


## Comparison of factor retention methods on binary data: A simulation study

Abdullah Faruk Kılıç 

Hacettepe University, Ankara, Turkey, abduhfarukkilic@gmail.com

İbrahim Uysal 

Bolu Abant İzzet Baysal University, Bolu, Turkey, ibrahimuysal@ibu.edu.tr



**ABSTRACT** In this study, the purpose is to compare factor retention methods under simulation conditions. For this purpose, simulation conditions with a number of factors (1, 2 [simple]), sample sizes (250, 1.000, and 3.000), number of items (20, 30), average factor loading (0.50, 0.70), and correlation matrix (Pearson Product Moment [PPM] and Tetrachoric) were investigated. For each condition, 1.000 replications were conducted. Under the scope of this research, performances of the Parallel Analysis, Minimum Average Partial, DETECT, Optimal Coordinate, and Acceleration Factor methods were compared by means of the percentage of correct estimates, and mean difference values. The results of this study indicated that MAP analysis, as applied to both tetrachoric and PPM correlation matrices, demonstrated the best performance. PA showed a good performance with the PPM correlation matrix, however, in smaller samples, the performance of the tetrachoric correlation matrix decreased. The Acceleration Factor method proposed one factor for all simulation conditions. For unidimensional constructs, the DETECT method was affected by both the sample size and average factor loading.

**Keywords:** *Exploratory factor analysis, Factor retention, Parallel analysis, Minimum average partial, DETECT*

## Faktör sayısını belirleme yöntemlerinin karşılaştırılması: Bir simülasyon çalışması

**ÖZ** Bu çalışmada faktör sayısının belirlenmesi amacıyla geliştirilen yöntemlerin simülasyon koşulları altında karşılaştırılması amaçlanmıştır. Bu amaç için faktör sayısı (1, 2 [basit]), örneklem büyüklüğü (250, 1000 ve 3000), madde sayısı (20, 30), ortalama faktör yükü (0.50, 0.70) ve kullanılan korelasyon matrisi (Pearson Momentler Çarpımı [PPM] ve Tetrakorik) simülasyon koşulu olarak araştırılmıştır. Her bir koşul için 1000 replikasyon yapılmış ve üretilen 24000 veri seti için PPM ve tetrakorik korelasyon matrisi üzerinden analizler gerçekleştirilmiştir. Araştırma kapsamında Paralel Analiz, Kısmi Korelasyonların En Küçüğü, DETECT, Optimal Koordinat ve İvmelenme Faktörü yöntemlerinin performansları doğru kestirim yüzdesi ve ortalama fark değerleri üzerinden karşılaştırılmıştır. Araştırma sonucunda hem tetrakorik hem de PPM korelasyon matrisiyle yürütülen MAP analizi en iyi performansı göstermiştir. PA da PPM korelasyon matrisiyle iyi performans göstermiş ancak küçük örnekleme tetrakorik korelasyon matrisiyle performansı düşmüştür. DETECT yöntemi tek boyutlu yapılarla örneklem büyüklüğü ve ortalama faktör yükünden etkilenmiştir.

**Anahtar Kelimeler:** *Açımlayıcı faktör analizi, Faktör sayısını belirleme, Paralel analiz, Kısmi korelasyonların en küçüğü, DETECT*

**Citation:** 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. DOI: 10.19128/turje.518636

## INTRODUCTION

When a construct of data set needs to be analyzed, factor analysis is one of the most common psychometric techniques (Osborne & Banjanovic, 2016). If researchers lack information regarding the construct, explanatory factor analysis (EFA) or unconstrained factor analysis is applied. If researchers have information regarding the construct, and if the compliance of the data sets with the construct in question is to be analyzed, confirmatory factor analysis (CFA) or constrained factor analysis is applied (Fabrigar & Wegener, 2012).

When EFA is applied, one of the obstacles is the determination of the number of dimensions. There are various methods to decide on the number of dimensions. For example, there is a method called the Kaiser K1, which accepts an eigenvalue number higher than 1 as the number of the dimension (Kaiser, 1960). In addition to this method, there is Parallel Analysis (PA), as proposed by Horn (1965), Minimum Average Partial test (MAP), as proposed by Velicer (1976), and the Scree Test, as proposed by Cattell (1966). The Scree Test has some limitations in terms of non-graphical solutions, and there are alternatives. Such methods include multiple regression, a t-value index, and the standard error of scree (SEscree) approach, as proposed by Zoski and Jurs (1993, 1996), the Scree Test Optimal Coordinate ( $n_{OC}$ ), and the Scree Test Acceleration Factor ( $n_{AF}$ ), as proposed by Raïche, Walls, Magis, Riopel, and Blais (2013). In addition to these methods, there is the NOHARM (Fraser & McDonald, 1988) method, which is based on latent trait theory and nonlinear harmonic approximations of the normal ogive error distribution. Additionally, DIMTEST (Nandakumar & Stout, 1993; Stout, 1987) and DETECT (Zhang & Stout, 1999) are non-parametric methods based on conditional covariances. Under the scope of this research, information about PA, MAP, DETECT, Optimal Coordinate ( $n_{OC}$ ), and Acceleration Factor ( $n_{AF}$ ) methods have been provided in detail.

Parallel Analysis (PA) determines the number of factors by generating a random variable for sample size N, and a p variable (Horn, 1965). In PA, eigenvalues obtained from the data set are compared with eigenvalues obtained from independent normal variables. Data sets, as well as the number of variables produced in PA are the same size as the researched data set. The number of factors in PA is decided by comparing the average eigenvalue obtained from the independent variables and eigenvalues obtained from the data set. The number of factors was proposed if the eigenvalue obtained from the data set was bigger than the mean of those obtained from the random uncorrelated data. (Ledesma & Valero-Mora, 2007). In PA, instead of this average eigenvalue, the median of an eigenvalue, or the 5th, 95th, or 99th percentile of an eigenvalue can be used (Cota, Longman, Holden, Fekken, & Xinaris, 1993; Glorfeld, 1995; Raïche et al., 2013). As Cota et al. (1993) proposed the use of the 95th percentile, in this research the 95th percentile was adopted.

The Minimum Average Partial (MAP) test was developed by Velicer (1976). The MAP test is based on principle component analysis (PCA). In an MAP test, after the PCA, a partial correlation matrix is formed. At first stage, the first fundamental components are separated from the correlation matrix obtained from the variables in the data set. Squares of off diagonal elements in the correlation matrix are calculated, and a partial correlation matrix is formed. At the second stage, two fundamental components are separated from the correlation matrix obtained from the variables in the data set. Squares of off diagonal elements in the correlation matrix are calculated, and, again, a partial correlation matrix is formed. This process is applied until one minus number of variables is attained. Average squared partial correlations obtained from these steps are ranged. The number of dimensions is defined as the number of steps necessary to analyze the smallest average squared partial correlation (Ledesma & Valero-Mora, 2007; O'connor, 2000; Velicer, 1976). Due to calculation, it is stated that, for each factor, the factor loading of at least two variables should be high (Zwick & Velicer, 1986).

The DETECT method is based on finding a positive conditional covariance between items that measure the same dimension and a negative conditional covariance between items that measure different dimensions (Zhang & Stout, 1999). This method can be applied with both a confirmatory and exploratory mode. With the confirmatory mode, the DETECT index is calculated over user-defined sections. With the exploratory mode, the DETECT method searches for partitions that will make the DETECT index maximum. The index calculated as a result of this method provides information regarding the multidimensionality of the test (Jang & Roussos, 2007; Kim, 1996). If the DETECT index is higher than 1.00, it can be interpreted as exhibiting a strong multidimensionality; if it is between 0.40 - 1.00, it can be interpreted as exhibiting a moderate multidimensionality; if it is between 0.20 - 0.40, it can be interpreted as exhibiting a weak multidimensionality, and if it is smaller than 0.20, it can be interpreted as exhibiting essential unidimensionality (Jang & Roussos, 2007; Zhang, 2007). In this research, the criteria selected was thus: smaller than 0.20 for one dimensional constructs, and larger than 1 for two dimensional constructs. For one dimensional constructs, if the DETECT index value is smaller than 0.20, it is accepted as one dimensional; for a DETECT index higher than 0.20, the construct was evaluated as multi-dimensional. For two dimensional constructs, values larger than 1 were considered. Thus, this research followed a more conservative approach.

In the Scree Plot method, Cattell (1966) combines a previous eigenvalue coordinate and the next eigenvalue coordinate with a trace line to determine the location of the screen. Thus, a point with immediate change is determined as an important number of a factor. In the Scree Test Optimal Coordinate ( $n_{OC}$ ) method, without any statistical test, the focal point is the elbow. For that, a two-point regression model is used. In this method, dimension size is decided based on if estimated eigenvalues from the regression model and observed eigenvalues are equal or higher. For a comparison of values (between observed and estimated eigenvalues), the K1 rules is adopted and eigenvalues higher than 1 are compared. If desired, the average eigenvalues obtained from the PA result or eigenvalue at 0.05, 0.95 quantile can be used (Raïche et al., 2013). This method can be expressed as:

$$n_{OC} = \sum_i I[(\lambda_i \geq 1) \& (\lambda_i \geq \hat{\lambda}_i)] \quad (1)$$

Here,  $\lambda_i$  represents observed eigenvalue, and  $\hat{\lambda}_i$  represents estimated eigenvalue.  $\hat{\lambda}_i$  Additionally, is named as the optimal coordinate (Raïche et al., 2013).

In the Scree Test Acceleration Factor ( $n_{AF}$ ) method, the acceleration factor represents the point on the coordinate where the slope of the curve changes abruptly. The value presented in Eq.1 and represented as  $\hat{\lambda}_i$  is calculated by taking a second derivative of the optimal coordinate. For more detailed information, Raïche et al., (2013) can be read. This method is expressed as:

$$n_{AF} = \sum_i I(\lambda_i \geq 1 \& i < k) \text{ with } k \equiv \operatorname{argmax}_j (af_j) \quad (2)$$

Here,  $\operatorname{argmax}_j (af_j)$  represents the maximum point of the second derivative of an optimal coordinate function (Raïche et al., 2013).

When the literature was reviewed, various studies regarding Parallel Analysis were identified (Buja & Eyuboglu, 1992; Cho, Li, & Bandalos, 2009; Cota et al., 1993; Dinno, 2014; Glorfeld, 1995; Green, Levy, Thompson, Lu, & Lo, 2012; Green, Redell, Thompson, & Levy, 2016; Green, Thompson, Levy, & Lo, 2015; Kaya Kalkan & Kelecioğlu, 2016; Raïche et al., 2013; Weng & Cheng, 2005; Xia, Green, Xu, & Thompson, 2019; Yang & Xia, 2015; Zwick & Velicer, 1986). It was observed that these studies mostly focused on parallel analysis processes rather than comparison of performance of factor retention methods. In addition, most of these studies were conducted on continuous or ordinal data. There was limited researches with binary data (Kaya Kalkan & Kelecioğlu, 2016; Weng & Cheng, 2005). Current study focused to binary data which mostly used to educational and social researches. It can be said that

the current study will contribute to the literature in this respect. In addition, there were studies about DIMTEST; however, in these studies, DIMTEST was often compared with NOHARM (Finch & Habing, 2005, 2007). But current study compares five factor retention methods (MAP, PA, optimal coordinate, DETECT and acceleration factor). Additionally, there were no studies comparing the methods adopted in the current study. For example Garrido, Abad, and Ponsoda (2011) examine only MAP procedure under different simulation conditions. But current study compares different factor retention methods under simulation conditions which frequently encountered in educational research. The current study differs from the others in the literature in terms of comparison of previously not compared factor retention methods. In other words, current study is important in terms of comparing methods not previously examined. Therefore, it is believed that the current research will contribute to the literature and the EFA that is commonly used in empirical studies (Conway & Huffcutt, 2003; Henson & Roberts, 2006) will be instructive to determine the number of factors.

## **METHODOLOGY**

As predictions can be made by considering probabilities (Gilbert, 1999), this research was designed as a Monte Carlo simulation study. Simulation studies provide advantages, as such studies enable a comparison between real parameters and estimated parameters (Feinberg & Rubright, 2016). This is because real parameters are unknown in empirical studies. Therefore, analyses are carried out on data sets of known real parameters and the real and estimated parameters are compared.

### **Simulation Conditions**

In this research, the simulation factors were identified as sample size, number of factors, test length, average factor loading, and type of correlation matrix. A full crossed pattern was used in this study. Since the focal point of this research was the tests that achieved a score of 1-0, binary data set were generated.

For the sample size factor, three conditions such as 250, 1000, and 3000 were determined. For the sample size conditions, a sample size of 250 was selected, as this size is both applicable and commonly used in studies (Beauducel & Herzberg, 2006; Hu, Bentler, & Kano, 1992). When 1000 people are selected as the sample size, the aim is to satisfy the sample size proposed in different studies for factor analysis (Comrey & Lee, 1992; Floyd & Widaman, 1995; Gorsuch, 1974; Guadagnoli & Velicer, 1988; Streiner, 1994). Generally, since estimated parameters rarely change after a sample size of 3000 (DeMars, 2010), this sample size was added to the research. Additionally, the effect of a sample size of more than 1000 people was analyzed to determine the number of factors acting on sample factor retention methods.

For number of factors, two conditions were included in this research. Here, the fact that achievement tests generally consisted of one dimension was considered (Bennett et al., 1990; Bennett, Rock, & Wang, 1991; Lissitz, Hou, & Slater, 2012; van den Bergh, 1990), and one-dimensional constructs were included in the research. However, it is impossible to analyze the performance of factor retention methods through the use of two-dimensional constructs. In addition, since it is possible that a method that operates with full accuracy in a unidimensional construct may fail to show the expected performance in a two dimensional one, two dimensional constructs were also included in this research. Yet, the relationship between these two dimensions was set around 0. Thus, the number of factor conditions was investigated as one and two (simple).

For the test length factor, 20 and 30 item conditions were determined, respectively. Turkish, mathematics and science tests consist of 20 items are used in the central examinations conducted by the

Ministry of Education in Turkey (MEB, 2019). Since it is rare for an achievement test with more than 30 items in the focal point, the research was limited to a 20 and 30 items condition.

The average factor loading was determined as 0.50 and 0.70. An average factor loading value below 0.50 is relatively rare for binary tests. This is because the lowest factor loading is proposed as 0.33 (Tabachnik & Fidell, 2012). When the factor loadings of other items were considered, a factor loading of 0.50 was added to the simulation conditions as the lower limit. Through an average factor loading of 0.70, the effect of a high relation between items and the factor on the factor retention method was evaluated. Attaining the same loading for all items was disregarded as a goal in this research. This is because the same factor loading for all items is rare. Therefore, the factor loadings of all items could differentiate, with only the average factor loading being kept within the desired conditions (e.g. for a 20-item test, the total factor loading was found, and an average factor loading was subsequently calculated by dividing by 20).

In this research, analyses were conducted for methods using correlation matrices (MAP and PA), including both the Pearson Product Moment (PPM) correlation matrix and the tetrachoric correlation matrix. Thus, the means by which such methods produce results according to correlation matrices was examined. Via this approach, the over factoring or under factoring effects of correlation matrices on methods were analyzed. In Table 1, the simulation factors and conditions are presented.

Table 1.  
*Simulation factor and conditions*

Constant factors			Changing Factors		
Data Type	Number of factors	Sample Size	Test Length	Average Factor Loading	Correlation Matrix
1-0	1	250	20	0.50	PPM
	2 (simple)	1000 3000	30	0.70	Tetrachoric

As seen from Table 1, a fully crossed pattern was used in this research. Accordingly, a total of  $2 \times 3 \times 2 \times 2 = 24$  conditions were evaluated: namely, the number of factors (2 conditions), sample size (3 conditions), test length (2 conditions), and average factor loading (2 conditions). For this purpose, a 24.000 data set was generated with 1.000 replications. Harwell, Stone, Hsu, and Kirisci (1996) proposed at least 25 replications. Robey and Barcikowski (1992) propose a formula for calculating the number of replications. The number of replications was determined to be 575 from the table created in accordance with this formula for research conditions. In order to increase the power of the research, 1.000 replicates were produced for this research. Both PPM and tetrachoric correlation matrix analysis on methods using correlation matrices were conducted for the same data set. However, no additional data set was produced for these conditions.

### Data Generation and Analysis

Binary (1-0) data was produced with Psych (Revelle, 2016) package in R program (R Core Team, 2018). For each simulation condition, 1.000 replications were applied. For this purpose, the first 1.000 seed values were randomly generated, and, for these seed values, data was generated for the simulation conditions.

The analyses of the generated data were analyzed using different R packages. For parallel analysis and MAP analysis, Psych (Revelle, 2016) package was used; for the DETECT method, sirt (Robitzsch, 2017) package was used; for the Acceleration Factor ( $n_{AF}$ ) and Optimal Coordinate ( $n_{OC}$ ) methods, nFactors (Raiche, 2010) package was used.

For Parallel Analysis (PA), principal factor solution was used as factoring method. The number of randomly generated correlation matrices in PA was determined as 50. Since the cut-off score in the

DETECT method was suggested as 0.20 (Jang & Roussos, 2007; Zhang, 2007), in this research, DETECT values higher than 0.20 for one dimensional data sets, and DETECT value higher than 1 for two dimensions were accepted as multi-dimensional. Thus, the test evaluation followed a more conservative approach. To evaluate performance of methods, both a real and proposed number of dimension were compared and the percentage of correct estimates (PCE) were obtained. For this purpose:

$$P_r = \begin{cases} 1 & \text{if Suggested Number of Dimensions} = \text{Actual Number of Dimensions} \\ 0 & \text{if Suggested Number of Dimensions} \neq \text{Actual Number of Dimensions} \end{cases} \quad (3)$$

function was used. In this function, r represented replication. Accordingly, the calculated percentage of correct estimates (PCE) can be given with;

$$\text{Percentage of Correct Estimates} = \frac{\sum_{r=1}^{1000} P_r}{1000} \cdot 100 \quad (4)$$

equation. This way, the percentage replication of each method's production of correct results was determined.

Additionally, the mean difference (MD) between the real and estimated number of factors was calculated. For this purpose:

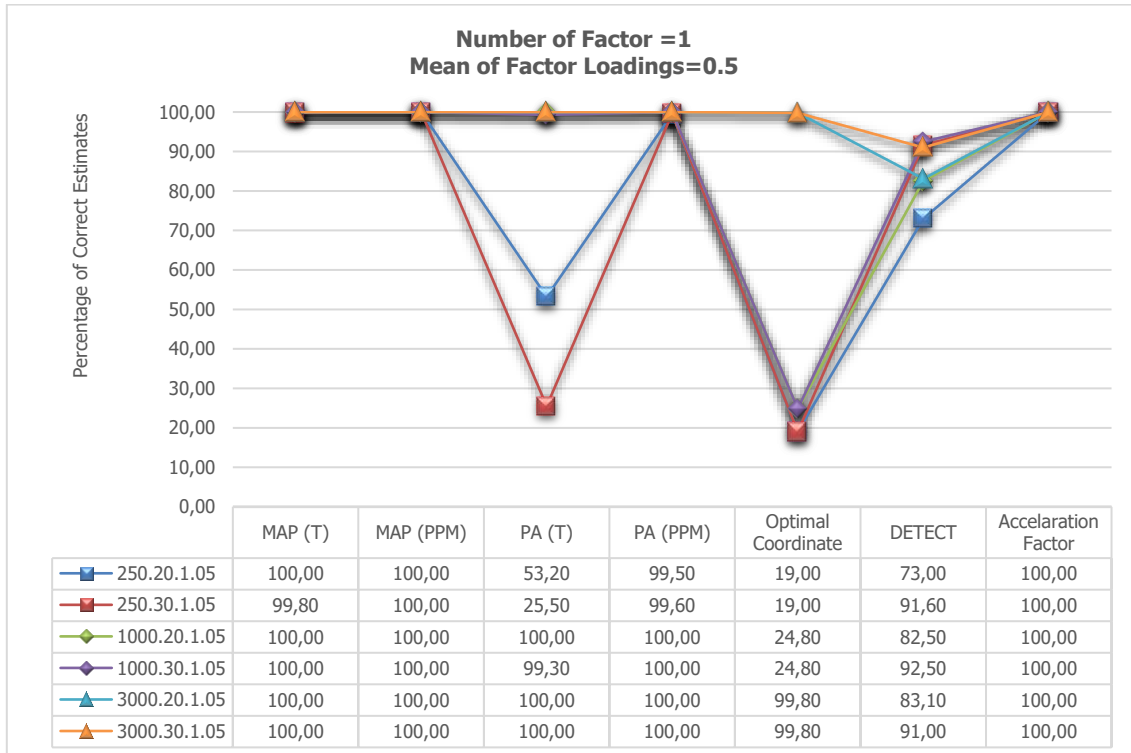
$$\text{Mean Difference} = \frac{\sum_{r=1}^{1000} (\hat{m} - m)}{1000} \quad (5)$$

equation was used. Here,  $\hat{m}$  proposed number of dimensions reflected the  $m$  real number of dimension. Since the number of replications was 1.000, the average was calculated, thus determining whether the methods overestimated or underestimated the number of factors compared to PCE values.

**FINDINGS**

In this section, the findings obtained from simulation conditions are presented.

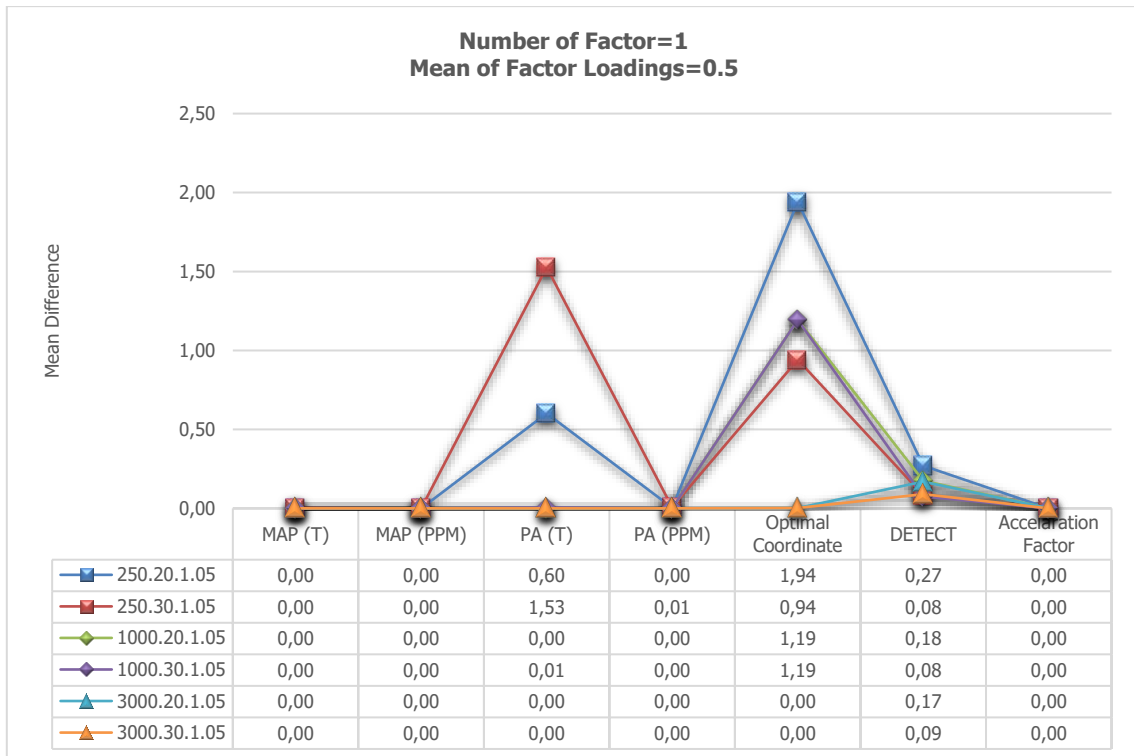
**Comparison of Simulation Conditions for One Factor Constructs**



Note: Coding System: First digit until the dot represented sample size, second digit represented number of item in the test, third digit represented number of factor, and fourth digit represented factor loading. (T): Tetrachoric correlation matrix, (PPM): Pearson Product Moment Correlation Matrix

Figure 1. Comparison of PCE values for simulation conditions when factor loading is 0.5 for one factor constructs

When Figure 1 was analyzed, the PCE values obtained through the factor retention method were observed. For an average factor loading of 0.5, when the results obtained for the sample size and the number of items method were considered, under MAP analysis with tetrachoric correlation matrix conditions (except for the sample size of 250, the number of items was 30 (99.8%)), all the conditions had 100% success rate. When MAP analysis was conducted with the PPM correlation matrix, the success rate was observed as 100%. However, when PA was conducted with the tetrachoric correlation matrix, it was impacted by the sample size. In case of the sample size of 250, PA had a significantly low percentage, while PCE had a rate of almost 100% for the sample sizes of 1000 and 3000. When PA was conducted with a PPM correlation matrix, a PCE value of almost 100% was attained for all sample sizes and numbers of items. When the Optimal Coordinate ( $n_{OC}$ ) method was investigated, for the sample sizes of 250 and 1.000, it attained significantly low percentages; while, for the sample size of 3.000, the PCE value was almost 100%. In the DETECT method, as the number of items increased, it can be stated that PCE value also increased. For conditions with 30 items, the DETECT method had a PCE value of approximately 90%; whereas for conditions with 20 items, the PCE value was between 73-83.10%. Sample size had an effect when the number of items was higher than 20. As the sample size increased, the PCE value increased. It was observed that the Acceleration Factor ( $n_{AF}$ ) method had a PCE value of 100% for all one factor constructs. Under simulation conditions with an average factor loading of 0.5 for one factor constructs, the MD value comparison is presented in Figure 2.



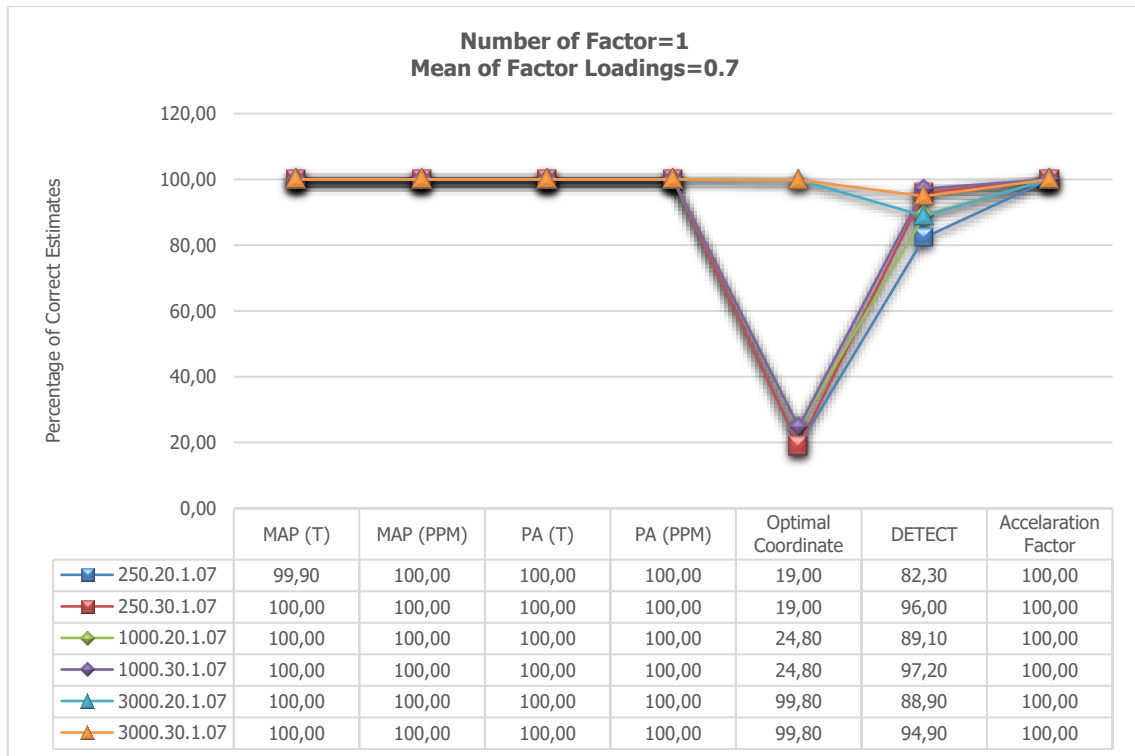
Note: Coding System: First digit until the dot represented sample size, second digit represented number of item in the test, third digit represented number of factor, and fourth digit represented factor loading. (T): Tetrachoric correlation matrix, (PPM): Pearson Product Moment Correlation Matrix

Figure 2. Comparison of MD values for simulation conditions when factor loading is 0.5 for one factor constructs

In Figure 2, when methods with PCE values other than 100% were analyzed, the bias value for MAP analysis was 0. When PCE values were investigated (the tetrachoric correlation matrix of PCE value of the MAP analysis), it was 99.8% for 250.30.1.05 condition. Accordingly, in 2 of 1.000 replications, it can be stated that the number of factors obtained was other than 1. In MD value, this situation caused differentiation in the third digit after the comma. When a tetrachoric correlation matrix was conducted on the PA 250 sample size, positive MD values were observed. Accordingly, when a tetrachoric correlation matrix was used, it can be said that the PA had the tendency to produce more factors. However, results of PA conducted with PPM indicated that the MD values were close to 0. When the MD values for Optimal Coordinate ( $n_{OC}$ ) method was investigated, it can be stated that the bias decreased as the sample size increased; yet, the biased results were positive, which meant that the number of factors was overestimated. Additionally, the Optimal Coordinate ( $n_{OC}$ ) method produced an unbiased estimation for sample size of 3.000. When the MD values for the DETECT method were investigated, it can be stated that the number of items was more effective on MD values. As in all methods, in the DETECT method, the number of factors was overestimated. In the Acceleration Factor ( $n_{AF}$ ) method, since a PCE value of 100% was obtained for all single factor constructs, the MD value was zero.

In Figure 3, when the average factor loading was 0.7, PCE values were compared for one factor construct simulation conditions.

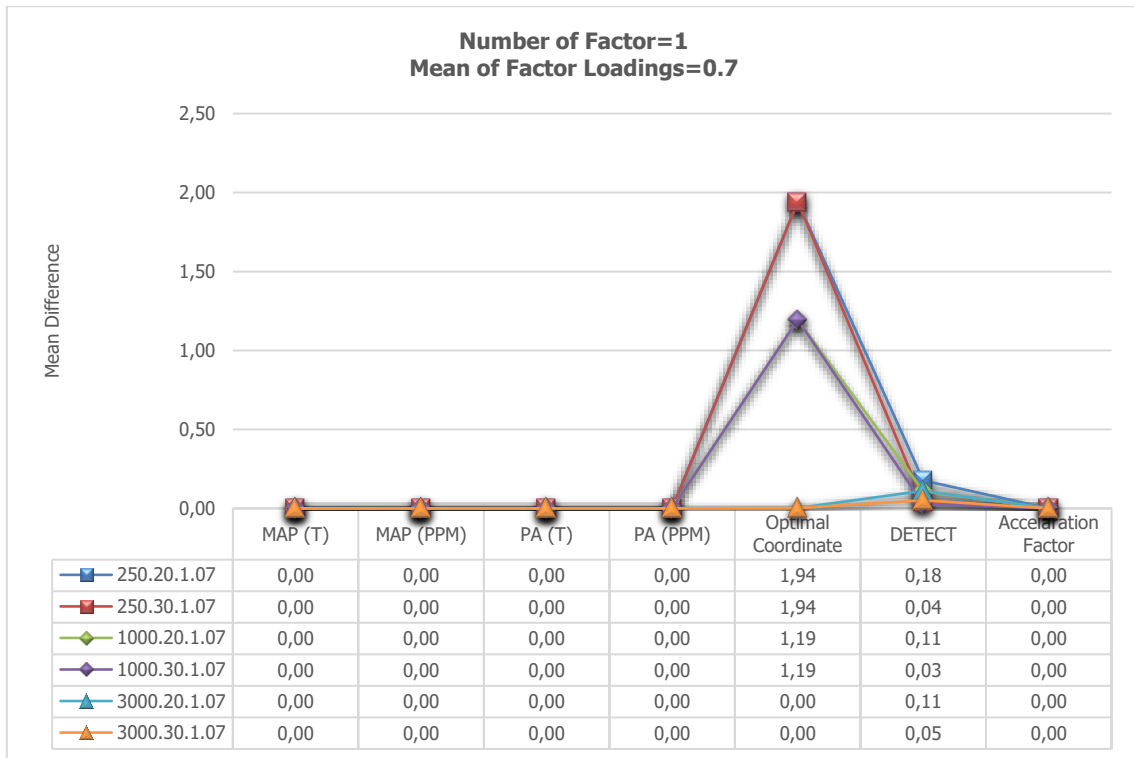




Note: Coding System: First digit until the dot represented sample size, second digit represented number of item in the test, third digit represented number of factor, and fourth digit represented factor loading. (T): Tetrachoric correlation matrix, (PPM): Pearson Product Moment Correlation Matrix

Figure 3. Comparison of PCE values for simulation conditions when factor loading is 0.7 for one factor constructs

When Figure 3 was investigated, compared to an average factor loading of 0.5, in methods with the condition of an average factor loading of 0.7, it was observed that PCE values increased. While all conditions except one (250.20.1.07) had PCE values of 100%, when MAP analysis was conducted with a tetrachoric correlation matrix and a PPM correlation matrix, a PCE value of 100% was obtained for all conditions. For this condition cluster, the PCE value was observed as 100% for PA results conducted with both a tetrachoric and PPM correlation matrix. When conditions of an average factor loading of 0.5 and 0.7 were analyzed together (Figure 1 and Figure 3), it can be stated that Optimal Coordinate ( $n_{OC}$ ) exhibits no differentiation for average factor loading. Under an average factor loading of 0.5 and the condition of a sample size of 3000, the PCE value was around 100%; whereas, with sample sizes of 250 and 1000, respectively, the PCE value was significantly low. In the DETECT method, when the average factor loadings increased, the PCE also increased. However, upon comparing Figure 1 and Figure 3, it can be stated that, as the number of items increased, the DETECT has become more effective at producing correct results. The PCE value was 94% when the DETECT method was most successful. In the Acceleration Factor ( $n_{AF}$ ) method, the PCE value was observed as 100%. Under simulation conditions with average factor loading of 0.7 for one factor constructs, the MD value comparison is presented in Figure 4.



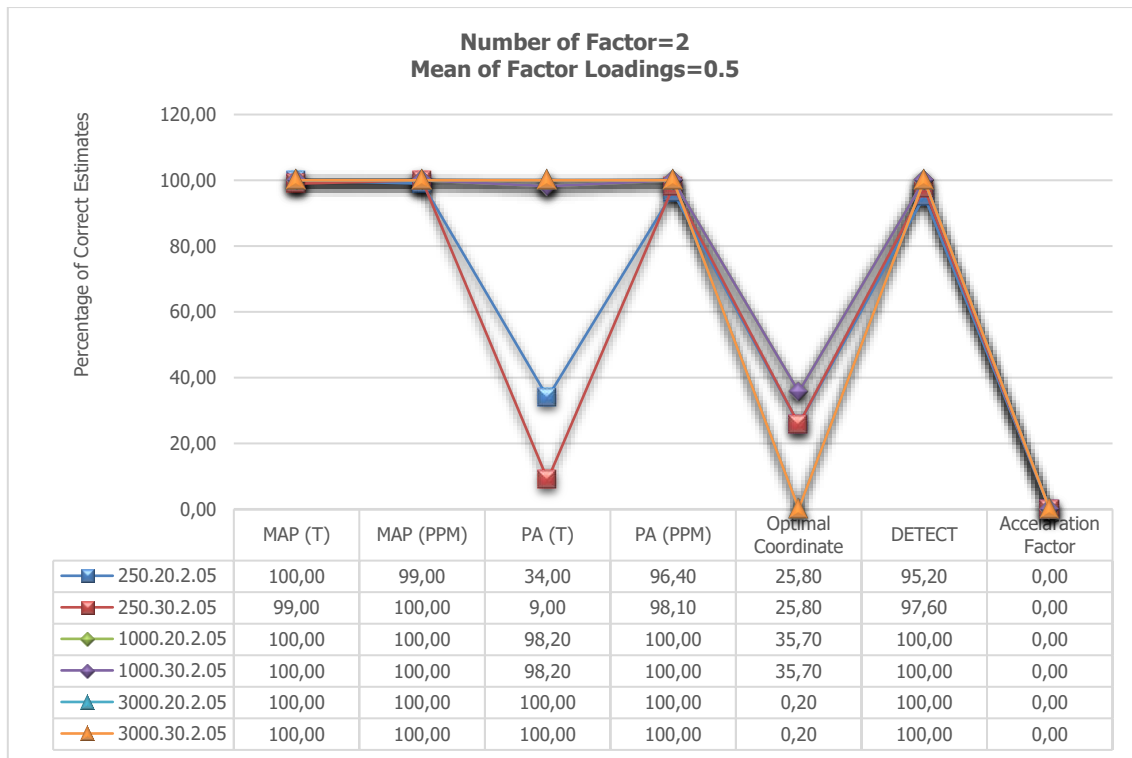
Note: Coding System: First digit until the dot represented sample size, second digit represented number of item in the test, third digit represented number of factor, and fourth digit represented factor loading. (T): Tetrachoric correlation matrix, (PPM): Pearson Product Moment Correlation Matrix

Figure 4. Comparison of MD values for simulation conditions when factor loading is 0.7 for one factor constructs

In Figure 4, when methods exhibiting PCE values other than 100% were analyzed, and since the MAP analysis only estimated the number of factors different to 1 in 1 of 1.000 replications, the PCE value was obtained as 99.90%. Therefore, the MD value differentiated in the third digit after the comma. Thus, it can be stated that the Optimal Coordinate ( $n_{OC}$ ) method, as with the DETECT method, has a tendency to overestimate the number of factors. However, in the Optimal Coordinate ( $n_{OC}$ ) method, while MD values decreased as the sample size increased, in the DETECT method, an increased number of item caused a rapid decrease in MD values. Additionally, the Optimal Coordinate ( $n_{OC}$ ) method produced unbiased results for the sample size of 3.000.

### Comparison of Simulation Conditions for Two Factors Constructs

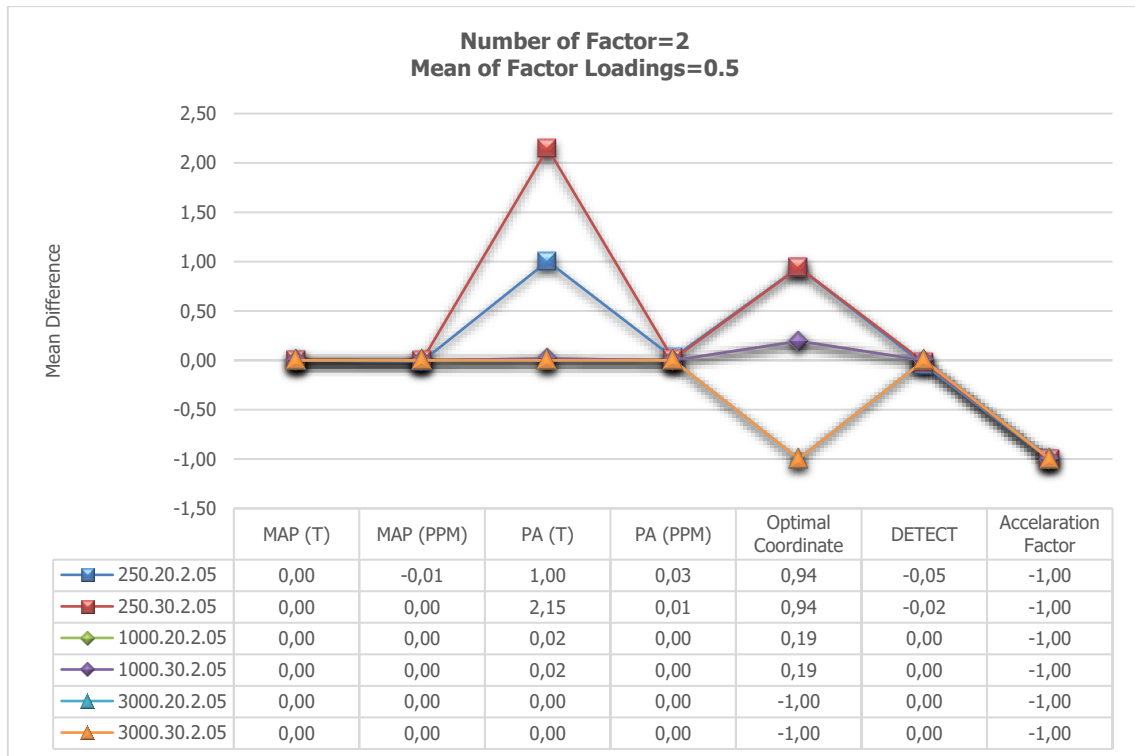
Under simulation conditions with average factor loading of 0.5 for two factors constructs, the PCE value comparison is presented in Figure 5.



Note: Coding System: First digit until the dot represented sample size, second digit represented number of item in the test, third digit represented number of factor, and fourth digit represented factor loading. (T): Tetrachoric correlation matrix, (PPM): Pearson Product Moment Correlation Matrix

Figure 5. Comparison of PCE values for simulation conditions when factor loading is 0.5 for two factors constructs

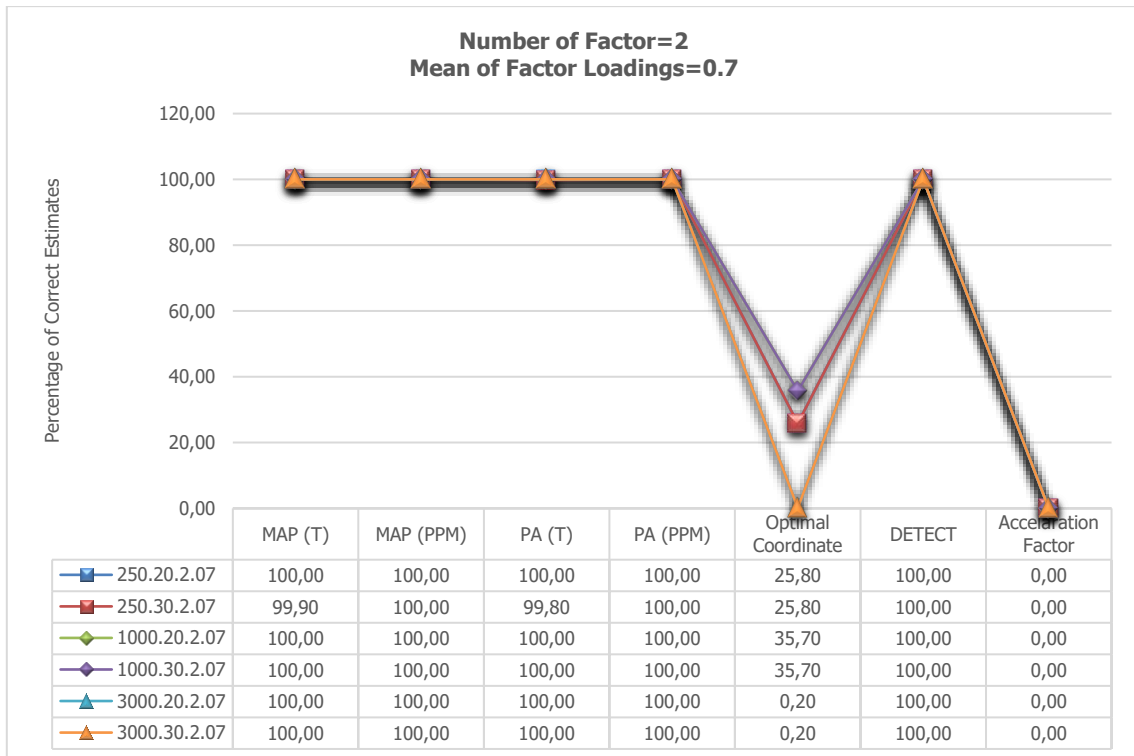
When Figure 5 was analyzed, simulation conditions with an average factor loading of 0.5 were compared in two factors constructs. When MAP analysis was conducted with a tetrachoric correlation matrix, the PCE value was around 100%. There was 99.0% PCE value in only one condition (250.30.2.05). When MAP analysis was conducted for the PPM correlation matrix, it was observed that the PCE value was 100%, excepting one condition (250.20.2.05). Accordingly, in the MAP analysis with the sample size of 250, it can be stated that the correct estimation was made with an error margin of 1%. When PA was conducted with a tetrachoric correlation matrix, for the sample sizes of 30 and 250 items, it had a PCE value of 9%. As the size of the sample increased, the PA conducted on a tetrachoric correlation matrix had a tendency to produced more accurate estimations. PA conducted for the PPM correlation matrix indicated a PCE value of around 100% for the 250 sample size, and 100% for the 1.000 and 3.000 sample sizes. For the simulation condition with a sample size of 250 and 20-item, PA results conducted with PPM showed a 96.4% PCE value. For the same sample size, as the number of items increased, the PCE value of the PA increased as well. In the Optimal Coordinate ( $n_{OC}$ ) method, it can be stated that there was no differentiation for both average factor loading and number of items. This is valid for both one factor (Figure 1 and Figure 3) and two factors (Figure 5 and Figure 7) constructs. contrary to one dimensional constructs, while PCE value of 0% was exhibited in the 3.000 sample size, PCE values of 25% and 35% were obtained for the 250 and 1.000 sample size, respectively. It can be stated that the DETECT method has higher PCE values for two dimensional constructs. When the sample size was 1.000 or more, the DETECT method had a PCE value of 100% for all conditions with 2 dimensions and an average factor loading of 0.5. In the Acceleration Factor ( $n_{AF}$ ) method, the PCE value was 0% for all conditions. When the dimension number suggested by this method was analyzed, it was observed that there was a one-dimension construct for all conditions. Under simulation conditions with an average factor loading of 0.5 for two factors constructs, the MD value comparison is presented in Figure 6.



Note: Coding System: First digit until the dot represented sample size, second digit represented number of item in the test, third digit represented number of factor, and fourth digit represented factor loading. (T): Tetrachoric correlation matrix, (PPM): Pearson Product Moment Correlation Matrix

Figure 6. Comparison of MD values for simulation conditions when factor loading is 0.5 for two factors constructs

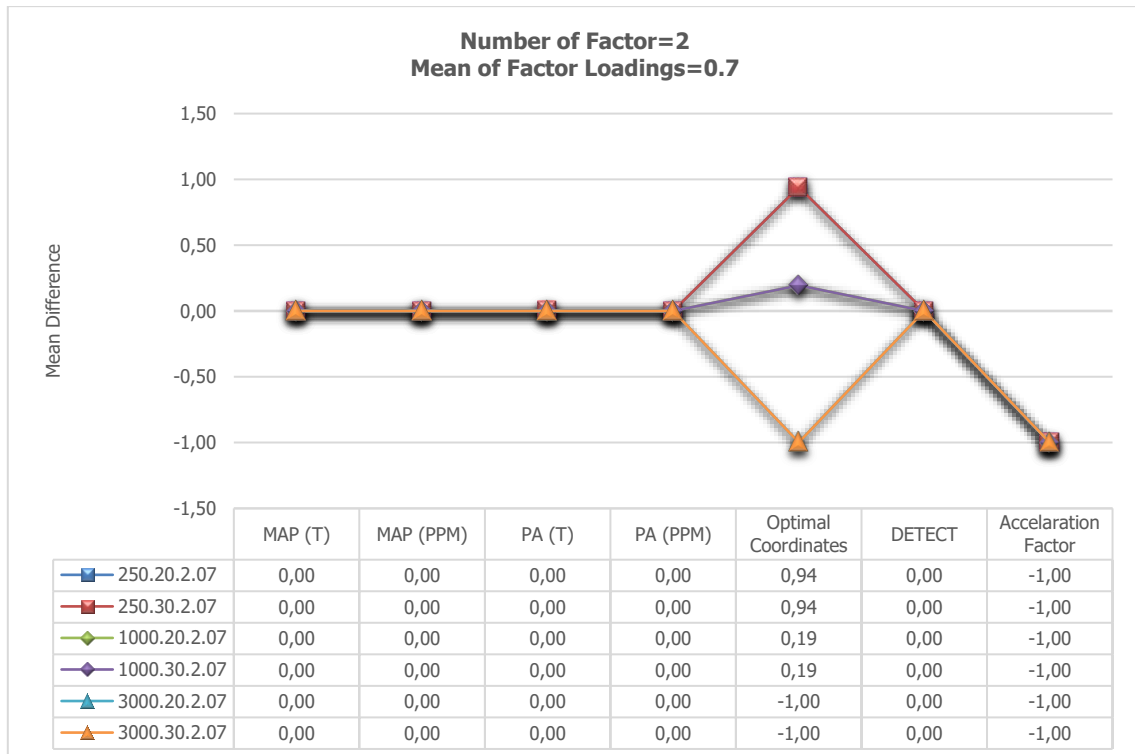
When the MD values presented in Figure 6 were analyzed, in MAP analysis conducted with PPM, the number of factors was underestimated for only one condition (250.20.2.05). For the sample size of 250, the results for PA conducted with a tetrachoric correlation matrix indicated that in a 30-item condition the PA had significantly high MD value. This is because PA estimated 2.15 times more factors than real number of factor. However, there was no such case when the analysis was conducted with PPM. In the Optimal Coordinate ( $n_{OC}$ ) method with a sample size of 250 and 1.000, the number of factors was overestimated; however, for the 3.000 sample size, the number of factors was underestimated. While the DETECT method had negative MD values for the sample size of 250, it produced unbiased results for the other samples. The acceleration Factor ( $n_{AF}$ ) method only proposed one factor for all conditions. Therefore, the MD value was -1. The PCE value analysis under two-dimensional construct simulation conditions with average factor loading of 0.7 is given in Figure 7.



Note: Coding System: First digit until the dot represented sample size, second digit represented number of item in the test, third digit represented number of factor, and fourth digit represented factor loading. (T): Tetrachoric correlation matrix, (PPM): Pearson Product Moment Correlation Matrix

Figure 7. Comparison of PCE values for simulation conditions when factor loading is 0.7 for two factors constructs

Figure 7 presented the PCE values obtained for conditions with two factors and an average factor loading of 0.7. When PA and MAP analysis were conducted with a tetrachoric correlation matrix, all conditions except for one (250.30.2.07) had a PCE value of 100%. When PA and MAP analysis were conducted with a PPM correlation matrix, all conditions had a PCE value of 100%. When both Figure 1 and 3 and Figure 5 and 7 were analyzed, it can be expressed that the Optimal Coordinate ( $n_{OC}$ ) exhibited no differentiation for both average factor loading and number of items. The same results were observed for conditions with average factor loading of 0.5. contrary to one dimensional constructs, while two dimensional constructs had 0% PCE value for the 3.000 sample size, they had 25% and 35% PCE values for the 250 and 1000 sample sizes, respectively. In the Optimal Coordinate ( $n_{OC}$ ) method, it can be stated that only sample size influenced two factor constructs. In the DETECT method, when average factor loadings increased, the PCE value also increased. The DETECT method had 100% PCE for this condition set. In the Acceleration Factor ( $n_{AF}$ ) method, the PCE value was 0. Accordingly, it can be stated that dimensionality for any data set was accurately estimated. MD value analysis under two-dimensional construct simulation conditions with an average factor loading of 0.7 is given in Figure 8.



Note: Coding System: First digit until the dot represented sample size, second digit represented number of item in the test, third digit represented number of factor, and fourth digit represented factor loading. (T): Tetrachoric correlation matrix, (PPM): Pearson Product Moment Correlation Matrix

Figure 8. Comparison of MD values for simulation conditions when factor loading is 0.7 for two factors constructs

When the MD values presented in Figure 8 were analyzed, for the Optimal Coordinate ( $n_{OC}$ ) method samples of 250 and 1.000, there was a positive MD value; whereas, for the 3.000 sample size, there was a negative MD value. The acceleration Factor ( $n_{AF}$ ) method estimated all data set as one factor. Therefore, all MD values obtained as -1. MD values of both PA and MAP conducted with tetrachoric and PPM correlation matrices were 0. Accordingly, it can be stated that PA and MAP create an unbiased estimation for this correlation matrix.

## DISCUSSION and CONCLUSION

When factor retention methods were compared, the results of this research indicated that, when MAP analysis was conducted on both tetrachoric and PPM correlation matrices, PCE values of 99% or more were observed for all conditions. Accordingly, it can be stated that MAP analysis may be used to determine dimensionality. Garrido, Abad, and Ponsoda (2011) stated that, for MAP analysis results conducted with (2, 3, 4, 5, 6, 7) variables with different category numbers for polychoric and PPM correlation matrices, a polychoric correlation matrix was more appropriate. Additionally, as factor loading increased, PCE values increased in MAP analysis. In our research, MAP analysis was the best method for all conditions. In Garrido, Abad, and Ponsoda (2011), this could be caused by the fact that the factor loading of each variable was accepted as equal. In our research, the average factor loading was kept constant; however, the factor loading of each item was differentiated. Zwick and Velicer (1986) compared PA, MAP, Cattell's scree test, Bartlett's chi-square test, and the K1 rule methods and reported that the PA and MAP method presented the best results. In our current research, the results are consistent with the literature.

When PA was conducted with a tetrachoric correlation matrix, it was affected by average factor loading and sample size. In small samples with an average factor loading of 0.5, the performance of PA decreased. Yet, as the sample increased, the performance of PA increased as well. These results are similar to those of Yang and Xia (2015) and Cho et al. (2009). While sample size had a significant effect on PA conducted with a tetrachoric correlation matrix, the same effect is invisible with a PPM correlation matrix. Guilford (1952) stated that, to calculate a tetrachoric correlation matrix, data should have a large sample size (at least 400) for binary data. The low performance of PA with the sample size of 250 may be linked with this.

Optimal Coordinate ( $n_{OC}$ ) showed a good performance for one dimensional data with a sample size of 3.000; however, the PCE value was below 50% for other conditions. This result is in line with Raïche et al. (2013). Raïche et al. (2013) worked with 36 and 72 variables, using 2 and 5 folds of variables for sample size (72, 180, 144, 360) and 0.5 and 0.8 average factor loadings. Results for the Optimal Coordinate ( $n_{OC}$ ) method's PCE value varied between 20% (72 variables, 144 sample size, 0.5 factor loading) and 82% (72 variables, 360 sample size, 0.8 factor loading).

The acceleration Factor ( $n_{AF}$ ) method estimated one dimension for all conditions under the scope of this research. In this case, it can be expressed that the Acceleration Factor ( $n_{AF}$ ) method presented no differentiation for factors under all conditions. Raïche et al. (2013) determined that the Acceleration Factor ( $n_{AF}$ ) method's PCE value varied between 17% and 50%. However, in this study, factor loadings were distributed over 3 components as 0.8 and 0.2. In our study, average factor loading was considered.

While the DETECT method had around 100% PCE value for the majority of two factor constructs, it was affected by both sample size and average factor loading for one factor constructs. Additionally, as the number of item increased, the DETECT method's performance increased as well. Similarly, as average factor loading increased, DETECT estimated more accurate results. van Abswoude, van der Ark, and Sijtsma (2004) stated that the DETECT method is affected by a sample's size and is more efficient for larger samples ( $n=2.000$ ). Additionally, DETECT determined multi-dimensionality accurately in the case of low correlation between factors. This is in line with other studies in the literature (Zhang, Yu, & Nandakumar, 2003).

Based on the conditions of this research, the following recommendations can be made; 1) both a tetrachoric and PPM correlation matrix of MAP can be used, 2) the PA method produced more accurate results with a PPM correlation matrix, but in the case of a tetrachoric correlation matrix, samples size should be considered, 3) instead of Optimal Coordinate ( $n_{OC}$ ) and Acceleration Factor ( $n_{AF}$ ) methods, PA or MAP methods can be preferred, and 4) average factor loading and samples size should be considered before using the DETECT method. Additionally, for this method, both the number of items and dimension number has an effect. It can be stated that this method may be used after collectively evaluating all conditions. Additionally, the combined use and evaluation of PA and MAP can be recommended. Since only simple construct data was used in this research, in future studies factor size could be increased and simulations that manipulate the correlation between factors could be applied.

## Acknowledgement

This study was presented in 6th. International Conference on Education, Zagreb, Croatia.

## 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](https://doi.org/10.1207/s15328007sem1302_2)
- Bennett, R. E., Rock, D. A., Braun, H. I., Frye, D., Spohrer, J. C., & Soloway, E. (1990). The relationship of

- expert-system scored constrained free-response items to multiple-choice and open-ended items. *Applied Psychological Measurement*, 14(2), 151–162. <https://doi.org/10.1177/014662169001400204>
- Bennett, R. E., Rock, D. A., & Wang, M. (1991). Equivalence of free-response and multiple-choice items. *Journal of Educational Measurement*, 28(1), 77–92. <https://doi.org/10.1111/j.1745-3984.1991.tb00345.x>
- Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research*, 27(4), 509–540. [https://doi.org/10.1207/s15327906mbr2704\\_2](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](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>
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). London: Lawrence Erlbaum Associates.
- Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods*, 6(2), 147–168. <https://doi.org/10.1177/1094428103251541>
- Cota, A. A., Longman, R. S., Holden, R. R., Fekken, G. C., & Xinaris, S. (1993). Interpolating 95th percentile eigenvalues from random data: An empirical example. *Educational and Psychological Measurement*, 53(3), 585–596. <https://doi.org/10.1177/0013164493053003001>
- DeMars, C. (2010). *Item response theory*. New York: Oxford University.
- Dinno, A. (2014). *Gently clarifying the application of Horn's parallel analysis to principal component analysis versus factor analysis*. Unpublished manuscript, School of Community Health, Portland State University, Oregon, USA. Retrieved from [http://doyenne.com/Software/files/PA\\_for\\_PCA\\_vs\\_FA.pdf](http://doyenne.com/Software/files/PA_for_PCA_vs_FA.pdf)
- Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory factor analysis*. New York: Oxford University.
- Feinberg, R. A., & Rubright, J. D. (2016). Conducting simulation studies in psychometrics. *Educational Measurement: Issues and Practice*, 35(2), 36–49. <https://doi.org/10.1111/emip.12111>
- Finch, H., & Habing, B. (2005). Comparison of NOHARM and DETECT in item cluster recovery: Counting dimensions and allocating items. *Journal of Educational Measurement*, 42(2), 149–169. <https://doi.org/10.1111/j.1745-3984.2005.00008>
- Finch, H., & Habing, B. (2007). Performance of DIMTEST- and NOHARM- based statistics for testing unidimensionality. *Applied Psychological Measurement*, 31(4), 292–307. <https://doi.org/10.1177/0146621606294490>
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7(3), 286–299. <https://doi.org/10.1037/1040-3590.7.3.286>
- Fraser, C., & McDonald, R. P. (1988). NOHARM: Least squares item factor analysis. *Multivariate Behavioral Research*, 23(2), 267–269. [https://doi.org/10.1207/s15327906mbr2302\\_9](https://doi.org/10.1207/s15327906mbr2302_9)
- 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>
- Gilbert, N. (1999). Simulation: A new way of doing social science. *American Behavioral Scientist*, 42(10), 1485–1487. <https://doi.org/10.1177/0002764299042010002>
- 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>
- Gorsuch, R. L. (1974). *Factor analysis*. Toronto: W. B. Saunders Company.
- Green, S. B., Levy, R., Thompson, M. S., Lu, M., & Lo, W.-J. (2012). A proposed solution to the problem with using completely random data to assess the number of factors with parallel analysis. *Educational and Psychological Measurement*, 72(3), 357–374. <https://doi.org/10.1177/0013164411422252>
- Green, S. B., Redell, N., Thompson, M. S., & Levy, R. (2016). Accuracy of revised and traditional parallel analyses for assessing dimensionality with binary data. *Educational and Psychological Measurement*, 76(1), 5–21. <https://doi.org/10.1177/0013164415581898>
- Green, S. B., Thompson, M. S., Levy, R., & Lo, W.-J. (2015). *Type I and Type II error rates and overall accuracy of the revised parallel analysis method for determining the number of factors*. *Educational and Psychological Measurement*, 75(3), 428–457. <https://doi.org/10.1177/0013164414546566>
- Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, 103(2), 265–275.
- Guilford, J. P. (1952). When not to factor analyze. *Psychological Bulletin*, 49(1), 26–37. <https://doi.org/10.1037/h0054935>
- Harwell, M., Stone, C. A., Hsu, T.-C., & Kirisci, L. (1996). Monte carlo studies in item response theory. *Applied Psychological Measurement*, 20(2), 101–125. <https://doi.org/10.1177/014662169602000201>



- Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research. *Educational and Psychological Measurement*, 66(3), 393–416. <https://doi.org/10.1177/0013164405282485>
- 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>
- Hu, L., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, 112(2), 351–362. <https://doi.org/10.1037/0033-2909.112.2.351>
- Jang, E. E., & Roussos, L. (2007). An investigation into the dimensionality of TOEFL using conditional covariance-based nonparametric approach. *Journal of Educational Measurement*, 44(1), 1–21. <https://doi.org/10.1111/j.1745-3984.2007.00024.x>
- 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>
- Kaya Kalkan, Ö., & Kelecioğlu, H. (2016). The effect of sample size on parametric and nonparametric factor analytical methods. *Educational Sciences: Theory & Practice*, 16(1), 153–171. <https://doi.org/10.12738/estp.2016.1.0220>
- Kim, H. R. (1996). A new index of dimensionality-DETECT. *The Pure and Applied Mathematics*, 3(2), 141–154.
- Ledesma, R. D., & Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research & Evaluation*, 12(2), 2–11.
- Lissitz, R. W., Hou, X., & Slater, S. C. (2012). The contribution of constructed response items to large scale assessment: Measuring and understanding their impact. *Journal of Applied Testing Technology*, 13(3), 1–50. Retrieved from <http://www.jattjournal.com/index.php/atp/article/view/48366/39234>
- MEB. (2019). Sınavla öğrenci alacak ortaöğretim kurumlarına ilişkin merkezî sınav başvuru ve uygulama kılavuzu. Retrieved from [https://www.meb.gov.tr/meb\\_iys\\_dosyalar/2019\\_04/03134315\\_Kilavuz2019.pdf](https://www.meb.gov.tr/meb_iys_dosyalar/2019_04/03134315_Kilavuz2019.pdf)
- Nandakumar, R., & Stout, W. (1993). Refinements of Stout's procedure for assessing latent trait unidimensionality. *Journal of Educational Statistics*, 18(1), 41–68. <https://doi.org/10.2307/1165182>
- 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>
- Osborne, J. W., & Banjanovic, E. S. (2016). *Exploratory factor analysis with SAS®*. Cary, NC: SAS Intitute Inc.
- R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org/>.
- Raiche, G. (2010). nFactors: An R package for parallel analysis and non graphical solutions to the Cattell scree test. Retrieved from <https://cran.r-project.org/web/packages/nFactors/nFactors.pdf>.
- Raiche, G., Walls, T. A., Magis, D., Riopel, M., & Blais, J.-G. (2013). Non-Graphical solutions for Cattell's scree test. *Methodology*, 9(1), 23–29. <https://doi.org/10.1027/1614-2241/a000051>
- Revelle, W. (2016). psych: Procedures for psychological, psychometric, and personality research. Evanston, Illinois. Retrieved from <https://cran.r-project.org/package=psych>
- Robey, R. R., & Barcikowski, R. S. (1992). Type I error and the number of iterations in monte carlo studies of robustness. *British Journal of Mathematical and Statistical Psychology*, 45, 283–288.
- Robitzsch, A. (2017). sirt: Supplementary item response theory models. Retrieved from <https://cran.r-project.org/package=sirt>
- Stout, W. (1987). A nonparametric approach for assessing latent trait unidimensionality. *Psychometrika*, 52(4), 589–617. <https://doi.org/10.1007/BF02294821>
- Streiner, D. L. (1994). Figuring out factors: The use and misuse of factor analysis. *Canadian Journal of Psychiatry*, 39(3), 135–140.
- Tabachnik, B. G., & Fidell, L. S. (2012). *Using multivariate statistics* (6th ed.). Boston: Pearson.
- van Abswoude, A. A. H., van der Ark, L. A., & Sijtsma, K. (2004). A comparative study of test data dimensionality assessment procedures under nonparametric IRT models. *Applied Psychological Measurement*, 28(1), 3–24. <https://doi.org/10.1177/0146621603259277>
- van den Bergh, H. (1990). On the construct validity of multiple-choice items for reading comprehension. *Applied Psychological Measurement*, 14(1), 1–12. <https://doi.org/10.1177/014662169001400101>
- Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41(3), 321–327.
- 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>
- Xia, Y., Green, S. B., Xu, Y., & Thompson, M. S. (2018). Proportion of indicator common variance due to a factor as an effect size statistic in revised parallel analysis. *Educational and Psychological Measurement*, 79(1), 75–107. <https://doi.org/10.1177/0013164418754611>
- 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>

- Zhang, J. (2007). Conditional covariance theory and DETECT for polytomous items. *Psychometrika*, 72(1), 69–91. <https://doi.org/10.1007/s11336-004-1257-7>
- Zhang, J., & Stout, W. (1999). The theoretical detect index of dimensionality and its application to approximate simple structure. *Psychometrika*, 64(2), 213–249. <https://doi.org/10.1007/BF02294536>
- Zhang, Y. O., Yu, F., & Nandakumar, R. (2003). *The impact of conditional scores on the performance of DETECT*. Paper presented at the annual meeting of the National Council on Measurement in Education. Chicago, IL. Retrieved from <https://files.eric.ed.gov/fulltext/ED478170.pdf>
- Zoski, K., & Jurs, S. (1993). Using multiple regression to determine the number of factors to retain in factor analysis. *Multiple Linear Regression Viewpoint*, 20(1), 5–9. Retrieved from [http://www.glmj.org/archives/MLRV\\_1993\\_20\\_1.pdf](http://www.glmj.org/archives/MLRV_1993_20_1.pdf)
- Zoski, K., & Jurs, S. (1996). An objective counterpart to the visual scree test for factor analysis: The standard error scree. *Educational and Psychological Measurement*, 56(3), 443–451. <https://doi.org/10.1177/0013164496056003006>
- 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>

## TÜRKÇE GENİŞLETİLMİŞ ÖZET

Veri setinin yapısı araştırılmak istenildiğinde faktör analizi en sık kullanılan psikometrik tekniklerden biridir. Araştırmacıların yapıyla ilgili bilgileri yoksa bu durumda açıklayıcı faktör analizi ya da sınırlandırılmamış faktör analizi yapılır. Eğer yapı hakkında araştırmacıların fikri varsa ve verilerin bu yapıya uyum gösterip göstermediği araştırılacaksa bu durumda da doğrulayıcı faktör analizi diğer bir deyişle sınırlandırılmış faktör analizi yapılmaktadır.

AFA için sıklıkla karşılaşılan güçlüklerden biri de boyut sayısına karar vermektir. Boyut sayısına karar vermek için geliştirilen birçok yöntem bulunmaktadır. Bu yöntemlere; Örneğin 1’den büyük özdeğer sayısını boyut sayısı kabul eden ve Kaiser’in K1 kuralı (Kaiser, 1960), Horn (1965) tarafından önerilen Paralel Analiz (PA), Velicer (1976) tarafından önerilen Kısmi Korelasyonların En Küçüğü (MAP) testi, Cattell (1966) tarafından önerilen Yamaç Grafiği yöntemleri de bulunmaktadır. Yamaç grafiğine alternatif olarak Zoski ve Jurs (1993, 1996) tarafından önerilen çok regresyon t-değeri ve yamacın standart hatası (SEscree) yaklaşımları, Raiche, Walls, Magis, Riopel ve Blais (2013) tarafından önerilen amaç grafiği optimal koordinatlar ve yamaç grafiği ivmelenme faktörü gibi yöntemler bulunmaktadır. Bu yöntemlerin yanında koşullu kovaryanslara dayanan ve nonparametrik bir yöntemler olan DIMTEST (Nandakumar & Stout, 1993; Stout, 1987) ve DETECT (Zhang & Stout, 1999) yöntemleri vardır. Mevcut araştırmada Paralel Analiz, Kısmi Korelasyonların En Küçüğü, DETECT, Optimal Koordinat ve İvmelenme Faktörü yöntemleri karşılaştırılmıştır.

İhtimalleri dikkate alınarak önerilerde bulunulabilmesi nedeniyle (Gilbert, 1999) bu araştırma, Monte Carlo simülasyon çalışması olarak tasarlanmıştır. Simülasyon çalışmaları gerçek parametreler ile kestirilen parametrelerin karşılaştırılmasını sağlaması nedeniyle avantaj sağlamaktadır (Feinberg & Rubright, 2016).

Araştırmada simülasyon faktörleri, örneklem büyüklüğü (250, 1000 ve 3000), faktör sayısı (tek ve iki [basit yapıda] boyutlu) test uzunluğu (20 ve 30 madde), ortalama faktör yükü (0.50 ve 0.70) ve kullanılan korelasyon matrisi (Pearson Momentler Çarpımı [PPM] ve tetrakorik) olarak belirlenmiştir. Araştırmada tamamen çaprazlanmış desen kullanılmıştır. Buna göre  $2 \times 3 \times 2 \times 2 = 24$  koşul üzerinde çalışmış ve 1.000 replikasyon yapılmıştır. Araştırmada ikili (1-0) yapıdaki veri kullanılmıştır.

Verinin üretimi için R yazılımındaki Psych paketi kullanılmıştır. Paralel analiz ve MAP analizi için Psych (Revelle, 2016) paketi, DETECT yöntemi için sirt (Robitzsch, 2017), ivmelenme faktörü ve optimal koordinat yöntemi için de nFactors (Raiche, 2010) paketleri kullanılmıştır.

PA için faktörleştirme yöntemi olarak temel faktör çözümlemesi kullanılmıştır. PA’da rassal olarak oluşturulan korelasyon matrisi sayısı ise 50 olarak belirlenmiştir. DETECT yönteminde kesme puanı olarak 0.20 önerildiğinden (Jang & Roussos, 2007; Zhang, 2007) araştırmada tek boyutlu veri seti için 0.20 üzerinde DETECT değeri elde edilen analizlerde çok boyutlu, iki boyutlu veri seti için ise DETECT değeri 1’den büyük olan sonuçlar çok boyutlu olarak kabul edilmiştir. Böylece testin değerlendirilmesinde daha tutucu bir yaklaşım izlenmiştir. Yöntemlerin performanslarının değerlendirilmesi için gerçek boyut sayısı ile önerilen boyut sayıları karşılaştırılarak doğru kestirim yüzdesi elde edilmiştir. Bunun için,

$$P_r = \begin{cases} 1, & \text{Önerilen Boyut Sayısı} = \text{Gerçek Boyut Sayısı} \\ 0, & \text{Önerilen Boyut Sayısı} \neq \text{Gerçek Boyut Sayısı} \end{cases} \quad (3)$$

fonksiyonu kullanılmıştır. Bu fonksiyondaki r, replikasyonu ifade etmektedir. Buna göre hesaplanan doğru kestirim yüzdesi ise;

$$\text{Doğru Kestirim Yüzdesi} = \frac{\sum_{r=1}^{1000} P_r}{1000} \cdot 100 \quad (4)$$

eşitliğiyle ifade edilebilir. Ayrıca gerçek ve kestirilen faktör sayıları arasındaki Ortalama Fark hesaplanmıştır. Bunun için;

$$\text{Ortalama Fark} = \frac{\sum_{r=1}^{1000} (\hat{m} - m)}{1000} \quad (5)$$

eşitliği kullanılmıştır. Burada  $\hat{m}$  önerilen boyut sayısını  $m$  gerçekteki boyut sayısını göstermektedir. Replikasyon sayısı 1000 olduğu için ortalaması alınmıştır. Böylece doğru kestirim yüzdesi değerinin yanında yöntemlerin faktör sayısını olduğundan daha az ya da daha fazla kestirip kestirmediği araştırılmıştır.

Araştırma sonucunda MAP analizi, hem tetrakorik hem de PPM korelasyon matrisiyle yürütüldüğünde araştırma kapsamındaki tüm koşullar için %99 ve üzerinde doğru kestirim yüzdesine sahip olduğu gözlenmiştir. Ayrıca faktör yükü arttıkça MAP analizinin doğru kestirim yüzdesi yükselmiştir. Buna göre MAP analizinin boyutluluk belirlemede kullanılabileceği söylenebilir.

PA, tetrakorik korelasyon matrisiyle yürütüldüğünde ortalama faktör yükünden ve örneklem büyüklüğünden etkilenmektedir. Küçük örneklemde ve ortalama faktör yükü 0.5 olduğunda PA'nın performansı düşmektedir. Ancak örneklem büyüdükçe PA'nın performansı da artmaktadır. Tetrakorik korelasyon matrisiyle yürütülen PA üzerinde örneklem büyüklüğü oldukça etkili iken PPM korelasyon matrisi üzerinde aynı etki mevcut değildir.

Optimal Koordinat yöntemi tek boyutlu veride örneklem büyüklüğünün 3000 olduğu durumda iyi performans göstermiş ancak diğer durumlarda doğru kestirim yüzdesi %50'nin altında kalmıştır. İvmelenme faktörü yöntemi araştırma kapsamındaki tüm koşullarda tek boyut önermiştir. Buna göre bu yönteminin araştırma kapsamındaki koşullar için, faktörleri ayırtamadığı söylenebilir.

DETECT yöntemi iki faktörlü yapıların büyük kısmında %100'e yakın PCE değerine sahipken tek faktörlü yapılarda örneklem büyüklüğü ve ortalama faktör yükünden etkilendiği gözlenmiştir. Ayrıca madde sayısının artması da DETECT'in performansında artışa neden olmaktadır. Ortalama faktör yükü arttıkça benzer şekilde DETECT daha doğru sonuçlar vermektedir.

Araştırmada yer alan koşullara dayalı olarak; 1) MAP analizini hem tetrakorik hem de PPM korelasyon matrisiyle kullanılabileceği, 2) PA yönteminin PPM korelasyon matrisiyle daha doğru sonuçlar verdiği ancak tetrakorik korelasyon matrisiyle yürütüldüğü durumda örneklem büyüklüğünün göz önüne alınması gerektiği, 3) Optimal koordinat ve ivmelenme faktörü yöntemlerinin yerine PA ya da MAP yöntemlerinin tercih edilmesi, 4) DETECT yönteminin ortalama faktör yükü ve örneklem büyüklüğü göz önünde bulundurularak kullanılması önerilmektedir.