Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Volume 22 Number 7, September 2017 ISSN 1531-7714

A Note on Using Eigenvalues in Dimensionality Assessment

Cengiz Zopluoglu, *University of Miami*
Ernest C. Davenport, Jr, *University of Minnesota, Twin Cities*

The dimensionality of a set of items is important for scale development. In practice, tools that make use of eigenvalues are often used to assess dimensionality. Parallel analysis is featured here as it is becoming an increasingly popular method for assessing the number of dimensions, and computational tools have recently been made available which will likely increase its use by practitioners. The current paper argues that methods that use eigenvalues to ascertain the number of factors may perform poorly under certain conditions, particularly for increasing levels of variable complexity and/or inter-factor correlations in the latent structure. A simulation study and an example are offered to substantiate this assertion.

Choosing the number of latent traits in measurement models has been extensively discussed in two separate but related contexts: exploratory factor analysis (EFA) and item response theory (IRT). In the EFA literature, choosing the number of latent traits has been debated in numerous papers under the heading of "factor retention criteria" or "number of common factors" (e.g., Cattell, 1966; Green, Levy, Thompson, Lu, & Lo, 2012; Kaiser, 1960). Studies for testing the unidimensionality assumption have dominated the IRT literature; however, these studies do not provide guidance beyond the first dimension once the unidimensionality assumption is rejected. If the unidimensionality assumption is rejected, one has no idea of the dimensionaliaty. The use of multidimensional item response theory (MIRT) models is increasing, but one needs an accurate assessment of the number of latent traits (a similar decision process as for EFA) to use these models. Different approaches have appeared in the literature to determine the number of latent traits in the context of MIRT models (Bock, Gibbons, Muraki, 1988; Cho, Li, Bandalos, 2009; Gessaroli & De Champlain, 1996; Gessaroli, De Champlain, & Folske, 1997; Maydeu-Olivares, 2001; Zhang, 1996).

In both the EFA and IRT literatures, eigenvalue examination is a common practice for deciding the number of latent traits in the model. Previous research indicate that three eigenvalue examination methods are frequently used in the EFA literature for determining the number of latent traits: the Kaiser-Guttman rule (KG; Guttman, 1954; Kaiser, 1960), the subjective scree test (Cattell, 1966), and parallel analysis (Horn, 1965; Green et al., 2012). Variance accounted for is the basis of the Kaiser-Guttmann rule. The variance for a complete system of non-redundant items equals the number of items. Thus, for a dimension to be stronger than average, it has to have an eigenvalue greater than one. The KG rule, then, chooses all dimensions with eigenvalues greater than one as "significant" dimensions. The scree test is really not a test, but a plot. It is a plot of the principal component eigenvalues against its dimension. Eigenvalues for "real" dimensions will be larger than average, however, after a point (dimension) all of the eigenvalues will be uniformly decreasing since the remaining variance accounted for by the eigenvalues is random. When the eigenvalues become uniformly decreasing the slope of eigenvalue to dimension becomes a constant. From this point on the plot will be simply a line. It is the

investigators task to determine (subjectively) at which dimension the scree plot becomes a line separating significant from random dimensions. Parallel analysis determines the number of valid dimensions by comparing the observed eigenvalues to those resulting from random matrices (of the same size). All observed eigenvalues larger than their corresponding random eigenvalues are considered "significant" and thus define a valid dimension.

In a review of 152 articles published in three psychology journals from 1975 to 1984, Ford, MacCallum, and Tait (1986) found that the KG rule was the most widely used method (21.7%), followed by the subjective scree test (11.2%) in determining the number of factors. In another review of 217 studies published in two psychology journals between 1991 and 1995, Fabrigar, Wegener, MacCallum, and Strahan (1999) reported that the KG rule was again at the top (16.5%), followed by the scree test (15.2%). In another review, the KG rule was again found to be the most widely used method in factor analytic studies (56.7%), followed by the scree test (35%), and parallel analysis (6.7%) of the published articles (Henson & Roberts, 2001). Note that the KG rule is the default criterion used in most factor analysis programs. Fabrigar et al. (1999) also reported that about 21.7% of the studies used multiple methods. However, the multiple methods they suggested were based on eigenvalues with the exception of descriptive fit indices commonly used in confirmatory factor analysis (although confirmatory factor analysis can provide inferential tests, descriptive indices are promoted due to its significance tests being a function of sample size).

Similarly, the IRT literature has recommendations for using eigenvalues computed from tetrachoric or polychoric correlations in dimensionality assessment. In an early seminal work, Hattie (1985) identified 87 indices for assessing the unidimensionality assumption. Some of these rules were based on eigenvalues. In like manner, Reckase (1979) suggested that the first principal component should account for at least 20% of the total variance in order to obtain acceptable unidimensional ability parameter estimates. This is also in line with a quote from Vernon (1950) quoted in Cronbach (1951). "For a test to be interpretable, however, it is not essential that all items be factorially similar. What is required is that a large proportion of the test variance be attributable to the principal factor running through the test." Using the ratio of the first and second eigenvalues as an index of unidimensionality (although no criterion was provided for interpreting this ratio) was also observed in the literature (Lord, 1980; Lumsden, 1957, 1961). Recently, modifications of parallel analysis using tetrachoric/polychoric correlations have appeared in the literature for exploring the unidimensionality assumption (Drasgow & Lissak, 1983; Finch & Monahan, 2008; Tran & Formann, 2009; Weng & Cheng, 2005) and the number of multiple latent dimensions (Cho, Li, & Bandalos, 2009) for dichotomously or polytomously scored data.

In this paper, we highlight potential implications of using eigenvalues in dimensionality assessment under certain conditions. We do this because eigenvalues are often considered when ascertaining dimensionality. In our discussion of eigenvalues we first highlight a mathematical necessity for the first eigenvalue. Although some expert methodologists and experienced researchers may have realized this feature of the first eigenvalue, we believe many researchers are still unaware of it as evidenced by the popularity of the eigenvalue > 1 rule and it being the default procedure in many factor analysis programs as well as the increasing recommendation of the potential of parallel analyses by factor analysis experts. Next, we discuss potential implications of this feature of the first eigenvalue in terms of dimensionality assessment using a simulation study.

A Closer Look at the First Eigenvalue

As previous reviews indicated, methods based on eigenvalue examination are popular and frequently used for dimensionality assessment in published studies. Among these methods, previous simulation studies also report that parallel analysis is the most viable option based on eigenvalues (Cho et al., 2009; Crawford, Green, Levy, Lo, Scott, Svetina, & Thompson, 2010; Dinno, 2009; Drasgow & Lissak, 1983; Finch & Monahan, 2008; Green et al., 2012; Hayton, Allen, Scarpello, 2004; Tran & Formann, 2009; Weng & Cheng, 2005). However, an interesting feature of the first eigenvalue seems to be overlooked in the literature when evaluating the performance of these methods. In our opinion, this feature was overlooked because many studies only used an independent cluster structure for their data simulation. In this section, we demonstrate why researchers should be cautious when using eigenvalues for assessing dimensionality.

Table 1 shows two factor structures of equal strength. The sum of the squared loadings for each factor is equal; the eigenvalue for each dimension is 1.47. Note that the total of the eigenvalues for each set of factors is also equal, 4.41 for both structures. The difference in the structures is that one is simple and the other is complex. Note that in both cases common variance accounts for just shy of 50% of the total variance, 9. Note, too, that there are two matrices representing the correlation of the factors; one with uncorrelated factors and the other with correlated factors, but with relatively small correlations, 0.2. If we let the factor matrix for the simple factors be F_1, the factor matrix for the complex structure be F_2, the correlation matrix with uncorrelated factors be R_1 and the minimally correlated factor matrix be R_2, then we can obtain the expected correlation matrices for the four conditions (two factor structures by two factor correlations) by simply multiplying the appropriate factor structure times the desired factor correlation matrix times the transpose of the same factor structure. One can then extract principal component eigenvalues from these four expected correlation matrices: $C_{SU}=F_1R_1F_1'$, $C_{SC}=F_1R_2F_1'$, $C_{CU}=F_2R_1F_2'$, $C_{CC}=F_2R_2F_2'$. C_{SU} is the correlation matrix with the simple factor structure and uncorrelated factors. Its first three eigenvalues are: 1.98, 1.98, and 1.98. C_{SC} is the correlation matrix with the simple factor structure and correlated factors. Its first three eigenvalues are: 2.57, 1.69, and 1.69. C_{CU} is the correlation matrix with the complex factor structure and uncorrelated factors. Its first three eigenvalues are: 4.14, 0.90, and 0.90. Finally, C_{CC} is the correlation matrix with the complex factor structure and correlated factors. Its first three eigenvalues are: 5.45, 0.68, and 0.68.

Readers familiar with the decision process for the number of dimensions in the factor analytic or item response theory frameworks should realize the potential implications of this feature of the first eigenvalue when eigenvalues are used to make a decision on dimensionality for a set of items. As the variable complexity in the structure and/or the correlation among factors increase, the first eigenvalue gets larger at the expense of the others. This should lead to a decrease in the number of predicted factors in the model when eigenvalues are used as the sole decision criterion. This decrease should be more pronounced as the variable complexity in the factor structure and/or the correlation among factors increase. The consequences for the KG rule which dominates the literature are obvious as one can easily see how many of the first three eigenvalues are greater than one. As variable complexity and/or factor correlation increases the number of predicted factors using the KG rule will decrease. Here, we consider the implications for parallel analysis since it has become more accessible to practitioners as computational tools have been made available in recent years (O'Connor, 2000; Patil, Vivek, Surendra, Mishra, & Danavan, 2007; Raiche, 2011; Revelle, 2014), and parallel analysis was recommended in the literature as a viable alternative in dimensionality assessment of latent structures. Note

Table 1. Simple and Complex Structures with and without Correlated Factors

	Simple (F1)				Complex (F2)			
λ_1	0.70	0.00	0.00		0.60	0.20	0.30	
λ_2	0.70	0.00	0.00		0.60	0.20	0.30	
λ_3	0.70	0.00	0.00		0.60	0.20	0.30	
λ_4	0.00	0.70	0.00		0.30	0.60	0.20	
λ_5	0.00	0.70	0.00		0.30	0.60	0.20	
λ_6	0.00	0.70	0.00		0.30	0.60	0.20	
λ_7	0.00	0.00	0.70		0.20	0.30	0.60	
λ_8	0.00	0.00	0.70		0.20	0.30	0.60	
λ_9	0.00	0.00	0.70		0.20	0.30	0.60	
$\sum\lambda^2$	1.47	1.47	1.47	4.41	1.47	1.47	1.47	4.41

Uncorrelated (R1)			Slightly Correlated (R2)		
1	0	0	1	0.2	0.2
0	1	0	0.2	1	0.2
0	0	1	0.2	0.2	1

that results for parallel analysis are not as obvious as one does not know the expected value of these eigenvalues for random matrices.

Implications for Parallel Analysis

Parallel analysis was originally proposed by Horn (1965) for principal component analysis and can be conceptualized as a more sophisticated way of implementing the KG rule. The eigenvalues from a correlation matrix for uncorrelated variables in the population should all be equal to one. In other words, the eigenvalues from an identity matrix would create a horizontal line at y=1 in a scree plot. The KG rule suggests keeping any dimension that accounts for more than chance variation. Parallel analysis similarly suggests retaining the components that have an eigenvalue greater than what would be expected due to chance, but it acknowledges sampling fluctuations in eigenvalues. Parallel analysis replaces the "one" in the KG rule with a cut-off criterion based on the empirical sampling distributions of the eigenvalues obtained from random data. What is meant by random data is a dataset where all correlations among the variables are zero at the population level.

Horn (1965) originally proposed comparing each sample data eigenvalue to the average of the empirical eigenvalue sampling distribution for the corresponding rank position and retaining the components that have larger sample eigenvalues than the average random data eigenvalue. The original procedure was also extended to common factor analysis by creating empirical sampling distributions of principal axis eigenvalues from the random data (Humphreys & Ilgen, 1969). Since it was proposed, some adjustments have been suggested for the original parallel analysis procedure to improve its efficiency. One practical concern was using the average of the empirical eigenvalue sampling distribution. The use of the average eigenvalue as a criterion implies that the original parallel analysis procedure operates at a significance level of 0.5, which is very generous in terms of the conventional hypothesis testing approach (Buja & Eyuboglu, 1992). This would tend to increase the probability of making a type I error (extracting a factor that actually should not be extracted) and makes the parallel analysis procedure tend to predict too many factors (Glorfeld, 1995). Therefore, it is suggested that the 95th or 99th percentiles of the empirical eigenvalue sampling distribution be used as cutoffs.

Another concern is related to the sensitivity of the eigenvalue sampling distribution to the distributional form used to generate multivariate data with uncorrelated variables. In most applications, the empirical eigenvalue distributions were derived from uncorrelated variables with a multivariate normal distribution. Whether these empirical eigenvalue sampling distributions are sensitive to non-normality was an open question. Simulation studies consistently showed that none of the distributional forms overestimate or underestimate the mean or quantiles of the random data eigenvalue sampling distributions. Both mean and centile estimates were stable across various distributional forms (Buja & Eyuboglu, 1992; Dinno, 2009; Glorfeld, 1995).

Recently, researchers argued that the 95th or 99th percentile eigenvalues generated from random data provide an appropriate null hypothesis only for the first eigenvalue, because the size of the later (noise) eigenvalues are influenced by the presence of the prior significant factors (Green et al., 2012; Turner, 1998). Beyond the first eigenvalue, the sampling distributions of the random data eigenvalues are not directly comparable to the sample eigenvalue estimates unless the previous significant factors have been modeled into the data generation process. Therefore, this would suggest using a separate simulation to test each eigenvalue as opposed to using only one simulation to test all eigenvalues at once. Green et al. (2012) proposed a revised version of parallel analysis that relies on successive simulations to test each eigenvalue independently by taking the magnitude of previous significant eigenvalues into account. Modifications of the parallel analysis procedure using tetrachoric/polychoric correlations have also appeared in the literature in order to determine the necessary number of latent traits when modeling categorical response data (Cho et al., 2009; Drasgow & Lissak, 1983; Finch & Monahan, 2008; Tran & Formann, 2009; Weng & Cheng, 2005).

As the magnitudes of the eigenvalues have been shown to be dependent on the increasing level of variable complexity and/or inter-factor correlations in the underlying structure, the performance of parallel or revised parallel analysis may not be optimal for factor structures displaying these attributes. For demonstration purposes, Table 2 presents four factor structures. Structure 1 represents an independent cluster structure in which each variable loads on only

one factor, while Structures 2 and 3 represent semi-complex structures, and Structure 4 represents a full complex structure with three common factors. For Structure 1, we also generated data by manipulating the correlations among common factors at levels of 0, 0.4, and 0.7. While the choice of correlations for the factors was arbitrary, it was also intentional. We chose 0 for no correlation (orthogonal solution), 0.40 for a moderate amount of correlation, and 0.70 for a more substantial amount of correlation in the factors. For the semi-complex and complex factor structures, we did not simulate correlation among the factors.

Factor Structure 1 was chosen because much of the previous research that has informed practice chose such simple structures with no item complexity or factor correlation. Factor Structure 2 shows each item loading on two dimensions, but with three sets of items loading on different pairs of dimensions, Factor Structure 3 is a bi-factor structure. In a bi-factor structure, all items load on a general factor – in this case F1, but there are also group factors, here F2 and F3. Note that at the primary factor level for this bi-factor structure only two factors will appear. However, hierarchical factor analysis would allow the original structure to be recovered (the general factor [F1] will be at level two). Finally, Factor Structure 4 is presented to specifically show that item inter-relations can manifest themselves at the factor level where the items are complex or at the factor correlation level where the items become less complex, but there are factor correlations. For example if we have a math test, it may be difficult to differentiate algebra and geometry items from an orthogonal solution. If we allow an oblique rotation, the items may become more differentiable, however, the factors will be correlated. We provide more on this later.

Using each of these four factor structures, we simulated 1000 datasets with manifest continuous outcomes under two different sample size conditions (N=100, 500) using a common factor model. Then, we applied parallel analysis using the 95th percentile of the random data principal axis eigenvalues to make a decision about the number of common factors for each simulated dataset. As seen in Table 3, the parallel analysis procedure correctly identified the number of factors in Structure 1 almost all the time regardless of sample size (99.2% of the time for N = 100 and 100% of the time for N = 500) when the correlation among factors was zero. When the correlation among factors was increased to 0.4, parallel analysis still yielded almost perfect results in identifying the number of factors in Structure 1 (98.3% of the time for N = 100 and 100% of the time for N = 500). When the correlation among factors was increased to 0.7, the results were not as good. Then parallel analysis was correct only 22.1% of the time (choosing 3 factors) for the datasets with sample size of 100. One factor was chosen 43.0% of the time and 2 factors 34.7% of the time. In contrast, for sample size of 500 parallel analysis still predicted with 100% accuracy.

For a semi-complex factor structure (Structure 2), the percentage of datasets in which the number of factors was correctly identified for the N=100 sample size condition dropped to 80.7%. Note that the suggested number of factors by the parallel analysis procedure was two for 17.6% of the generated datasets. In contrast, parallel analysis still predicted with 100% accuracy for sample size of 500. For a bi-factor structure (Structure 3), parallel analysis suggested a two factor solution most of the time regardless of sample size (85.7% of the time for N = 100 and 100% of the time for n = 500). For a complex factor structure (Structure 4), the procedure was correct only 0.7% of the time (predicting 3 factors) while predicting one factor 71.2% of the time for the N=100 sample size condition. For the N = 500 sample size condition, the correct classification increased to 35.8%, while predicting two factors the remaining 64.2% of the time.

It may be argued that the poor results presented above occurred due to a technical flaw in the application of the traditional parallel analysis procedure as outlined in the literature (Green et al., 2012; Turner, 1998). As mentioned before, the sampling distribution of the random data eigenvalues is not directly comparable to the sample eigenvalue estimates beyond the first eigenvalue, unless the previous significant eigenvalues have been modeled into the data generation process to test later eigenvalues. For increasing levels of variable complexity or inter-factor correlations, the increase in the first eigenvalue leaves less room remaining for the second or third eigenvalue, so that they were smaller than the corresponding random data eigenvalues. Green et al. (2012) proposed a revised version of parallel analysis that relies on successive simulations to test each eigenvalue independently by taking the magnitude of previous significant eigenvalues into account. Theoretically, the revised version of parallel analysis should help in the above

scenarios. For this reason, the same simulation procedure was repeated for Structures 1, 2, 3, and 4 and the number of factors for each dataset was identified using the revised parallel analysis procedure. The results from the revised parallel analysis procedure are given in Table 4. Although the revised parallel analysis procedure showed noteworthy improvement over the traditional parallel analysis procedure, there was still a significant amount of replications with incorrectly identified number of factors, particularly when the sample size is smaller. The results imply that identifying the number of major factors with parallel analysis is less than optimal, and there is a risk of under-prediction under certain conditions. In factor analysis, under-extraction may have serious consequences as discussed in the literature (Fava & Velicer, 1992, 1996; Wood, Tataryn, & Gorsuch, 1996).

To summarize, the first eigenvalue becomes dominant while the rest of the eigenvalues get closer to zero when the variable complexity and/or correlation among factors increase. Consequently, this phenomenon influences the implementation of parallel analysis. It should be noted that the complexity in Structures 2 and 4 was caused by fairly insignificant loadings; values of 0.2 or 0.3 that researchers tend to ignore in most practical applications. This shows that it doesn't take much variable complexity for dimensionality assessment procedures based on eigenvalues to under-predict the number of major factors. We realize that we only produced results for parallel analysis, but other procedures such as the Kaiser-Guttman rule will tend to produce results in the same direction.

Conclusions

Both parallel analysis and revised parallel analysis are recommended in the literature and seem to be the most viable alternatives among the methods to ascertain dimensionality that rely on eigenvalues. The increasing availability of computational tools for these methods makes them more likely to be used in applied research. However, we argue that they may perform poorly in certain conditions. The poor results are not directly related to the methods, but due to the fact that the interpretations of eigenvalues are ambiguous under certain conditions. When variable complexity and/or inter-factor correlations increase, the first eigenvalue gets larger at the expense of the others. Due to the large first few eigenvalues, the remaining later eigenvalues may be smaller than random data eigenvalues. As a result, parallel analysis and revised parallel analysis tend to predict fewer substantial dimensions

Inter-relationships of items from related latent dimensions can manifest themselves as factor complexities (items loading on multiple factors) or as

Table 2. Population latent factor structures used for the simulation

	Factor Structure 1			Factor Structure 2			Factor Structure 3			Factor Structure 4		
Item	F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3
1	0.7	0	0	0.7	0	0.3	0.7	0.4	0	0.7	0.3	0.3
2	0.6	0	0	0.6	0	0.2	0.6	0.3	0	0.6	0.2	0.2
3	0.7	0	0	0.7	0	0.3	0.5	0.4	0	0.7	0.3	0.3
4	0.6	0	0	0.6	0	0.2	0.7	0.3	0	0.6	0.2	0.2
5	0.7	0	0	0.7	0	0.3	0.6	0.4	0	0.7	0.3	0.3
6	0	0.7	0	0.3	0.6	0	0.5	0.3	0	0.3	0.6	0.2
7	0	0.6	0	0.2	0.5	0	0.7	0.4	0	0.2	0.5	0.3
8	0	0.5	0	0.3	0.4	0	0.6	0.3	0	0.3	0.4	0.2
9	0	0.7	0	0.2	0.6	0	0.5	0	0.4	0.2	0.6	0.3
10	0	0.6	0	0.3	0.5	0	0.7	0	0.3	0.3	0.5	0.2
11	0	0	0.7	0	0.3	0.5	0.6	0	0.4	0.2	0.3	0.5
12	0	0	0.6	0	0.2	0.4	0.5	0	0.3	0.3	0.2	0.4
13	0	0	0.5	0	0.3	0.3	0.7	0	0.4	0.2	0.3	0.3
14	0	0	0.7	0	0.2	0.5	0.6	0	0.3	0.3	0.2	0.5
15	0	0	0.6	0	0.3	0.4	0.5	0	0.4	0.2	0.3	0.4

items on separate dimensions that are related (factors correlated). Thus, a complex structure with orthogonal factors, such as Structure 4 in Table 2, can be equivalent to a much simpler factor structure with oblique factors after rotation. Note that it is the researcher's choice to allow or disallow factor correlations. Thus, the researcher's choice will influence the result we see and subsequently our understanding of reality. After an oblique rotation of Structure 4, we get factor F3 below with its corresponding factor correlation matrix.

$$\mathbf{F}_3 = \begin{pmatrix} 0.76 & 0.04 & 0.04 \\ 0.70 & -0.02 & -0.02 \\ 0.76 & 0.04 & 0.04 \\ 0.70 & -0.02 & -0.02 \\ 0.76 & 0.04 & 0.04 \\ 0.13 & 0.64 & -0.04 \\ -0.01 & 0.50 & 0.17 \\ 0.20 & 0.36 & 0.02 \\ -0.05 & 0.64 & 0.13 \\ 0.17 & 0.50 & -0.01 \\ -0.01 & 0.15 & 0.51 \\ 0.21 & 0.01 & 0.37 \\ 0.06 & 0.22 & 0.24 \\ 0.17 & -0.03 & 0.51 \\ 0.03 & 0.19 & 0.37 \end{pmatrix} \qquad \boldsymbol{\varphi} = \begin{pmatrix} 1 & & \\ 0.67 & 1 & \\ 0.68 & 0.67 & 1 \end{pmatrix}$$

In this case, it may be argued that the factors become redundant as the correlations among factors increase, and parallel analysis correctly advises the analyst to retain fewer factors. The number of dimensions reported by the parallel analysis procedure can be thought of as a lower bound to the number of dimensions and becomes a measure of nonredundancy of dimensions. Another alternative explanation could be that high inter-factor correlations imply a second-

Table 4. The percentage of decisions using revised parallel analysis procedure

	Number of Identified Factors				
	1	2	3	4	5
N =100					
Factor Structure 1 (ϕ=0)			95.30%	4.49%	0.21%
Factor Structure 1 (ϕ=.4)			96.50%	3.30%	0.20%
Factor Structure 1 (ϕ=.7)			95.10%	4.60%	0.30%
Factor Structure 2 (ϕ=0)		11.60%	85.10%	3.20%	0.10%
Factor Structure 3 (ϕ=0)	0.50%	97.50%	2.00%		
Factor Structure 4 (ϕ=0)	21.80%	73.60%	4.50%	0.10%	
N =500					
Factor Structure 1 (ϕ=0)			95.40%	4.60%	
Factor Structure 1 (ϕ=.4)			94.30%	5.70%	
Factor Structure 1 (ϕ=.7)			94.50%	5.40%	0.10%
Factor Structure 2 (ϕ=0)			95.10%	4.90%	
Factor Structure 3 (ϕ=0)		96.20%	3.70%	0.10%	
Factor Structure 4 (ϕ=0)		21.40%	77.60%	1.00%	

Table 3. The Percentage of decisions using parallel analysis procedure

	Number of Identified Factors			
	1	2	3	4
N =100				
Factor Structure 1 (ϕ=0)			99.20%	0.80%
Factor Structure 1 (ϕ=.4)		1.10%	98.30%	0.60%
Factor Structure 1 (ϕ=.7)	43.00%	34.70%	22.10%	0.20%
Factor Structure 2 (ϕ=0)	0.20%	17.60%	80.70%	1.50%
Factor Structure 3 (ϕ=0)	14.30%	85.70%		
Factor Structure 4 (ϕ=0)	71.20%	28.10%	0.70%	
N =500				
Factor Structure 1 (ϕ=0)			100.00%	
Factor Structure 1 (ϕ=.4)			100.00%	
Factor Structure 1 (ϕ=.7)			100.00%	
Factor Structure 2 (ϕ=0)			100.00%	
Factor Structure 3 (ϕ=0)		100.00%		
Factor Structure 4 (ϕ=0)		64.20%	35.80%	

order general factor, and thus the eigenvalues and parallel analysis correctly identifies this strong second-order general factor, while ignoring the first-order factors. While we acknowledge these alternative perspectives, we note the practical challenges. Suppose that the underlying structure is a bifactor structure as in Structure 3 in Table 2. If a researcher uses eigenvalues and parallel analysis as decision makers, it's very likely that the parallel analysis would recommend two dimensions as shown in our simulations. If the researcher extracts two factors and then use an oblique rotation, there would be a nice independent cluster structure with two correlated dimensions. However, the interpretation of this two dimensional solution would be different than the interpretation of the original bi-factor structure. Therefore, we urge researchers to carefully consider and interpret the results of parallel analysis or revised parallel analysis procedures.

We also note that some scholars already indicated their reservations with regard to use of eigenvalues in dimensionality assessment. McDonald (1981) stated that "…it's important to recognize that there is no direct relationship between the proportion of variance due to the first common factor and the presence or absence of additional common factors. Certainly, it is easy to invent realistic numerical examples in which a multidimensional set of items has higher first-factor variance than a unidimensional set. (p. 112)" Similarly, Mulaik (2010) made a numerical example where there were seven factors with inter-factor correlations in the .50s and a few cross-loadings and parallel analysis indicated only three dimensions. He also noted "…Not only does the first eigenvalue depend on the magnitude of the loadings on the factors, but also on the correlations among them. So, those Monte Carlo studies of the PA method that only study simple orthogonal factors do not realistically portray situations frequently encountered with real data. (p.191)" To conclude, practitioners should be careful in interpreting the number of dimensions indicated by parallel analysis, revised parallel analysis, or any other method that relies on the magnitude of eigenvalues in dimensionality assessment. The resulting number of factors should be taken as a lower bound to dimensionality, and one should conduct a thorough exploratory analysis by considering alternative plausible models with oblique/orthogonal (to detect higher order factors) multiple dimensions (to detect other plausible models with differing number of factors that may be more interpretable).

References

Bock, R. D., Gibbons, R., & Muraki, E. (1988). Full-Information Item Factor Analysis. *Applied Psychological Measurement*, 12(3), 261-280.

Buja, A., & Eyuboglu, N. (1992). Remarks on Parallel Analysis. *Multivariate Behavioral Research*, 27(4), 509-540.

Browne, M. W., Cudeck, R., Tateneni, K., & Mels, G. (2004). CEFA: Comprehensive exploratory factor analysis, Version 2.00 [Computer software and manual]. The Ohio State University, Columbus.

Cattell, R. B. (1966). The Scree Test For the Number of Factors. *Multivariate Behavioral Research*, 1(2), 245-276.

Cho, S.J., Li, F., & Bandalos, D. (2009). Accuracy of the Parallel Analysis Procedure With Polychoric Correlations. *Educational and Psychological Measurement*, 69(5), 748-759.

Crawford, A. V., Green, S. B., Levy, R., Lo, W.J., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of Parallel Analysis Methods for Determining the Number of Factors. *Educational and Psychological Measurement*, 70(6), 885-901.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.

Dinno, A. (2009). Exploring the Sensitivity of Horn's Parallel Analysis to the Distributional Form of Random Data. *Multivariate Behavioral Research*, 44, 362-388.

Drasgow, F., & Lissak, R. I. (1983). Modified Parallel Analysis: A Procedure for Examining the Latent Dimensionality of Dichotomously Scored Item Responses. *Journal of Applied Psychology*, 68(3), 363-373.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299.

Fava, J. L., & Velicer, W. F. (1996). The effects of underextraction in factor and component analyses. *Educational and Psychological Measurement*, 56(6), 907-929.

Finch, H., & Monahan, P. (2008). A Bootstrap Generalization of Modified Parallel Analysis for IRT Dimensionality Assessment. *Applied Measurement in Education*, 21, 119-140.

Floyd, F. J. and Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessments*, 7(3), 286-299.

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: a critical review and analysis. *Personnel Psychology*, 39(2), 291-314.

Gessaroli, M. E., De Champlain, A. F., & Folske, J. C. (March, 1997). Assessing Dimensionality Using a Likelihood-Ratio Chi-Square Test Based on a Non-linear Factor Analysis of Item Response Data. Paper Presented at the Annual Meeting of the National Council on Measurement in Education, Chicago, IL.

Gessaroli, M., & De Champlin, A. (1996). Using an approximate chi-square statistic to test the number of dimensions underlying the responses to a set of items. *Journal of Educational Measurement*, 157-179.

Glorfeld, L. W. (1995). An Improvement on Horn's Parallel Analysis Methodology for Selecting the Correct Number of Factors to Retain. 55(3), 377-393.

Green, S. B., Levy, R., Thompson, M. S., Lu, M., & Lo, W. (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.

Guttman, L. (1954). Some necessary and sufficient conditions for common factor analysis. *Psychometrika*, 19, 149-161.

Hattie, J. (1985). Methodology Review: Assessing Unidimensionality of Tests and Items. *Applied Psychological Measurement*, 9(2), 139-164.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor Retention Decisions in Exploratory Factor Analysis: A Tutorial on Parallel Analysis. *Organizational Research Methods*, 7(2), 191-205.

Henson, R. K., & Roberts, J. K. (February, 2001). A Meta-Analytic Review of Exploratory Factor Analysis Reporting Practices in Published Research. Paper Presented at the Annual Meeting of the Southwest Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED449227).

Holzinger, K. J. & Swineford, F. (1939). A study in factor analysis: The stability of a bi-factor solution. Supplementary Educational Monographs, No. 48. Chicago:Department of Education, University of Chicago.

Horn, J. L. (1965). A Rationale and Test for the Number of Factors in Factor Analysis. *Psychometrika*, 30(2), 179-185.

Humphreys, L. G., & Ilgen, D. R. (1969). Note on a Criterion for the Number of Common Factors. *Educational and Psychological Measurement*, 29(3), 571-578.

Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, 20(1), 141-151.

Lord, F. M. (1980). Applications of item response theory to practical testing problems. Hillsdale, New Jersey: Erlbaum.

Lumsden, J. (1957). A factorial approach to unidimensionality. *Australian Journal of Psychology*, 9, 105-111.

Lumsden, J. (1961). The construction of unidimensional tests. *Psychological Bulletin*, 58, 122-131.

Maydeu-Olivares, A. (2001). Multidimensional Item Response Theory Modeling of Binary Data: Large Sample Properties of NOHARM Estimates. *Journal of Educational and Behavioral Statistics*, 26(1), 51-71.

McDonald, R. P. (1981). The dimensionality of tests and items. British *Journal of Mathematical and Statistical Psychology*, 34(1), 100-117.

Mulaik, S. A. (2010). Foundations of factor analysis (2nd ed.) Chapman & Hall/CRC.

Muthén, L., & Muthén, B. (1998-2010). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

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*, 32(3), 396-402.

Patil, Vivek H., Surendra N. Singh, Sanjay Mishra,and D. Todd Donavan (2007), "Parallel Analysis Engine to Aid Determining Number of Factors to Retain [Computer software]. Available from http://smishra.faculty.ku.edu/parallelengine.htm

Preacher, K. J. and MacCallum, R. C. (2003). Repairing Tom Swift's Electric Factor Analysis Machine *Understanding Statistic* 2(1) 13 – 43. http://www.unc.edu/~rcm/papers/tomswift.pdf .

Raiche, G. (2011). nFactors: an R package for parallel analysis and nongraphical solutions to the Cattell scree test. R package version 2.3.3

Reckase, M. D. (1979). Unifactor Latent Trait Models Applied to Multifactor Tests: Results and Implications. *Journal of Educational Statistics*, 4(3), 207-230.

Revelle, W. (2014) psych: Procedures for Personality and Psychological Research, Northwestern University,

Evanston, Illinois, USA, http://CRAN.R-project.org/package=psych Version= 1.4.8.

Tran, U., & Forman, A. (2009). Performance of Parallel Analysis in Retrieving Unidimensionality in the Presence of Binary Data. *Educational and Psychological Measurement*, 50-61.

Turner, N. E. (1998). The Effect of Common Variance and Structure Pattern on Random Data Eigenvalues: Implications for the Accuracy of Parallel Analysis. *Educational and Psychological Measurement*, 58(4), 541-568.

Vernon, P.E. (1950). An application of factorial analysis to the study of test items. British *Journal of Psychology, Stat. Section*, 3, 1-15.

Weng, L.J., & Cheng, C.P. (2005). Parallel Analysis with Unidimensional Binary Data. *Educational and Psychological Measurement*, 65(5), 697-716.

Wood, J. M., Tataryn, D. J., & Gorsuch, R. L. (1996). Effects of under- and overextraction on principal axis factor analysis with varimax rotation. *Psychological Methods*, 1(4), 354 – 365.

Zhang, J. (1996). Some fundamental issues in item response theory with applications. Dissertation Abstract International, 57(11).

Citation:

Zopluoglu, Cengiz & Davenport Jr., Ernest C. (2017). A Note on Using Eigenvalues in Dimensionality Assessment. *Practical Assessment, Research & Evaluation*, 22(7). Available online: http://pareonline.net/getvn.asp?v=22&n=7

Corresponding Author

Cengiz Zopluoglu
Department of Educational and Psychological Studies
University of Miami
Max Orovitz Building, 333 A
1570 Levante Ave, Coral Gables, FL 33146

email: c.zopluoglu [at] miami.edu