The Rotation Problem

Steven Marcel Bißantz

"The Rotation problem" is an excerpt from my master thesis "Exploratory Dimensionality Analysis in the Social Sciences – State of the Art and Alternative Approaches to Combat the Problem of Identifying Latent Dimensions". More precisely, it aims to shed light on the third of the three major problems in exploratory factor analysis (EFA): The rotation problem. Leaving the communality problem and the number of factor problem aside, the excerpt focuses on (a) illuminating the topic from a researcher's perspective while (b) developing an understanding of a more technical view on rotation – the robot's perspective. Following the object under investigation in the social sciences, I identify bad defaults and propose decent alternatives. This guide will ultimately help researchers to navigate between dozens of well-known but often poorly understood alternatives to rotate in EFA.

Keywords: Three Major Problems, Rotation Problem, Exploratory Factor Analyis

Extracting factors using Maximum Likelihood Factor Analysis (MLFA) yields–conditional on the data–the most plausible values for the above reproduction problem. But there is usually a concern with the proposed loading matrix. Even though the result suffices for factor-analytic robots, it is not so for researchers. Why? Indeed, the solution describes the association between factors and indicators; but researchers often cannot interpret it straightaway. The loading picture is often too diffuse. Or, to put it another way, the loading matrix has no *simple structure*[1].

Transformation overcomes the problem. A transformation of the loading matrix is visually referred to as ROTATION. Rotations function as interpretation aids for researchers trying to produce unambiguous loading pictures while maintaining the extracted relationships between variables in the result. Their outcome usually leads to a clear overall picture of the factor-item correlations. Ultimately, rotations foster clarity by facilitating a broader understanding of the loading pattern.

The simpler (i. e., more interpretable) structure can be achieved in two ways; research can either utilize orthogonal or oblique rotation techniques. Both will be discussed after shedding light on the researcher's and robot's perspectives.

Researcher's perspective

To get an idea of what it means to rotate, imagine a Cartesian coordinate system (Figure 1). For ease of sake, imagine the space is two-dimensional. With two dimensions, there are two axes X and Y. Rename X and Y using the labels of the previous MLFA-result, for instance, "honor" and "pride". In addition, restrict the range of values for X and Y to stay within the one-minus-one interval. Why minus one to one? Because the items are now defined in terms of their loadings and loadings, in turn, are factor-item *correlations*. Let's say the item "You would praise a man who acts aggressively to insult" is located at $I : (0.5, 0.6)$. If the values 0.5 and 0.6 represent the item's correlations with each factor, one can plausible assume "You would praise a man who acts aggressively to insult" correlates moderately to highly with "honor" and highly with "pride". Generalizing this logic, one can map the entire loading matrix into the two-dimensional coordinate system. The factor-item correlations become factor coordinates.

But does the above item, "You would praise a man who acts aggressively to insult", now relate to "honor" or "pride"? That is just it; since both values are approximately equal, one cannot easily say. As mentioned

---

The LaTeXcode and the pdf-version of the writing sample are freely accessible on https://github.com/sbissantz/smip_application. For correspondence concerning the article please open an issue on Github.

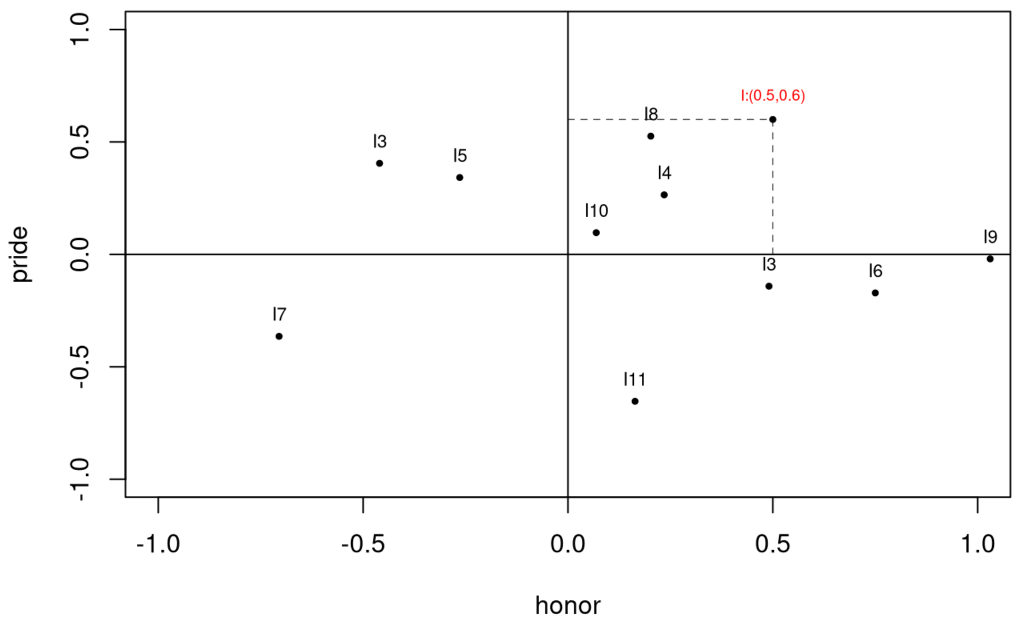

Figure 1

Factor Coordinate System – a visualization of a loading matrix. The loadings allow mapping the ten simulated variables in the coordinate system. The average loading is $\bar{\lambda} = .55$ *and the between-factor correlation is* $\phi_{12} = 0.2$

above, tackling the issue implies translating the intermediate solution into something simple. In the graphical example, the question formulates as follows: How to reallocate the axis in the coordinate system to generate the desired simple(r) loading structure? The researcher has two choices: First, to "spin the axes", using orthogonal rotation techniques, or to apply an oblique transformation. Both methods differ in adaptability to the object under investigation (Oui) in the social sciences. While orthogonal rotations maintain the right angle between the factor axes, transformation allows them to move independently. The latter implies that the factors are allowed to correlate, which will become a significant detail hereafter.

Let's pick up on understanding orthogonal and oblique techniques first. With Orthogonal rotation the elementary question is how to reallocate the axis that a simple (i. e., easy-to-interpret) loading-structure results. To give an analogy, think of a spinning rotor. The rotor as a whole represents the factor-coordinate system. The bolt (i. e., the origin) holds the rotor in place so that the blades (i. e., the axes) can spin. When the rotor blades turn, they maintain a fixed 90-degree angle. The same holds for factor rotation. If the axes in a factor-coordinate system turn, the blades move counterclockwise, preserving the right angle. Mainly, one refers to the procedure as "orthogonal rotation". The metaphor breaks, however, with Oblique rotations. Here, one does not spin the axes, preserving the right angle between them. The crux is to allow the axis to abandon perpendicularity explicitly. Intuitively, one should think of the two-dimensional oblique case as a clock (i. e., a factorial coordinate system) in which the hands (i. e., the factors) move independently. Thus, the "time" becomes substantial (i. e., the angle between the factors). It corresponds to an additional piece of information: the between-factor correlations.

Second, the link between graphical rotation and matrix algebra should become clear. While a visual overview is helpful to get started, it makes it difficult to fully grasp the comprehensive modeling approach. The fact holds especially for modern oblique techniques. To distinguish the graphical and algebraic approach the term Oblique transformation (Revelle, 2021) will be used. With the preliminaries set, one question remains for the following section: What happens inside the machinery of the factorial robot? Let's switch perspectives to get right under the hood.

Robot's perspective

From a robot's point of view, rotation problems are a matter of solving the transformation equation – under different constraints. Any solution builds upon the Transformation equation, formalizing as:

$$\hat{\Lambda}_r = \hat{\Lambda} T \tag{1}$$

Notice that the equation is just a formalized version of the issue outlined above. Namely, finding a set of axes that produces a simple(r) loading structure ($\hat{\Lambda}_r$). More precisely, the solution just *appears* simple(r) (i. e., more accessible for the researcher). The reason is that the relationships between variables are left untouched. More technically, the robot searches a transformation matrix (T) rearranging the estimated information ($\hat{\Lambda}$) but leaving the model's fit to the data untouched (Mair, 2018). In this light, the formal equivalent of spinning the axis is a multiplication with this matrix (T). Consequently, the transformation matrix qualifies as the pivot point in aboves equation. It is key to understand transformation procedures. Developing an understanding of the transformation next will answer the question of what happens inside the machinery.

In the orthogonal case, to multiply the plotted matrix of factor loadings ($\hat{\Lambda}$) with a transformation matrix (T), moves the axes counterclockwise until the desired result ($\hat{\Lambda}_r$) is reached. Thereby, the robot searches for a solution with two properties: Obviously, the transformation needs to solve the transformation problem (Equation 1). But even more important is the premise of preserving the right angle between the factors. The crux is to employ a set of orthogonal vectors because they preserve the right angle – rotating the factors (counterclockwise) by 90 degrees. In mathematical terms, assuming $TT^t = I$, the orthogonalization preservation malizes as follows:

$$\begin{aligned} P_r &= \Lambda_r \Lambda_r^t + \Psi \\ &= \Lambda T (\Lambda T)^t + \Psi \\ &= \Lambda \underbrace{TT^t}_{I} \Lambda^t + \Psi \\ &= \Lambda \Lambda^t + \Psi \\ &= P \end{aligned} \tag{2}$$

In the oblique case, solving the rotation problem implies resolving the perpendicularity of the factor axes (algebraically, $I \neq TT^t$). Giving up on the assumption of factor orthogonality, one literally preserves the matrix of factor correlations (Φ). The matrix of factor correlations is part of the fundamental equation of factor analysis ($P = \Lambda \Phi \Lambda^t + \Psi$) and contains the correlations between factors. For instance, if $\Phi_{\phi_1;\phi_2} = 0.6$, where ϕ_1 represents "honor" and ϕ_2 represents "pride", the data include a high correlation between "honor" and "pride". Using orthogonal rotation instead, factor-correlations are assumed to be 0, and Φ is omitted[2].

Solutions. To solve the rotation problem, researchers have two options [3]: orthogonal rotation and oblique transformation. When it comes down to deciding on a particular one, researchers should critically evaluate if it is reasonable to ignore the between-factor correlations. The basis of their decision must be the correspondence with the Oui. Researchers have to internalize that their decision has inevitable consequences on their inferences. Orthogonal rotation keeps the perpendicularity of the factor axes – another appearance of factor independence. *The decision between orthogonal and oblique rotations has a profound effect on the* Oui. *I would argue that it is much stronger than selecting within the sphere of rotation criteria.*

To highlight the conclusion structurally, all criteria are nested regarding their origins in the following sections (i. e., orthogonal rotation or oblique transformation). Note that today, there are over two dozen rotation procedures (Gorsuch, 2015, pp. 214) – too many to discuss them all. However, discussing bad defaults and finding decent alternatives will be an inevitable part of the following sections.

Bad defaults

Among orthogonal rotation techniques like Quartimax (Carroll, 1953), Equamax (Saunders, 1962) , and

ORTHOMAX (Harman, 1970); VARIMAX (Kaiser, 1958) is the most popular choice in applied research (Costello & Osborne, 2005; Ford et al., 1986; Loo, 1979).

Kaiser (1958)'s VARIMAX criterion proposes a solution for the above rotation problem, maximizing the variance of the squared loadings for each factor. This means VARIMAX is out to boost the (squared) correlations between factors and items. In doing so, the transformation makes the factor axis rotate by 90 degrees counterclockwise, preserving the right angle between the factor vectors. The matrix spins the axes, so to speak. Accordingly, Kaiser's VARIMAX rotation subsumes under the orthogonal rotation techniques.

But why is it a bad default? To grasp the problem, one must understand how the orthogonality of factor-vectors relates to the assumption of independent factors. Mathematically, they connect through the cosine. The cosine concentrates the 360-degree spectrum to a one-minus-one range. Therefore, the angle, a visualization of their relationship, can be transferred into a well-known sphere of relatedness – their correlation. As a result, if factor-vectors are perpendicular, they should not correlate. Indeed, since cos(90) equals 0, they do not. Thus, maintaining the 90-degree angle in rotations preserves the right angle and consequently factor independence. There was a cursory note on the previous finding in VARIMAX's explanation. VARIMAX maximizes the variance of the squared loadings. Thinking back on the syringes in the extraction problem, VARIMAX acts accordingly; it maximizes information density within each factor – but at the cost of omitting between factor correlations. Correspondingly, the result remains factor independence.

However, there is usually no reason for the researcher to assume the independence of factors – especially not by default. Costello and Osborne (2005, p. 3) urges furthermore to expect variation among factors because "behavior is rarely partitioned into neatly packaged units that function independently of one another". The quote should sound familiar, recalling the characteristic features of the OUI (e. g., its blurry boundaries). Latent dimensions nearly always share some bits of information. Accessing them through independent factors is a recipe for doom, inducing ignorance bias (Loo, 1979). The finding is in line with Sass and Schmitt (2010), who conducted a comparative investigation of different rotation criteria (within the framework of exploratory factor analysis). They encourage researchers to reflect on their rotation choices because it has a meaningful impact on the manifestation of the factor structure.

The habitual use of orthogonal rotations, like VARIMAX, is not harmless. It is reasonably a bad default. It can prove a dangerous undertaking in exploratory investigations, especially if factors demand to correlate with each other, but chosen options hinder them in doing so (Loo, 1979). That's the reason why switching to decent alternatives is mandatory.

Decent alternatives

Recall the goal of rotation and transformation procedures. They aim for transferring the initial robot's results into something simple (i. e., easy to understand). To judge simplicity among criteria, simplicity indexes come into play (see, e. g. Bentler, 1977; Kaiser, 1974; Lorenzo-Seva, 2003). Although they differ in terms of their mathematical formulation, conceptually SIMPLICITY INDEXES aim to find a solution with each item indicating the least number of factors. In this sense, simplifying the result maximizes interpretability.

The reason to use one of the most popular oblique transformation procedures, like OBLIMN and PROMAX (Hendrickson & White, 1964) is twofold. The first argument is their correspondence with the object under investigation (OUI) in the social sciences. The second is that they score high on one of the above's simplicity criteria. It seems like a win-win situation. On the one hand, all solutions aim to maximize factor simplicity; on the other, they avoid factor independence. So all of them are decent alternatives. But when it comes down to finding the most decent alternative, OBLMIN, as well as PROMAX, do not provide the simplest result. Indeed, they all score high on *some* indexes. But Lorenzo-Seva (2003) argues that these are flawed ones.

The problem is that no criteria concentrate on the loading itself. Because all tackle the loadings merely indirectly, they miss the target, which is the simplest loading structure doable, as a result. Bentler's index, for example, builds upon the *colums* of the matrix of factor loadings. The loading simplicity, in turn, focuses solely on the values of the loadings directly.

So what is the high-performance rotation criterion according to the more accurate loading simplicity (LS) index? SIMPLIMAX (Kiers, 1994). In Lorenzo-Seva (2003)'s simulation study the criterion not only outperformed more elaborate versions of OBLIMIN and PROMAX.

Lorenzo-Seva (2003) furthermore demonstrated the Simplimax algorithm delivers a transformation matrix pooling the communality of each item on the fewest number of factors. So the rotated loading pattern is indeed as simple as possible. The factor-loadings are either zero or as far from zero as doable [4].

In conclusion, Simplimax is the most decent alternative for two reasons. Given the LS-index, it proved to deliver the simplest (i. e., most interpretable) solution possible while keeping track of the characteristic features of the Oui.

Research recommendations

The decision between orthogonal and oblique methods is weightier than choosing a particular criterion. Concerning the object under investigation, oblique transformations prove a better standard for applied researchers. Factor independence is an exertion rather than the rule, and oblique rotations allow modeling factor dependencies explicitly. They permit factors to correlate with one another while still simplifying the patterns in the loading matrix as much as possible. Oblique transformations do no harm the Oui. With orthogonal rotation, the same is not unconditionally true.

Despite their shortcoming, orthogonal rotations are a common choice in the social sciences. Smuggling additional information into an exploratory investigation, like factor independence, proves a profound problem in applied research. Loo (1979) stresses this point. As one of a few, he reviewed the (clinical) literature, questioning the appropriateness of the assumption of independence. Loo (1979) found researchers fall back on orthogonal rotation procedures regularly. But noteworthy is, second, in most reviewed cases, the assumption of independence was unreasonable. Over twenty years later, orthogonal rotations are still at the top of common-go to methods in applied research (Costello & Osborne, 2005; Ford et al., 1986). It is hard to estimate how much research is flawed by the implicit assumption of orthogonality[5].

In this light, it is questionable to me why researchers like Bortz and Schuster (2016) are concerned with a loss in compression when going oblique. Again, *the goal in the social sciences is to compress meaningfully, not maximally*. Although allowing the factors to correlate, induces some redundancy; it will often prevent the researcher from having to deal with highly distorted results (Sass & Schmitt, 2010).

However, one should not exclude orthogonal rotations from exploratory investigations. Previous sections suggest only to *start with the least, not the most rigid model assumption*. Learning from data implies adjusting a model to the data, not imposing a well-known model on the data. In the rotation or transformation context, it means to model inter-factor dependencies by default and learn about them on the go. The conclusion is in line with Muthén (1984) and Sass and Schmitt (2010) who recommend rotation procedures providing a simple solution, but not at the cost of inducing incompatibility between methods and the object under investigation. The bottom line is to find the most simple result possible. But researchers must avoid the current practice of finding simplicity at any cost.

Again, orthogonal rotations are not worthless in applied social-science research. Researchers can use them strategically. For example, to highlight the added value of resolving to assume independence. Research should become experimental, trying different techniques and evaluating their scientific use by learning from model differences and similarities (Tabachnick & Fidell, 2007, p. 642). No matter if results are equal across trials or completely different, in an exploratory stage of the investigation, every finding is valuable information about the models and how they see the data. That information should never be excluded, suppressed, or abandoned on a preliminary ground. Researchers should report them to improve domain knowledge.

Additionals

There are scenarios in which orthogonal solutions are a reasonable choice. For example, if compressing information more fully is prior. The small additive thus guides the application of orthogonal rotation techniques for practical researchers. In general, Varimax is reasonable if factor dependence is a minor issue or if the researcher expects more than a single general factor to underlay a set of items (Gorsuch, 2015, p. 195). But if the researcher anticipates a single factor, Varimax becomes problematic. Why? Because it distributes the variance across factors and thus dampens the tendency of a single factor to occur in the result (Sass & Schmitt, 2010). As a result, if the researcher anticipates a single general factor, Quartimax is the better choice (Mair, 2018). Despite that Equamax, combines Quartimax and Varimax, portioning the variance more evenly across factors (Gorsuch, 2015, p. 214). If it comes down to a single choice for a particular criterion, Gorsuch (2015) recommends VARIMAX because, in visual inspections, they produce interpretable results and prove invariant across a wide range of circumstances.

The second add-on is devoted to the GESTURE OF PROPOSING. In applied research, dimensionality analysis is often thought of and taught as if there has to be a definitive result. But especially in the early stages of exploratory investigations, there might be more than one plausible solution compatible with the constraints (e. g., data). In case of doubts, researchers should start to propose different plausible solutions to the research community (for an exception, see Timmerman et al., 2017). The conclusion is in line with Gorsuch (2015, p. 224) who states that any achievable result in an early stage of theory development is an intermediate, a hypothesis, for (follow-up) investigation yet to come. As will be shown, his suggestion holds, especially when it comes down to determining number of factors in the next section.

Notes

1. Thurstone (1935, 1947, 1969) developed a set of principles on how to rearrange the loading matrix. Thinking in zeros and ones, one aims to redistribute both of them in the loading matrix to achieve an unequivocal loading pattern. At best, every item loads only on a single factor. If this is true for every item, the outcome has a simple structure: Each is attributable to (i.e., loads highly on) a single factor. For Carroll (1953), Jennrich and Sampson (1966), and Tucker (1955), and others, this must have been a statistical fairytale. They all criticize that a simple structure is hardly ever reached. Nonetheless, today there are various implementations to come close to the ideal. Some of them are part of the following sections.

2. To give additional background, let's have a look at the pattern-structure-matrix distinction. The structure encompasses all common information in the correlation matrix. (1) The information between items and factors, as well as (2) the information between factors. The decomposition is used in the fundamental equation ($P = \Lambda\Phi\Lambda^t + \Psi$). If the factors do not share any additional information, their correlation is zero ($\forall\phi_{i,j|i\neq j} = 0$) and Φ becomes the identity matrix I. In this case, the structure is reproducible using only factor-item information. Finally, the structure (S) is reproducible as a function of the factor-item information (Λ) and factor-factor information (Φ). Algebraically, it formalizes as: $S = \Lambda\Phi$. So rotating orthogonal sets of the correlation between factors, the off-diagonals become zero, Φ equals I, and S reduces to Λ. The structure matrix and pattern matrix are identical. The key question in the following will be how realistic the assumption of independence is in applied social science research.

3. Since the pioneering work of Thurstone in the 1940s, visual rotation are also part of the exploratory analyzers toolkit. The problem with this rotation, however, is selecting factor radians. Carroll (1953), for example, critiques that rotation visually induces some kind of arbitrariness in the procedure. Researchers are suspected of missing a concise "objective" argument on which to evaluate the position of the axes. As a result, researchers usually turn towards analytic criteria. What often remains unnoticed: Visual rotation allows a pretty good approximation of a simple structure Gorsuch (2015, p. 202). To combine the best of both worlds, nowadays, analytic criteria predominate, which orient towards providing a mathematical gateway to a simple structure

4. The workhorse in Kiers (1994)'s algorithm is $\sigma(T, \Lambda_r) = ||\Lambda T - \Lambda_r||^2$, with pattern matrix Λ, transformation T, and target Λ_r. The crux is to find the so-called BEST SIMPLE TARGET, which means to solve for a transformation (T) that allows rotating the pattern matrix Λ, such that the rotated matrix (Λ_r) has a given number (p) of zero entries (see: Kiers (1994) for more details).

5. Gorsuch (2015) prompts towards the rare case of Guilford (1981). Replacing orthogonal with oblique transformation, he started reanalyzing his past research, finally resolving the assumption of independence.

References

Bentler, P. M. (1977). Factor simplicity index and transformations. *Psychometrika*, *42*(2), 277–295. https://doi.org/10.1007/BF02294054

Bortz, J., & Schuster, C. (2016). *Statistik für Human- und Sozialwissenschaftler*. Springer. https://www.ebook.de/de/product/26244409/juergen_bortz_christof_schuster_statistik_fuer_human_und_sozialwissenschaftler.html

Carroll, J. B. (1953). An analytical solution for approximating simple structure in factor analysis. *Psychometrika*, *18*(1), 23–38. https://doi.org/10.1007/BF02289025

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, *10*, 1–10. https://doi.org/10.7275/jyj1-4868

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: a critival review and analysis. *Personnel Psychology*, *39*(2), 291–314. https://doi.org/10.1111/j.1744-6570.1986.tb00583.x https://doi.org/10.1111/j.1744-6570.1986.tb00583.x

Gorsuch, R. L. (2015). *Factor analysis*. Routledge.

Guilford, J. P. (1981). Higher-Order Structure-Of-Intellect Abilities. *Multivariate Behavioral Research*, *16*(4), 411–435. https://doi.org/10.1207/s15327906mbr1604_1

Harman, H. H. (1970). *Modern factor analysis.* The University of Chicago Press.

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

Jennrich, R. I., & Sampson, P. F. (1966). Rotation for simple loadings. *Psychometrika*, *31*(3), 313–323. https://doi.org/10.1007/BF02289465

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*(3), 187–200. https://doi.org/10.1007/BF02289233

Kaiser, H. F. (1974). A note on the equamax criterion. *Multivariate Behavioral Research*, *9*(4), 501–503. https://doi.org/10.1207/s15327906mbr0904_9

Kiers, H. A. (1994). Simplimax: Oblique rotation to an optimal target with simple structure. *Psychometrika*, *59*(4), 567–579. https://doi.org/10.1007/BF02294392

Loo, R. (1979). The orthogonal rotation of factors in clinical research: A critical note. *Journal of Clinical Psychology*, *35*(4), 762–765. https://doi.org/10.1002/1097-4679(197910)35:4<762::AID-JCLP2270350414>3.0.CO;2-M

Lorenzo-Seva, U. (2003). A factor simplicity index. *Psychometrika*, *68*(1), 49–60. https://doi.org/10.1007/BF02296652

Mair, P. (2018). Classical test theory and factor analysis. *Modern psychometrics with r* (pp. 1–34). Springer. https://doi.org/10.1007/978-3-319-93177-7_1

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, *49*(1), 115–132. https://doi.org/10.1007/BF02294210

Revelle, W. (2021). Constructs, Components, and Factor models and Reliability. *An introduction to psychometric theory with applications in r* (pp. 145–239). Online publication. https://personality-project.org/r/book/

Sass, D. A., & Schmitt, T. A. (2010). A comparative investigation of rotation criteria within exploratory factor analysis. *Multivariate Behavioral Research*, *45*(1), 73–103. https://doi.org/10.1080/00273170903504810

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Pearson/Allyn & Bacon.

Thurstone, L. L. (1935). *The Vectors of Mind.* University of Chicago Press.

Thurstone, L. L. (1947). *Multiple - factor analysis.* University of Chicago Press.

Thurstone, L. L. (1969). *Multiple-factor analysis: a development and expansion of The vectors of mind*. The University of Chicago Press.

Timmerman, M. E., Lorenzo-Seva, U., & Ceulemans, E. (2017). The number of factors problem. *The wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development* (pp. 305–324). John Wiley & Sons Ltd. https://doi.org/10.1002/9781118489772.ch11

Tucker, L. R. (1955). The objective definition of simple structure in linear factor analysis. *Psychometrika*, *20*(3), 209–225. https://doi.org/10.1007/BF02289018