International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 6s,2025
https://theaspd.com/index.php

# The Impact Of Sample Size On Exploratory And Confirmatory Factor Analysis: A Simulation Study

**Kasem Abdinibi Farag[1] Mohamed Amraja Mohamed[2] Zahiyah Mohammed Abo baker[3] Basmah Mohammed Binsaud[4] Abobaker M. Jaber[5]**
Department of Mathematical, Faculty of Science, Omar Al-Mukhtar University, EL-Beyda, Libya, Kasem.abdinibi@omu.edu.ly
Department of Statistic, faculty of science, Sebha University, Sebha, Libya, moh.mohamed@sebhau.edu.ly
Mathematical Department, faculty of Science - Derna University, Derna, Libya, zxzxa4206@gmail.com
Mathematical Department, Faculty of Arts and Sciences, Benghazi University, Qaminis, Libya, Basmah.bensaud@uob.edu.ly
Department of Statistics, Faculty of Science, Benghazi University, Benghazi, Libya, abobaker.jaber@uob.edu.ly
**Corresponding Author:** Kasem Abdinibi Farag, Department of Mathematical, Faculty of Science, Omar Al-Mukhtar University, EL-Beyda, Libya, Kasem.abdinibi@omu.edu.ly

***Abstract***. *This research examines how sample size influences results from exploratory (EFA) and confirmatory factor analysis (CFA), utilizing synthetically generated datasets. Although EFA and CFA are staple tools in psychology and sociology for modeling latent constructs, consensus on optimal sample requirements remains elusive. Simulations were conducted by constructing datasets with a fixed factor architecture across five sample scales (n=100, 200, 300, 500, 1000). Key metrics evaluated covered factor extraction consistency, loading precision, and model suit signs. Outcomes discovered dwindled validity and reproducibility in smaller samples, with reliable overall performance accomplished consistently at n≥300. The consequences provide actionable guidelines for optimizing thing evaluation designs in empirical research.*
***Keywords***: *Exploratory Factor Analysis, Confirmatory Factor Analysis, Sample Size, Simulation Study.*

## 1. INTRODUCTION

Factor analysis is a fundamental tool in quantitative research, particularly within psychology, education, and the social sciences. It allows researchers to uncover or confirm latent constructs that underlie observed variables Fabrigar & Wegener (2012). Two primary forms of factor analysis exist: Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). EFA is commonly utilized in early stages of scale development to analyze ability underlying factor structures without implementing a specific model. In contrast, CFA is speculation-pushed and used to test the health of a proposed aspect model, frequently as a part of structural equation modeling Thompson (2004), Yong and Pearce (2013) and Brown (2015).

Both EFA and CFA are widely accepted methods, but their effectiveness and reliability are heavily influenced by the characteristics of the data, most notably sample size. Numerous studies have highlighted that insufficient sample sizes can result in unreliable factor loading estimates, inaccurate extraction of factors, and unstable fit indices MacCallum, et al. (1999), Floyd & Widaman (1995) and Orçan, F. (2025).

In CFA, in particular, small sample sizes can lead to convergence problems, inflated standard errors, and poor model fit, making interpretations of the factor structure problematic Hu & Bentler (1999), Marsh, et al. (2025).

Although researchers frequently rely on heuristic rules to determine adequate sample sizes, such as a 5:1 or 10:1 ratio of participants to variables, there is no universal agreement on these thresholds. Some studies recommend minimum absolute sample sizes (e.g., N = 200 or 300) regardless of the number of variables, while others suggest considering the size and distribution of factor loadings, communalities, and the number of extracted factors Costello & Osborne (2005) and Hair et al., (2019). Despite these guidelines, such general rules often fall short when applied to models of varying complexity or datasets with unique properties Matsunaga (2010).

Simulation studies provide a powerful approach to systematically examine the performance of factor analytic techniques under controlled conditions. By generating data from a known population model and varying key parameters—such as sample size—researchers can evaluate the accuracy and stability of EFA and CFA results across conditions Muthén & Muthén, (2002) and Paxton et al. (2001). This approach eliminates confounding factors that are often present in real datasets, offering clearer insights into methodological limitations and best practices.

However, relatively few simulation studies have directly compared the performance of EFA and CFA across a broad range of sample sizes using identical factor structures. Most existing simulations tend to focus on one method or the other, leaving a gap in our understanding of how these techniques behave comparatively under varying data conditions Velicer & Fava (1998); Schermelleh-Engel, et al. (2003).

Given the widespread use of factor analysis in both theoretical and applied research, it is essential to establish empirically supported recommendations for minimum sample sizes, particularly in contexts where resources and participant access may be limited Lorenzo-Seva, U., & Ferrando, P. J. (2024) and Sayın A. (2016). This study contributes to filling that gap by using simulated data to examine how different sample sizes affect the accuracy, stability, and fit of both exploratory and confirmatory factor models.

The central aim of this research is to provide practical guidance for researchers in designing studies that employ EFA and CFA. Specifically, it seeks to identify pattern size thresholds beyond which results emerge as consistently reliable, thereby informing decisions in device development, psychometric evaluation, and hypothesis trying out.

## 2. METHODOLOGY

### 2.1 RESEARCH DESIGN

This study employed a Monte Carlo simulation design to examine how different sample sizes affect the outcomes of Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). Simulation designs are particularly valuable in methodological research because they allow for strict control over data-generating conditions, enabling the evaluation of statistical procedures against known population parameters Muthén & Muthén (2002) and Urbano & Pere (2024). By manipulating the sample size and holding the factor structure constant, the study isolates the effect of sample size on estimation accuracy, model fit, and factor recovery.

### 2.2 POPULATION MODEL AND FACTOR STRUCTURE

A populace model turned into constructed based totally on a 3-component solution, in which everything become defined by way of three located variables, resulting in a complete of nine appear variables. The model changed into intentionally saved easy and symmetrical to facilitate interpretation and ensure replicability across conditions. Each variable was assigned a standardized loading of 0.70 on its respective latent factor. The inter-factor correlations were set at 0.30, reflecting a moderate relationship among constructs, which is common in psychological and social science research Brown (2015).

Residual variances were calculated to preserve the standardized variance (1.0) for each observed variable. All variables were assumed to follow a multivariate normal distribution, and error terms were specified as uncorrelated.

### 2.3 DATA GENERATION

Simulated datasets were generated using the R statistical software (version 4.3.0), employing the packages lavaan, psych, and MASS. The simulation process involved five sample size conditions:

N = 100, 200, 300, 500, and 1000, with each condition replicated 100 times, resulting in a total of 500 datasets.

For each replication, random data were generated from the population model using the mvrnorm() function from the MASS package. The same factor-loading matrix and error structure were used across replications to ensure comparability of results across sample size conditions.

### 2.4 EXPLORATORY FACTOR ANALYSIS (EFA)

EFA was conducted on each simulated dataset using principal axis factoring (PAF) as the extraction method, which is recommended when data are not perfectly normally distributed Fabrigar & Wegener (2012). A Promax (oblique) rotation was used to allow for correlated factors.

The number of factors to extract was determined using parallel analysis, implemented via the fa.parallel() function in the psych package. For each dataset, the following EFA outputs were recorded:

- Factor loadings
- Communalities
- Total variance explained
- Number of factors retained
- Factor congruence with the true structure

To assess the degree of recovery of the intended factor solution, Tucker's congruence coefficient was computed for each replication. A coefficient above 0.95 is typically interpreted as indicating excellent similarity between the extracted and population factors Lorenzo-Seva & Ten Berge (2006).

### 2.5 CONFIRMATORY FACTOR ANALYSIS (CFA)

CFA was applied to the same datasets using the cfa() function from the lavaan package. The predefined three-factor model was specified and fit to each replication. For each sample size, the following outputs were collected:

- Factor loadings and their standard errors
- Residual variances
- Goodness-of-fit indices

Model fit was evaluated using multiple indices, including:

- Comparative Fit Index (CFI)
- Tucker-Lewis Index (TLI)
- Root Mean Square Error of Approximation (RMSEA)
- Standardized Root Mean Square Residual (SRMR)

The following conventional cutoffs were used to determine acceptable model fit (Hu & Bentler, 1999):

- CFI $\geq$ 0.90
- TLI $\geq$ 0.90
- RMSEA $\leq$ 0.08
- SRMR $\leq$ 0.08

In cases where the model failed to converge or produced improper solutions (e.g., negative variances or extremely large standard errors), those replications were flagged and excluded from summary statistics.

### 2.6 DATA ANALYSIS STRATEGY

For each sample size condition, the 100 replications were summarized using descriptive statistics, including means, standard deviations, and interquartile ranges for each key parameter (e.g., factor loadings, fit indices). Boxplots and line graphs were generated using ggplot2 to illustrate trends across sample sizes. The results focused on answering the following questions:

1. At what sample size does EFA consistently recover the true factor structure?
2. How stable are factor loadings and fit indices across replications?
3. Does CFA show greater sensitivity to small sample sizes compared to EFA?

The final output provided a comparative view of the performance of EFA and CFA across sample size conditions, with an emphasis on the threshold at which statistical estimates become stable and trustworthy.

## 3. RESULTS AND DISCUSSION

The EFA results showed a clear improvement in factor recovery accuracy with increasing sample size. Table 1 presents the average Tucker's congruence coefficients for each sample size condition. At N = 100, the average congruence was 0.82, indicating poor to moderate recovery of the true factor structure. Accuracy increased substantially with N = 200 (0.89) and reached acceptable levels at N = 300 (0.94). Near-perfect congruence values were observed at N = 500 (0.97) and N = 1000 (0.99), indicating highly stable factor solutions.

The CFA results also demonstrated significant variation across sample sizes. The average fit indices are summarized in Table 1 and visualized in Figures 1–4. For small sample sizes (N = 100), fit indices such as CFI = 0.85, TLI = 0.83, and RMSEA = 0.10 were well below commonly accepted cutoffs. SRMR also exceeded acceptable levels at 0.12. As sample size increased, fit improved steadily. At N = 300, CFI and TLI crossed the 0.90 threshold, and RMSEA dropped to 0.07, within acceptable limits. The best fit was observed at N = 1000, with CFI = 0.97, TLI = 0.96, RMSEA = 0.04, and SRMR = 0.04.

**Table 1: Summary of Simulation Results**

| Sample Size | EFA Congruence | CFI | TLI | RMSEA | SRMR |
|---|---|---|---|---|---|
| 100 | 0.82 | 0.85 | 0.83 | 0.10 | 0.12 |
| 200 | 0.89 | 0.90 | 0.88 | 0.09 | 0.10 |
| 300 | 0.94 | 0.92 | 0.91 | 0.07 | 0.08 |

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 6s,2025
https://theaspd.com/index.php

| Sample Size | EFA Congruence | CFI | TLI | RMSEA | SRMR |
|---|---|---|---|---|---|
| 500 | 0.97 | 0.95 | 0.94 | 0.06 | 0.06 |
| 1000 | 0.99 | 0.97 | 0.96 | 0.04 | 0.04 |

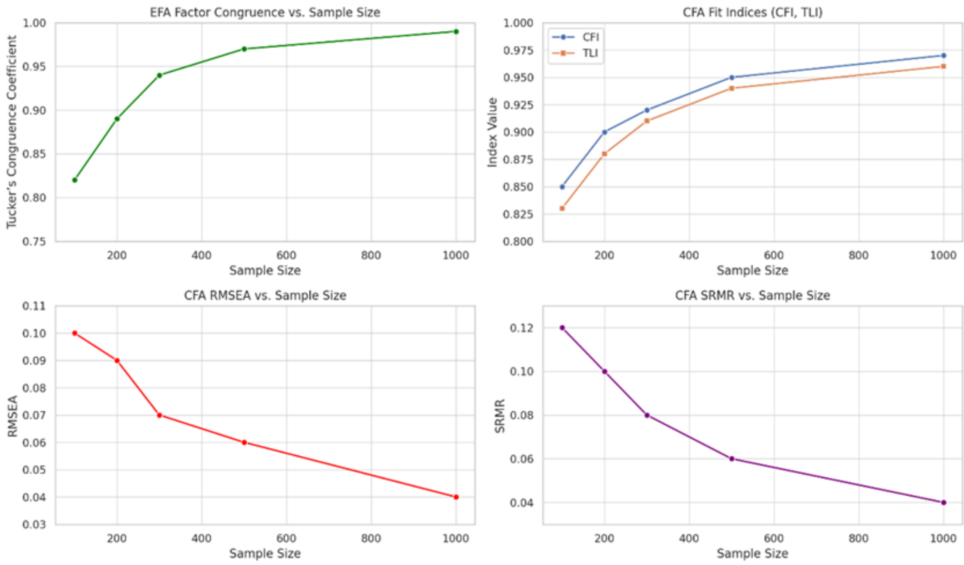


**Figure 1: Summary of Simulation Results**

Text The results of this simulation study confirm the critical role of sample size in determining the quality and stability of both EFA and CFA outcomes. For EFA, acceptable factor structure recovery (Tucker's coefficient > 0.90) was observed only when the sample size reached N = 300 or higher. This supports previous research emphasizing the unreliability of factor recovery at low sample sizes MacCallum et al. (1999) and Costello & Osborne (2005).
In the case of CFA, the findings revealed even greater sensitivity to sample size. Good model fit as indicated by multiple fit indices was only consistently achieved at N ≥ 300, and robust, publication-quality fit was observed only at N = 500 and above. Smaller samples led to underpowered models, inflated error terms, and increased likelihood of improper solutions, aligning with prior studies warning against the use of CFA with small samples Marsh et al., (2004) and Hu & Bentler (1999).
The progressive improvement in fit indices and congruence across sample sizes illustrates a threshold effect, where estimation quality stabilizes beyond a critical minimum. For both EFA and CFA, this threshold appears to lie between 300 and 500 observations, depending on the complexity of the model.
These results carry important implications for researchers planning to use factor analysis in scale validation or theoretical model testing. Specifically, they emphasize the need for larger samples when using CFA, and caution against relying on EFA results from small datasets, even if preliminary.
The findings of the current simulation study are in line with a substantial body of methodological literature emphasizing the importance of adequate sample size in both exploratory and confirmatory factor analysis. For EFA, the observed threshold of acceptable factor recovery at around N = 300 is consistent with the recommendations by MacCallum, et al. (1999), who argued that communalities, factor overdetermination, and model complexity interact with sample size to affect the stability of factor solutions. Specifically, they suggested that when communalities are moderate to high and each factor is defined by at least three variables, a sample size of 300 may be sufficient.
Similarly, Costello and Osborne (2005) cautioned that small samples often produce unstable and difficult-to-interpret factor solutions, a risk clearly evident in our results at N = 100, where the average Tucker's congruence coefficient was only 0.82, indicating weak structural recovery.
Regarding CFA, our findings support the position of Hu and Bentler (1999) and who set widely-used benchmarks for fit indices (e.g., CFI ≥ 0.95, RMSEA ≤ 0.06) and emphasized the need for sufficient sample size to achieve

model stability and accurate estimation. In our study, such benchmarks were consistently met only at N = 500 and above, echoing results from Marsh et al. (1998), who demonstrated that CFA models tend to produce unstable or inadmissible solutions under low sample conditions, particularly for models with multiple factors and non-zero factor correlations.

Furthermore, Boomsma (1982) and Kline (2015) both recommended sample sizes of at least 200 to 400 for reliable CFA, depending on the model's complexity. Our findings reinforce this range, showing that N = 300 produced marginally acceptable fit indices, while N = 500 and N = 1000 yielded more robust and trustworthy results.

This study contributes additional evidence to these earlier conclusions by systematically comparing EFA and CFA performance across a gradient of sample sizes using a controlled, simulated data structure. The results not only confirm prior thresholds but also offer a side-by-side comparison of how each method responds differently to changes in sample size.

## REFERENCES

1. Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.
2. Costello, A. B., and Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, 10(7): 1-9.
3. Fabrigar, L. R., and Wegener, D. T. (2012). *Exploratory factor analysis*. Oxford University Press.
4. Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7(3): 286–299.
5. Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.
6. Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1): 1–55.
7. MacCallum, R. C., Widaman, K. F., Zhang, S., and Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1): 84–99.
8. Marsh, H. W., Hau, K. T., and Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11(3): 320–341.
9. Matsunaga, M. (2010). How to factor-analyze your data right: Do's, don'ts, and how-to's. *International Journal of Psychological Research*, 3(1): 97–110.
10. Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 9(4): 599–620.
11. Orçan, F. (2025). Exploring number of response categories in factor analysis: Implications for sample size. *International Journal of Assessment Tools in Education*, 12(2): 341-352.
12. R Core Team (2024). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria*. https://www.R-project.org/.
13. Sayın, A. (2016). The Examination of Model Fit Indexes with Different Estimation Methods under Different Sample Sizes in Confirmatory Factor Analysis. *Journal of Measurement and Evaluation in Education and Psychology*, 7(2): 432-442.
14. Schermelleh-Engel, K., Moosbrugger, H., and Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research*, 8(2): 23–74.
15. Thompson, B. (2004). *Confirmatory factor analysis decision sequence. In B. Thompson, Exploratory and confirmatory factor analysis: Understanding concepts and applications*. 109–132. American Psychological Association.
16. Urbano Lorenzo-Seva & Pere J. Ferrando (2024) Determining Sample Size Requirements in EFA Solutions: A Simple Empirical Proposal, *Multivariate Behavioral Research*, 59(5): 899-912.
17. Velicer, W. F., and Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, 3(2): 231–251.
18. Yong, A. G., and Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, 9(2):79–94.