Original Research Article

Educational and Psychological Measurement
1–16

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/00131644241308528
journals.sagepub.com/home/epm
Sage

A Comparison of the Next Eigenvalue Sufficiency Test to Other Stopping Rules for the Number of Factors in Factor Analysis

Pier-Olivier Caron[1]

Abstract

A plethora of techniques exist to determine the number of factors to retain in exploratory factor analysis. A recent and promising technique is the Next Eigenvalue Sufficiency Test (NEST), but has not been systematically compared with well-established stopping rules. The present study proposes a simulation with synthetic factor structures to compare NEST, parallel analysis, sequential χ^2 test, Hull method, and the empirical Kaiser criterion. The structures were based on 24 variables containing one to eight factors, loadings ranged from .40 to .80, inter-factor correlations ranged from .00 to .30, and three sample sizes were used. In total, 360 scenarios were replicated 1,000 times. Performance was evaluated in terms of accuracy (correct identification of dimensionality) and bias (tendency to over- or underestimate dimensionality). Overall, NEST showed the best overall performances, especially in hard conditions where it had to detect small but meaningful factors. It had a tendency to underextract, but to a lesser extent than other methods. The second best method was parallel analysis by being more liberal in harder cases. The three other stopping rules had pitfalls: sequential χ^2 test and Hull method even in some easy conditions; the empirical Kaiser criterion in hard conditions.

Keywords

exploratory factor analysis, number of factors, Monte Carlo simulation

[1]Université TÉLUQ, Montréal, Québec, Canada

Corresponding Author:
Pier-Oliver Caron, Département des Sciences Humaines, Lettres et Communication, Université TÉLUQ, 5800, Rue Saint-Denis, Bureau 1105, Montréal, Québec, Canada H2S 3L5.
Email: pocaron19@gmail.com

Determining the correct number of factors remains a long-standing challenge in exploratory factor analysis. Factor analysis is used to reduce the dimensionality of data sets and, when no specific factorial structure is known a priori, deciding how many latent factors to extract is a decisive question. Underextraction results in substantial bias on factor loadings, whereas overextraction can lead to factor splitting.

There has been a wide variety of techniques, called stopping rules, developed over the years to determine the correct number of factors. A recent and promising technique which has already shown excellent performances in previous studies is the Next Eigenvalue Sufficiency Test (NEST; Achim, 2017) which has been shown to be robust to cross-loadings (Brandenburg & Papenberg, 2024), to have a low false-positive rate (Achim, 2017, 2021) to be accurate in circumplex models (Brandenburg & Papenberg, 2024), to be strongly theoretically aligned with factor analysis conceptualization (Brandenburg & Papenberg, 2024), and, being a nonparametric approach, to be more flexible and broadly applicable across different types of data sets (Achim, 2021). In the work by Achim (2017), NEST was already shown to be better at identifying the number of dimensions than other stopping rules, such as parallel analysis (PA; Horn, 1965), revised PA (Green et al., 2012, 2015), minimum average partial correlation (Velicer, 1976), and comparison data (Ruscio & Roche, 2012) over a limited set of scenarios. It was also shown to have a clear advantage over exploratory graph analysis in the work by Brandenburg and Papenberg (2024).

Despite this promising start, the advantages of NEST in deciphering the number of dimensions have remained relatively undiscussed. Among four recent extensive comparisons (Auerswald & Moshagen, 2019; Finch, 2023; Lim & Jahng, 2019; Neishabouri & Desmarais, 2024), none included NEST as a potential contender. Therefore, it is unclear how well NEST performs compared with a wide range of stopping rules in extensive simulation studies. Given the very few comparisons of NEST to other methods, the purpose of the current study is to compare NEST performance against the recommended stopping rules (Auerswald & Moshagen, 2019; Finch, 2023; Lim & Jahng, 2019; Neishabouri & Desmarais, 2024), which are PA (Horn, 1965), sequential χ^2 model test (SMT; Lawley, 1940), Hull method (HULL; Lorenzo-Seva et al., 2011), and empirical Kaiser criterion (EKC; Braeken & van Assen, 2017). They will serve as a benchmark for the current study.

Method

This simulation was carried out in **R** (R Core Team, 2023). The method NEST, PA, and EKC were homemade functions. The HULL and the sequential χ^2 model tests were from the `EFAtools` package (Steiner & Grieder, 2020). The `MASS` package (Venables & Ripley, 2002) was used for data generation.

Methods to Decide the Number of Retained Factors

Parallel Analysis. PA (Horn, 1965) compares the empirical eigenvalues to the average eigenvalues derived from a random multivariate normal distribution with uncorrelated variables (the identity matrix). The random samples are designed to have the same number of observations and variables as the actual data set. The criterion for extracting factors in PA is that their eigenvalues must exceed the average eigenvalue of the random samples.

Although, some studies (for instance, Glorfeld, 1995) argued for the 95th percentile of sample eigenvalues instead of traditional PA (average, or 50th percentile), Auerswald and Moshagen (2019) and Lim and Jahng (2019) found that the traditional PA performed better than 95th PA. Meanwhile, Finch (2023) and Neishabouri and Desmarais (2024) only used and recommended traditional PA. Thus, only the mean variant of the PA method will be used herein.

The Next Eigenvalue Sufficiency Test. Like PA, NEST uses a synthetic correlation matrix and resamples over a large amount (over a thousand) of data sets with the same number of subjects as the target data set. The main difference is that NEST sequentially uses synthetic correlation matrices containing k factors from which the $(k+1)^{th}$ sampled eigenvalues are recorded.

Contrary to PA which only uses the identity matrix for all eigenvalues (a well-known issue with PA, Turner, 1998), NEST updates the synthetic correlation matrix at every step. At $k=0$, NEST uses, like PA, the identity correlation matrix to generate these samples. If the $(k+i)^{th}$ empirical eigenvalue is higher than the 95th percentile of the $(k+1)^{th}$ sampled eigenvalues, the synthetic correlation matrix is updated to contain the k-factor model's loadings. In the common factor model defined as

$$\mathbf{X} = \Lambda\boldsymbol{\xi} + E \tag{1}$$

where $\mathbf{X}$ are the manifest variables, Λ are the loadings, $\boldsymbol{\xi}$ are the factor scores, and E is the residual error. If the manifest variables are standardized, we can express the correlation matrix as

$$\mathbf{R} = \mathbb{E}(\mathbf{X}\mathbf{X}'). \tag{2}$$

From the loadings, $\mathbf{L}$, and the uniquenesses, $\mathbf{P} = \mathbb{E}(ee')$, factor analysis is the model for the correlation matrix, $\mathbf{R}$, of $\mathbf{X}$,

$$\mathbf{R} = \mathbf{L}\mathbf{L}' + \mathbf{P}. \tag{3}$$

When $k > 0$, the synthetic correlation matrix, $\mathbf{R}$, is based upon Equation 3, that is, from loadings, $\mathbf{L}$ and communalities, $\mathbf{P}$, extracted from the k-factor model. The synthetic correlation matrices are full correlation matrices. At this point, it is clear that $k=0$ is the specific case where $\mathbf{L}$ is empty (no factors, no loadings), and $\mathbf{P}$ is an identity matrix.

There are numerous ways to estimate the loadings $\mathbf{L}$, such as maximum likelihood (ML), principal axis factoring, minimum rank factor analysis, or any other factor analysis method. As such, NEST is similar to revised PA (Green et al., 2012) which first proposed to generate artificial data to produce an empirical probability for the

$(k+1)^{\text{th}}$ eigenvalues from a k-factor model. In its computation though, revised PA uses the squared multiple regressions as the communality estimates and the associated eigenvalues, which reduced its performance (Achim, 2017).

Once the synthetic correlation matrix is updated, there is again a resampling over a large number of data sets from which the $(k+1)^{th}$ sampled eigenvalues are recorded. The $(k+i)^{th}$ empirical eigenvalue is compared to the 95th percentile of the $(k+1)^{th}$ sampled eigenvalues. This procedure continues until the test fails to reject the $(k+1)^{th}$ eigenvalues from the 95th percentile sampled eigenvalues obtained from the k-factor model. When the test fails, k factors are deemed sufficient to account for the data set.

Sequential χ^2 Model Test. Common factor models are often evaluated using the likelihood ratio test statistic (Lawley, 1940) with ML estimation. This test assesses whether the model's implied covariance matrix is equal to the population covariance matrix. The test statistic follows an asymptotic χ^2 distribution if the observed variables conform to a multivariate normal distribution and other underlying assumptions are met (Bollen, 1989).

To determine the appropriate number of factors in the model, the likelihood ratio test can be sequentially applied, starting with a zero-factor model. If the χ^2 test statistic is statistically significant (e.g., $p<.05$), it suggests that a model with one additional factor (unidimensional factor model) should be estimated and tested. This iterative process continues until a nonsignificant result is obtained, indicating the identification of the appropriate number of common factors.

Hull Method. The HULL (Lorenzo-Seva et al., 2011) is an approach inspired by the Hull heuristic. This method, akin to nongraphical versions of Cattell's scree plot, seeks to identify an *elbow* as evidence for the appropriate number of common factors. The HULL utilizes goodness-of-fit indices in relation to the model degrees of freedom instead of relying on eigenvalues. It is based on the goodness-of-fit index (better fit equals better models) and the viability of the model (more complex models with lower fit index are unviable). The *elbow* is determined as the point where, concerning the change in the model's degrees of freedom, there is a substantial increase in model fit compared with a lower number of factors, while the model fit is only marginally lower compared with a higher number of factors. This criterion is established by considering all viable fit values in relation to both their preceding and subsequent fit values.

Empirical Kaiser Criterion. The EKC (Braeken & van Assen, 2017) is an approach that incorporates random sample variations of the eigenvalues in the Kaiser–Guttman criterion. On a population level, both criteria are equivalent. On the sample level, the criterion is based on the distribution of eigenvalues of an identity matrix, which asymptotically follows a Marchenko–Pastur distribution as

$$\lambda_0 = \left(1 + \sqrt{p/n}\right)^2$$

for the first and then corrected for the next ones

$$\lambda_j = \max\left(\frac{p\sum_{i=0}^{j}\lambda}{p-j-1}\left(1+\sqrt{p/n}\right)^2, 1\right)$$

The first empirical eigenvalues above the criteria are retained. Like the Kaiser–Guttman criterion, the value of one is the minimum to consider a factor as meaningful.

Simulations

Simulations were similar to previous studies (Auerswald & Moshagen, 2019; Caron, 2019; Finch, 2023; Lim & Jahng, 2019; Peres-Neto et al., 2005) using synthetic factorial structures. To compare the methods, we incorporated a wide range of data conditions that are challenging, but realistic in psychological and biological research. Data were generated using random multivariate normal distributions from factorial structures based on 24 variables with the number of factors (ξ) ranging between one and eight factors, that is, common denominators of 24. The population loadings (δ) were .40, .50, .60 .70, and .80 in a given scenario; correlations between factors (ρ) were .00, .10, .20, and .30. The first eight eigenvalues of every factor model are presented in the Online Supplementary Material. Three sample sizes (N), 120, 240, and 480 were considered to reflect the 5, 10, and 20 subjects per item. In total, 360 scenarios were repeated 1,000 times. The performance of the stopping rules were evaluated in terms of accuracy, the percentage of correct identification of the number of factors ($p(z=\xi)*100$), and bias, the tendency to over- or underestimate ($\bar{z}-\xi$).

Results

The results section is divided into three parts: easy (one and two factors), intermediate (three and four factors), and hard cases (six to eight factors). All figures are structured similarly. The x-axis is the sample size (N.); the colored lines are the different methods (see the legend). Each column represents a different population loading in the given condition. Rows are divided by two conditions, the correlations between factors (ρ) and the number of factors (ξ) in the scenario. The y-axis represents either the accuracy (ranging from 0 to 100) or bias (ranging from −8 to 8).

Easy Cases

Figure 1 illustrates easy conditions and shows that PA, NEST, and EKC had the best performances, having close to 100% correct identification. This result was obtained regardless of sample sizes, loadings, correlations between factors, and the number of factors. SMT and HULL were not the most accurate methods, both having correct identification close to 80%, ranging between 70% and 90% depending on the conditions. Figure 2 shows that SMT had a tendency to overestimate, whereas HULL had

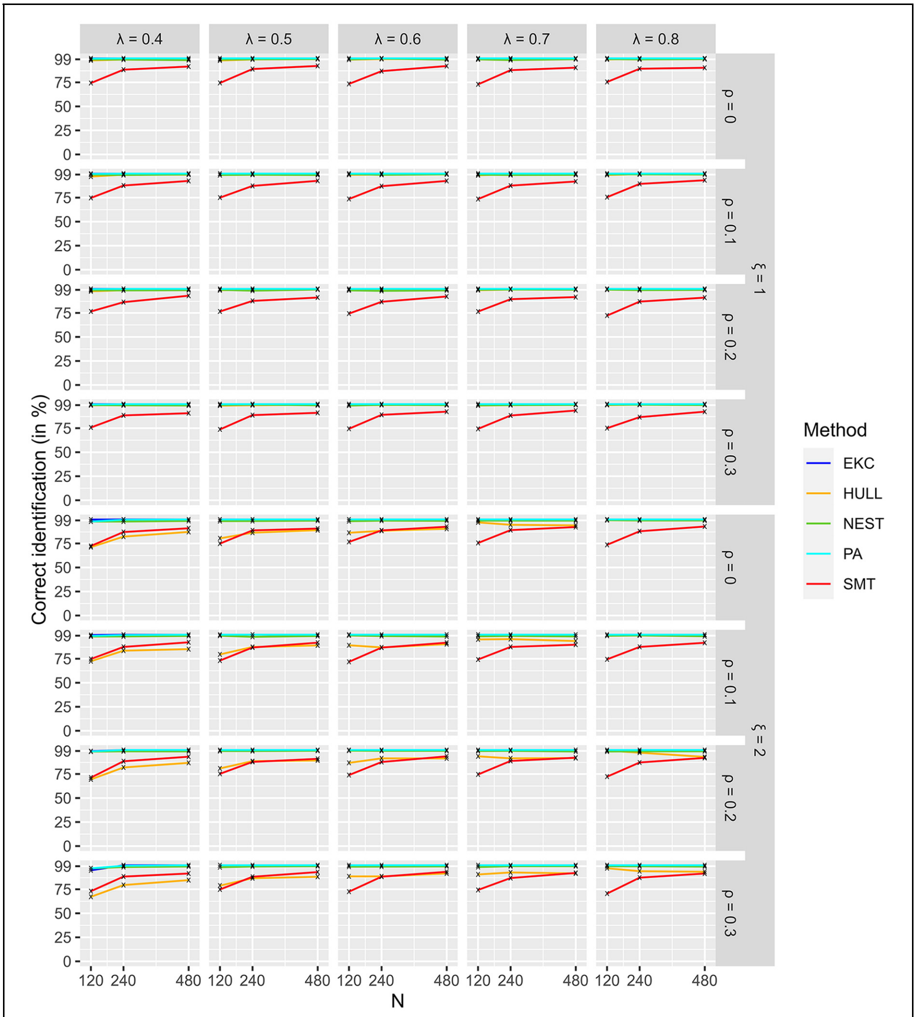

Figure 1. Identification of Dimensionality for Conditions With 1 and 2 Factors (Easy).
Note. EKC = empirical Kaiser criterion; NEST = Next Eigenvalue Sufficiency Test; PA = parallel analysis; SMT = sequential χ^2 model test; ρ = correlations between factors; ξ = the number of factors; λ = loading; N = sample size.

a tendency to underestimate the number of factors in the $\xi = 2$ condition. Overall, the results depicted in Figure 1 indicate that the conditions were so easy that loading values and correlations between factors did not seem to have a substantial impact on the performances.

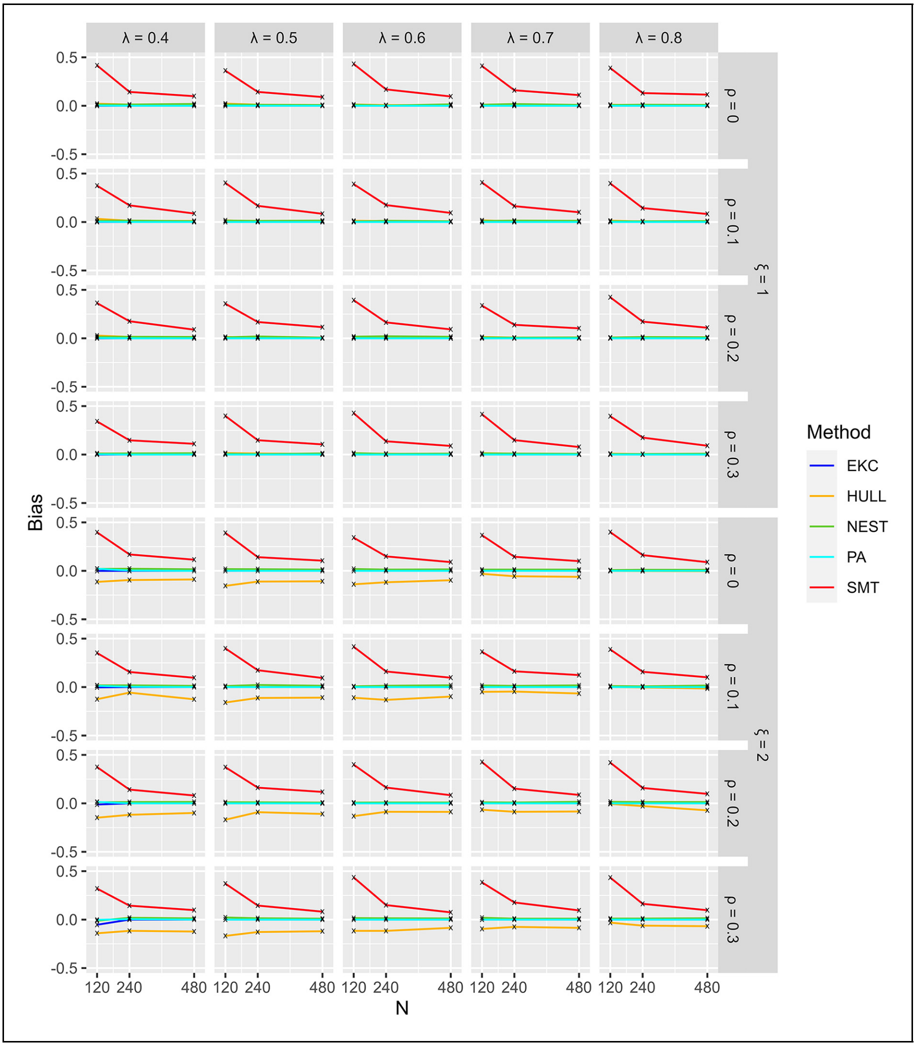

Figure 2. Bias for Conditions With 1 and 2 Factors (Easy).
Note. EKC = empirical Kaiser criterion; NEST = Next Eigenvalue Sufficiency Test; PA = parallel analysis; SMT = sequential χ^2 model test; ρ = correlations between factors; ξ = the number of factors; λ = loading; N = sample size.

Intermediate Cases

Figure 3 shows that PA and NEST had the best overall performances in the intermediate condition. The worst performances were in the harder conditions pictured in Figure 3, that is, $\lambda < .5$ and $N = 120$. Otherwise, they showed excellent results. In

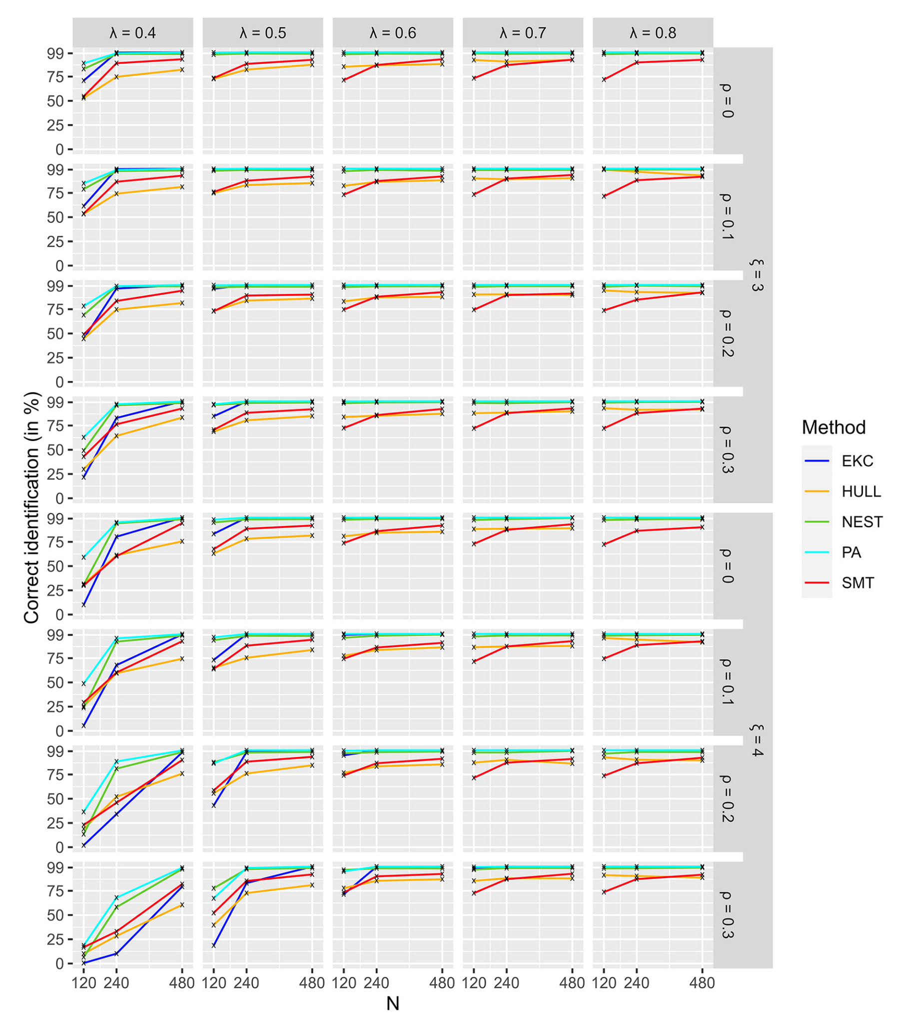

Figure 3. Identification of Dimensionality for Conditions With 3 and 4 Factors (Intermediate).

Note. EKC = empirical Kaiser criterion; NEST = Next Eigenvalue Sufficiency Test; PA = parallel analysis; SMT = sequential χ^2 model test; ρ = correlations between factors; ξ = the number of factors; λ = loading; N = sample size.

intermediate conditions ($\xi = 3$ or 4) shown in Figure 3, SMT and HULL still struggled in the 70% to 90% correction identification. Again, correlations between factors did not seem to have a substantial impact on the performances. Loading values (λ) started to show some influences at the low values of .40 and .50 (harder cases). EKC had

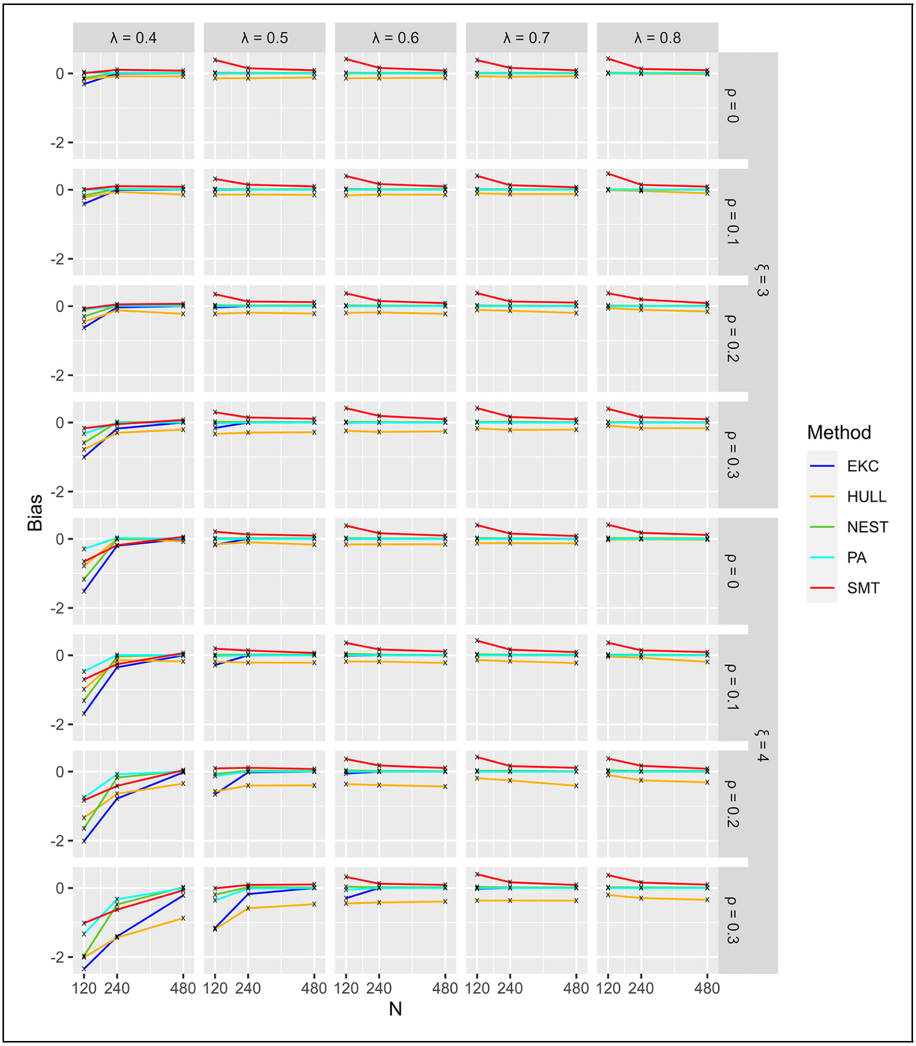

Figure 4. Bias for Conditions With 3 and 4 Factors (Intermediate).

Note. EKC = empirical Kaiser criterion; NEST = Next Eigenvalue Sufficiency Test; PA = parallel analysis; SMT = sequential χ^2 model test; ρ = correlations between factors; ξ = the number of factors; λ = loading; N = sample size.

poorer performance than SMT and HULL at low sample sizes ($n = 120$). According to Figure 4, SMT and HULL had the same pattern. HULL underestimated every condition, whereas SMT overestimated in most conditions, except the low sample size ($n = 120$) and low loadings ($\lambda = .40$) where all stopping rules underestimated.

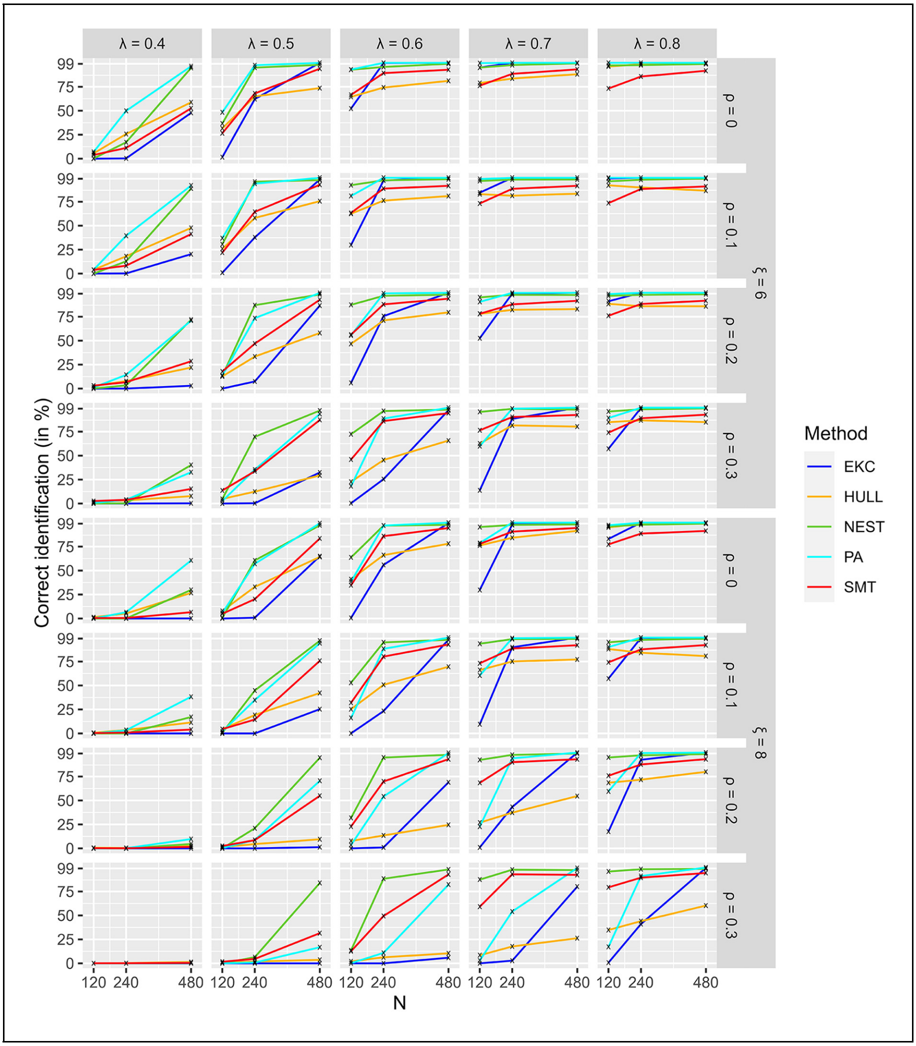

Figure 5. Identification of Dimensionality for Conditions With 6 and 8 Factors (Hard).
Note. EKC = empirical Kaiser criterion; NEST = Next Eigenvalue Sufficiency Test; PA = parallel analysis; SMT = sequential χ^2 model test; ρ = correlations between factors; ξ = the number of factors; λ = loading; N = sample size.

Hard Cases

Figure 5 presents the hardest conditions when the number of factors is $\xi = 6$ and 8. NEST and PA stand out the most in identifying the number of factors compared with other stopping rules. They were not perfect though as they were still incorrectly identifying the dimensionality in very hard conditions, especially when $\lambda < .50$ and $n < 240$. Nevertheless, they were the closest to the correct identification on average compared with other stopping rules. In the hardest case of $\lambda = .4$, PA appeared superior to NEST under the $\xi = 8$ condition, but this advantage is expected to vanish as sample size increases, as shown in the $\xi = 6$, $\lambda = .4$ conditions. As it is already known, PA struggles in oblique structure $\rho > 0$ (Caron, 2019), which was noticeable in the $\lambda = .6$ conditions. NEST was much less influenced by the correlations between factors. Finally, in the most challenging conditions, SMT had an generally good performance, never better than NEST, but sometimes better than PA when $n = 120$. EKC and HULL had the worst performances in these conditions.

Figure 6 shows that all methods had a tendency to underestimate, as expected. Given the already-known fact that SMT overestimated, see Figures 2 and 4, it is not surprising that the stopping rules are less biased than the others.

Discussion

The purpose of the current study was to present the stopping rules NEST and compare its performance against four recommended stopping rules. Overall, NEST showed the best performance, especially in difficult conditions. NEST was good at identifying small but meaningful factors (very small eigenvalues). It was followed by PA who sometimes outperformed NEST, partially due to PA being more liberal by using the average sampled eigenvalues, whereas NEST used the more conservative 95th percentile. This, however, was only beneficial in the $\lambda = .4$ conditions where it gave an advantage to PA but was less effective when $\lambda > .4$. In those cases, the performances between NEST and PA were nearly equal or NEST was better. EKC had very good performance in easy conditions, but worsened as conditions became harder. These harder conditions had factors with eigenvalues close to or below one once sampling error was added, which EKC cannot detect by definition. Thus, EKC cannot be used when small factors are theoretically considered relevant. Finally, SMT and HULL performed poorly in easy scenarios, achieving close to 80% correct identification, whereas the other three had close to 99%. SMT showed an overestimation bias in these conditions which proved helpful in harder conditions where its performance almost matched NEST's performance. HULL had the worst performance overall, in the easy, intermediate, or hard conditions.

The current study extends previous work by investigating recommended stopping rules, ensuring that EFA methodologies are up-to-date with the latest techniques. By focusing on the performance of stopping rules under difficult conditions, which have been less frequently addressed (Caron, 2019), the study demonstrates the robustness of NEST in challenging scenarios, encouraging its broader adoption in EFA. Comparing NEST with recommended stopping rules ensures that simulations results can be analyzed, disentangled, and synthesized effectively. There are always new methods to investigate and compare, such as the Comparison Data Forest (Goretzko & Ruscio, 2024; combining machine learning techniques with comparison data

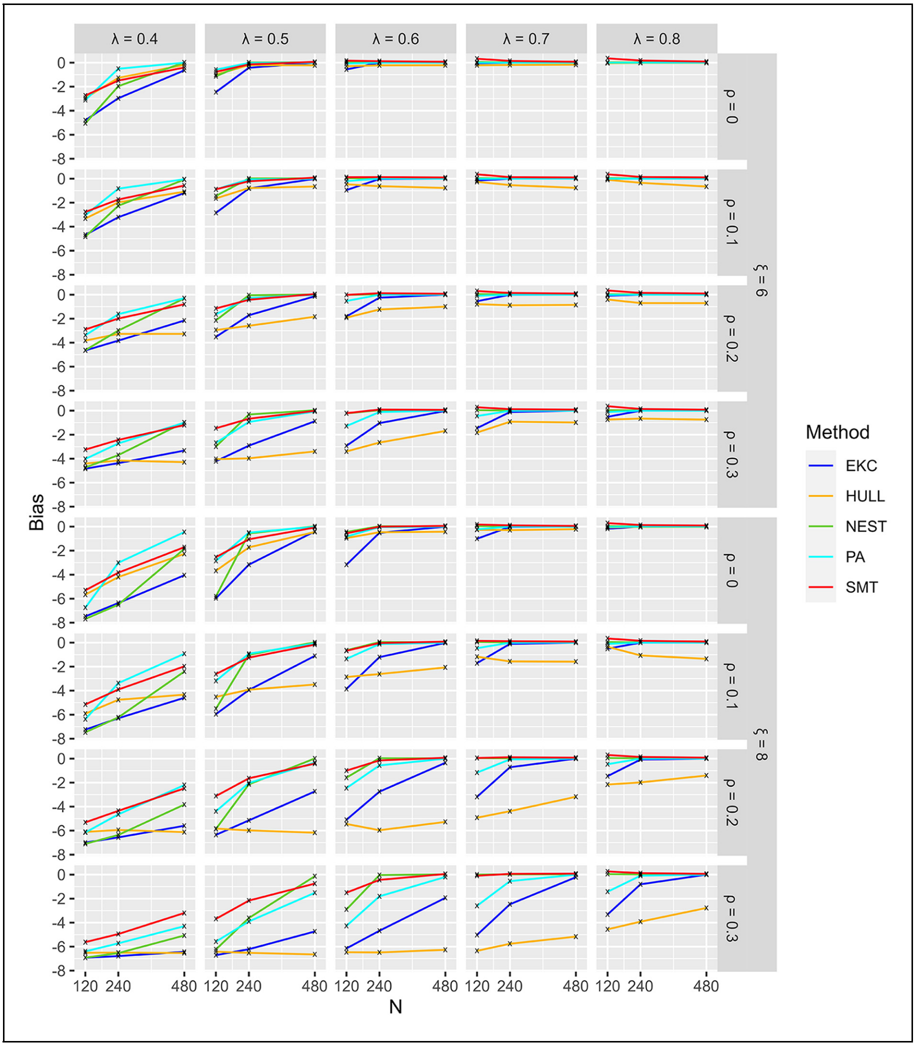

Figure 6. Bias for Conditions With 6 and 8 Factors (Hard).

Note. EKC = empirical Kaiser criterion; NEST = Next Eigenvalue Sufficiency Test; PA = parallel analysis; SMT = sequential χ^2 model test; ρ = correlations between factors; ξ = the number of factors; λ = loading; N = sample size.

method), methods based on minimizing out-of-sample prediction error across EFA models (Haslbeck & van Bork, 2024), and signal cancelation factor analysis (Achim, 2024), which recovers each factor with at least two unique indicators by canceling their common factor signal. Future studies should be carried out to investigate the state of the art in determining the number of factors in factor analysis.

Limits

The current simulations have some limits. First, factor structures had uniform loadings across factors. That means that eigenvalues of each factor were the same in the population correlation matrix. Similarly, the correlations between factors were the same for all factors, which was captured by the first eigenvalues at the population level. These limits seem warranted to avoid an overcrowded design that would be harder to interpret, but it would be an interesting step further for comparing stopping rules. Another limit is that the data generation technique used multivariate Gaussian variables only. A future study could investigate the effect of asymmetrical distribution and, if necessary, implement corrections to increase performance of the stopping rules.

Another conceptual limit regarding NEST, but all techniques based on eigenvalues, such as PA and EKC also, is that the p eigenvalues are based on $p(1-p)/2$ correlations, which leads to overidentification: there are infinitely many correlation matrices that can produce a given set of eigenvalues. Most of these correlation matrices lack a meaningful factorial structure, despite that stopping rules would still yield a number of factors. To avoid such issues, researchers must construct scales in a thoughtful manner. It is of substantial importance to retrieve meaningful factor structures. That is why we emphasized earlier the importance of assessing factor retention methods under real-world conditions.

Conclusion

The purpose of the current study was to compare the performance of NEST against four recommended stopping rules (PA, HULL, EKC, and SMT). Overall, NEST showed the best performance, especially in challenging conditions where it had to detect small but meaningful factors. PA followed closely, sometimes outperforming NEST, mainly due to being more liberal. While most techniques performed well in easy scenarios, NEST particularly stood out in difficult ones. Some limitations of NEST were addressed. Future studies should investigate and compare stopping rules with more realistic and varied factor structures, such as non-Gaussian variables, ordinal variables, different importance of factors, and various correlation patterns between factors.

Acknowledgments

The author would like to thank André Achim for his helpful comments on an earlier draft of this paper.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Consent for Publication

The author agrees that the work is ready for publication.

ORCID iD

Pier-Olivier Caron https://orcid.org/0000-0001-6346-5583

Supplemental Material

Supplemental material for this article is available online.

References

Achim, A. (2017). Testing the number of required dimensions in exploratory factor analysis. *The Quantitative Methods for Psychology*, *13*(1), 64–74. https://doi.org/10.20982/tqmp.13.1.p064

Achim, A. (2021). Determining the number of factors using parallel analysis and its recent variants: Comment on Lim and Jahng (2019). *Psychological Methods*, *26*(1), 69–73. https://doi.org/10.1037/met0000269

Achim, A. (2024). *Signal cancellation factor analysis*. https://doi.org/10.31234/osf.io/h7qwg

Auerswald, M., & Moshagen, M. (2019). How to determine the number of factors to retain in exploratory factor analysis: A comparison of extraction methods under realistic conditions. *Psychological Methods*, *24*(4), 468–491. https://doi.org/10.1037/met0000200

Bollen, K. A. (1989). *Structural equations with latent variables*. Wiley.

Braeken, J., & van Assen, M. A. L. M. (2017). An empirical Kaiser criterion. *Psychological Methods*, *22*(3), 450–466. https://doi.org/10.1037/met0000074

Brandenburg, N., & Papenberg, M. (2024). Reassessment of innovative methods to determine the number of factors: A simulation-based comparison of exploratory graph analysis and Next Eigenvalue Sufficiency Test. *Psychological Methods*, *29*(1), 21–47. https://doi.org/10.1037/met0000527

Caron, P.-O. (2019). Minimum average partial correlation and parallel analysis: The influence of oblique structures. *Communications in Statistics: Simulation and Computation*, *48*(7), 2110–2117. https://doi.org/10.1080/03610918.2018.1433843

Finch, W. H. (2023). A comparison of methods for determining the number of factors to retain with exploratory factor analysis of dichotomous data. *Psych*, 5, 1004–1018. https://doi.org/10.3390/psych5030067

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*, 377–393. https://doi.org/10.1177/0013164495055003002

Goretzko, D., & Ruscio, J. (2024). The comparison data forest: A new comparison data approach to determine the number of factors in exploratory factor analysis. *Behavior Research Methods*, *56*(3), 1838–1851. https://doi.org/10.3758/s13428-023-02122-4

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., 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

Haslbeck, J. M. B., & van Bork, R. (2024). Estimating the number of factors in exploratory factor analysis via out-of-sample prediction errors. *Psychological Methods*, *29*(1), 48–64. https://doi.org/10.1037/met0000528

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

Lawley, D. N. (1940). The estimation of factor loadings by the method of maximum likelihood. *Proceedings of the Royal Society of Edinburgh*, *60*(1), 64–82. https://doi.org/10.1017/S037016460002006X

Lim, S., & Jahng, S. (2019). Determining the number of factors using parallel analysis and its recent variants. *Psychological Methods*, *24*(4), 452–467. https://doi.org/10.1037/met0000230

Lorenzo-Seva, U., Timmerman, M. E., & Kiers, H. A. L. (2011). The Hull method for selecting the number of common factors. *Multivariate Behavioral Research*, *46*(2), 340–364. https://doi.org/10.1080/00273171.2011.564527

Neishabouri, A., & Desmarais, M. C. (2024). An ensemble approach to determine the number of latent dimensions and assess its reliability. *Communications in Statistics—Simulation and Computation*. Advance online publication. https://doi.org/10.1080/03610918.2024.2328166

Peres-Neto, P. R., Jackson, D. A., & Somers, K. M. (2005). How many principal components? Stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics & Data Analysis*, *49*, 974–997. https://doi.org/10.1016/j.csda.2004.06.015

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ruscio, J., & Roche, B. (2012). Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychological Assessment*, *24*(2), 282–292. https://doi.org/10.1037/a0025697

Steiner, M. D., & Grieder, S. (2020). EFAtools: An R package with fast and flexible implementations of exploratory factor analysis tools. *Journal of Open Source Software*, *5*(53), Article 2521. https://doi.org/10.21105/joss.02521

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*, 541–568. https://doi.org/10.1177/0013164498058004001

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*(3), 321–327. https://doi.org/10.1007/BF02293557

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. Springer. https://www.stats.ox.ac.uk/pub/MASS4/