöntemlerinin doğru tahmin yüzdesi değerleri nasıldır?
- Simülasyon koşullarına göre faktör sayısı belirleme yöntemlerinin göreli yanlılık değerleri nasıldır?

YÖNTEM

Boyutluluk belirleme yöntemlerinin karşılaştırıldığı bu çalışma bir Monte Carlo simülasyonudur. Simülasyon çalışmalarında istenilen özelliklere (dağılım, faktör yükü ya da madde sayısı gibi) göre üretilen veri setleri ilgilenilen yöntemlerle analiz edilerek sonuçları karşılaştırılır.

Simülasyon Koşulları

Çalışmada simülasyon faktörleri olarak dağılım, örneklem büyüklüğü, faktör başına düşen madde sayısı, kategori sayısı ve ölçme modeli belirlenmiştir. Çalışmada her bir koşul için 100 replikasyon yapılmıştır.

Simülasyon koşullarından verilerin dağılımı koşulunda veri seti çarpıklık katsayısı ±2.5 olacak şekilde çarpık hale getirilmiştir. Örneklem büyüklüğü için 200 ve 1000 koşulları belirlenmiştir. Kategori sayısı koşulunda 3 ve 5 kategori bulunmaktadır. İncelenen model koşulunda tek boyutlu, 2 faktör ve faktörler arası korelasyon 0.00 ve 2 faktörlü faktörler arası korelasyonun 0.30 olduğu koşulları incelenmiştir. Ortalama faktör yükü 0.40 ve 0.70 olarak manipüle edilmiştir.

Veri Analizi

Verileri üretmek için için R paket programında (R Core Team, 2022) bulunan lavaan paketi (Rosseel, 2012) kullanılmıştır. Optimal paralel analiz ve HULL yöntemi için EFA.MRFA paketi (Navarro-Gonzalez & Lorenzo-Seva, 2021), MAP analizi için EFA.dimensions (O'Connor, 2022), EGA yöntemleri için EGAnet paketi (Golino & Christensen, 2020) kullanılmıştır. Karşılaştırmalı faktör ormanı (comparison factor forest) yöntemi için ise Goretzko & Bühner (2022) tarafından paylaşılan kodlar kullanılmıştır.

TARTIŞMA, SONUÇ VE ÖNERİLER

Çarpık dağılımlarda faktör sayısı belirleme yöntemlerinin performanslarının incelendiği bu çalışma sonucunda tüm koşulların doğru kestirim yüzdesi değerlerinin ortalaması dikkate alındığında en yüksek ortalamaya HULL yönteminin sahip olduğu görülmüştür. Aynı zamanda en düşük göreli yanlılık (RB) ortalaması da HULL yöntemindedir. Ancak tüm koşullarda yeterli performansı gösteren bir yöntemin olmadığı söylenebilir. Diğer bir deyişle her koşulda doğru sonucu verecek bir yöntem bulunmamaktadır. Ancak genel olarak incelendiğinde factor forest yönteminin faktör başına düşen madde sayısının az ve örneklem büyüklüğünün yüksek olduğu iki faktörlü yapılarda kullanılmasının uygun olabileceği söylenebilir. Benzer şekilde Goretzko & Ruscio (2023) yaptığı çalışmada factor forest yöntemi karşılaştırmalı veri (comparison data [CD]) yöntemine göre tek boyutlu yapılarda daha yanlı sonuçlar göstermiştir. Factor forest yönteminin bu araştırmada diğer yöntemlerle daha tutarsız sonuçlar vermesi nedeniyle tek başına faktör sayısı belirleme yöntemi olarak kullanılmasının 1. tip hata oranını arttıracağı ve testin gücünü azaltacağı düşünülmektedir. Bu nedenle Goretzko & Ruscio (2023) tarafından bu yöntemin CD yöntemiyle birlikte kullanılabileceği önerisi bundan sonraki çalışmalarda dikkate alınabilir.

Bu çalışmada diğer faktör sayısı belirleme yöntemlerine göre, MAP (n=56) ve optimal PA (n=55) yöntemleri incelenen simülasyon koşulları içinde diğer yöntemlere göre daha çok koşulda yeterli performansı göstermiştir. Bu yöntemleri ise; Revize MAP (n=53), HULL (n=51), EGA(Glasso) (n=48), factor forest (n=33) ve EGA(TMFG) (n=31) yöntemleri izlemektedir. Bu çalışmada değişkenlerin çarpıklık katsayısı ±2.5 olacak şekilde belirlenmiştir. Çarpıklık katsayısının gerçek verilerde tahmin edilen alt ve üst sınırın da üzerinde olmasına rağmen koşulların yarısından fazlasında yeterli performansı gösteren MAP ve optimal PA'nın çarpık dağılımlara diğer yöntemler göre daha dayanıklı olduğu söylenebilir. Bu çalışmanın sonuçları, Optimal PA'nın zorlu koşullar altında doğru sonuçlar verdiği alanyazındaki bulguları (Golino vd., 2020; Nájera vd. 2021) desteklemektedir. EGA(TMFG) ve EGA(Glasso) yöntemleri örneklem büyüklüğünün 200 olduğu koşullar için uygun bir faktör belirleme yöntemi olduğu söylenemez ancak büyük örneklemler (n≥1000) ve iki faktörlü yapılarda kullanılabilecek bir yöntemdir. EGA(TMFG) ve EGA(Glasso) faktör sayısı belirleme yöntemleri kullanılırken benzer sonuçlar vermesi yöntemler arasındaki uyum anlamına gelir ve birlikte kullanımı faktör sayısını doğru belirleme şansını arttırabilir. Gelecekteki çalışmalarda sürekli veri setleri ya da iki kategorili verilerle yöntemlerin performanslarının karşılaştırılması araştırmacılara önerilebilir. Diğer taraftan uygulayıcılara; i) tek başına factor forest yöntemini kullanmamaları, ii) optimal PA ve HULL yöntemlerinin önerilerini daha fazla önem vermeleri önerilebilir.