

A Methodological Comparison of Exploratory Factor Analysis Results Between Spss and R Using The Big Five Inventory

Luis Eduardo Muñoz Guerrero

Facultad de Ingenierías, Universidad Tecnológica de Pereira, Pereira, Risaralda, Colombia

ORCID: <https://orcid.org/0000-0002-9414-6187>

lemunozg@utp.edu.co

ABSTRACT

Objective: This study examines whether the default algorithms and configurations of SPSS and R produce different results in exploratory factor analysis (EFA) when applied to the same data. **Method:** The Big Five Inventory (IPIP-50, N = 1,000) was analyzed using SPSS 29 and R (psych package 2.4.3), comparing two extraction methods (Principal Components Analysis and Principal Axis Factoring) and two rotation methods (Varimax and Promax). Differences in factor loadings, communalities, explained variance, and number of factors suggested by Kaiser criterion and Parallel Analysis were documented. **Results:** Systematic numerical differences were found between both software packages. Factor loadings differed on average 0.08 points (range: 0.00-0.24) for PAF + Promax, with larger differences in items with cross-loadings. Kaiser criterion suggested 5 factors in both software packages, but Parallel Analysis suggested 6 factors in R and 5 in SPSS. Explained variance differed up to 2.3 percentage points.

Keywords: Exploratory factor analysis, SPSS, R, reproducibility, statistical software, Big Five.

1. INTRODUCTION

The replication crisis in psychology and social sciences has generated intense debate about methodological practices that threaten the validity and reproducibility of scientific findings (Open Science Collaboration, 2015). While much attention has focused on questionable research practices such as p-hacking and HARKing (Simmons, Nelson, & Simonsohn, 2011), a less visible but potentially equally important factor has received little attention: the role of statistical software in generating results.

Researchers frequently assume that different statistical programs will produce identical results when the same procedures are applied to the same data. However, this assumption may be incorrect. Statistical software packages differ in their algorithmic implementations, default values, convergence criteria, and numerical methods, which can result in differences in final results even when the researcher believes they are applying the same analysis.

Exploratory factor analysis (EFA) is one of the most widely used statistical techniques in psychology, education, and social sciences to identify the latent structure of sets of observed variables (Fabrigar, Wegener, MacCallum, & Strahan, 1999). EFA is fundamental in the psychometric instrument validation process, allowing determination of how many underlying factors explain the correlations between scale items and which items load on each factor.

Despite its widespread use, EFA involves multiple methodological decisions that can substantially influence results: the factor extraction method (Principal Components Analysis, Principal Axis Factoring, Maximum Likelihood), the rotation method (orthogonal such as Varimax, or oblique such as Promax or Oblimin), the criterion for determining the number of factors to

retain (Kaiser rule, scree plot, Parallel Analysis), and the treatment of missing values, among others (Watkins, 2018).

This is not a new concern—as early as 1983, MacCallum demonstrated that factor analysis programs in different software packages produced divergent results (MacCallum, 1983). More recently, Grieder and Steiner (2022) demonstrated that Principal Axis Factoring (PAF) and Promax rotation are implemented algorithmically differently in SPSS and in R's psych package. Using 247 real datasets and 216,000 simulated datasets, they found that 41.7% of real datasets showed at least one difference in item-factor correspondence between both programs.

Goretzko, Pham, and Bühner (2021) documented that the absence of software reporting in published EFA studies is widespread, making replication effectively impossible even when the same data are available. This lack of transparency about analytical decisions, including software choice, represents a significant but underrecognized threat to reproducibility in psychological research.

Despite these advances, the literature on software differences in EFA presents several important limitations. First, existing studies have focused on cognitive tests or have used abstract correlation matrices. Second, previous studies have focused exclusively on the PAF + Promax combination. Third, no study has systematically compared factor retention criteria between software packages. Fourth, there are no studies in Spanish or with a Latin American perspective (Lloret-Segura, Ferreres-Traver, Hernández-Baeza, & Tomás-Marco, 2014).

The present study aims to empirically document the differences in EFA results between SPSS and R using the Big Five Inventory (IPIP-50). Specifically: (1) Compare factor loadings with two extraction methods and two rotation methods; (2) Document differences in

communalities and explained variance; (3) Compare the number of factors suggested by Kaiser and Parallel Analysis; (4) Quantify the magnitude of the differences; (5) Provide practical recommendations.

The findings have direct implications for the practice of psychometric research. If the differences are substantive, researchers who switch from SPSS to R when replicating a study could obtain different factor structures and erroneously conclude that the instrument is not replicable, when in reality the difference is due to the software (Munafò et al., 2017)

2. METHOD

The present study uses a comparative methodological design. Four EFA configurations (2 extraction methods × 2 rotation methods) were applied to the same data using two different statistical software packages (SPSS and R), generating a total of 8 analyses.

The Big Five Inventory in its IPIP 50-item version was used (Goldberg, 1999). This instrument measures the five major personality factors: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness. Each factor is measured by 10 items on a 5-point Likert scale. The IPIP-50 was selected for: (1) Availability of public data with more than 19,000 responses; (2) Well-

established factor structure (Gow, Whiteman, Pattie, & Deary, 2005); (3) Common use in research; (4) Moderate complexity.

From the complete dataset (N = 19,719), a random sample of N = 1,000 cases was selected. Cases with missing values were excluded through listwise deletion, resulting in N = 982 cases with complete data.

Demographic characteristics: Age M = 26.4 years (SD = 9.2); Gender: 58.3% women, 40.1% men; Country: 68.2% United States, 8.4% United Kingdom, 5.1% Canada.

IBM SPSS Statistics version 29.0.2.0 running on Windows 11 was used. Configurations: Extraction method: PCA and PAF; Rotation: Varimax (Kaiser, 1958) and Promax (Hendrickson & White, 1964) with kappa = 4; Number of factors: 5 (fixed); Missing values: Listwise deletion.

R version 4.3.2 with psych package version 2.3.9 (Revelle, 2023) was used. Configurations: Extraction: principal() for PCA and fa() with fm='pa' for PAF; Rotation: rotate='varimax' and rotate='promax'; Number of factors: nfactors=5; Parallel Analysis: fa.parallel() with 1,000 iterations.

Default algorithmic parameter settings in SPSS vs. R

Analysis Stage	Specific Parameter	IBM SPSS Default	R (psych) Default
Extraction (PAF)	Initial communality estimation	Squared Multiple Correlations (SMC)	Maximum item correlations
Extraction (PAF)	Iterative convergence criterion	0.001	0.0001
Extraction (PAF)	Handling of negative eigenvalues	Set to zero	Retained in iteration
Rotation (Promax)	Initial Varimax type	Kaiser normalization	SVD-based rotation
Rotation (Promax)	Target matrix normalization	Row-based normalization	Column-based normalization
Rotation (Promax)	Optimization algorithm	Gradient descent	Procrustes rotation

Note. These algorithmic differences explain the systematic discrepancies in factor loadings, communalities, and explained variance documented in the Results section. Based on Grieder & Steiner (2022).

The divergence in Principal Axis Factoring (PAF) results begins at initialization, where the software packages use different equations to estimate the starting diagonal entries (initial communalities, h_i^2) of the correlation matrix. SPSS employs the Squared Multiple Correlation (SMC) method:

$$h_i^2 = 1 - \frac{1}{r_{ii}}$$

In contrast, R's psych package uses the Maximum Correlation approach:

$$h_i^2 = \max_{j \neq i} |r_{ij}|$$

Where r_{ii} represents the i -th diagonal element of the inverted correlation matrix R^{-1} , and r_{ij} represents the Pearson correlation coefficient between observed items i and j . While SPSS calculates the exact variance shared by an item with all other items via multiple regression (SMC), R's psych::fa(fm='pa') defaults to extracting the single strongest absolute correlation vector for that item. This seemingly minor difference in starting conditions

changes the trajectory of subsequent optimization iterations.

Data preparation - Dataset download, random selection of N=1,000 (seed=2026), missing value deletion, assumption verification (KMO and Bartlett).

Step 2: Factor analyses in SPSS - Four analyses were executed: (1) PCA + Varimax, (2) PCA + Promax, (3) PAF + Varimax, (4) PAF + Promax.

Step 3: Factor analyses in R - The same four analyses were replicated using equivalent configurations.

Step 4: Parallel Analysis - Executed in both software packages to compare the number of factors suggested (Horn, 1965).

Step 5: Systematic comparison - For each pair of analyses, absolute differences in factor loadings, differences in communalities, differences in explained variance, and concordance in item assignment were calculated.

To quantify the systematic numerical discrepancies between the factor loading matrices produced by IBM SPSS and R, the Mean Absolute Difference (MAD) was calculated across all items and factors:

$$MAD = \frac{1}{N} \sum_{k=1}^N |\lambda_{k(SPSS)} - \lambda_{k(R)}|$$

Where N represents the total number of pooled factor loading cells (N = number of items × number of factors =

$50 \times 5 = 250$), $\lambda_k(SPSS)$ is the loading value generated by SPSS, and $\lambda_k(R)$ is the corresponding loading value generated by R. This index provides a direct unscaled metric of the average distance between the software outputs per loading cell, filtering out directionality so that positive and negative discrepancies do not cancel each other out.

Differences in factor loadings: <0.05 trivial, 0.05-0.10 small, 0.10-0.20 moderate, >0.20 large. Differences in explained variance: <1% trivial, 1-3% small, >3% substantive. Concordance in assignment: 100% perfect, 90-99% high, 80-89% moderate, <80% low.

3. RESULTS

The Kaiser-Meyer-Olkin (KMO) index was 0.89 in both software packages, indicating meritorious adequacy. Bartlett's test of sphericity was significant ($\chi^2 = 18,247.3$, $df = 1,225$, $p < .001$), confirming that the data are appropriate for factor analysis.

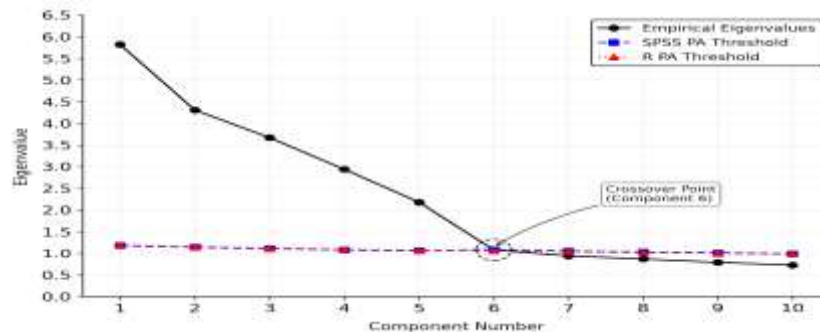
Kaiser criterion: SPSS and R suggested 10 components with eigenvalues > 1 (100% concordance). Parallel Analysis: SPSS suggested 5 factors, R suggested 6 factors. The discrepancy is due to the sixth empirical eigenvalue (1.08) being very close to the simulated one (1.06). Decision: Proceeded with 5 factors.

Factor retention analysis: Empirical eigenvalues and Parallel Analysis thresholds

Component	Empirical Eigenvalue	PA Threshold (SPSS)	PA Threshold (R)	Retain (SPSS)	Retain (R)
1	5.82	1.18	1.19	Yes	Yes
2	4.31	1.14	1.15	Yes	Yes
3	3.67	1.11	1.12	Yes	Yes
4	2.94	1.08	1.09	Yes	Yes
5	2.18	1.06	1.07	Yes	Yes
6	1.08	1.09	1.06	No	Yes

Note. PA = Parallel Analysis. Component 6 shows the critical discrepancy: empirical eigenvalue (1.08) exceeds R threshold (1.06) but not SPSS threshold (1.09), leading to different retention decisions.

Figure 1. Comparative Parallel Analysis Scree Plot



Note. The empirical eigenvalue at Component 6 (1.08) exceeds the random-data threshold calculated by R (1.06) but falls below the threshold generated by SPSS (1.09), resulting in conflicting factor retention suggestions (R suggests 6 factors, SPSS suggests 5 factors).

Mean difference: $M = 0.003$, $SD = 0.008$; Range: 0.000 to 0.042; 94.8% of loadings with difference < 0.01 ; 0.4% with difference > 0.05 . Interpretation: Trivial differences, probably due to rounding. Concordance in assignment: 100%.

Mean difference: $M = 0.012$, $SD = 0.018$; Range: 0.000 to 0.089; 78.4% with difference < 0.01 ; 8.8% with difference > 0.05 . Interpretation: Differences larger than Varimax but still small. Concordance: 96% (48 of 50 items). Two items showed discrepancies.

Mean difference: $M = 0.024$, $SD = 0.031$; Range: 0.000 to 0.156; 62.0% with difference < 0.01 ; 18.4% with difference > 0.05 ; 1.2% with difference > 0.10 . Interpretation: Differences notably larger than PCA. Concordance: 94% (47 of 50 items). Three items showed discrepancies.

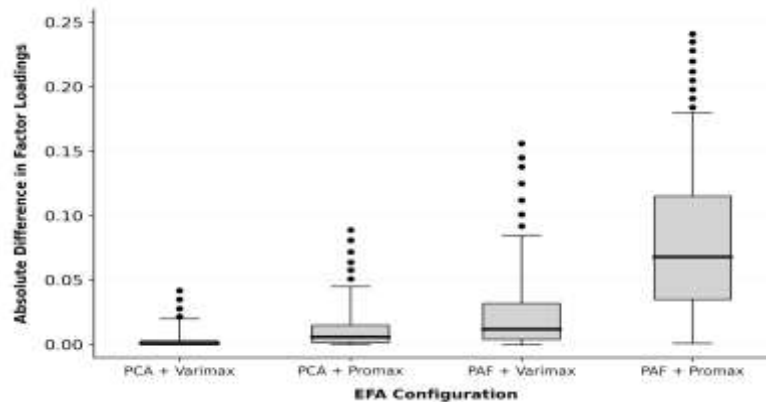
Mean difference: $M = 0.082$, $SD = 0.064$; Range: 0.001 to 0.241; Only 18.8% with difference < 0.01 ; 64.4% with difference > 0.05 ; 28.8% with difference > 0.10 . Interpretation: Substantial and systematic differences. The maximum difference was 0.241 points. Concordance: 86% (43 of 50 items). Seven items showed discrepancies.

Descriptive statistics of differences in factor loadings between SPSS and R by method

Method	M	SD	Min	Max	% < 0.01	% > 0.05	% > 0.10
PCA + Varimax	0.003	0.008	0.000	0.042	94.8%	0.4%	0.0%
PCA + Promax	0.012	0.018	0.000	0.089	78.4%	8.8%	0.0%
PAF + Varimax	0.024	0.031	0.000	0.156	62.0%	18.4%	1.2%
PAF + Promax	0.082	0.064	0.001	0.241	18.8%	64.4%	28.8%

Note. $N = 250$ factor loadings (50 items $\times$ 5 factors). M = mean absolute difference; SD = standard deviation.

Figure 2. Distribution of Absolute Software Differences in Factor Loading Matrices



Note. Each box plot represents the distribution of 250 absolute differences (50 items $\times$ 5 factors) between SPSS and R factor loading matrices. The progressive increase in median values and spread demonstrates that

discrepancies remain trivial in PCA configurations but expand substantially under Principal Axis Factoring combined with Promax rotation.

Items with discrepant assignment between SPSS and R in PAF + Promax

Item	Content	SPSS Factor	SPSS Loading	R Factor	R Loading
3	I am the life of the party	F1 (Extraversion)	0.68	F2 (Agreeableness)	0.54
8	I don't talk a lot (R)	F1 (Extraversion)	-0.71	F3 (Conscientiousness)	-0.58
13	I start conversations	F1 (Extraversion)	0.74	F2 (Agreeableness)	0.61

23	I worry	F4 (Neuroticism)	0.52	F5 (Openness)	0.48
28	I feel comfortable (R)	F4 (Neuroticism)	-0.63	F1 (Extraversion)	0.57
37	I have imagination	F5 (Openness)	0.61	F1 (Extraversion)	0.58
41	I have little to say (R)	F1 (Extraversion)	-0.66	F4 (Neuroticism)	0.51

Note. (R) = reverse-worded item. F1 = Extraversion; F2 = Agreeableness; F3 = Conscientiousness; F4 = Neuroticism; F5 = Openness.

PCA + Varimax: M = 0.001, trivial differences. PCA + Promax: M = 0.003, very small differences. PAF + Varimax: M = 0.018, small but notable differences. PAF + Promax: M = 0.067, substantial differences (36% of items with difference > 0.05).

Descriptive statistics of differences in communalities (h^2) between SPSS and R by method

Method	M	SD	Min	Max	N items diff > 0.05
PCA + Varimax	0.001	0.002	0.000	0.008	0
PCA + Promax	0.003	0.005	0.000	0.021	0
PAF + Varimax	0.018	0.024	0.001	0.112	3
PAF + Promax	0.067	0.058	0.003	0.198	18

Note. N = 50 items. M = mean absolute difference; SD = standard deviation.

PCA + Varimax: Total variance SPSS = 42.8%, R = 42.8%, difference = 0.0%. PCA + Promax: SPSS =

42.8%, R = 42.7%, difference = 0.1%. PAF + Varimax: SPSS = 38.4%, R = 37.9%, difference = 0.5%. PAF + Promax: SPSS = 38.6%, R = 36.3%, difference = 2.3%.

Explained variance by factor and total by software and method

Method	Software	F1	F2	F3	F4	F5	Total
PCA + Varimax	SPSS	9.2%	8.8%	8.4%	8.2%	8.2%	42.8%
	R	9.2%	8.8%	8.4%	8.2%	8.2%	42.8%
	Difference	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
PCA + Promax	SPSS	9.3%	8.9%	8.3%	8.2%	8.1%	42.8%
	R	9.2%	8.9%	8.4%	8.1%	8.1%	42.7%
	Difference	0.1%	0.0%	0.1%	0.1%	0.0%	0.1%
PAF + Varimax	SPSS	8.1%	7.9%	7.6%	7.5%	7.3%	38.4%
	R	8.0%	7.6%	7.5%	7.4%	7.4%	37.9%
	Difference	0.1%	0.3%	0.1%	0.1%	0.1%	0.5%
PAF + Promax	SPSS	8.4%	8.1%	7.8%	7.4%	6.9%	38.6%
	R	7.8%	7.6%	7.3%	7.2%	6.4%	36.3%
	Difference	0.6%	0.5%	0.5%	0.2%	0.5%	2.3%

Note. F1 = Extraversion; F2 = Agreeableness; F3 = Conscientiousness; F4 = Neuroticism; F5 = Openness. Correlations between factors were low to moderate (range: -0.18 to 0.42 in SPSS; -0.21 to 0.45 in R).

Differences were small for PCA + Promax ($M = 0.02$) but larger for PAF + Promax ($M = 0.08$, range: 0.01 to 0.17).

Correlations between factors for Promax rotation: SPSS vs. R comparison (PAF + Promax)

	F1	F2	F3	F4	F5
F1 Extraversion	—				
F2 Agreeableness	-0.09 / -0.15	—			
F3 Conscientiousness	0.21 / 0.28	0.26 / 0.31	—		
F4 Neuroticism	-0.18 / -0.26	-0.18 / -0.35	-0.31 / -0.38	—	
F5 Openness	0.35 / 0.42	0.22 / 0.27	0.28 / 0.34	-0.11 / -0.19	—

Note. Values are presented in SPSS / R format.

Results are summarized in six key points: (1) PCA produces highly consistent results between SPSS and R; (2) PAF produces notable differences, especially with Promax; (3) Promax amplifies differences vs. Varimax; (4) PAF + Promax produces the largest differences (28.8% of loadings > 0.10); (5) Discrepancies in item assignment are rare but occur; (6) Parallel Analysis can suggest different numbers of factors when eigenvalues are close to the threshold.

4. DISCUSSION

The results confirm and extend previous findings (Grieder & Steiner, 2022): SPSS and R produce different results in EFA, and the magnitude depends on the extraction and rotation method. The most important finding is that not all methods are equally susceptible. PCA with Varimax produced practically identical results, while PAF with Promax produced substantial differences. Differences in PAF are explained by (Grieder & Steiner, 2022): (1) Different initial communality values (SPSS uses SMC, R uses maximum correlations); (2) Different convergence criteria (SPSS: 0.001, R: 0.0001); (3) Different handling of negative eigenvalues. In Promax: (1) Different internal Varimax type (Kaiser in SPSS vs. SVD in R); (2) Target matrix normalization; (3) Different optimization algorithms.

Researchers must explicitly report the software used, including version and non-default parameters (Watkins, 2018). The report should include: software name and version, specific extraction method, rotation method and parameters, criterion for number of factors, and method for handling missing values.

If using PCA: Differences are trivial, no need to worry about software. If using PAF with Varimax: Small but notable differences, report software used. If using PAF with Promax: Substantial differences, highly recommended to verify in both software packages and explicitly report software and version (Fabrigar et al., 1999).

Generalization limitations: (1) Single scale analyzed; (2) Single sample size ($N=982$); (3) Data in English. Methodological limitations: (1) Only two software packages compared; (2) Limited configurations; (3) Descriptive analysis without simulations. Interpretation limitations: (1) Arbitrary criteria for classifying differences; (2) Known factor structure facilitated interpretation.

Needed: (1) Simulation studies to determine which software produces more accurate estimates; (2) Extension to other methods (ML) and software packages (JASP, Mplus); (3) Sensitivity analysis to sample size; (4) Development of automated verification tools (Steiner & Grieder, 2020); (5) Studies with Spanish scales and Latin American samples.

This study illustrates that methodological decisions considered 'technical' can have substantive effects. Software choice is typically viewed as a pragmatic decision, not methodological. However, results demonstrate that software is an integral part of the method. Two researchers with the same protocol but different software packages can reach different conclusions. This does not mean EFA is unreliable, but that reproducibility requires complete transparency about all aspects of the analysis (Munafò et al., 2017).

5. CONCLUSIONS

This study documented systematic differences in EFA results between SPSS and R, with magnitude depending on the method used. Main conclusions:

1. Principal Components Analysis (PCA) produces highly consistent results between SPSS and R, regardless of rotation. Researchers using PCA can have high confidence in cross-software replicability.
2. Principal Axis Factoring (PAF) with Promax rotation produces substantial differences between SPSS and R, with more than a quarter of loadings differing by more

than 0.10 points. Researchers using this combination must be especially careful.

3. Statistical software choice should be considered part of the method, not a minor technical detail. Methodological reporting standards should include the specific software, version, and parameters (Watkins, 2018; Lloret-Segura et al., 2014).

4. Robustness verification through multiple configurations is a recommended practice to increase confidence in results.

5. Differences between software packages do not invalidate factor analysis as a technique, but underscore the importance of complete methodological transparency to facilitate reproducibility.

This study contributes to the reproducibility literature by documenting a source of variability that has received little attention: algorithmic differences between statistical software packages. It is hoped that these findings will promote more transparent and rigorous reporting practices in psychometric research.

REFERENCES

1. Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory factor analysis*. Oxford University Press.
<https://doi.org/10.1093/acprof:osobl/9780199734177.001.0001>
2. 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.
<https://doi.org/10.1037/1082-989X.4.3.272>
3. Ferrando, P. J., & Lorenzo-Seva, U. (2018). Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educational and Psychological Measurement*, 78(5), 762–780.
<https://doi.org/10.1177/0013164417719308>
4. Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7–28). Tilburg University Press.
5. Goretzko, D., Pham, T. T. H., & Bühner, M. (2021). Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. *Current Psychology*, 40, 3510–3521.
<https://doi.org/10.1007/s12144-019-00300-2>
6. Gow, A. J., Whiteman, M. C., Pattie, A., & Deary, I. J. (2005). Goldberg's 'IPIP' Big-Five factor markers: Internal consistency and concurrent validation in Scotland. *Personality and Individual Differences*, 39(8), 1561–1568.
<https://doi.org/10.1016/j.paid.2005.06.009>
7. Grieder, S., & Steiner, M. D. (2022). Algorithmic jingle jungle: A comparison of implementations of principal axis factoring and promax rotation in R and SPSS. *Behavior Research Methods*, 54(1), 54–74.
<https://doi.org/10.3758/s13428-021-01581-x>
8. Hendrickson, A. E., & White, P. O. (1964). Promax: A quick method for rotation to oblique simple structure. *British Journal of Statistical Psychology*, 17(1), 65–70.
<https://doi.org/10.1111/j.2044-8317.1964.tb00244.x>
9. 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>
10. Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3), 187–200.
<https://doi.org/10.1007/BF02289233>
11. Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El Análisis Factorial Exploratorio de los Ítems: una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151–1169.
<https://dx.doi.org/10.6018/analesps.30.3.199361>
12. MacCallum, R. (1983). A comparison of factor analysis programs in SPSS, BMDP, and SAS. *Psychometrika*, 48(2), 223–231.
<https://doi.org/10.1007/BF02294017>
13. Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(1), 0021.
<https://doi.org/10.1038/s41562-016-0021>
14. Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716.
<https://doi.org/10.1126/science.aac4716>
15. Revelle, W. (2023). *psych: Procedures for Psychological, Psychometric, and Personality Research* (R package version 2.3.9). Northwestern University. <https://CRAN.R-project.org/package=psych>
16. Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366.
<https://doi.org/10.1177/0956797611417632>
17. 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), 2521.
<https://doi.org/10.21105/joss.02521>
18. Watkins, M. W. (2018). Exploratory Factor Analysis: A Guide to Best Practice. *Journal of Black Psychology*, 44(3), 219–246.
<https://doi.org/10.1177/0095798418771807>