University of Koblenz-Landau

Master Thesis

Exploratory Dimensionality Analysis in the Social Sciences

State of the Art and Alternative Approaches to Combat the Problem of Identifying Latent Dimensions

Author:
Steven Bissantz

Supervisors:
Dr. Inga Wagner
Jonas Steinbrenner

A thesis submitted in fulfillment of the requirements for the degree of Master of Arts

in the

Social Sciences
(i. e., Social Sciences & Communication)

September 2, 2021

Declaration of Authorship

I, Steven Bissantz, declare that this thesis titled, "Exploratory Dimensionality Analysis in the Social Sciences" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

"*I may be wrong and you may be right, and by an effort, we may get nearer to the truth* "

Sir Karl Popper

UNIVERSITY OF KOBLENZ-LANDAU

Abstract

Social Sciences

(i. e., Social Sciences & Communication)

Master of Arts

State of the Art and Alternative Approaches to Combat the Problem of Identifying Latent Dimensions

by Steven BISSANTZ

Identifying dispositions in social collectives is a unifying goal in the social sciences. Therefore, researchers explore large data sets to get inevitable clues of latent dimensions. But often bad defaults predominate exploratory investigations (e. g., Cronbach's alpha, PAFA, varimax, K1, and the scree test). All of them are problematic, being out of date or incompatible with the object under investigation (OUI) in the social sciences – or both. Smuggling in the incompatible assumption of independent dimensions, for example, researchers frequently distort their dimensionality assessments with low-performing common go-to tools. Therefore, identifying the state of the art has thus two prerequisites; first, finding up-to-date options, which are, second, compatible with the OUI in the social sciences. Starting from scratch, profound knowledge on the concept of dimensionality is developed together with a detailed understanding of the nuts and bolts of each criticized as well as the proposed method. Moving bottom-up furthermore restores the researcher's sensitivity for the characteristic features of the OUI and imparts skills on how to handle them. This in-depth approach alleviates method-blindness fallacy and ignorance bias in applied research by replacing predominating bad defaults with decent alternatives (e. g., coefficient omega, MLFA, simplimax, the hull method, and the sequential χ^2-test). Ultimately, factor analysis and exploratory Likert scaling remain as bespoke algorithms to combat the problem of identifying latent dimensions.

Zusammenfassung: Dispositionen in sozialen Kollektiven ausfindig zu machen, ist ein universales Ziel in den Sozialwissenschaften. Aus diesem Grund erkunden Forscher große Datensätze auf der Suche nach empirischen Anhaltspunkten für das Vorhandensein latenter Dimensionen. Oftmals greifen sie bei den explorativen Analysen jedoch auf problematische Standards zurück (z.B. Cronbach's alpha, PAFA, Varimax, K1 und Scree-Test). Das Problem mit diesen Standards ist, dass sie entweder oft veraltet oder mit dem sozialwissenschaftlichen Untersuchungsgegenstand unvereinbar sind – oder beides. Durch die Wahl inadäquater Standardmethoden verzerren Forscher häufig ihre Analyseergebnisse, beispielsweise durch die Annahme der Unabhängigkeit latenter Dimensionen. Folglich gibt es zwei Voraussetzungen, um neue Methoden zu evaluieren; erstens, die eingesetzten Verfahren müssen auf dem neusten Stand der Technik und, zweitens, mit dem Untersuchungsgegenstand kompatibel sein. Dazu wird in dieser Arbeit zunächst das Dimensionalitätskonzept entwickelt und ein tiefgreifendes Verständnis für das entsprechende Problem mit der jeweilige Methode sowie ihrer vorgeschlagenen Alternative vermittelt. Mit dieser Vorgehensweise soll das Feingefühl für die charakteristischen Eigenschaften des Untersuchungsobjektes, ebenso wie Fertigkeiten im Umgang mit ihm vermittelt werden. Diese Vorgehensweise soll Trugschlüssen durch Methodenblindheit (eng.: method-blindness fallacy) und Ignoranzverzerrung (eng.: ignorance bias) vorbeugen, indem problematische Standards durch angemessene Alternativen ersetzt werden (z.B. Koeffizient Omega, MLFA, Simplimax, die Hull-Methode und der sequentielle χ^2-Test). Schlussendlich sind es die explorative Faktorenanalyse und die explorative Likertskalierung, die es als maßgeschneiderte Alternativen ermöglichen latente Dimensionen in den Sozialwissenschaften ausfindig zu machen.

Acknowledgements

I would like to say a special tank you to my first supervisor Dr. Inga Wagner. Thank you for your consistent support and guidance during the last two years. Previous accomplishments would not have been possible without you. I would also like to thank Jonas Steinbrenner. The conversations were vital in inspiring me to think outside the box. I am gratefully indebted to all your valuable comments in the process. Finally, I must express my very profound gratitude to my parents and to my partner, for providing me with unfailing support and continuous encouragement throughout my years of study.

Contents

List of Figures

List of Abbreviations

CFI	**C**omparative **F**it **I**ndex
CTT	**C**lassical **T**est **T**heory
EFA	**E**xploratory **F**actor **A**nalysis
ELS	**E**xploratory **L**ikert **S**caling
FA	**F**actor **A**nalysis
LRT	**L**ikelihood **R**atio **T**est
MLFA	**M**aximum **L**ikelihood **F**actor **A**nalyis
MRFA	**M**inimum **R**ank **F**actor **A**nalyis
OLS	**O**rdinary **L**east **S**quares
OUI	**O**bject **U**nder **I**nvestigation
PA	**P**arallel **A**nalysis
PAFA	**P**rincipal **A**xis **F**actor **A**nalyis
PCA	**P**rincipal **C**omponent **A**nalysis
PCM	**P**rincipal **C**omponent **M**ethod
SMT	**S**equential **M**odel **T**est
UDA	**U**ni **D**imensionality **A**nalysis
UVA	**U**nderlying **V**ariable **A**pproach

List of Symbols

X^t	Transposed of X
$\hat{X}$	Estimate of X
$\bar{X}$	Average of X
P	Model-implied Correlation Matrix
R	Correlation Matrix
R^*	Reduced Correlation Matrix
T	True Score
U	Singular Matrix
X	Data set
$X_{i=1,\ldots,m}$	Observable Indicators
$X_{i,j}$	Single Indicators
α	Cronbach's Alpha
ζ	Latent Dimension
Λ	Loading Matrix
λ	Loading
λ^*	Random eigenvector
Φ	Factor Correlation Matrix
ϕ	Between-Factor Correlation
ξ	Uniqueness
ρ	Correlation Coefficient
ρ_{TX}	Item-to-Total Correlation
ρ^*_{TX}	Corrected Item-to-Total Correlation
Σ	Struture
Ψ	Uniqueness
ω	Coefficient Omega

For Julia...

1 Introduction

One of the major bridges that connect various disciplines in the social sciences is widespread interest in multi-faceted constructs. They are part of research questions, utilized to develop theories, embedded in explanations, or used to describe even more complex social phenomena. To get clues of their existence and access them precisely is an essential part of common research practice in the social sciences. Take "honor" as an example. A whole research paradigm developed around the phenomenon, aiming to conceptualize and finally measure it (Beck, 1996; H. H. Jackson et al., 1997; Nisbett, 2018; Saucier et al., 2016; Shackelford, 2005; van Osch et al., 2013). But there is nothing special to "honor". The argument can be easily applied to other concepts. Take "pro-environmental behavior" for instance. Both examples poke right inside the intricate problems a social scientist has to confront: *identify and measure complex social phenomena with respect to their own particularities.*

But what is actually the OBJECT UNDER INVESTIGATION (OUI) in social science, and what is peculiar to it? First, the OUI in the social sciences, things like "honor", can be referred to as dispositions or latent dimensions. A DISPOSITION in social science is a (relatively) stable tendency to react. It does not primarily matter if the reaction manifests cognitively or leads to affective or behavioral consequences; but that the pattern is traceable and remains constant across a wide range of situations and circumstances. A broader range of typical examples includes areas such as abilities, skills, and expertise, as well as attitudes and personality traits. In other words, the main goal in applied social sciences research is to access and capture the (relatively) stable patterns that emerge from social interactions. Or, tying in with the above, the aim is to describe and explain dispositions in social collectives.

To answer the second, the particularity question, dispositions have special demands. They result from the OUI's inherent features – its CHARACTERISTIC DEPENDENCIES. Take "honor" as an example again. The extent to which honor is present in a society strongly hinges on the social collective; furthermore, it has turned out to be a time-dependent phenomenon (Nisbett, 2018; Shackelford, 2005; van Osch et al., 2013). This testifies to a special time and cultural dependency and induces a concept-specific continuity. Dispositions develop along a continuum, they can vary over time and manifest differently across cultures (Bornstein, 2018; Corballis & Traub, 1970; Nesselroade, 1972; Ram & Grimm, 2015; Tisak & Meredith, 1989, 1990; Wickrama et al., 2020). But they stand out through an additional feature. Dispositions have BLURRY BOUNDARIES, which means they share common attributes with other latent dimensions. They overlap with them, so to speak. Consequently, demarcating attributes is often a tough issue, because the boundaries between them are blurry. Think intuitively of "patriotism" and "nationalism". With little elaboration, one will agree that they are not easily distinguishable (for an attempt, see V. M. Costa, 2018). Thus, both tend towards blurry boundaries.

A serious problem in applied research arises from neglecting the inherent properties of the OUI. Ignorance has shown to impose bias (Loo, 1979). The so-called IGNORANCE BIAS is often a result of bad (i. e., unreflected) default behavior. Methods are often applied without seeking for compatibility with the object under investigation. Loo (1979) conducted pioneer work, reviewing if the application of go-to methods in applied research coincides with the characteristic dependencies of the OUI. He found that often researcher's (default) choices smuggled in some set of assumptions that are unreasonable under most circumstances (i. a., independence of latent dimensions). One could also say, researchers are blind to the characteristic features of the OUI from the background of their BAD DEFAULTS. Think of the term "bad" here as a proxy for the incompatibility between the applied method and the characteristic features of the OUI. In this light, the shown METHOD-BLINDNESS FALLACY often results in IGNORANCE BIAS.

Even though decent alternatives exist, less effort in the research community is devoted to reconciling statistical procedures with the characteristic features of the OUI. This thesis tries to fill in the gap. It is now time to eliminate bad defaults and replace them with recent and decent alternatives. The realignment will boost the quality of applied research and lead to a considerate upgrade of the modern researcher's toolbox. With respect to Loo (1979)'s findings, it is long overdue to scrutinize if the applied methods are appropriate for empirical research in the social sciences (i. e., in accordance with its OUI). Too long, it was ignored to seek compatibility between research methods and the object under investigation. However, this thesis sets out to address this issue. To attract and shift attention towards greater convergence of research methods and the

OUI is the major concern of a more in-depth approach. More precisely, this requires (1) understanding to evaluate the most common statistical techniques, (2) reviewing their flaws and (3) assess their capability to identify latent dimensions, while keeping (4) sight of the OUI's characteristic features. Correspondingly, the features define the criterion for applicability of a method in this work. A subsequent critical reflection should recover sensitivity and contribute to developing new skills. Tools for exploratory purposes and theory development need to be tailored to the use in applied social science research. Ultimately, a series of bespoke tools will be proposed as decent alternatives. Replacing bad defaults will preempt method-blindness fallacy and additionally alleviate ignorance bias.

Outline

Turning intentions into actions requires a more in-depth approach. Therefore, a robust methodological foundation building upon some intense thoughts on the object under investigation was already set up, *before* now starting to work with it (ch. 1). Chapter 2 provides some requisite know-how for the chapters to come. The three upcoming chapters will be devoted to a particular type of dimensionality analysis, chapter 3 focuses on unidimensionality analysis and reliability analysis. Chapter 4, is about the common go-to method for multidimensionality analysis in the social sciences – exploratory factor analysis. Solutions for the big three problems, namely the communality problem, the rotation problem, and the number of factor problem will be elaborated. A significant theoretical objection against multidimensional analysis finally opens up the opportunity to encounter a multiple unidimensional scaling approach (ch. 5). All chapters are accompanied by R code examples to facilitate practical application in self-instructions. The final chapter, chapter 6, summarizes and discusses previous key findings.

2 Primer

Before particular strategies will be presented to combat the problem of capturing dispositions in social collectives, a primer on scale development is prefixed to prepare the upcoming chapters. It elaborates on general details, formalism, and some terminology (e. g., goals, scale development processes, and the latent variable logic).

An important goal in the social sciences is to assess social dispositions. Since almost all of them, like honor, are not directly observable, many fall into the category of latent dimensions. LATENT DIMENSIONS are not directly observable variables. They lead to the INTRICATE PROBLEM in the social sciences, namely that the OUI is hardly ever directly observable. Trying to cope with the intricate problem, dimensionality analysis takes the role of a means to infer latent dimensions. More precisely, it provides researchers with empirical arguments for their presence based on which inference is finally done. A crucial task in applied research is thus to capture the latent dimension, or, technically speaking, to quantify or measure it. This is where scales come in handy. SCALES are measurement instruments trying to access latent dimensions with manifest (i. e., actually observable) variables. These are also known as OBSERVABLE INDICATORS. The most common type of observable indicator in the social sciences is the questionnaire item (e. g., "You would praise a man who acts aggressively to an insult" – strongly agree, somehow agree, . . . , somehow disagree, strongly disagree; see Saucier et al., 2016). Researchers often combine multiple ones on a scale to develop reliable instruments capable to assign each participant a numerical value which pins down his or her position on the underlying latent dimension. Overall, researchers receive a numerical outline of the social collective in form a distribution of participants' values, quantifying their object of interest.

Scales are the pivot point of dimensionality assessment. There are basically two entry points for researchers to develop them. In the first one, they start from concrete considerations of a latent dimension (e. g., honor) and invent questionnaire items that are reasonable to capture the latent dimension. For evaluation, data are collected, and the most appropriate items are selected to pin down participants' values on the underlying latent dimension. The second entry point starts with a data set and moves from exploration to measurement (de Leeuw, 2005; Mair, 2018; Mair & Leeuw, 2015; McIver & Carmines, 1981). Thereby, item responses need to be explored for underlying patterns that structure the observable indicators before the participant's value on the underlying latent dimension is finally determined. In the following, the focus is only on the second scenario. So let's flesh it out hereafter.

Being confronted with a data set (X)[1], which is actually a collection point for multiple content-related questions $(X_{i=1,\ldots,m})$, the researcher has to explore the data set. Data exploration in dimensionality assessment means finding out about how many latent dimensions underlay the data. Often used synonymously is the expression "to analyze how the data are structured". Either way, there are at least three plausible scenarios: First, there is no particular structure recognizable, which often means no latent dimension causes the observed item responses, or no latent dimension underlies the data. If they can be traced back to a single cause, one speaks of UNIDIMENSIONALITY. multidimensionality, in turn, implies multiple latent dimensions which produce the observed item patterns. In later chapters, further distinction will be made between (ordinary) multidimensionality and MULTIPLE UNIDIMENSIONALITY, but for now, it is more important to shed light on the assumed causal mechanism that interlinks indicator (X_i) and latent dimension (ζ). In form of a question: Why do latent dimensions *cause* the observed item response patterns? The answer can be given, following the LATENT VARIABLE LOGIC (Wardrop & Loehlin, 1987), which additionally clarifies the ultimate use of dimensionality analysis – LATENT VARIABLE INFERENCE:

Let's assume there is a latent dimension (ζ), which manifests in a series of observable indicators (X). If so, each indicator (X_i) contains information on the latent dimension (ζ) and thus is somehow related to it: $X_i \sim \zeta$. However, from the intricate problem one knows, as the latent dimension is unknown, so must be the relationship with it (Bandalos, 2018, ch. 8): $X_i \overset{?}{\sim} \zeta$. Nonetheless, all indicators arise from the same generative process – they are produced by the same latent variable – hence share common attributes. The common attributes, in turn, manifest in a particular structure (Σ) which includes the relationship ρ among indicators (X_i, X_j):

$$\zeta \to \Sigma : \rho_{X_i,X_j} \tag{2.1}$$

Correspondingly, the entire structure $(\Sigma \in R)$ is like a measurable fingerprint to unlock the mystery of how many latent dimensions underlay the data. As a result, observed inter-item relations suffice to infer a latent dimension by assuming its presence (Thissen et al., 1989). The handiest tool for dimensional assessment is thus an assembly point for inter-item relations: the correlation matrix (R).

There is a huge number of possibilities to measure interrelations (Adelson et al., 2019, ch. 5), most common, however, is Pearson's correlation coefficient (ρ). Measuring the relationships between items via Person's coefficient, the correlation matrix (R) is the place where all linear inter-item associations assemble. It is the foundation of dimensionality analysis because it is the spot where researchers get an inevitable (empirical) clue for latent dimensions, inspecting relationships between observable indicators. It is the place where the latent dimensions (ζ) become visible, thus accessible, and, in the end, traceable. Therefore, the correlation matrix tags the starting point of most dimensional assessments (Gregory, 2014, p. 159). Unidimensional analysis (UDA), exploratory factor analysis (EFA) as well as exploratory Likert scaling (ELS) prove this to be true in the upcoming chapters.

3 Unidimensionality Analysis

A set of indicators is UNIDIMENSIONAL if exactly one latent dimension causes its structure. But how does the dimension (ζ) leave its fingerprint on the correlation matrix (R)? Well, it reveals itself through consistently high inter-item correlations (ρ) among the (m) indicators:

$$\forall X_{i=1,\dots,m} \in R : \rho_{X_i,X_j} \wedge \rho_{X_k,X_l} \gg 0 \tag{3.1}$$

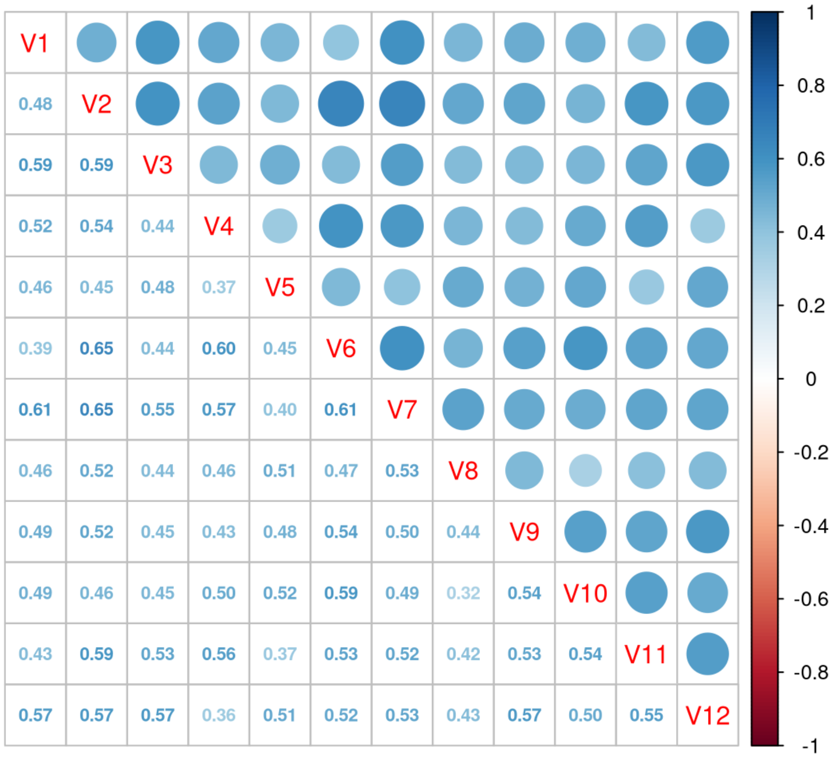

FIGURE 3.1: Unidimensional Item Response Pattern – 12 simulated variables with item-factor correlations $\lambda_{i=1,\dots,12} = 0.7$

.

Recall, all indicators have joint attributes due to the common generative process. Since there is only a single latent dimension affecting the observed response patterns, the indicators only share common features, which manifest in the strength of their inter-item correlations. Likewise, unidimensionality is an attribute of a scale (i. e. a collection of observable indicators). Combing the two insights leads straight to the definition of classical test theory[2] (CTT; Lord et al., 1968), which defines SCALES as unidimensional sets of items (DeVellis, 2006).

Besides, CTT can help to tackle further issues in connection to measure complex social phenomena. As it was shown, accessing collective dispositions like "honor" is not an easy undertaking. Obstacles like the intricate problem in the social sciences and the peculiarities of the OUI stand in the researcher's way, hampering his or her possibility to gain easy access. Ultimately, it is not exaggerated to say; *measurement of complex social phenomenon somehow dooms to measure with error*. A way to get rid of the error is to quantify and control for it. A tool for this purpose is CTT. Hereafter, reliability analysis following CTT is presented as one way to quantify the error of investigation and to select appropriate indicators to pin down participants' values on the underlying latent dimension.

3.1 Reliability analysis

The burden of complex social science measurement in terms of CTT is to access a participant's true score (T) on the latent dimension from error-prone measurement (X). In a nutshell, CTT urges repeating measurement to raise confidence in the assessment. Why? Because every attempt to measure the actual score (T) of a latent dimension (ζ) is a conjecture of the score researchers actually observe (X) plus some error (ϵ):

$$\zeta : T = X + \epsilon \tag{3.2}$$

The increase in confidence then arises from averaging across repeated (error-prone) measurement. When adding small fluctuations on average they dampen one another (McElreath, 2020, p. 73). As the error component is presumably a composition of random fluctuations (e. g., minor shortcomings when answering the questionnaire, like misreading an item), they offset each other. So we can expect the true score of a latent dimension to push through:

$$E(\epsilon) = 0 \Rightarrow \zeta : E(X) = T \tag{3.3}$$

But the assumption ($E(\epsilon) = 0$) is often unrealistic, there always will be some error of measurement (σ_E) one can count on and thus has to estimate (Mair, 2018). So let's continue discussing bad defaults and decent alternatives in assessment.

3.1.1 Bad defaults

The aforementioned error of measurement (σ_E) is of great interest in test theory because it is inversely correlated to the RELIABILITY of measure – its consistency across trials. Thus, reducing the error of measurement yields a better overall estimate of the participant's true score. The most common estimate of scale reliability in applied research is Cronbach's alpha (α; Cronbach, 1951). CRONBACH'S ALPHA provides an estimate for the lower bound of reliability for a scale evaluating the internal consistency of a set of items (Bandalos, 2018, p. 152). INTERNALLY CONSISTENT in CTT implies a set of items sharing strong linear relationships on average. So following the latent variable's logic, the indicators assumably prompt towards a latent dimension and can thus infer its presence.

The actual inference, however, must seem like a stroke of genius. Recall the intricate problem in social science; there is only one set of observable indicators (X) to access the true score (T) of the latent dimension (ζ). From this follow, the relation between the true score and the observed indicators is unknown: $T \overset{?}{\sim} X$. Cronbach's alpha bridges the gap relying upon the average inter-item correlation ($\bar{\rho}$) to provide an *estimate* for the reliability of a scale[3]:

$$\alpha = \frac{n\bar{\rho}}{1 + \bar{\rho}(n-1)} \tag{3.4}$$

So falling back on internal consistency, Cronbach's alpha provides a gateway to scale reliability (Bandalos, 2018, pp. 163). It is thus often used as an empirical argument to justify a latent dimension.

Nevertheless, in over over four decades, a vast body of evidence has grown in applied research, questioning the use and showing the problems of Cronbach's alpha (Cole et al., 2007; Fleishman & Benson, 1987; Green et al., 1977; Green & Yang, 2009; Hayes & Coutts, 2020; McNeish, 2018; Revelle & Zinbarg, 2009; Sijtsma, 2009; Ten Berge & Sočan, 2004; Trizano-Hermosilla & Alvarado, 2016). Sijtsma (2009), for example, takes the view of many of his colleagues, denying Cronbach's alpha to be an appropriate lower bound and disclaims its ability to estimate reliability (see also Barbaranelli et al., 2015; Revelle & Zinbarg, 2009; Sijtsma, 2009; Sijtsma & Van Der Ark, 2015). Cho and Kim (2015) aptly summarizes the problem with alpha: it is well known but poorly understood. Following Trizano-Hermosilla and Alvarado (2016), the era of Cronbach's alpha is finally over. There are more decent alternatives, Hayes and Coutts (2020) state, which easily overcome its flaws and provide a more adequate means to evaluate the reliability of a scale. Ultimately, McNeish (2018) boils down the zeitgeist among experts on using Cronbach's alpha for reliability assessment in one statement: "Thanks alpha, we'll make it from here".

3.1.2 Decent alternatives

More decent alternatives to Cronbach's alpha discussed in the literature are the greatest lower bound (*glb*; P. H. Jackson and Agunwamba, 1977) or coefficient omega (ω; McDonald, 1978, 1999). Both methods have been shown to outperform Cronbach's alpha (Sijtsma, 2009; Trizano-Hermosilla & Alvarado, 2016). However, following Revelle and Zinbarg (2009) recommendation of the greatest lower bound seems untenable, since coefficient omega has shown to triumph over the greatest lower bound in direct comparison. Consequently, the focus will be devoted to coefficient omega. Coefficient omega is actually a bundle, comprising two components ω_t (McDonald, 1978) and ω_h (McDonald, 1999). Please note, both estimates actually require prior knowledge on factor analysis, so the explanation falls briefly and may require a reread.

For now, think of a factor (ξ) as an empirical argument of an unobserved variable (ζ) including information on multiple observed indicators ($X_{i=1,\ldots,m}$). For the sake of reliability analysis, ω_h is more relevant than ω_t, because it depends on the sum of the squared loadings ($\Lambda\Lambda^t$) of the dominant factor (ξ) – not all factors (Revelle, 2021a). So ω_h focuses on how much of the common variation in the data is actually absorbed by a single factor. This is important because, if a single factor predominates the items, they share high inter-item relationships, which can be attributed to the influence of a single latent dimension. But care must be taken when applying this strategy because, as Mair (2018) notes, the concept is ultimately too weak to assess unidimensionality since the allocation (i. e., general and sub-factors) is neither encompassing nor sufficient.

3.1.3 Research recommendations

Since many studies have proved coefficient omega's superiority over Cronbach's alpha, one should switch to coefficient omega. Although a more recent investigation of Hayes and Coutts (2020) could testify its benefits over alpha, the authors note that coefficient omega is not flawless. The finding, however, should not surprise. Even though it is a more precise estimate of reliability, it is still approximate. However, as a generalization of Cronbach's alpha, coefficient omega includes the classical measurement as a special case, which is the reason for its applicability across a wider range of circumstances. For the background of empirical evidence, Hayes and Coutts (2020) finally agree with previous investigations of their colleagues and bemoan the absence of coefficient omega in commercial software. So even though coefficient omega is not absolute, it appears to be the better the default.

3.1.4 Additionals

Despite scale reliability, test-theoreticians evaluate the reliability of items themselves (DeVellis, 2006). CTT's strategy is to assess the item's correlation to the total score. The concept is called ITEM-TO-TOTAL CORRELATION (ρ_{TX}) and can be derived following the idea of internal consistent inter-item correlations:

Internally consistent items are on average strongly associated with one another. Thus, if an item fits into the cluster, it its high inter-item correlations (on average). So assessing the indicator itself, one can evaluate how well an item performs on compared to all the other variables on an emerging scale. That is, each item is contrasted with the total (i. e., "item-to-total"). If each item is reasonable to exert same influence (i. e., items have equal weight), the total is usually the bundle's sum ($T = \sum X_i$) or its mean ($T = 1/n \sum X_i$). Finally, when calculating the item-to-total correlation, the item is generally excluded from the total score (T_{X/x_j}), to improve the estimate. Technically, a part-whole correction is used to eliminate variance inflation in the result (McIver & Carmines, 1981, p. 102). This defines the so-called CORRECTED ITEM-TO-TOTAL CORRELATION (ρ^*_{TX}) which is predominantly used in applied research to select appropriate items (I. H. Hwang, 1970).

3.2 Measurement

There are many ways to develop a unidimensional measurement instrument. The Guttman approach (Guttman, 1944), or Georg Rasch's procedure (Rasch, 1960) are just two examples. Besides, Rennis Likert's summated scaling method (Likert, 1932) is predominant among classical test theoreticians. Since exploratory Likert scaling (ELS) draws on Likert's approach (and multidimensionality analysis can also be conducted using it[4]), focus will be devoted to this method in particular.

The Likert-approach makes at least three important assumptions (McIver & Carmines, 1981; van Alphen et al., 1994): (1) The assumption of equal weights across indicators since all items are viewed as interchangeable units from an infinite population of items – the "item universe", (2) the assumption of a single latent dimension underlying the response patterns, and (3) that trace-lines are monotonic; so an increase in the underlying latent dimension is assumed to increase a participant's general tendency for item agreement. All assumptions can be examined with the aforementioned methods. A good example of the congruence between the correlation approach of CTT and Likert scaling is the corrected item-to-total correlation. Why? Because any item is evaluated in terms of its compatibility with the total score of the emerging scale. So on the go, items have ever since been chosen to optimize this criterion.

Lastly, a participant's value needs to be assigned to pin down his or her location on the underlying latent dimensions. In particular, the previously developed scale is used in this undertaking. Because the measurement instrument passed for former quality checks (i. e., reliability analysis), classical test theoreticians judge the so-developed scale to be sufficient for this purpose. Following the aforementioned development process, lastly a sum or mean score locates participants on the underlying latent dimension.

3.3 Large data set challenge

Unidimensionality analysis is appealing as it contributes to clarity and unequivocalness by fostering understanding. The researcher is capable to evaluate a particular set of items and select among those being reasonable to capture a *single* latent dimension – assessing their correlation structure. Consequently, unidimensionality analysis is an indispensable tool for theory development (DeVellis, 2017; Hattie, 1985; Klein et al., 2014; McIver & Carmines, 1981; Shively, 2017; Zeller & Carmines, 2009).

However, in applied research, the setting is often vastly different. Usually, a large sample from a population of content-related items awaits the researcher, which is not expected to include a single latent dimension. A survey is a good example. A researcher can still try to inspect the correlation matrix going over combinations of items to assess their structure. But with an increasing number, the ordinary approach will soon converge towards its limits. Although improvements in visualization allow adjustments of overall performance, eventually one has to give up on optimization and hand the work over to an algorithmic procedure. This is the LARGE DATA SET CHALLENGE the ordinary unidimensional approach is helpless against. Simulated data in Figure 3.2 visualize the problem and show the difficulty of detecting corresponding item patterns. There, the limit of visual inspection becomes obvious.

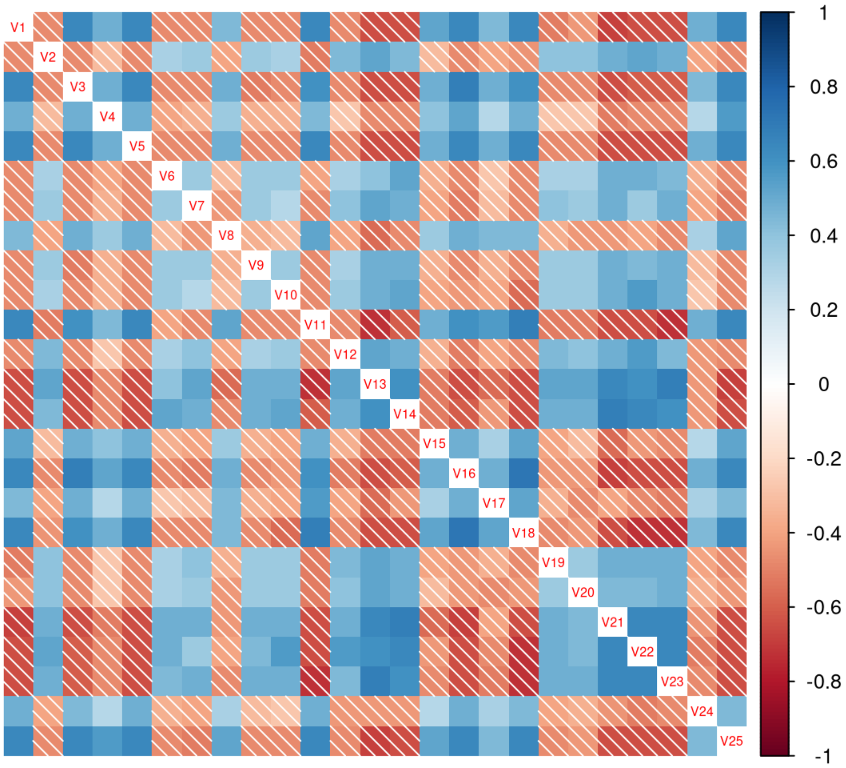

FIGURE 3.2: Large-Data-Set Challenge — 25 simulated variables and the difficulty of identifying highly correlated item clusters visually

In the following, two algorithms will be presented that take up the large data set challenge. One is a "dimensional extension" of the unidimensional framework, called "multiple-unidimensionality". In research practice, however, its multidimensional competitor, exploratory factor analysis, leads the field.

3.3.1 Example code

The data set used for the following examples are a subset of the Big Five personality traits (in particular: extraversion). It comes from the openpsychonometrics website. 19719 participants answered 10 questions, like "I am the life of the party" on a five-point Likert scale ([1]: Very Accurate, . . . , [5] Very Inaccurate). For the sake of understanding and transparency, some noise ($m = 2$) is added using `fabricatr` (Blair et al., 2021), and the item references (i. e., column names) are kept. The starting point is the correlation matrix (R), which is made up using the item data set (X) and visualized through the eponymous function from the `corrplot` package (Wei & Simko, 2021). To perform a reliability analysis in CTT, one can use `alpha()` and `omega()` from the `psych` package (Revelle, 2021b). If researcher aims for taking into account the ordinal nature of the data, they can stick with an adapted version of coefficient alpha developed Likert scales (Zumbo et al., 2007). `ordinal_alpha()` is part of the `jogRu`package (Gründl, 2021).

Data preparation

```
# Load the data
url <- "https://quantdev.ssri.psu.edu/sites/qdev/files/dataBIG5.csv"
d <- read.csv(url, header=TRUE)
# Exclude sociodemographics
X <- d[, -(1:7)]
# 0 is NA
X[X == 0] <- NA
# Extract extraversion
X <- X[, grep("E", colnames(X))]
\begin{verbnobox}

\subsubsection*{Add noise}

\begin{verbnobox}[\tiny\arabic{VerbboxLineNo}\small\hspace{3ex}]
# Create two noisy Likert scale variables
len <- length(X$E7)
# Pattern to define a 5-Point Likert scale
# Note: Use a normal distribution (UVA)
breaks <- c(-Inf, -1.5, -0.5, 0.5, 1.5, Inf)
# Note: noise come from a normal distribution
set.seed(1234)
noise_1 <- fabricatr::draw_ordered(x = rnorm(len), breaks = breaks)
noise_2 <- fabricatr::draw_ordered(x = rnorm(len), breaks = breaks)
# Assemble the items in a data set
items <- data.frame(X, N1 = noise_1, N2 = noise_2)
```

Unidimensionality analysis

```
# Correlation matrix (RA)
R <- cor(items, use="pairwise.complete.obs")
# visualization of the correlation matrix
```

```
corrplot::corrplot(R,  method='square', type = 'upper',
                     diag = FALSE)
```

Reliabilty analysis

```
# Tag the reversed-scored
keys <- c(1, -1, 1, -1, 1, -1, 1, -1, 1, -1, -1, 1)
# reliability analysis
uda <- psych::alpha(items, keys = keys)
# Total statistics
uda$total
# Item statistics
uda$item.stats
# Calculate coefficient omega for the item bundle
psych::omega(items[, -c(11,12)])
# Propose a item bundle
extra <- items[, -c(11,12)]
# Additional: calculate a Cronbach's alpha for Likert scales
# Note: Extension of CTT (Underlying Variable Approach)
# jogRu::ordinal_alpha(extraversion)
```

Measurement

```
# Sumscore
sumscore <- rowSums(extra, na.rm = TRUE)
# Meanscore
meanscore <- rowMeans(extra, na.rm = TRUE )
```

4 Multidimensionality analysis

When an area of content-related items is pervaded by more than a single latent dimension, one speaks of MULTIDIMENSIONALITY. Multidimensionality expands the concept of unidimensionality literally to multiple dimensions. This is done, assuming all (sub-)dimensions act simultaneously to produce a particular structure (Jacoby, 1991, pp. 36). The so-called ASSUMPTION OF EQUAL CONTRIBUTIONS will be of much greater interest in the next chapter. For now, the question is how many dimensions underlay a particular set of content-related indicators. Unlike the unidimensional case where justifying the absence or presence of a single dimension was prior, now multiple scenarios are plausible (e. g., $k = 1, 2, 3, \dots, K$).

To pick up on common ground, let's view how different multidimensional scenarios translate to the correlation matrix (R). Remember why this is important; a correlation matrix is a place where a latent dimension leaves its mark. Hence, any tool for latent-variable inference extracts empirical arguments from parts of the correlation matrix. Cronbach's alpha, for example, draws heavily on the average inter-item correlation. In the following, factor analysis will even try to reproduce common parts of it. The correlation matrix is fundamental and understanding its patterns is essential to grasp the flaws and limits of the tools falling back on it.

With this in mind, recall the unidimensional pattern from chapter 3. A single underlying latent dimension (ζ) implies a consistently high correlation among the indicators (X). So NON-DIMENSIONALITY is the opposite, consistently low correlations among the indicators (Figure 4.1). In an extreme case, the correlation matrix equals the identity matrix: $R = I$. The identity matrix (I) is a matrix with ones in the on-diagonals and zeros in the off-diagonals. This is important to know, because if the correlation matrix equals the identity matrix ($R = I$), items do not correlate with each other ($\rho_{i,j} = 0$).

Despite the unidimensional case, the correlation matrix can hold multiple highly correlating item bundles: $k = 1, 2, \dots, K$. The pattern is shown in Figure 4.2 and

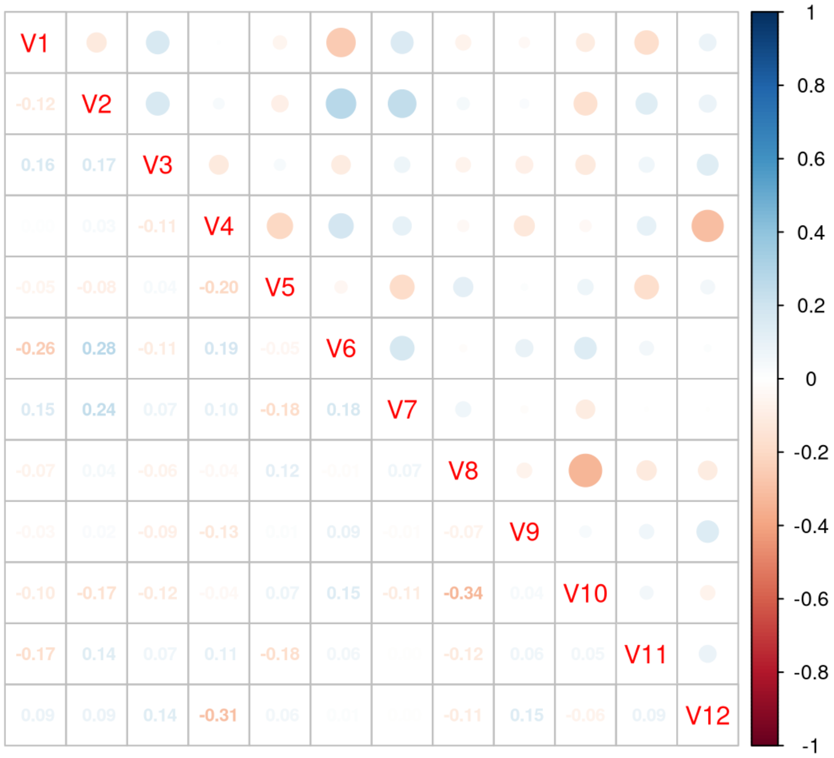

Figure 4.1: Non-Dimensional Item Response Pattern – 12 simulated variables sharing few attributes. The item-factor correlation is $\lambda_{i=1,\dots,12} = 0.1$

a mix of the aforementioned. Assembling multiple unidimensional patterns, filling the gaps with the ones previously shown, the correlation matrix is characterized by high within-group correlations and low between-group correlations:

$$\rho_{x_{i|k}, x_{j|k}} \gg \rho_{x_{i|k}, x_{j|l}} \tag{4.1}$$

The three item response patterns outlined so far cover a wide of circumstances in applied research. But especially the last one ($\rho_{x_{i|k}, x_{j|k}} \gg \rho_{x_{i|k}, x_{j|l}}$) will be of great importance in the upcoming chapters. It determines the ease of latent-variable inference in large data sets: *The more a structure looses the in-between-group pattern, the harder to identify the clusters and the harder to infer a latent dimension*. Why? Well, assume the contrary; if the between-group variation decreases, clusters resemble one another and become more equal. This means, their boundaries dissolve until the pattern vanishes:

$$\rho_{x_{i|k}, x_{j|k}} \approx \rho_{x_{i|k}, x_{j|l}} \tag{4.2}$$

The problem with such a pattern is that neither the robot nor the researcher can attribute given influences to a particular latent dimension. This hinders inference, because clusters cannot be meaningfully distinguished.

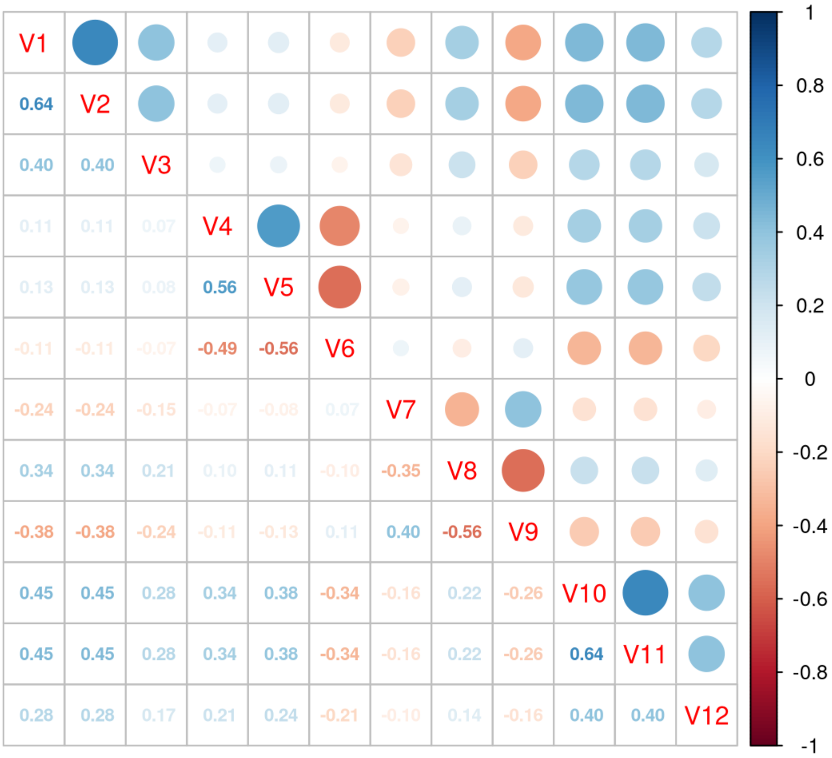

FIGURE 4.2: Between-Within-Group Pattern – 12 simulated variables with average loading $\bar{\lambda} = 0.7$ and average between-factor correlation $\bar{\phi} = 0.45$

Now recall the OUI's established characteristics. This so-called VANISHING EFFECT is exactly how the blurry boundary feature translates to the correlation matrix. Since latent dimensions often have common attributes, they share information, and the influence on an indicator cannot be uniquely assigned to any of them in particular. The last statement substantiates why, from an empirical point of view, complex social phenomena are hard to separate. Due to their characteristic features, they produce structural patterns which can be hardly distinguished. Those, in turn, hamper the possibility to tell latent dimensions apart. However, a method trying to master both obstacles (i. e., vanishing effect and large data set challenge) is factor analysis [5].

4.1 Factor analysis

Demarcating factor analysis (FA) mainly forces to separate it from principal component analysis (PCA). PCA is mainly used for descriptive purposes and metric data. It is superior when aiming to minimize the loss of information, compressing a set of variables ($X_{i=1,\dots,m}$). But in the social sciences, focus of the investigation is merely always on meaningful compression rather than full compression (Revelle, 2021a). This is why social scientists are nearly always concerned with COMPRESSION IN MEANINGFUL WAYS; that is to compress and infer latent dimensions. The principle will be expanded in the researcher's perspective on FA. For now, let it just be said, the common factor model is the formal basis of FA and will be subsequently specified in the fundamental theorem. It is the more suitable option for the undertaking, which is why the focus will be fully on FA[6].

Factor analysis comes in two major flavors: confirmatory and exploratory. The CONFIRMATORY version (CFA) is used to test theory-implied relationships about the common factor model. If, for example, the culture of honor phenomenon is assumable a construct of four sub-dimensions (e. g., masculine honor, feminine honor, social honor, and family honor; see Mosquera et al., 2008; Souza et al., 2017), researchers can test this proposed structure against data. One can also say, the researcher tries to *confirm* a proposed structure with the data at hand[7]. In this thesis, however, the exploratory version takes center stage.

EXPLORATORY FACTOR ANALYSIS (EFA) can be intuitively understood as determining the number of factors in an "exploratory manner" (Mair, 2018). The starting point of any EFA is the correlation matrix (R). Even though improvements in visualization allow to survey the structure of the correlation matrix, most of the time researchers draw on algorithmic procedures to meet the large data set challenge. Nowadays, the most comprehensively implemented algorithmic procedure is factor analysis. Since researchers usually hand work over to an algorithm, the analogy of a robot is used in the following. So think of factor analysis as commissioning robots. In particular, the robot's task is to help the researcher inspecting the correlation matrix since his or her capabilities of visual inspections are often quickly reached. To tune the search procedure, researchers direct the robot with a particular set of instructions. Some of them will be discussed in the chapter and checked for compliance with the OUI in the social sciences. But first, let's carve out the researcher's and the robot's perspective on factor analysis.

4.1.1 Researcher's perspective

To get into factor analysis, it is helpful to understand its aim in the social sciences. In applied social research, FA is mainly used to infer latent dimensions (ζ). This should remind the reader of the latent variable logic from chapter 2. However, now the view will be more specific on factor analysis and how the procedure contributes to this goal.

Which role does the factor-analytic approach play in latent-variable inference? Well, FA tries to provide the researcher with an empirical argument for (at least) one latent dimension. Replicating the common information of the correlation matrix, the key argument is a reduced set of synthetic variables (ξ) called FACTORS. In other words, they summarize the monotonic inter-item relationships in a condensed factorial form. The job of the researcher is now to infer a latent dimension from factors. This is done by assumption. Assuming that the latent dimension causes a given pattern in the correlation matrix, most of the inter-item correlations are assumably not random, but results from a common generative process – the latent dimension. Hence, each factor stores some of the common features in the correlation matrix, which labels the factor a parsimonious meaningful summary of the observed response patterns. Conversely, a specific pattern is often enough for latent variable inference – assuming the presence of the latent dimension.

Moreover, latent-variable inference using FA allows graduating beyond an indicator (e. g., "You would praise a man that acts aggressively in an insult") towards an encompassing social construct (e. g., honor). Thus, the result of an FA (ξ) will be said to be theoretically more important than the input product (X). In conclusion, *factor analysis compresses the input meaningfully to exceed its meaning in the result*:

$$X_{i=1,\dots,m} \xrightarrow{comp.} \xi_{k:k\ll m} \xrightarrow{exce.} \zeta \tag{4.3}$$

This is at the heart of LATENT-VARIABLE INFERENCE. In the light of the previous finding, factor analysis is indeed a latent variable approach in compliance with the principle of compression in meaningful ways.

4.1.2 Robot's perspective

The last section outlined the process of utilizing factors for latent-variable inference from a researcher's perspective. However, there is one crucial point missing – the robot's perspective. Recall, researchers usually use algorithmically driven factorial robots to search of highly correlating item groups. Accordingly, researchers have to understand how to commission their robot for a proper search.

Being able to supervise a robot and critically evaluate the quality of its results is key in applied research, as Gregory (2014, p. 165) emphasizes:

> "Unfortunately, factor analysis is frequently misunderstood and often misused. Some researchers appear to use factor analysis as a kind of divining rod, hoping to find gold hidden underneath tons of dirt. But there is nothing magical about the technique. No amount of statistical analysis can rescue data based on trivial, irrelevant, or haphazard measures. If there is no gold to be found, then none will be found; factor analysis is not alchemy. Factor analysis will yield meaningful results only when the research was meaningful to begin with."

The quote prompts towards the GIGO perspective on factor analysis – garbage in, garbage out. Researchers need to think intensively about the quality of their data to produce feasible inferences. Factor analysis cannot and won't protect inferences from poor data. Apparently, Gregory (2014) nudges to understand FA to get the magic out of the procedure. His statement goes hand in hand with the aforementioned perspective on dimensionality analysis. Namely, to grasp the problem and evaluate one's defaults to alleviate bias. With this in mind, let's get into the nuts and bolts of factor analysis, understanding the procedure by taking the robot's view on FA:

$$P \rightarrow R \tag{4.4}$$

The formula is nothing more than a formal instruction to tackle above's inference problem. Factor analytic robots ask how to smoothly replicate the common information in the correlation matrix (R) using a model (P) that owns a reduced set of synthetic variables. A suitable example to solve the problem is called the COMMON FACTOR MODEL (CFM). CFM is specified in the so-called FUNDAMENTAL EQUATION of factor analysis – and merely a reformulation of the input equation $X_i = \Lambda\xi + \epsilon$ Mair (2018):

$$P = \Lambda\Phi\Lambda^t + \Psi \tag{4.5}$$

Let's take this one step at a time, trying to understand both equations by designating the known bits. View both equations simply as (1) a formal outline of components (e. g., P, Λ, X_i, ξ) which are (2) available to the robot to (3) tackle above's replication issue ($X_{i=1} \rightarrow \xi_k \rightarrow \zeta$) from (4) a specific point of view ($P \rightarrow R$). With this in mind, look at both equations:

First, in the input equation ($X_i = \Lambda\xi + \epsilon$) one can surmise the vector of observable indicators (X_i), which stores participant's responses (e. g., ""strongly agree,..., strongly disagree") to questions, like "You would praise a man that acts aggressively to an insult". Additionally, "honor" is part of ξ – the vector of factors. Second, Λ is of interest. It is the matrix of factor loadings, administering the strength of relationship between factors and items. To demystify this terminology, instead of saying an item "loads on", one can say "correlates with" a given factor (Gregory, 2014, p. 159). In this sense, a loading ($\lambda_i \in \Lambda$) of let's say $\lambda_i = 0.7$, implies that the item "You would praise a man that reacts aggressively to an insult" correlates highly with "honor". If the item does not strongly prompt towards any other factor, it can be tagged as a good indicator for the latent dimension. In other words, "honor" manifests strongly in the item "You would praise a man that reacts aggressively to an insult", because both correlate strongly ($\lambda_i = \rho_{X_i,\xi_j}$). Since the matrix of factor loadings furthermore includes all item-factor relations it will prove as a handy tool in later evaluations.

To summarize, factor analysis offers a concrete formal solution for the intricate problem in the social sciences. Indeed, it extends the correlation approach from chapter 3 to meet the large data set challenge. Why? Because the common factor model puts the researcher in the position of using observable indicators to get a vital clue for latent dimensions – factors. The inference equation can thus be re-specified in light of previous findings:

$$X_{i=1,\ldots,m} \xrightarrow{P} \xi_{k:k\ll m} \xrightarrow{Asm.} \zeta \tag{4.6}$$

Put into words, using the common factor model (P) the data (X_i) are meaningfully compressed and become factors (ξ) which are assumed ($Asm.$) to be empirical indications for the latent dimension (ζ). Now, there is evidence on how to solve the intricate problem within FA. In this light, it is worth shifting attention towards factor analysis. Specifically, the upcoming sections will lie out how the researcher-robot duo tackles the intricate problem in the social sciences. As a result, answer will be given on how to combat the problem of identifying latent dimensions. But before one must face three major problems.

4.1.3 Three major problems

Tackling the intricate problem within FA implies wrangling the three major problems. First, one has to overcome the communality problem and choose an appropriate method to extract factors. Then, there is the rotation problem, which imposes a burden of choosing between dozens of implementations. Another battle in social sciences is the choice between orthogonal rotation and oblique transformations. However, the toughest nut to crack is the number of factor problem, which literally requires answering the tricky question of how many factors to retain. Along the way, many other obstacles, like sample size and interpretation, await the researcher. Especially the latter hurdle applies to theory development because chosen indicators have to be meaningfully upgraded or inferred. All of this will become clear throughout the chapter.

Nonetheless, under the aforementioned outline (exploratory) factor analysis first seems like a hurdle race and the question pops up why FA is the go-to method for exploratory investigations in the social sciences. But undoubtedly it is; and it has proved widely applicable across multiple circumstances in applied research (Brillinger et al., 2004; Dwivedi et al., 2006; Majors & Sedlacek, 2001; Tang, 2000) and disciplines (Cox & Dale, 2001; Pitombo et al., 2011; Williams et al., 2010; Yang et al., 2003). Thus, reasons must be found to justify its dominance.

A trivial one may be the absence of decent alternatives. chapter 5 and 6 deal with this consideration. But if decent alternatives are indeed missing, researchers should clear the hurdles. That is, familiarizing themselves with well-suited solutions for the aforementioned obstacles. Conversely, FA can be an optimal method to learn from sources of incomplete information – data. The last-mentioned perspective is taken on now. Furthermore, over four decades of applied research offer a rich pool of critics and alternatives to cope with each aforesaid obstacle. But concerning the OUI, skepticism must be held high along the way, and researchers need to be critical of any proposed solutions. Carefully each procedure must be assessed to prevent method blindness fallacy, and avoid ignorance bias.

Communality problem

The communality problem results from the common factor model's indeterminacy. More precisely, the problem arises from the indeterminacy of the fundamental equation, which defines the model. An equation is said to be INDETERMINATE when it holds more unknowns than knowns. To give an example, the equation, $2 + y = x$ is indeterminate. It holds more unknowns (i. e., x and y) than knowns (i. e., 2). The problem with an indeterminate model is it cannot be solved. One can also say, there is no closed-form solution, but a range of possibilities (e. g., $x = 2 \vee 3, y = 0 \vee 1$).

The common factor model, as specified by the fundamental equation, contains more unknown bits than knowns and is thus indeterminate (Schönemann & Steiger, 1978; Schönemann & Wang, 1972). The known bits are the observed item responses on the indicators, which are embedded in the correlation matrix (R). The unknown bits are the matrix of factor loadings and the communalities. Both will be defined and elaborated in the following sections. But for now, just note and memorize two things: First, following the stated definition, the common factor model is indeterminate. Second, the listed components (e. g., factor loadings and communalities) are the ingredients of the common factor model (P), which is capable to tackle the replication issue ($P \rightarrow R$). But so far, no answer was given on why indeterminacy is problematic. Mair (2018, p. 24) puts the communality issue in a nutshell:

> "The problem is that in order to estimate the communalities, we need the loadings. Conversely, in order to estimate the loadings, we need the communalities."

Mair (2018) proposes to *estimate* the solution; and an estimator will be shown which yields the most plausible values to reproduce the values in the correlation matrix. But an alternative strategy has to be discussed in advance. It is still possible to solve any indeterminate equation – namely by assumption. If a researcher sets $y = 3$, in the indeterminate equation $2 + y = x$ the equation can be solved. Assuming $y = 3$, the closed-form solution is 5. This is a toy example, but it suffices to get the problem and options to solve the upcoming extraction problem. Let's dive deeper into the topic, starting with the robot's perspective.

Robot's perspective As Mair (2018) notes, solving the communality issue means estimating an indeterminate common factor model. As he said, two components are of great importance: the matrix of factor loadings (Λ) and, in addition, the communalities. The COMMUNALITIES are the first part at the right-hand side of the fundamental equation ($P = \Lambda\Phi\Lambda^t + \Psi$) and define the exploratory power of the common factor model. In addition, there is another component, Ψ, which is called UNIQUENESS. The uniqueness captures the unexplainable bits – more precisely, item-specific variation and measurement error. The uniqueness thus represents information leftovers in the replication process.

Let's resume with the problem; since the common factor model is indeterminate, the robot cannot simply solve the fundamental equation. Focusing on the definition of the communalities while recalling Mair (2018)'s outline, the problem becomes obvious: The loadings are an inevitable (Λ) part of the communalities ($\Lambda\Phi\Lambda^t$). To provide a result, the robot has two strategies: (1) keep and estimate or (2) simplify and solve. So what to do? In the first case, a factorial robot estimates the most plausible values for the loading matrix ($\hat{\Lambda}$) and the communalities ($\hat{\Lambda}\Phi\hat{\Lambda}^t$) given the data. In the second case, it neglects between-factor correlations (Φ) to "determinate" the equation to return a result. In this case, the communalities simplify to $\Lambda\Lambda^t$[8]. The two strategies are what the robot offers – choosing one of them is a matter for the researcher.

Researcher's perspective Above's discussion identified the primary concern of the communality issue: the model's indeterminacy. Recall, indeterminacy is a property of an equation that applies to the fundamental equation and imposes the problem to calculate a model containing more unknowns than knowns. From a researcher's perspective, tackling the communality problem now requires "nothing more than" selecting one of the aforementioned strategies under which a solution will be provided. The first approach simplifies the fundamental equation by assumption (i. e., neglecting between-factor correlations). In the second approach, the researcher commissions the robot to estimate the most plausible solution given the data.

The problem seems easy to overcome, but over four decades of researchers proved the communality issue to be a complicated problem. Care must be taken when choosing a method to extract factors because the second approach smuggles (factor) independence into the analysis. But assuming independence is problematic since the assumption is incompatible with the OUI in the social sciences. Moreover, assuming independence predetermines the result at an early stage of the investigation. Researchers actually deprive themselves of the possibility to find out about factor independence. In any case, independence is involved. So all proposed solutions must be critically assessed and the reader is advised to watch out for method-blindness fallacy and expect ignorance bias.

Solutions The following two important approaches in applied research are outlined: (1) principal axis factor analysis (PAFA) and (2) maximum likelihood factor analysis (MLFA). PAFA will be a relevant topic of discussion for a single reason. It uses the principal component method (PCM) to provide a closed-form solution for the aforementioned problem under the assumption of factorial independence Literally, that's what it takes to *solve* the fundamental equation. PAFA is the classic way of extracting factors[9]. It dominated the research landscape in the 1980s, before more accurate and resource-hungry methods like MLFA received entry in the social sciences (see e. g., Überla, 1971).

Bad defaults The crux of PRINCIPAL AXIS FACTOR ANALYSIS (PAFA) is the idea of maximizing "information reconnaissance". Thereby, factors are extracted successively, keeping as much structural information from a reduced correlation matrix (R^*) as possible. For the sake of completeness, the REDUCED MATRIX is just the ordinary correlation matrix (R) in which values in the diagonals are replaced with communalities (Revelle, 2021a). If one sets the communalities to 1, the reduced matrix equals the original matrix ($R = R^*$)[10].

However, more important for the following discussion is the extraction procedure: the PRINCIPAL COMPONENT METHOD (PCM). Think of PCM as drawing out liquid with syringes from a glass of water. In the first step, one absorbs as much of the water (i. e., information) from the glass (i. e., reduced matrix) as possible. So the first syringe (i. e., first factor) comprises as much water from the glass (i. e., information from the reduced correlation matrix) as possible. After extracting the first syringe, the second one (i. e., factor two) absorbs as much of the remaining water drops from the glass as possible. Thus, the second factor processes the left-out information of factor one. This procedure is repeated until k filled syringes are on the table and no more water (i.e., information) is in the glass (i. e., reduced matrix). Now it is obvious why the method can be said to minimize the loss in compression. It maximizes the information density (i. e., water) within each factor (i. e., syringe)[11].

However, concerning the OUI, the procedure is problematic. It produces independent factors (Rummel, 1967). Notice that the second syringe does not contain any drop of water which is already present in syringe one – vice versa. Thus, as the absorbed information is independent, so are the factors[12].

But usually, there is no reason for the researcher to assume independence between factors. This holds especially for the default in an early stage of an exploratory investigation. Remember, the OUI tends towards blurry boundaries. Latent dimensions usually share some bits of information among factors. Referring to examples like "honor" and "pride", or "nationalism" and "patriotism", factorial independence is an exertion rather than a rule. Forcing factors to be independent often disrespects the characteristic features of the OUI in the social sciences. Ruthless application of this procedure is not harmless. It distorts results and can prove as a recipe for doom in applied research (Hayton et al., 2004; Taherdoost et al., 2014). In the end, it is reasonable to label it a bad default and its widespread use reveals a method-blindness fallacy that furthermore imposes ignorance bias in applied social science research (Heise, 1973). This urges switching to decent alternatives.

Decent alternatives The decent alternative mentioned hereafter is Maximum Likelihood Factor Analysis (MLFA). This is a vastly different and a more complicated approach to solve the communality problem. Here, the aim is to provide a conceptual understanding rather than fleshing out the mathematical nuances. This section is outsourced[13].

Conceptually, MLFA is pretty tangible. With the data at hand, the robot tries to estimate the most plausible (i. e., the most likely) values that are capable to optimally reproduce the common information in the (reduced) correlation matrix. In a more approach-based description, one can also say it solves the problem by maximizing the likelihood of the data (Revelle, 2021a). The robot is furthermore commissioned under the assumption of multivariate normality. So if the assumption of multivariate normal-distributed residuals is reasonable, maximum likelihood estimates are an attractive solution for the reproduction problem.

Research recommendations Überla (1971) already mentioned the principal component method as a go-to method in applied research. He also noted some of its flaws, especially the problem of independence. Almost twenty years later, however, Hayton et al. (2004) and Taherdoost et al. (2014) found the problem of independence still a topical subject in applied work. Researchers still resort to methods, inducing independence by default. The widespread use of bad defaults indicate method-blindness fallacy, imposing ignorance bias. Again, in many cases, the OUI is incompatible with the assumption. Similarly, it distorts result. In sum, PCM predetermines the results way too in an exploratory investigation. Its use should be strongly restricted, plus omitted by default. In this light, MLFA is the recommended way to start the theory development process.

Notwithstanding, PCM should not be considered generally “bad”. A bad default default does not imply the method is itself worthless. But given what we know about the OUI, it is too rigorous to start with. However, as part of model comparison, it can be one of the most valuable practical strategies to learn from data. By contrasting models with varying basic assumptions and predicting how they will behave often generates greater model and data insights. Thus, using

PCM in model comparison can prove as a valuable source of information in the learning process[14].

In the end, even if the assumption of independence may not be verifiable, learning about the model and understanding how it sees the data helps to grasp flaws and better understand its results. As most researchers use algorithmic procedures, understanding the robot is an essential prerequisite to gain access to the OUI. Beyond that, understanding is probably the best tool to combat method-blindness fallacy in applied research and thus to alleviate ignorance bias.

Additionals A practical concern with maximum likelihood estimation is its tendency to run into convergence issues (Santos Silva & Tenreyro, 2011). Since most implementations do not allow incorporating prior information, weak loading patterns and low communalities are especially problematic with a low number of subjects within each factor. This is the so-called NUMBER OF SUBJECT PROBLEM (Revelle, 2021a). It involves the question of how many subjects are needed to get stable estimation results; but the last two decades of applied research missed providing an unequivocal answer (de Winter & Dodou, 2012; D. L. Jackson, 2001). However, maybe there is a single unsatisfying one-size-fits-all recommendation: More is always better. To be precise, first, researchers should get as many subjects as possible, preferably tweaking the subject-to-factor ratio (Costello & Osborne, 2005; Kline, 1994; MacCallum et al., 2001; Pearson & Mundfrom, 2010; Zhao, 2009). Second, they should get numerous "good" indicators, where, third, "good" means high communalities (Hogarty et al., 2005; Jung & Lee, 2011; MacCallum et al., 2001). This assessment may seem vague but comes as close as one can get to determine appropriate and applicable requirements for a wide range of factor analyses.

Nonetheless, from an applied researcher's perspective, the recommendations probably sound like statistical fairytales. Often researchers have no choice but to address problems with the data at hand. So how to deal with emerging (convergence) issues? An easy judgment rule for MLFA is to not trust a result that has not converged. Parameter estimates are too volatile. Under those circumstances, generalized ordinary least squares (OLS) estimation procedures (minimizing the error of reproduction) are good alternatives to the ML approach. The same is true if "normality of errors" is no reasonable assumption (Cudeck & MacCallum, 2012). OLS approaches mark a great place for the reader to restart (see, e. g. Mair, 2018; Revelle, 2021a). However, in the absence of convergence issues, when it is reasonable that errors follow a normal distribution, maximum likelihood estimation (or one of its robust extensions) is a good default.

Rotation problem

When extracting factors with MLFA, the robot delivers the most plausible values (given the data) for reproduction problem. Although the result often suffice robots, it is not so for researchers. Why? Even though the solution describes the association between factors and indicators ($\hat{\Lambda}$); the result is often not in a shape that can be interpreted straightaway. The loading picture is often too diffuse. Or, in other words, the loading matrix is not in a simple structure[15].

Transformation often overcomes the problem. A transformation of the loading matrix is typically referred to as ROTATION. Rotations are literally interpretation aids for researchers, trying to produce unambiguous loading pictures while maintaining the extracted relationships between variables in the result. Their outcome usually yields a clear overall picture of the factor-item correlations ($\hat{\Lambda}_r$). Ultimately, rotations foster clarity, by facilitating the broader understanding of the patterns themselves. A simpler (i. e., 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 4.3). For ease of sake, imagine the space is two-dimensional. With two dimensions, there are two axes X and Y. Re-label X and Y with the labels of two factors, 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 ($\hat{\lambda}_{honor}$; $\hat{\lambda}_{pride}$), one can say "You would praise a man who acts aggressively to insult" correlates moderately to highly with "honor" and highly with "pride". By generalizing this logic, the entire loading matrix ($\hat{\Lambda}$) can be mapped into a two-dimensional space using factor coordinates.

But does the item now relate to "honor" or "pride"? This is just it; since both values are approximately equal, one cannot easily say. As mentioned above, tackling the issue implies translating the intermediate into something simple. In

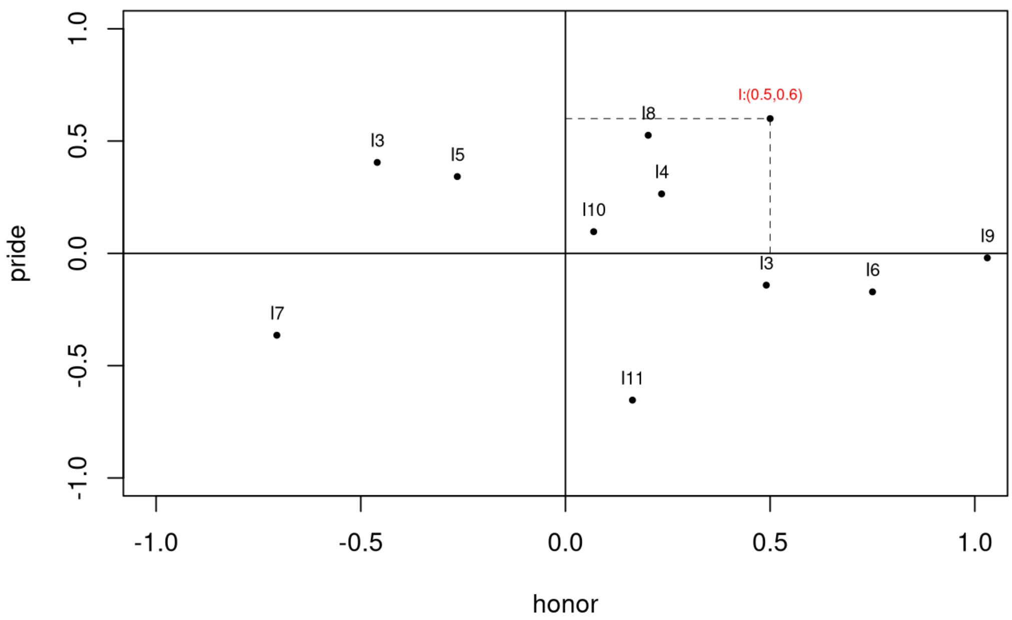

FIGURE 4.3: Factor Coordinate System – a visualization of a loading matrix. The loadings allow mapping the 10 simulated variables in the coordinate system. The average loading is $\bar{\lambda} = .55$ and the between-factor correlation $\phi_1 2 = 0.2$

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? There are mainly two choices: either to "spin axes", using orthogonal rotation techniques that maintain the right angle between axes; or to apply an oblique transformation, which allows the factor axis to move independently. Hence, in the latter, the factors can correlate with each other.

First, for understanding ORTHOGONAL ROTATION, the key question is how to reallocate the axis that a simple (i. e., easy-to-interpret) loading-structure results. Think of the solution in terms of a spinning rotor. The rotor represents the factor coordinate system. The bolt (i.e., the origin) holds the rotor in place so that the blades (i.e., axes) can spin. When the rotor blades turn, they maintain a fixed 90-degree angle. The same holds true for factor rotation. When spinning the axes in a factor coordinate system, the blades move counterclockwise, preserving the right angle between them. Mainly, one refers to the procedure as "orthogonal rotation".

The metaphor breaks, second, in the oblique case. Here, one does not spin the axes, preserving the right angle between. The crux is to explicitly allow the axis to abandon perpendicularity. Intuitively, the two-dimensional oblique case is better thought of as a clock (i. e., a factorial coordinate system) in which the hands (i. e., the factors) may move independently. Thus, the "time" becomes relevant, (i. e., the angle between both factors), because it provides additional information on between-factor correlations.

Next, it will be shown how graphical rotation builds up on matrix algebra. Indeed, a visual outline is helpful to start with, though impede with a full understanding of the comprehensive modeling approach. This holds especially true for modern oblique techniques. To distinguish the graphical and algebraic approach the term "oblique transformation" (Revelle, 2021a) will be uses afterwards. Nevertheless, the question remains; what happens inside the machinery? So let's change perspectives to get under the hood.

Robot's perspective As it will be shown, from a robot's point of view, rotation problems are a matter of solving the transformation equation under different constraints. Any solution, however, builds upon the TRANSFORMATION EQUATION, formalizing as:

$$\hat{\Lambda}_r = \hat{\Lambda} T \tag{4.7}$$

The first thing to notice is, once more, the equation is just the formalized versions of the aforesaid obstacle. Namely, finding a set of axes that produces a simpler loading structure ($\hat{\Lambda}_r$). More precisely, the solution just *appears* simpler (i. e., more accessible for the researcher), because the actual relationships between variables are left untouched. To put it another way, the robot searches a transformation matrix (T) that rearranges the estimated information ($\hat{\Lambda}$), but does not change the model's fit to the data (Mair, 2018). In this light, the formal equivalent of spinning the axis is a multiplication with a transformation matrix (T). As such, the transformation matrix is the pivot point in the above's equation and key to understanding transformation procedures in general. But what particularly happens inside the machinery?

In the orthogonal case, the plotted matrix of factor loadings ($\hat{\Lambda}$) is multiplied with a transformation matrix (T), which moves the axes counterclockwise until the desired result ($\hat{\Lambda}_r$) is reached. Thereby, the robot has to find a solution with two important properties: First, obviously, the transformation (T) should be able to solve the transformation problem ($\hat{\Lambda}_r = \hat{\Lambda}T$). But even more important, second, is the premise to preserve the right angle between factors. Mathematically, a corresponding solution can be achieved with a set of orthogonal vectors, because their orthogonality allows to maintain the right angle – rotating the factors (counterclockwise) by 90 degrees. This is done assuming $TT^t = I$[16].

In the oblique case, solving the rotation problem implies resolving the perpendicularity of the axes (algebraically, $I \neq TT^t$). By giving up on the assumption of factor orthogonality, one preserves and estimates 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[17].

Solutions As it has already been said; when solving the rotation problem, researchers have mainly two choices [18]: orthogonal rotations or oblique transformations. When it comes down to a decision, researchers have to evaluate if it is reasonable to ignore the between-factor correlations or if they need to be modeled explicitly. This is crucial to the OUI in the social sciences because the first option keeps perpendicularity, which is another appearance of factor independence. Thus, *the decision between orthogonal and oblique rotations has a much bigger effect on the OUI than choosing within the sphere of orthogonal or oblique criteria*. That is why the particular solutions are clustered with respect to their origins (i. e., orthogonal or oblique). Today, there are over two dozen rotation procedures (Gorsuch, 2015, pp. 214) – too many to discuss all of them. However, some of them (i. a., the bad default, and the most decent alternative) are elaborated in the following sections.

Bad defaults Among orthogonal rotation techniques like quartimax (Carroll, 1953), equamax (Saunders, 1962)[19], 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 (T) for the rotation problem $\hat{\Lambda}_r = \hat{\Lambda}T$, by maximizing the variance of the squared loadings for each factor (ξ). This means varimax is out to boost the (squared) correlations between factor and items ($\hat{\Lambda}\hat{\Lambda}^t$). In doing so, the transformation (T) makes the factor axis rotate by 90 degrees counterclockwise, preserving the right angle between the factor-vectors[20]. T spins the axes (counterclockwise), 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, it is useful to understand how orthogonality of factor-vectors relates to the assumption of independent factors. Mathematically, they are connected through the cosine. The cosine concentrates the 360-degree spectrum to a one-minus-one range. Therefore, the angle, which shows their relationship graphically, can be transferred in a well-known sphere of relatedness – their correlation. As a consequence, if factor-vectors are perpendicular, they should not correlate. And indeed, since cos(90) equals 0, they do not. Thus, maintaining the 90-degree angle while rotating preserves the right angle and consequently factor independence. A cursory note on previous findings was already given in varimax' 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 ($\hat{\Lambda}\hat{\Lambda}^t$) – but at the cost of omitting between factor correlations ($\Phi = I$). Correspondingly, the result is the same: factor independence.

But again, there is usually no reason for the researcher to assume independence between 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., blurry boundaries). Latent dimensions nearly always share some bits of information. Accessing them through independent factors is a recipe for doom, which furthermore leads to ignorance bias (Loo, 1979). This is in line with Sass and Schmitt (2010), who conducted a comparative investigation of different rotation criteria within EFA. They encourage researchers to critically reflect on their rotation choices since they have a significant impact on the manifestation of the factor structure. The habitual use of orthogonal rotations like varimax as the default behavior is not harmless. It is reasonably a bad default and can prove as a dangerous undertaking, especially if factors demand to correlate with each other, but chosen options hinder them to do so (Loo, 1979). That's the reason why switching to decent alternatives is mandatory.

Decent alternatives Recall, the main goal of rotation and transformation procedures. They aim for transferring the initial 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 all SIMPLICITY INDEXES aim to find a solution with each item indicating the least number of factors. In this sense, simplifying the result maximizes interpretability[21].

Usage of the most popular oblique transformation procedures, like oblimn and promax (Hendrickson & White, 1964) can be justified by scoring well on one of above's criteria. Ultimately, all solutions aim to maximize factor-simplicity and avoid factor independence. So all of them are indeed decent alternatives. But when it comes down to find the most decent alternative, oblmin, as well as promax, do not provide the simplest result possible. Indeed, they all score high on some indexes; but those are the flawed ones (Lorenzo-Seva, 2003)[22]. According to Lorenzo-Seva (2003)'s more accurate loading simplicity (LS) index, simplimax (Kiers, 1994) is a decent choice, since it outperforms more elaborate versions of oblimin and promax[23].

In the end, by scoring high on the LS index, the simplimax algorithm delivers a transformation matrix (T) pooling the communality of each item on the fewest number of factors possible. This is why the rotated loading pattern ($\hat{\Lambda}_r$) is as simple as possible – factor-loadings ($\hat{\lambda}_{i=1,\dots,m}$) are either zero or as far from zero as possible. In conclusion, simplimax is the most decent alternative. 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 To emphasize it once more, *concerning the OUI in the social sciences, choosing between orthogonal an oblique technique is weightier than choosing a particular criterion within each framework*. As it was argued, deciding for oblique transformations is a better default. 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.

Orthogonal rotations are ever a common choice in the social sciences. Thus, smuggling in additional information (i. e., factor independence) in an exploratory investigation turns out to be a serious 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 – given background knowledge – 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). Thus, it is difficult to quantify, how many published research papers are flawed by the implicit assumption of orthogonality[24].

For this background, it seems questionable why researchers such as Bortz and Schuster (2016) are concerned with a loss in compression when going oblique. Again, the major goal in the social sciences is to compress meaningfully, not maximally. So even though some informational redundancy is induced, allowing the factors to correlate with each other; it may prevent the researcher to cope with highly distorted results (Sass & Schmitt, 2010).

But despite their shortcomings, orthogonal rotation methods should not be excluded 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 (Φ) by default and learn about them on the go. This 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 still find the simplest result possible. But researchers must avoid the current practice of finding simplicity at any cost.

Orthogonal rotations are not worthless in applied social-science research. One can use them strategically. For example, to highlight the added value of resolving the assumption of independence. Research should become experimental, trying different techniques and evaluate 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. Those kinds of information should not be omitted, suppressed or abandoned on a preliminary ground. It should be reported to improve domain knowledge[25].

Additionals There might be scenarios in which orthogonal solutions are a reasonable choice. Thus the small additive provides some guidance on how to use orthogonal rotation techniques in applied research. 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, 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 one 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, authors like Gorsuch (2015) recommend varimax, because in visual inspections, they produce interpretable results and prove invariant across a wide range of circumstances.

The second addition 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). This is in line with Gorsuch (2015, p. 224) who states that any attainable 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 true, when it comes down to determine *the* number of factors.

Number of factor problem

Conceptually the easiest and at the same time the most striking problem in factor analysis is the number of factor problem. The obstacle is literally, to find the appropriate number of factors to extract. It is so striking because misspecification distorts results (Velicer et al., 2000). Two scenarios proved troublesome in applied research: (1) UNDEREXTRACTION – pulling out too few factors and (2) OVEREXTRACTION – pulling out too many of them.

Fava and Velicer (1996), for instance, show underextraction evoking biases when inferring results from distorted patterns ($\hat{\Lambda}$). Furthermore, pulling out too few factors leads to OMITTED FACTOR BIAS. Hayton et al. (2004) argue omitting factors induces a loss of information, which compares to a loss in theoretical significance. But typically the most common decision-making tools in applied research tend to overextract, they add. Therefore, pulling out too many factors seems far more common in applied research.

Overextraction is problematic because overshooting the number of factors introduces spurious ones, leaving researchers with over-saturated models. As a result, model outcomes are often too complicated to understand (Fabrigar et al., 1999; Wood et al., 1996; Zwick & Velicer, 1982, 1986a). Although overextraction is more common in applied research, both misspecifications blur insights and impedes gaining a true understanding of underlying structures. Distorted solutions furthermore hamper the ability to interpret results meaningfully and thus mislead latent-variable inference. This is why a lot of research has gone into solving the number of factors problem. The outcome has been a bunch of RETENTION CRITERIA proposing a plausible range for the number of factors to extract. With a focus on bad defaults choices and decent alternatives, some of them will be presented after shedding light on the researcher's and robot's perspective of the problem.

Researcher's perspective Rephrasing above's problem description, the number of factors can be said to affect the accuracy to reproduce the information in the correlation matrix directly. Researchers find themselves in a SIMPLICITY-ACCURACY TRAP: Whereas a smaller number of factors contributes to parsimony, a larger number of factors contributes to a more accurate description of the phenomenon of interest (Revelle, 2021a). At maximum, the number of parameters ($|\xi|$) and indicators ($|X|$) are equal and the information in the model-implied correlation matrix (P) is simply a "factorial re-description" of the given structure in the input correlation matrix (R): $|\xi| = |X| \Rightarrow P = R$. In other words, there is no compression but a bunch of factors mimicking indicators to re-describe the structure. But if researchers seek latent-variable inference, they need to compress meaningfully. So the key question is how to navigate between simplicity and accuracy?

Even though advice is hard to give advice, probably *truth* lies somewhere in between a full model and a null model. Researchers may avoid "structural mimicry" by choosing the full model, allowing to transmit all structural information, but at the cost of simply re-describing what is already known. At the same time, they may question the use of overly simple models, because large data sets are usually pervaded by – often even more than two – latent dimensions.

Ultimately, researchers must realize that *samples are flawed, incomplete representations of true patterns and data generating processes* – latent dimensions. Why? Because then it might become obvious that not every piece of information in the correlation matrix is worth extracting. Researchers should aim for the common features – the ones which can be found across samples. From this point of view, FULL ENCRYPTION of the sample distorts results by learning too much from the sample, whereas NULL-ENCRYPTION distorts results by ignoring relevant features of the data (McElreath, 2020, p. 192). Researchers must navigate between the two extremes. Thus, from their perspective, the number of factor problem is rather an optimization problem. Researchers have to find the minimum number of common factors decrypting the maximum amount of information.

Robot's perspective From a robot's perspective, the researcher's problem (i.e., to find the minimum number of common factors decrypting the maximum amount of information) reads as follows:

$$\min p : P \to R \tag{4.8}$$

The first thing to grasp is that the number of factor problem is almost independent of the factor analytic robot. Ordinary FA does not solve for the minimum number of factors (p), so to say[26]. In the following, it will become clear; the information on how many factors to consider mostly flows from external sources into the analysis. This is where statistical and mathematical criteria set in.

However, at first, the finding clarifies the role of the robot in FA. It delivers a result (P) to reproduce the common information in the correlation matrix (R) only given pre-specified input (i). Correspondingly, it is the researcher proposing a possible number of factors (p) to retain. The robot only delivers. It is just a servant inspecting the correlation matrix under some given set of constraints (e. g., $p = i$). As a consequence, the formalization of the problem re-renders as:

$$R \to P \,|\, p = i \tag{4.9}$$

Against the backdrop, FA follows a master-servant model in which the appropriate number of factors is found in a trial-and-error approach. Researchers must realize their important position in factor analysis. They are masters, instructing their robotic servants as follows: (1) The researcher requests a result (P) with a particular value (i) for the number of factors to extract (p), (2) the robot delivers the requested result ($P|i$). (3) The researcher evaluates the outcome ($P|i$). If (4) the researcher approves the solution ($p = i$) the search is over; else (4=1) they insert another possible number of factors to extract ($p = j | j \neq i$) and restart the process.

Under the aforementioned process, researchers can come close to an appropriate number of factors to extract. They iterate through possible solutions manually, following the logic; if there is insufficient accuracy in the result, another factor should increase the goodness of approximation (Gorsuch, 2015, p. 152). So researchers increase the number of factors successively, improving the model's goodness of fit: $p_i > p_j \rightarrow Pi > Pj$. The ideal place to stop is when improvements become negligible. If this point is reached, they solved the optimization problem optimally; finding the most parsimonious and at the same time most sufficient result possible ($\min p$).

Even though the outlined strategy solves the parsimony-accuracy trap theoretically, executing the strategy becomes a Herculean task – especially for increasingly large data matrices. One has to request and evaluate $p = m$ possible solutions. Thus, pragmatically seen, some guidance proves helpful. This is where mathematical and statistical criteria come in. They help to find a plausible *range* of possible solutions for the number of factors to extract. Focusing on bad defaults and decent alternatives, some of them are presented right away.

Solutions There are numerous implemented solutions to solve the number-of-factor problem (see e. g., Revelle, 2021a; Timmerman et al., 2017). Revelle (2021a) tells the anecdote of Henry Kaiser who invented *a* criterion every morning before breakfast; knowing that the problem is one of inventing criteria for *the* solution. On that account, the focus will be shifted to only a subset of criteria shown to provide sufficient results.

Bad defaults The most common ad hoc criteria used in applied research are the scree test (Cattell, 1966) and the eigenvalues-greater-than-one (K1) criterion [@Kaiser1960]. Both methods propose a cut-point for the number of factors to retain and are still today's defaults in statistical software like SPSS. As such, they are common go-to methods in applied research (Hayton et al., 2004; Taherdoost et al., 2014).

K1 The idea of K1 (Kaiser, 1960) is to extract principal components as long as each of them pools together more information than a single item in the data set. More technically, it stops iterative extraction of further components[27] only if their variance – the sum of the squared factor loadings ($\Lambda\Lambda^t$) – falls below the variance of a standardized item. That's where the "eigenvalue-greater than-one criterion" gets its name from; eigenvalues of standardized items have value "1". Correspondingly, K1 can be said to extract principal factors as long as one of them exceeds an eigenvalue of 1, too.

Reviewing past literature[28], the downfall of K1 can be justified based on its erratic and abysmal suggestions (Lee & Comrey, 1979; Tucker et al., 1969). More particular, it showed mixed tendencies to underextract and overextract (Linn, 1968; Yeomans & Golder, 1982). Its poor performance is often attributed to small sample sizes per factor (Browne, 1968) and low communalities (Hakstian & Cattell, 1982; Zwick & Velicer, 1982, 1986a). However, Revelle and Rocklin (1979) as well as Velicer and Jackson (1990) underline K1's tendency to overextract. Revelle (2021a) furthermore proposes to divide K1's suggestion by three, since results turn out to be more realistic with the adjustment.

Besides its tendency to overextract, K1 suggests a number of factors upon a model including the assumption of independence (Lorenzo-Seva et al., 2011; Timmerman et al., 2017). As a consequence, researchers should expect deviating results if they set up a common factor model under those constraints. But in absence of systematic investigations, it remains unclear if (and how) the assumption of independence leads to actual differences in recommendations for the number of factors to retain. Theoretically, however, deviations might arise if the assumption of independence does not hold. Why? Because, a suggestion based on a model resting on independence can differ vastly from a model which resolves the assumption explicitly. These deviations, in turn, should influence the result (p) – which is the suggestion for the number of factors to retain[29].

In sum, K1 turn out to be a bad default for applied social science research. Using it irresponsible as a common go-to approach to determine the number of factor witnesses a method blindness fallacy which is reasonable to induce ignorance bias.

Scree test In Cattell's scree test (Cattell, 1966) researchers examine the scree plot to determine the number of factors to retain. The SCREE PLOT is a graphical visualization in which the eigenvalues (λ) of the correlation matrix (R) are sorted in descending order and plotted against an index ($I = 1, \ldots, m$). Thus, it is sometimes called an eigenvalue-index plot[30]. The result is shown in Figure 4.4. To test for the number of factors to keep, the researcher watches out to find the cliff which divides rock and scree. This means to examine the plot for a drop in the eigenvalues – the so-called "elbow" – allowing to differentiate between systematic and random bits of variation visually. The proposed number of factors to retain is given by the index-value *before* the line flattens (Gorsuch, 2015, p. 176).

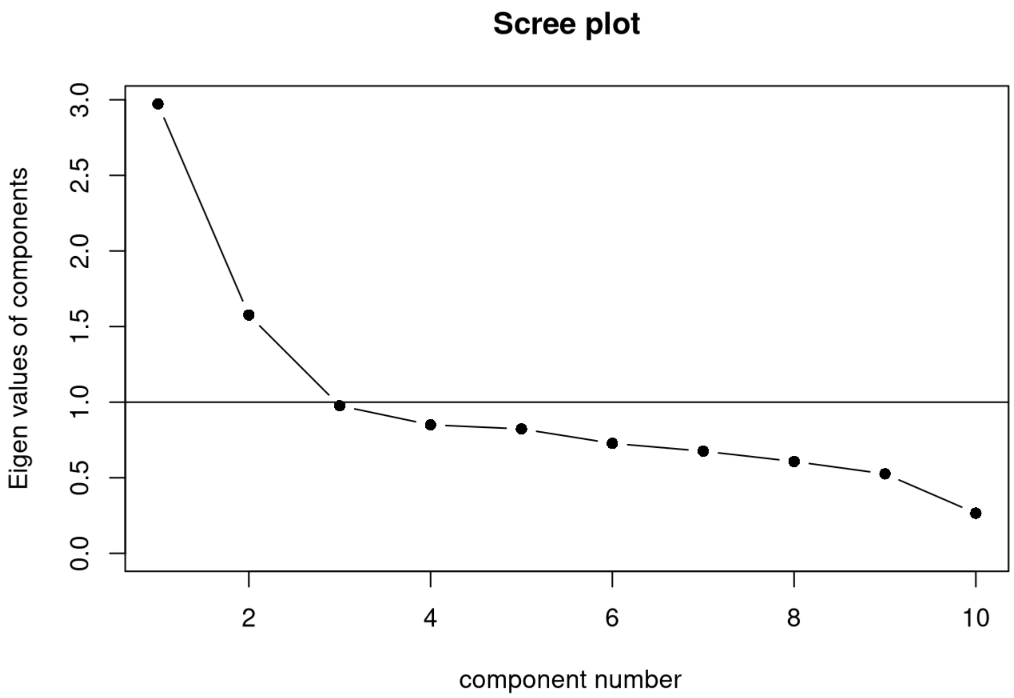

FIGURE 4.4: Scree plot of 10 simulated variables with average loading $\bar{\Lambda} = .55$ and between-factor correlation $\phi_{12} = 0.2$. The horizontal line furthermore includes the suggestion of K1.)

From an applied researcher's perspective, a scree test proves troublesome if there is (1) either no cliff, allowing to tell apart the rock from the scree, or (2) if the eigenvalue-landscape looks like a mountain scenery, with multiple rocks, cliffs, and screes. This is in line with Hayton et al. (2004), who argues scree plots are manageable if there are powerful factors, but turn ambiguous in the absence of a clear drop-off or with multiple cut points. But if the number of observations and communality values increase, some ambiguity vanishes (Horn & Engstrom, 1979). However, the previous fact cannot outweigh another in complex scenarios. scree plots are often inconclusive; especially if different inspectors evaluate more complicated plots (Crawford & Koopman, 1979; Zwick & Velicer, 1986a). Thus, it is questionable if the scree tests allow for unity in the decision of how many factors to retain. Concerning the problem of misspecification Zwick and Velicer (1986a) further notice the scree test's bias towards overextraction. But because it is not directly influenced by the number of items, it misbehaves much less than K1 (Ledesma & Valero-Mora, 2007).

Nonetheless, the scree test also bases on eigenvalues, hence shares its flaws with K1. The problem manifests particularly in terms of the OUI. scree tests suggest their number of factors to retain based on a model assuming independence between factors. Although deviations are not empirically evident, one should expect them[31].

In conclusion, scree plots disqualify as default retention criterion. Their flaws label them suboptimal as a go-to choice, mainly because of their erratic and abysmal suggestions. Furthermore, they propose solutions based on a model, which is incompatible with the OUI's characteristic features. Their frequent use in the social sciences testifies existing method blindness and validates ignorance bias in applied research.

Decent alternatives In applied science, probably the greatest achievement in the last three decades was the rapidly growing use of simulation-based approaches. Concerning factor analysis, parallel analysis (Horn, 1965) is a good example. Parallel analysis (PA) is an upgrade of Kaiser's idea, incorporating sampling variation through a random process (Auerswald & Moshagen, 2019; Saccenti & Timmerman, 2017). A basic PA comprises 4 steps (Hayton et al., 2004): First, it produces (K) random duplicates of the original data set $(X_{n \times m})$. Second, PA determines the random eigenvalues $(\lambda^*_{j=1...J})$ from the (K) random duplicates $(X^*_{k=1...K})$. Third, it calculates the mean and, more recently, the 95th-percentiles (Glorfeld, 1995), of all eigenvalues $(\lambda^*_{j=1...J})$. Four, it compares the eigenvalues from the original data set $(\lambda_{j=1...J})$ with the (implied boundaries) of the random eigenvalues $(\lambda^*_{j=1...J})$. Finally, factors are rejected if their eigenvalues are smaller than the simulated ones.

About possible flaws, (Turner, 1998) noted the tendency to underextraction. However, the more recently overall performance of PA (with 95th-percentiles) has shown to be good (Dinno, 2009; Lorenzo-Seva et al., 2011; Peres-Neto et al., 2005; Steger, 2006). This is in line with, Velicer et al. (2000) who replicated the results of Zwick and Velicer (1986a) in a comprehensive simulation study, ordinary PA does comparatively well. But recall that PA is an extension of Kaiser's method. As such, it is a PCA-based approach, even though it is commonly used to determine the number of factors Timmerman et al. (2017).

More recent developments try to overcome previous issues and provide more widely applicable and more accurate parallel analyzes in applied research. Using PA-MRFA (Timmerman & Lorenzo-Seva, 2011), for instance, one tries to get two birds with one stone, (a) identifying the number of common factors in (b) a series

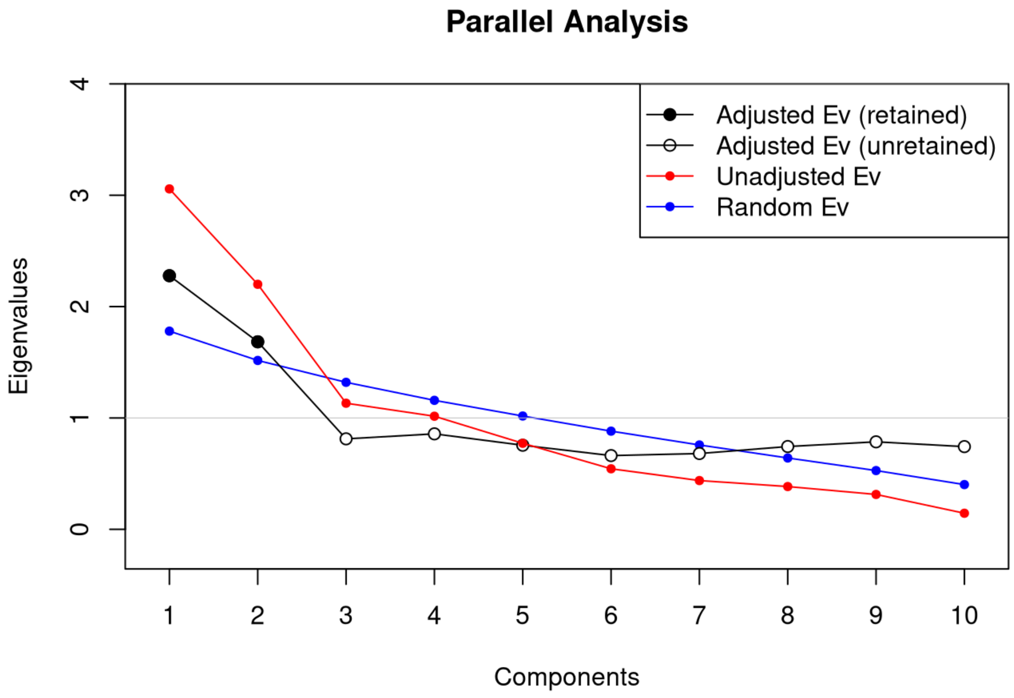

FIGURE 4.5: Parallel Analysis – trying to decide how much factors to retain, based upon 10 simulated variables with average loading $\bar{\Lambda} = .55$ and between-factor correlation $\phi_1 2 = 0.2$. The blue line show the simulated eigenvalues.

of polytomous items. Instead of the common PCM approaches (i.e., PCA and PAFA), PA-MRFA draws on minimum rank factor analysis (MRFA) (Shapiro & Ten Berge, 2002). In a nutshell, the procedure tries to find the number of factors minimizing the unexplained bits of the common variance. More precisely, PA-MRFA does not fiddle with eigenvalues (λ, λ^*). It draws on more meaningful comparisons focusing on the explained common variance portion of random and empirical data (Timmerman & Lorenzo-Seva, 2011).

To additionally include the ordinal nature of most single stimulus data (e. g., 1: Fully disagree, . . . , 5: Fully disagree) polychoric correlations can be used with PA and FA. Think of polychoric correlations as bespoke correlation-measure for polytomous variables. They promise more accurate results (Timmerman & Lorenzo-Seva, 2011); but assume an underlying normal distribution for the observed variables in the data set[32] (Jöreskog, 1978; Muthén, 1984). A strategy integrating all the aforesaid improvements is the hull method.

The hull method The hull method (Lorenzo-Seva et al., 2011) is a recent development to determine the number of common major factors in EFA. As

FIGURE 4.6: Hull Method – The FA results bases upon 10 simulated variables with average loading $\bar{\lambda} = .55$ and between-factor correlation $\phi_1 2 = 0.2$.)

such, it is a bespoke alternative to the aforementioned PCA-based approaches. It builds upon a numerical model selection approach based on a convex hull (Ceulemans & Kiers, 2006), trying to pin down the number of factors by numerical inspection of a scree-like plot (i. e., X: degrees of freedom, Y: goodness-of-fit measure, Figure 4.6). Relying on calculation instead of visual inspection, the hull solution can overcome subjectivity in the choice of how many factors to keep. To cope with overextraction, it sets an upper bound calculated in an ordinary parallel analysis plus one additional factor (Lorenzo-Seva et al., 2011). The "numerical elbow" is calculated at the point where increasing the number of factors leads only to a marginal, local increase in the goodness of fit. The approach is thus an elegant mathematical solution to navigate between simplicity and accuracy.

The hull method incorporates different goodness-of-fit measures, defining a particular hull variant. According to (Lorenzo-Seva et al., 2011) most successful (with average hit rates over 80-90%), are (1) Hull-CFI, which bases on the comparative fit index (Bentler, 1977) as well (2) Hull-CAF, which builds upon the Kaiser-Meyer-Olin index (Kaiser, 1974). In combination with maximum-likelihood (or unweighted-least-squares) extraction procedures, Hull-CFI should be the method of choice in applied research (Timmerman et al., 2017). A recent Monte Carlo investigation furthermore testified to previous findings, and proved it to be a handy retention criterion in a researcher's toolbox (Auerswald & Moshagen, 2019).

A concluding remark, Lorenzo-Seva et al. (2011) notes the hull method to perform best with a sufficient number of indicators per factor ($m_{\xi_i} > 5$). If this is not possible, ordinary parallel analysis with a 95th-percentile threshold (Glorfeld, 1995) is a decent alternative. Besides, (Auerswald & Moshagen, 2019) found high hit rates in orthogonal unidimensional scenarios, regardless of the sample size.

Sequential χ^2-Test A well-suited test for MLFA integrated into most statistical software is the LIKELIHOOD RATIO TEST (LRT). The LRT evaluates the fit of the model under the null hypothesis of a model implied correlation matrix resembling the population correlation matrix (Auerswald & Moshagen, 2019). Starting from a null model, the test is repeated, increasing the number of parameters successively. In the limits ($n \to \infty$) the test statistic furthermore follows a χ^2 distribution (Auerswald & Moshagen, 2019), which is why it is called SEQUENTIAL χ^2 MODEL TEST (SMT; Auerswald and Moshagen, 2019). At some point, SMT will yield an insignificant result ($p < 0.05$). If so, increasing the number of parameter stops and the model which produced the last significant result finally contains the number of factors to retain.

However, one should also familiarize with SMT's flaws. Replicating the results of earlier investigations, Ruscio and Roche (2012) show its tendency to overextract. Besides, a more recent study of Green et al. (2015) also shows its tendency to underextract. Thus, recent findings are inconclusive. However, in the latest systematic investigation of Auerswald and Moshagen (2019), SMT outperforms other approaches as the hull method – especially in situations with correlated factors. Based on their Monte Carlo study Auerswald and Moshagen (2019) promote the use of SMT in applied research. Another minor peculiarity is SMT's arbitrary 5% threshold. But the comment seems negligible given its verifiable performance under those constraints. Despite that fact, no serious violations were found concerning the OUI's characteristic features. In conclusion, the sequential model test proves a decent alternative in comparison to more rigid retention criteria.

Research Recommendations Let's start with the non-recommended bits. Today, one can to hence should avoid K1. Patil et al. (2008), for instance, is very clear about it, demanding to abandon the Kaiser criterion in theory development. Velicer et al. (2000) follow the claim and additionally request to eliminate K1 as the default option in commercial stats software. Preacher and MacCallum (2003) looking forward to this decision, because they continuously try "[r]epairing Tom Swift's electric factor analysis machine". More particularly, (Revelle, 2021a) criticizes statistical software like SPSS, because it still relies heavily on outdated criteria. From today's view, the use of K1 can be mainly justified on historical reasons (see e. g., Überla, 1971, p. 92). Back in the 1960s, more elaborate and accurate approaches were computationally far too expensive. Today, with an average of 3 to 4 cores and modern i5-processors, K1 can be seen as a hero of a past war. The veteran should be praised with honor, but its time is over.

However, neither the scree test, PA, the hull method nor SMT proved faultless. Consequently, no one can be unreservedly recommended, especially if the researcher seeks out for determining the "correct" number of factors. But this is just it. Often the discussion is lead as if there is a one-and-only solution. Proposing multiple possible propositions is scarce in applied research. Timmerman et al. (2017), as one of a few, demonstrated more than one model to be compatible with the data. So, in general, there is nothing wrong with disagreement between criteria. Disagreement must be seen as an invitation to poke inside and understand model differences. Maybe there is more to it than misspecification. Consequently, model comparison is, once more, the recommended way to learn about models and data.

To summarize, although neither method was sufficient under any circumstances, some tools (e. g., hull method and PA-MRFA) proved more useful than others (e. g., K1 and scree test). Ultimately, there is not a single correct way to determine the number of factors. Nevertheless, there is a well-adapted opportunity to get the most out of any criteria, compensating for its weaknesses through another one. The gambit is called COMBINATION RULES (Auerswald & Moshagen, 2019; Timmerman et al., 2017). By combining different methods, researchers try to find an optimal composition of retention criteria, providing high hit rates across multiple scenarios. Auerswald and Moshagen (2019), for example, found SMT and Hull to form a good duo in applied research. With respect to Lorenzo-Seva et al. (2011)'s investigations, however, one should set up a trio. PA-MRFA is a bespoke complement fitting well into the common-factor-model framework.

Additionals A problem researchers will encounter with PA-MRFA is convergence issues. The recommended way to treat them is, again, skepticism. When parameter values haven't been set, the results are highly volatile. Researchers should trust none of them. Bumping up the number of iterations and re-run often helps; but if not, Timmerman et al. (2017) recommend switching back to Pearson correlations with a mean threshold.

Besides the aforementioned, there are a couple of alternative approaches which weren't mentioned in the previous section. The very simple structure criterion (VSS; Revelle and Rocklin, 1979), Revered parallel analysis (R-PA; Green et al., 2012), the empirical Kaiser criterion (Braeken & Van Assen, 2017), and comparison data (CD; Ruscio and Roche, 2012 are just a few examples. All seem promising. Therefore, they will tag a great place for the interested reader to restart. The only outlined as additional hereafter is VSS.

Very simple structure criterion In a nutshell, the *Very Simple Structure Crtierion* evaluates the loss in quality of a factor solution when it is "degraded" to the simple structure researchers assume to be there (Revelle & Rocklin, 1979). To put it another way, think of VSS as a retention criterion basing upon the "Pippi-Longstocking principle"; researchers create the loading matrix, just the way it suits them. The procedure (deliberately) fades out minor loadings – the ones which fall below a certain limit – to construe loading matrices as simplified versions of themselves. The simplicity of any given solution is regulated by the so-called "complexity" (C). The COMPLEXITY is the number of non-zero loadings per factor:

$$C = j \Leftrightarrow \xi : \exists \lambda_{i=1,\dots,j} \tag{4.10}$$

By manipulating it, researchers can literally design their desired solution, plugging in the number of non-zero loadings per factor (i. e., how simple the result should be). VSS then searches for a simplified model with the given constraints: First, it contains a pre-specified number of zero loadings which are, second, able to reproduce the original information in the correlation matrix. Consequently, the approach is practically useful. Why? Because by monitoring the simplified model's fit, researchers can adapt to a solution that tweaks simplicity and accuracy at the same time.

4.2 Factor scores (Measurement)

The last step in a dimensionality analysis is to determine participant's values on the latent dimension(s). In factor analysis, those are called FACTOR SCORES. In advance of the upcoming sections, note that the focus of the thesis is on the exploratory part of factor analysis. So the topic will not be tackled with full strength. Grice (2001a) provides a neat introduction on how to evaluate and estimate factor scores. Additionally, DiStefano et al. (2009) and Grice (2001b) provide a comprehensive comparison of methods to calculate them.

There are at least two common approaches to assign factor scores ($\hat{F}$): (1) summated factor scores and (2) different regression and correlation-preserving approaches. First, recall Likert's method from the previous chapter. Ultimately, a sum score was used to assign values after items proved reliable, according to CTT. By repeating the procedure as many times as there are latent variables, the strategy can be adjusted accordingly. The result is usually a bunch of sum or mean scores, with equally weighted (sometimes standardized) items called UNITY WEIGHTS. However, the method neglects the estimated loadings completely. Omitting the loadings can be troublesome because items with relatively low loadings can partly gain significant weight. According to given circumstances, this might be a reasonable assumption, or it may not be.

In summary, Tabachnick and Fidell (2007) aptly call sum or mean scores the "quick and dirty" estimate of factor scores:

$$\hat{F} = \sum_{i=1}^{n} X_i \tag{4.11}$$

They are comparatively easy to calculate, although not the most we can get out of the analysis. (p. 650) Their advantage over regression estimates, for example, is to express variation in the result (DiStefano et al., 2009). But if one aims for a more accurate estimate of participants' true scores on a latent dimension, regression approaches and correlation-preserving methods have proved superior (Andersson & Yang-Wallentin, 2021). Hence their name – REFINED METHODS (DiStefano et al., 2009). Conceptually speaking, factor scores are produced by pushing the information (X) back out through the model using $\hat{\Lambda}$ to re-weight participant's scores ($\hat{F}$) on the latent dimension (ζ)[33]:

$$\hat{F} = f(\hat{P}) \tag{4.12}$$

One of the most elaborate approaches is the one of Ten Berge and Hofstee (1999). Leaving mathematical details aside for now, the approach is a well-suited method for applied social science research. Why? Because it provides a solution, taking into account the characteristic dependencies of the OUI. This is due to the fact that Ten Berge and Hofstee (1999)'s estimates are a generalized approach for oblique factors, preserving their between-factor correlations (Revelle, 2021a). In addition, the procedure readjusts each item according to its loadings, incorporating a weighting scheme that regulates item influence conditional on the data.

However, care must be taken, including factor scores in further analyzes (Mair, 2018). Recall; parameters cannot be estimated uniquely. Correspondingly, since the model cannot be uniquely defined in terms of its parameters, so are its predicted scores. The model's indeterminacy inherits to factor scores, so to speak. More precisely, the model's indeterminacy leads to a FACTOR-SCORE INDETERMINACY (Grice, 2001b). As a consequence, all approaches calculating factor scores actually produce estimates ($\hat{F}$) for participants' values on the latent dimension (Revelle, 2021a).

A shortcoming of regression approaches is their inherent complexity (Gorsuch, 2015, p. 283). Gorsuch (2015), for example, bemoans that all items are included in the solution to estimate factor scores. So there might be the risk of fall for irregular features of the sample. In factor analysis, the problem is referred to as CAPITALIZING ON CHANCE. Gorsuch (2015) suggests trimming the model's possibility to capitalize on chance, by cutting the number of included items assigning equal weights. But at the same time, the suggestion misses an explanation on how to select the most appropriate ones. For the background of previously mentioned findings (i. a., where regression approaches outperformed sum-scores; Andersson and Yang-Wallentin, 2021), the suggestions seem untenable. However, the problem of capitalizing on chance remains and seems not easily solvable. However, in a larger research context, the problem will be identified and diminished. In the end, the possibility to capitalize on chance is the trap-door of exploration and must be seen as a burden of learning from incomplete sources of variation – data. This is what Sir Karl Popper once noted, saying, "I may be wrong, and you may be right, and by an effort, we may get nearer to the truth".

4.3 Example code

The data set used is, again, the subset of the Big Five personality traits and come from the openpsychonometrics website. 19719 participants answered 50 questions, like “I am the life of the party” on a five-point Likert scale ([1]: Very Accurate, . . . , [5] Very Inaccurate). To address the number of factor problem, the hull method via `hullEFA()` and parallel analysis using minimum rank factor analysis via `parallel.MRFA()` was used from the `EFA.MRFA` package (Navarro-Gonzalez & Lorenzo-Seva, 2021). In addition, the sequential χ^2 model test `SMT()` from the `EFA.dimension` package was applied (O’Connor, 2021). To fit the models with different rotation or transformation methods, `fa` from the `psych` package (Revelle, 2021b) was deployed with maximum likelihood factor extraction. In addition, there is an example for the scree test and principal axes factor analysis.

Data preparation

```
# URL
url <- ‘‘https://quantdev.ssri.psu.edu/sites/qdev/files/dataBIG5.csv’’
d <- read.csv(url, header = TRUE)
# Exclude: sociodemographics
X <- d[, -(1:7)]
# 0 is NA
X[X==0] <- NA
# Complete case analysis (CC)
X_cc <- X[complete.cases(X),]
 # Correlation matrix
 R <- cor(X)
```

Multidimensionality analysis

```
# Polychoric correlation matrix
R_poly <- psych::polychoric(X)$rho

# Number of factor problem
# Eigenvalue criterions
psych::scree(R_poly)
# Hull method
EFA.MRFA::hullEFA(X_cc, maxQ=5, extr = ‘‘ML’’,
                  index_hull = ‘‘CAF’’,  display = TRUE,
                   graph = TRUE, details = TRUE)
 # Parallel analysis using minimum rank factor analysis
```

```
EFA.MRFA::parallelMRFA(X_cc, graph=TRUE)
# Sequential chi-sq test
EFA.dimensions::SMT(X, corkind=’’polychoric")
# Very simple structure
n.obs <- nrow(R_poly)
psych::vss(R_poly, fm = ‘‘ml’’, n.obs = n.obs)
# Rotation problem & communality problem
# Maximum likelihood factor analysis with varimax rotation
MLFA.vmax <- psych::fa(R_poly, nfactors = 2,
                      rotate = ‘‘varimax’’, fm = ‘‘ml’’)
# Maximum likelihood factor analysis with simplimax rotation
MLFA.simpl <- psych::fa(R_poly, nfactors = 2,
                      rotate = ‘‘simplimax’’, fm = ‘‘ml’’)
# Factor correlations
round(MLFA.simpl$Phi,digits=3)
# Loadings plot
# factor loadings
lambda.vmax <- MLFA.vmax$loadings
lambda.simpl <- MLFA.simpl$loadings
# Add labels
labels <- names(X)
plot(lambda.vmax, type=’’n")
text(lambda.vmax,labels=labels,cex=.5)
text(lambda.simpl, labels = labels, cex=.5, col="red")
abline(h=0, v=0, lty=2)
# Principal axis factor analysis
PAFA.varimax <- psych::fa(R_poly, nfactors = 2,
                      rotate = ‘‘varimax’’, fm = ‘‘pa’’)
```

Factor scores (Measurement)

```
F_tB <- psych::fa(X, nfactors = 2, rotate = "simplimax",
                  cor = "poly", fm = "ml",
                  scores = ‘‘tenBerge’’, missing = TRUE,
                  impute = ‘‘median’’)$scores
```

5 Multiple unidimensionality analysis

As introduced in the last chapter, if an area of contend-related items is pervaded by more than one latent dimension, one speaks of multidimensionality. Additionally, it was assumed that all underlying (sub-)dimensions act simultaneously to produce a particular outcome. When exploring large data sets, this assumption can be problematic. Jacoby (1991, p. 36) states:

> "Multidimensional solutions assume that all of the dimensions operate simultaneously in contributing to the observed differences between the scaled objects [...]. [But] these assumptions can be problematic. Even if a subset a of objects possess K 'objective' characteristics, there is no particular reason that *all* of the characteristics are used to differentiate among all of the objects."

Correspondingly, in a large data sets multiple separate dimensions (ζ, η) can be at work, where each dimension contributes to a particular subset of the data (van Schuur, 1989). For example:

$$\zeta \sim X_{1,\ldots,j} \wedge \eta \sim X_{j,\ldots,m} \tag{5.1}$$

In this case, one speaks of MULTIPLE UNIDIMENSIONAL REPRESENTATIONS (Jacoby, 1991) of the outcome because each particular subset of the data is associated with a distinct latent variable (ζ, η). In other words, there is some kind of cross-influence of separate latent dimensions.

But how does it relate to the previous discussion? Well, FA returns a multidimensional solution (Asún et al., 2016; Fischer, 1974). Hence, it includes the ASSUMPTION OF SIMULTANEOUS CONTRIBUTIONS. The critic from a multiple unidimensional point of view is that the assumption may be untenable. Regarding a multiple unidimensional pattern, there is an alternative explanation for the observed patterns in the correlation matrix, requiring fewer presuppositions. Therefore, the multiple unidimensional approach accuses FA of proving a bad (i. e., too pre-suppositional) default to meet the large data set challenge. Logically, one must search alternatives.

5.1 Exploratory Likert scaling

A decent alternative for the complex multidimensional factor analysis from a multiple unidimensional standpoint is EXPLORATORY LIKERT SCALING (ELS; Müller-Schneider, 2001). ELS is an implementation of the multiple unidimensional scaling approach, trying to overcome flaws of multidimensional investigations. In order to get potential benefits out of the framework, one needs to better understand the problem of multidimensionality and what tags it a bad default.

5.1.1 Bad defaults

To grasp the problem of multidimensionality and get the benefits of ELS , one needs to understand the appeal to unidimensionality. That is why unidimensionality is so precious for exploratory purposes and its key role in theory development – conceptual clarity. But unidimensionality is not just valuable for exploration; it is also an inevitable part of measurement (Hattie, 1985, p. 156). After the value of unidimensionality has become clear, it will be obvious why extending this approach seems to be a handy tool for the large data set challenge.

The appeal to unidimensionality

In exploration and theory development, there seems to be an ongoing appeal to unidimensionality (DeVellis, 2017; Klein et al., 2014; McIver & Carmines, 1981; Shively, 2017; Zeller & Carmines, 2009). One of the fundamental questions is how this appeal is justified, especially as some authors judge multidimensional techniques to be superior? (see e. g., M. Costa, 2003; Kruskal & Wish, 1978; Stone & Wright, 1999). The major reason is given in the ISOMORPHISM ARGUMENT (McIver & Carmines, 1981, pp. 12):

> “The final and most important reason why unidimensional scaling models continue to be of substantial interest is that they are isomorphic with the primary type of concepts devised by social scientists. Shively (1980), for example, has argued that social scientists should strive to develop and use unidimensional concepts because they are more susceptible to theory-relevant research (also see Clausen and

> Van horn, 1977). Multidimensional concepts, on the other hand, typically hamper such research because they are too ambiguous in terms of their meaning, too difficult to measure in a clear and precise manner, and too theoretically oriented themselves. Their complexity and ambiguity renders them less optimal for the development and assessment of social science theories. In other words, using unidimensional scaling models to measure unidimensional concepts puts the theory construction and the measurement strategy on the same analytical level."

McIver and Carmines (1981) advocate a PRINCIPLE OF UNEQUIVOCALNESS. Unidimensional constructs are practically more tangible; in fact, they are easier to understand. This may sound trivial at first, but it is what matters when exploring a series of items – to gain a true understanding of interrelations. Thus, understanding should be prior whenever exploratory purposes are present. In case no theoretical guidance is provided, unequivocal concepts and unidimensional scales could aid and foster the goal of explicit statements about an OUI. Due to their inherent complexity, multidimensional solutions seem to be a necessary downstream consideration for McIver and Carmines (1981) and their colleagues. They would value approaches like exploratory Likert scaling (ELS) and emphasize its use as a decent alternative over more complex multidimensional frameworks like EFA.

5.1.2 Decent alternatives

EXPLORATORY LIKERT SCALING (ELS; Müller-Schneider, 2001)[34] is a multiple unidimensional scaling approach and operates on a bottom-up item selection process that integrates characteristic values of classical test theory (i. e., Cronbach's alpha and the corrected item-to-total correlation). It is mainly used to discover internally consistent variable clusters; which means item bundles with a low between-group variance and a high within-group variance ($\rho_{x_{i|k},x_{j|k}} \gg \rho_{x_{i|k},x_{j|l}}$).

Robot's perspective

At the heart of ELS is the so-called CRYSTALLIZATION PRINCIPLE (Mokken, 1971), used to start and expand a scale bottom-up. First, the robot inspects the correlation matrix to find the highest positively correlating pair of items. By "crystallizing them", which means to summarize them particularly (i. e., using a sum-score or mean-score) the nucleus of the emerging Likert scale is generated.

Second, out of the remaining items, the robot picks the one, which correlates highest with the total score. If the item's (k) correlation with the sum score ($\rho_{T,X_{k\setminus\{i,j\}}}$) is larger than a pre-defined lower bound ($\min \rho_{T,X}$), the item is added to the emerging scale. This defines the INCLUSION CRITERION:

$$\min \rho_{T,X} < \rho_{T,X_{k\setminus\{i,j\}}} \Rightarrow k \in T' := \sum_{i=1}^{k} X_i \tag{5.2}$$

Lastly, three, the second step is repeated until the threshold of the lower bound is finally reached ($\min \rho_{T,X}$). If so, the first scale is completed. Looking past the formulas, the algorithm is just the automated version of the unidimensional scale-development process discussed in chapter 1.

However, to meet the large data set challenge, the algorithm needs to be extended, hence become a MULTIPLE SCALING APPROACH (Mokken, 1971). To adapt to it, the robot simply loops over the previous ones: build the core, extend it, finish the scale; *repeat*. This way, the robot successively construes more and more unidimensional Likert scales. Or, to put it another way, it finally assembles multiple unidimensional ones. Thus, the last step is just the instruction to repeat the process (i. e., steps 1 and 2) until either the lower bound is reached or the correlation matrix is fully partitioned into multiple unidimensional Likert scales. Speaking in an analogy, the robot just acts like the pigeons in the Cinderella movie. Compatible ones (i. e., items with a high correlation) go into pots (i. e., scales) bad ones (i. e. items with a low correlation) go into the crop. Repeat. The pigeons (i. e., algorithm) stop if they cannot compatible lenses (i. e., highly correlating items) because all of them have either been sorted or none of them meets the inclusion criterion.

One may have already guessed the problem with so-constructed scales. Each item is exclusively allocated to a single scale. More precisely, each item occurs in every scale exactly once. But with respect to the blurry boundaries feature, those kinds of scales are problematic. Remember; latent dimensions often share common attributes. ELS' solution is to let scales overlap. OVERLAP scaling bases on the principle to *reconsider* the use of every item in the correlation matrix, which is not already part of a disjoint scale. What that means, is the robot (1) takes a disjoint scale fragment, (2) checks which items it already includes, and (3) duplicates the data set - but omits the ones which are already part of the fragment. If the item meets the inclusion criterion, the robot (4) expands the scale. By applying the 4-point strategy to every disjoint scale, items can occur in more than one of them, because the inclusion criterion is often met multiple times.

In summary, the robot starts producing disjoint scales. But since they are incompatible with the OUI, the disjoint scales are viewed only as intermediates or SCALE FRAGMENTS. Completing them means expanding each fragment into an overlap. By reconsidering the use of every item, ELS completes not only a scale but furthermore aligns the result with the OUI's tendency towards blurry boundaries. The remaining question for the upcoming sections is how to set the above threshold ($\min \rho_{T,X}$).

Researcher's perspective

From the researcher's perspective, using ELS as the implementation of a multiple-unidimensional approach requires choosing two lower bounds. The first one is used to produce disjoint scale fragments ($\min \rho_{T,X}^{disj.}$). The second one is needed for the overlap ($\min \rho_{T,X}^{ovlp.}$). To foster the identification of homogeneous item clusters, there is a strategy on how to set these values. The first couple of runs are used to find the disjoint kernels with the highest internal consistency. The lower bound ($\min \rho_{T,X}^{disj.}$) is increased successively until a sufficient maximum is reached. The subsequent runs with the second lower bound ($\min \rho_{T,X}^{ovlp.}$) complement the maximal consistent item fragment, hence, it is an expansion process. Researchers should set the second boundary to the minimum threshold they are willing to tolerate for inclusion.

5.1.3 Research recommendations

So far, exploratory Likert scaling is neither widely recognized in the research community nor widely used. Thus, it is hard to give research recommendations based on solid empirical evidence or simulation studies. But the author gained some insights into ELS, implementing the according R package `elisr` (Bißantz, 2021) with the author of ELS. Thus, some restricted practical advice on how to set the lower bounds can be given. In practice, it has proven to be effective to raise the first lower bound ($\min \rho_{X,T}^{disj.}$) until the maximal number of scale fragments result. But determining the second lower bound ($\min \rho_{X,T}^{ovlp.}$) is much more difficult. Researchers should set the second boundary to the minimum threshold they are willing to tolerate for inclusion and report those values. Why? Because the values for those thresholds depend upon the particular context and setting. Up to now, there is only one clear guideline for the use of ELS. The second boundary is typically much lower than the first one:

$$\min \rho_{T,X}^{disj.} \gg \min \rho_{T,X}^{ovlp.} \tag{5.3}$$

In addition, it should be noted that using existing ELS variants might be only of limited applicability since they fully rely on CTT. From the perspective of Borsboom (2005, p. 32) the classical true score approach "[...] is like an alien in a B-Movie: no matter how hard you beat it up, it keeps coming back". Kohli et al. (2015) probably share Borsboom (2005)'s aversion, showing that only modern extensions of CTT – subsumed under the UNDERLYING VARIABLE APPROACH (UVA) (i. e., assuming underlying normally distributed variables) – perform well enough to compete against modern item response theory models. So at this point ELS (as well as 'elisr') is out-to-date. However, upcoming versions will expand the classical approach and include contemporary alternatives such as polychoric correlations and adapted versions of Cronbach's alpha tailored for Likert scales Zumbo et al. (2007).

5.2 Measurement

Measurement in ELS is conducted using the sum-score. This makes sense as scales are built by maximizing the item-to-total correlation, hence optimizing ELS' scaling criterion. But note that measurement and item selection in ELS are intertwined; they go hand in hand. Scales are finished on the run, so to speak. The items are detected, analyzed, and merged in one go. ELS is thus a comprehensive framework combining detection, item selection, and scale measurement in one application. If the researcher aims to, first, fully automate unidimensionality scale assessment to meet the large data set challenge and is, second, willing to trust in CTT as well as summated rating scales, ELS is the way to go.

5.3 Example Code

Once more, data comprises participant's responses to a questionnaire on the Big Five personality traits and come from the openpsychonometrics website. 19719 participants answered 50 questions, like "I am the life of the party" on a five-point Likert scale ([1]: Very Accurate, . . . , [5] Very Inaccurate). To analyze the data `disjoint()` and `overlap()` from the R package `elisr` (Bißantz, 2021) were used.

Data preparation

```
# URL
url <- ‘‘https://quantdev.ssri.psu.edu/sites/qdev/files/dataBIG5.csv’’
d <- read.csv(url, header = TRUE)
# Exclude: sociodemographics
X <- d[, -(1:7)]
# 0 is NA
X[X==0] <- NA
# Build disjoint scale fragments: ITT_c=.55
# NA: Pairwise.complete.obs
```

Multiple unidimensionality analysis

```
msdf <- elisr::disjoint(X, mrit_min = .6)
# Complete the scale
# NA: Pairwise.complete.obs
elisr::overlap(msdf, mrit_min = .5)
```

6 Concluding remarks

The thesis started with a compatibility issue, namely that applied methods and the OUI in the social sciences are often incompatible with one another. Due to the (implicit) assumption of independence, researchers frequently smuggle an incompatible assumption into their analysis, which undermines the blurry boundary feature of their OUI. But as argued in the course of this paper, complex social phenomena often share common attributes, suggesting to expect correlations among them. Reckless application of common go-to methods, ignoring object-specific features, leads to distorted results – a method-blindness fallacy imposing ignorance bias. The OUI's characteristic features themselves suggest bad defaults and decent alternatives for the exploration process and, thus for dimensionality analysis in general. Therefore, the aim to evaluate the state of the art and find alternative approaches obliged to reconcile proposed methods for compliance with the OUI. The proposed solution was a bottom-up approach with two requirements: (1) developing the characteristic features of the OUI, and (2) elaborating profound knowledge of the methods themselves, which allows to assess and criticize them. The general strategy in this thesis consequently builds upon the U-E-R-A approach: *understand the problem, evaluate bad defaults, replace them with decent alternatives to alleviate bias.*

Integrating U-E-R-A approach, this work moved from an understanding of the OUI to bad defaults and proposed decent alternatives, as well as research recommendations to alleviate bias. Moreover, the sequence of chapters and sections is structured to mimic the ongoing debate of appropriate dimensionality assessment in the social sciences. It thus worked up the topic programmatically and systematically.

First and programmatically, unidimensionality analysis was outlined and its problem to meet the large data set challenge. Second, multidimensionality analysis was suggested and the issue of assuming simultaneous contributions across dimensions from a multiple unidimensional standpoint. Finally, multidimensionality analysis was proposed as a decent alternative. In addition to the between-framework view, the debate moves on systematically within each approach. On this occasion, the aforementioned compatibility issue was included. This made it possible to select appropriate criteria and methods to combat the problem of identifying latent dimensions – while keeping track of the OUI's characteristic features. In total, five bad defaults were detected and replaced with decent alternatives: (1) In unidimensionality assessment, when running reliability analysis, Cronbach's alpha has shown to be a bad default and should be replaced with coefficient omega. (2) Carrying out multidimensional assessments with factor analysis is best performed using (2a) MLFA to solve the communality issue, and (2b) simplimax to solve the rotation problem. Based on the arguments presented, PAFA and varimax rotation should be avoided, because they impute the incompatible assumption of independence into the analysis. (2c) Tackling the number of factor problem is more complex. In any case, one should avoid K1 and scree-test to determine the number of factors, because they are prone to either overextraction or underextraction – or both. The decent alternative proposed is a combination of hull method and sequential χ^2-Test. Additionally, one could top up with the very simple structure criterion. (3) To finally pin down the participants' values on the latent dimension, one may replace sum-scores and use factor scores instead. Especially the solution of @TenBerge1999 seems promising because it incorporates between-factor correlations when estimating participants' values on the latent dimension. It thus complements the proposed extraction and rotation strategy.

6.1 Discussion

The summary highlights the most important guidelines to conduct an exploratory dimensionality analysis. However, there is one crucial point left on the table worth discussing. In the debate on multidimensionality analysis, using FA emerged as bad default from the comparison with multiple unidimensionality and exploratory Likert scaling. Designating FA as bad default would have dramatic consequences since it is the common go-to method for dimensionality analysis in the social sciences. Are numerous papers yet flawed by the assumption of equal contributions? There are at least three significant arguments against this conclusion.

First, multiple unidimensionality makes an equally strong assumption. Recall its critique; multidimensional solutions assume each (sub-)dimension to operate *simultaneously* in contributing to the observed differences between the scaled objects. Their counterargument was that in large data sets there might be multiple *separate* dimensions at work, where each dimension contributes to a particular subset of the data. However, the counterargument includes an assumption, too. It is true that the multiple-unidimensional solution avoids the assumption of equal contributions though, but at the cost of assuming multiple separate contributions. Which one is more appropriate? It depends. But a priori, it is reasonable that neither one should get preference over the other. In the early stages of exploratory investigations, when prior knowledge is absent, both are equally plausible. The best thing researchers can do is model comparison. Its results are valuable information about the investigation – either by proposing different plausible results, which are compatible with the data, or evaluating model differences trying to explain variation in the results. In contrast, selecting a particular model suppresses those insights from the research community, hence hampers the benefit to rule out potential flaws of an investigation. But researchers should take all help they can get when learning from incomplete sources of information – data.

Second, researchers do not need to decide between options. In large data sets, it might be helpful to combine the use of multidimensional and multiple unidimensional approaches to get the most out of the analysis. There is no reason to assume both of them to be mutually exclusive. In large data sets, there might be dimensions that contribute equally to produce the observed differences between the outcomes as well as separate ones. So omitting either one neglects its benefit to explain certain parts of the whole structure. Again, there are often multiple plausible solutions compatible with the data. Learning from them requires disclosing all investigation efforts, not only parts. Selecting a model in an exploratory investigation is a decision made way too early in the "stream of information" that could justify it (Gelman & Rubin, 1995; McElreath, 2020).

Third, one could tackle the distinction between (multiple) unidimensionality and multidimensionality, as (Jacoby, 1991, pp. 81) suggests:

> " [...] [T]he distinction between 'unidimensional' and 'multidimensional' scaling models is essentially an artificial one. [...] In every scaling analysis, the objective is to model the variability in the observations as accurately as possible. This may require several dimensions, or a single dimension may suffice. The choice between them is an empirical question. "

So probably one should move beyond the goal of conducting dimensionality analysis onto a modeling approach that solely commits itself to detect interesting sources of variation. As such, factor analysis is a comprehensive modeling tool.

By drawing on decent alternatives instead of bad defaults, FA can be harmonized with the characteristic features of the OUI and stays in line with the principle of compression in meaningful ways. Various implementations allow to fine-tune applications and adapt to a wide range of circumstances in applied research. Abandoning the usage of FA for artificial reasons is nonsense. The same applies to ELS. The problem with ELS from a practical point of view is merely that its current variants and implementations are underdeveloped. Nonetheless, a modern update of the algorithm will make it possible to detect interesting sources of variation in the data.

6.2 Conclusion

In the end, giving up on the artificial distinction between unidimensionality on one side and multidimensionality on the other side allows keeping two bespoke algorithms in applied research (i. e., FA and ELS). Both are in line with the characteristic features of the OUI, capable to meet the large-data-set challenge, and able to detect relevant sources of variation in the data. As such, they are well-tailored tools in applied research. Both should extend a modern researcher's toolbox. Comparing their results or combining their use allows for further insights into the data. If it comes down to a final research recommendation, ELS as wells as FA can be justified. Modern extensions of FA are still state of the art, but a modern update of ELS will be an additional innovation to combat the problem of identifying latent dimensions.

Notes

1. The symbols are now introduced to smoothly prepare the formalism in the upcoming chapters. But any explanation can be understood without them. If they are inconvenient, just ignore them. The symbols will often allow concertizing the explanation, specifying the referenced element.

2. Since ELS completely builds upon CTT, focus in UDA is fully on CTT. In applied research, there is a big ideological debate on whether to use it (Borsboom, 2005; Hattie, 1985). But concerning its performance, meanwhile, a big empirical body of evidence developed, showing that the results are often quite comparable (Ergüven, 2014; Fan, 1998; D.-Y. Hwang, 2002; Kohli et al., 2015; MacDonald & Paunonen, 2002; Xitao, 1998). However, most studies miss giving convincing arguments why this is the case. In a more recent study, Kohli et al. (2015) showed that only modern extensions, subsumed under the underlying variable approach (UVA), are comparable with modern item response theory models. These UVA-extended CTT models are most performant. Under the so-called UVA assumption (i. e., underlying normally distributed variables) polychoric correlations will be outlined in chapter 4.

3. Actually, the starting point of the derivation is the measurement standard error. It can be estimated assuming tau-equivalence, as follows: $\sigma_\epsilon = \sigma_X \sqrt{1-\rho^2_{XT}}$. However, the relationship $X \sim T$ is unknown, since T is unknown. As a consequence, one has to bridge the gap with known bits of information. Since only X is known, assuming each item to represent a single test, there are m parallel tests for which $X = \sum_{i=1}^{m} X_i$. So one can approximate σ_ϵ by $\alpha = k/k-1(1-\sum_{i=1}^{k}\sigma^2_{X_i}/\sigma^2_X)$ (Mair, 2018). Note that this derivation shows how reliability is linked to and accessed by the internal consistency of a set of items.

4. In next chapter, using a regression framework to estimate factor scores ($\hat{F}$) is argued to be superior to pin down a participants value on the latent dimension.

5. As a little side note; the true score approach of CTT relates closely to exploratory factor analysis. Mair (2018), for example, interlinks both approaches in equation 1.1 and 2.2.

6. There are several differences between PCA and FA. For instance, PCA tries to minimize the loss of compression when reproducing information in the correlation matrix (R). But the question is if researchers should really expect the variables to capture all the information in the correlation matrix. Or, to put it another way, if the variables can fully explain the variation in the data. In the social sciences, it seems more realistic to model only the *common* features of a sample. That's where common factor analysis (FA) comes in. FA tries to explain

the *common* parts of the correlation matrix; which means (1) covariation (i. e., off-diagonals) plus (2) the so-called COMMUNALITIES. The communalities are the common portion of the variance in the on-diagonals. What further distinguished the common factor model (CFM) from PCA is the leftover, called "uniqueness". The uniqueness holds unexplainable bits, things like measurement error and item-specific variation. These are usually suppressed in standard PCA models. Against the background of the complex measurement (i.e., being doomed to measure with error) CFM's advantages over PCA seem convincing. One major reason for the confusion between PCA and FA arises from PCM (i.e., principal component method) or PFM (i. e., principal factor method). The method will be part of later sections. For now, simply memorize, there is a difference between both approaches (i. e., PCA and FA) and a particular extraction method within the approach (i. e., PCM and PFA).

7. Brown (2015) gives a comprehensive and applied introduction in his book.

8. $\Lambda\Phi\Lambda^t$ is a common way to define the communalities (Yong & Pearce, 2013). Literally, the expression reads as "(sum of the) squared loadings". The definition follows from CFM under the assumption of independence Assuming independent factor forces their correlation to be zero: $\forall\phi_{i=1,\dots,m} \in \Phi : \phi_{i\neq j} \Rightarrow \phi_{i,j} = 0$ Accordingly, Φ becomes the identity matrix: I. The identity matrix contains 1's in the on-diagonals and 0's in the off-diagonals. It is the neutral element of matrix multiplication. As a consequence, any matrix M multiplied with I is $M \times I = M$. Thus, if $\Psi = I$, the right-hand side of the fundamental equation $\Lambda\Phi\Lambda^t$ renders to $\Lambda\Lambda^t$

9. The principal component method (PCM) is the reason for major confusion about FA and PCA. Commercial software like SPSS choose the classic their default extraction method in EFA. As a consequence, singular value or eigenvalue decomposition and factor extraction became intertwined. But they shouldn't go hand in hand. Stata and R, for instance, take the distinction into account. Stata, separate models within tabs. Likewise, R includes them in different packages. The next sections are devoted to the problem of PCM in social sciences. In recourse, it may become clear why SPSS chose a bad default.

10. Experimenting with the communality values allows understanding the differences and similarities between PAFA and PCA. To get a better intuition of their relation, Revelle (2021a) nudged to play the COMMUNALITY GAME: In PAFA, the communalities can be determined. If the researcher iterative increases the value to one, PAFA and PCA yield approximately equal results. If the communalities equal 1, PCA can be said to be identical to PAFA. Why? First, on one hand decompose the underlying correlation matrix with an EIGENVALUE DECOMPOSITION (or the input matrix with a singular-value decomposition). So they are equal regarding that fact. But on the other hand they use different variants of correlation matrices. Principal component analysis applies the eigenvalue decomposition to the whole correlation matrix R, whereas PAFA uses the reduced matrix. In simplified terms, they relate to each other as follows: $diag(R) = \mathbf{1}$; $diag(R^*) < \mathbf{1}$. Accordingly, if the reduced matrix iterates its diagonal towards 1, both resemble one another: $diag(R^*) \rightarrow 1 \Rightarrow R^* \approx R$. Lastly, as already mentioned, both use an EVD to decompose their matrices. Applying the EVD to an arbitrary correlation matrix (R) yields: $R = U\Lambda U^t$ (Mair, 2018). U is a left singular matrix, U^t the transposed right-singular matrix and Λ the factor loading matrix. A principal component C

can now be defined as product of the singular matrix and the root of the eigenvectors ($\sqrt{\lambda}$): $C = U\sqrt{\lambda}$ (Revelle, 2021a). This simplifies R to the product of two component matrices: $R = CC^t$. Accordingly, for the reduced matrix: $R^* = FF^t$. To summarize, if the correlation matrices resemble one another, so do their results: $diag(R^*) \to 1 \Rightarrow FF^t \approx CC^t$. This is in line with the findings of Harris (1964), for example, who showed that PCA and PAFA provide almost identical results under the principal extraction method. In conclusion, if the correlation matrix resembles one another, the results are equal. If they do not, the results differ (Gorsuch, 2015, p. 129).

11. Technically speaking, the procedure provides a solution that successively maximizes the (squared) factor-item correlations: $\Lambda\Lambda^t$ As a result, the between factor correlations are literally thrown out. Why? Think of the syringes. Maximizing the information density within each factor minimizes the information shared between factors. All soak water from a common source of information. One must decide how to distribute the water across syringes (i. e., information across factors). PCM's choice is, once more, maximizing information reconnaissance. Ultimately, the robots maximize the sum of squares of the factor structure and thus the information kept from the residual matrix after the first factor has been extracted (Gorsuch, 2015, p. 101)

12. Conceptually, factor independence can be understood following fundamental theorem: $P = \Lambda\Lambda^t + \Psi$. Again, assuming factor independence, the correlation between factors is 0. The matrix of between factor correlations becomes the identity matrix: $\Phi = I$. Under the assumption, the model equation simplifies to $P = \Lambda\Lambda^t + \Psi$. Researchers can easily get rid of the uniqueness, too. Expecting the variables to pick up all the information in the correlation matrix, the uniqueness drops out and P renders to R. Why? Because the previous assumption necessitates to get all the information in the correlation matrix – not just the common ones. Accordingly, P simplifies twice: $R = \Lambda\Lambda^t$. At any case, the explanation suffice understanding the process. However, the description misses a concise argument. Independence actually results from EVD and comes in the form of U: $R = U\Lambda U^t$. Remember, U is a singular matrix. The singular matrix comprises orthogonal vectors, which are independent of each other. In sum, independence results from how the information are decomposed. It furthermore expresses in the result, as it was shown with the syringes.

13. Technically, the ML problem is solved by generalizing the ordinary least squares approach (Jöreskog, 1978). In fact, the robot searches for a parameter vector θ so that $\mathrm{S}(\theta) \to \min \frac{1}{2} tr(SP^{-1} - I)^2$ (S: sample correlation matrix). Minimizing the expression can be shown to maximizing the likelihood of the solution given the data (Revelle, 2021a). For this reason, maximum likelihood estimation yields the most plausible values (θ) to reproduce the common parts of the structure in the correlation matrix.

14. Note that by varying different estimation approaches, one does usually iterate through various error assumptions (Jöreskog, 1978).

15. Thurstone (1935, 1947, 1969) developed a set of principles to guide the rearrangement of the loading matrix ($\hat{\Lambda}$). Thinking in zeros and ones, one aims to redistribute the zeros and ones 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 literally a simple structure. Each item is clearly attributable to (i.e., loads highly on) a single factor: $\exists \xi_i, \xi_j : \lambda_{a_i}, \lambda_{b_i}, \lambda_{c_i} \gg \lambda_{a_j}, \lambda_{b_j}, \lambda_{c_j}$. For Carroll (1953), Jennrich and Sampson (1966), and Tucker (1955), and others, this must-have been one of those statistical fairytales. They all criticize that a "Simple Structure" will be hardly ever reached. Nonetheless, nowadays there are various implementations to come close to the ideal. Some of them will be discussed afterward.].This is at the heart of the rotation problem. The key question is how to transfer the robot's intermediate to something simple – which means more interpretable ($\hat{\Lambda}_r$).

16. Once introduced, orthogonality maintains in a system. This holds especially true, modeling the information with the eponymous rotation. Starting (1) from the fundamental theorem, using (2) the transformation equation and (3) the orthogonality constraint, one can easily show how orthogonality preserves. Accordingly, the derivation is called the "law of conservation of orthogonality":

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

17. To shed light on this, let2019s look at the pattern-structure-matrix distinction. In research literature, the terms STRUCTURE MATRIX and PATTERN MATRIX are omnipresent. To build an intuition, 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. Here, the structure can be reproduced using only factor-item information. So in the end, the structure (S) is reproducible as a function of the factor-item information (Λ) and factor-factor information (Φ). Algebraically, this can be written 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. Structure matrix and pattern matrix are identical. The key question in the following will be how realistic the assumption of independence really is in applied social science research.

18. A technique which was always part of exploratory analyzes since the prior work of Thurstone in the 1940s is visual rotation. The problem with it is that selecting factor radians (θ) imposes some kind of arbitrariness in the rotation procedure (Carroll, 1953). Researcher are often accused to miss a concise "objective" argument on which to evaluate the position of the axes. Hence, researchers usually turn towards analytic criteria. What often remains unnoticed is, visual rotation allows a pretty good approximation of a simple structure Gorsuch (2015, p. 202). However, nowadays, the use of analytic criteria predominates the social sciences. Most of them are oriented towards providing a mathematical gateway to a simple structure. Thus, today's focus is on how to minimize a matching criterion

19. The original paper from Saunders "Transvarimax: Some properties of the ratiomax and equamax criteria for blind orthogonal rotation" is not accessible. Drawing on (Gorsuch, 2015), the proposed method was documented in a paper delivered at the American Psychological Association meeting in 1962. Additional information about equamax can be found in (Kaiser & Kaiser, 1974).

20. Technically speaking, maximizing the sum of squared loadings ($\hat{\Lambda}\hat{\Lambda}^t$) on each factor ($\xi$) actually leads to a transformation matrix (T), which multiplies the columns of the matrix of factor loadings ($\hat{\Lambda}$) by radians (θ) (Revelle, 2021a). Thereby, they rotate (counterclockwise) by θ.

21. Technically speaking, the transformation (T) should distribute the communality of each item across the least number of factors (ξ). This generates the easiest loading picture possible. Why? Because each factor-loading $\hat{\lambda}$ is either zero or as far from zero as possible (Lorenzo-Seva, 2003).

22. Conceptually, the argument of Lorenzo-Seva (2003) is that other indexes partially miss the target (i. e., the simplest result possible). They deal with the loading merely indirectly. Bentler's index, for example, builds upon the columns of the matrix of factor loadings (Λ). The loading simplicity, on the other hand, puts the focus on the values of the loadings directly.

23. The workhorse in Kiers (1994) 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.

24. 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.

25. A four-step procedure of fitting, evaluating, understanding, and communicating deviations with the results is often part of so-called SENSITIVITY ANALYZES (Saltelli, 2002). Sensitivity analysis plays a special role if the object under investigation is of social or political significance.

26. In recent developments, like BAYESIAN EXPLORATORY FACTOR ANALYSIS (BEFA; Conti et al., 2014, the algorithm proposes a set of solutions rated according to their plausibility. This means BEFA rates the different number of factors according to their posterior probabilities. Regarding MLFA, the result also includes the most plausible values given the constraint. But constraints are furthermore data and prior knowledge.

27. Mair (2018) notes that the K1-solution is based on an eigenvalue decomposition of the correlation matrix (R). Correspondingly, the Kaiser criterion provides a solution based on the principal component (PC) model. As a result, it iteratively extracts *principal components* which

explains successively less variation in the data – until the eigenvalue of a standardized variable is reached. For that reason, Timmerman et al. (2017) subsume K1 under the PCA-based criteria.

28. Because of its flaws, nowadays, there is almost no research effort put into the Kaiser criterion. Reviewing the literature, there are only improved versions of the classical criterion, like the empirical Kaiser criterion (Braeken & Van Assen, 2017), under review (see Auerswald & Moshagen, 2019).

29. The argument is derivable, following the fundamental equation: $P = \Lambda\Phi\Lambda^t + \Psi$. If a model ($P'$) assumes factor-independence ($\Phi = I$) and P' simplifies to the independent version of the fundamental equation: $P' = \Lambda I \Lambda^t \Psi$. In that sense, P' equals P the more Φ equals I, which means, the more factor independence is reasonable. Formally, this can be written as: $P \approx P'|\Phi \to 0$. The argument is, since, the number of factors (p) determines the form of the model (P) deviations between the models ($P \not\approx P'$) should imply deviation in the number of factors to retain ($p \not\approx p'$). Formally: $\Phi \not\approx I \Rightarrow P \not\approx P' \Rightarrow p \not\approx p'$.

30. Timmerman et al. (2017) subsume the scree test under the PCA-based approaches, because it bases on eigenvalue decomposition of the correlation matrix (R), too. This means, the scree test, as well as K1, provide their solutions based on the principal component model. More precisely, the model iteratively extracts components that explain successively less and less variation in the data. When there is no unexplained bit of variation left, all eigenvalues (λ) are sorted in decreasing order and plotted along with an index ($I = 1, \ldots, m$). Note that, because of the decomposition strategy, there are as many eigenvalues ($\lambda_{i=1,\ldots,m}$) as there are items in the data set.

31. So far, the argument why PCA-based and CFA-based model differ regarding the number of factor to retain based upon factor independence. But as shown, the differences exceed the problem of independence. Both use (1) different models and rely (2) on different sources of information. Zwick and Velicer (1986b) did already prompt towards differences between the two strategies in terms of the implied number of factors to retain. They conclude, PCA will propose far more components than CFM do factors. The result should not surprise. Once more, PCA attempts to explain all the information in the correlation matrix. For this reason alone, there is more information to explain, logically demanding a larger number of factors.

32. This assumption has mathematical roots. (Timmerman et al., 2017) pointed out that assuming the ordered categorical indicators to be the product of underlying normally distributed variables turns the polychoric correlation into an MLE for the Pearson correlation between latent variables.

33. More precisely, the regression approach is a two-step procedure (Mair, 2018): (1) combine the correlation matrix (R) and loading matrix ($\hat{\Lambda}$) to estimate weights ($\hat{B}$): $\hat{B} = R^{-1}\hat{\Lambda}$ (2) Use the data (X), standardize them (Z) and push them out through the model using the score coefficient matrix ($\hat{B}$) to estimate the factor scores ($\hat{F}$): $\hat{F} = Z\hat{B}$

34. An upcoming paper of Thomas Müller-Schneider concerning exploratory Likert Scaling is planned at the end of 2021. It will elaborate on this topic in much greater detail.

A Code to Reproduce the Graphics

Graphic 1

```
set.seed(359)
# Define factor loadings
loads <- rep(0.7, 12)
# Population correlation matrix (I)
# tau <- psych::sim.congeneric(loads=loads)
# Sample correlation matrix
R <- psych::sim.congeneric(loads=loads,N=50)
# Visualize the correlation matrix
corrplot::corrplot.mixed(R, number.cex=.7)
```

Graphic 2

```
# Random number generator
set.seed(359)
# Define factor loadings
loads <- rep(0.6, 12)
# Population correlation matrix
tau <- psych::sim.congeneric(loads=loads)
# Sample correlation matrix
tau_sample <- psych::sim.congeneric(loads=loads,N=50)
# Rounded correlation matrix
 round(tau_sample,2)
```

Graphic 3

```
# Random number generator
set.seed(359)
# Define factor loadings
```

```
loads <- rep(0.1, 12)
# Population correlation matrix (I)
# tau <- psych::sim.congeneric(loads=loads)
# Sample correlation matrix
R <- psych::sim.congeneric(loads=loads,N=50)
# Visualize the correlation matrix
corrplot::corrplot.mixed(R, number.cex=.7)
```

Graphic 4

```
# Random number generator
set.seed(0405)
# Determine 0-vector
zeros <- rep(0,12)
# Define a measurement model for x
# Note: values ~ loadings
fx <-matrix(c(.8, .8, .5, zeros, .7, .8, -.7, zeros,
              -.5, .7, -.8, zeros, .8, .8, .5), ncol = 4)
# Define the structure matrix
# On-diagonals: 1
phi <- diag(rep(1, 4))
# Between factor correlations
phi[1, 2] <- phi[2, 1] <- 0.2 ; phi[1, 3] <- phi[3, 1] <- 0.6
phi[1, 4] <- phi[4, 1] <- 0.7 ; phi[2, 3] <- phi[3, 2] <- 0.2
phi[2, 4] <- phi[4, 2] <- 0.6 ; phi[3, 4] <- phi[4, 3] <- 0.4
# Visualize the measurement model
# psych::structure.diagram(fx, phi,
                  main = ''Measurement model for X'')
# Produce the correlation matrix
R <- psych::sim.structure(fx, phi)$model
# Visualize the correlation matrix
corrplot::corrplot.mixed(R, number.cex=.7)
```

Graphic 5

```
# Random number generator
set.seed(213)
# Determine 0-vector
N <- 20
loadings <- sample(seq(-.8,.8,0.01), N, replace=TRUE)
```

```
# Define a measurement model for x
# Note: values ~ loadings
fx <-matrix(loadings, ncol = 2)
# Define the structure matrix
 # On-diagonals: 1
 phi <- diag(rep(1, 2))
 # Between factor correlations
 phi[1, 2] <- phi[2, 1] <- -0.2
 # Visualize the measurement model
 # psych::structure.diagram(fx, phi,
                    main = ‘‘Measurement model for X’’)
 # Produce the correlation matrix
 R <- psych::sim.structure(fx, phi)$model
 # Visualize the correlation matrix
 #corrplot::corrplot.mixed(R, number.cex=.7)
 fa <- psych::fa(R, nfactors = 2)
 plot(NULL, xlab = ‘‘honor’’,  ylab = ‘‘pride’’,
      ylim = c(-1,1), xlim = c(-1,1))
 points(fa$loadings[,1], fa$loadings[,2], pch=20, cex=.7)
 text(fa$loadings[,1], fa$loadings[,2]+.1,
      paste0(‘‘I’’, 2:(N/2)+1), cex = .7)
 l1 <- .5 ; l2 <- .6
 points(l1, l2, pch=20, cex=.7)
 text(l1, l2+.1, ‘‘I:(0.5,0.6)’’, cex = .6, col="red")
 lines(c(l1,l1), c(0,l2) ,lty=2, lwd=.5)
 lines(c(0,l1), c(l2,l2) ,lty=2, lwd=.5)
 abline(h = 0, v = 0)
```

Graphic 6

```
# psych::structure.diagram(fx, phi,
                   main = "Measurement model for X")
# Produce the correlation matrix
R <- psych::sim.structure(fx, phi)$model
# Visualize the correlation matrix
#corrplot::corrplot.mixed(R, number.cex=.7)
fa <- psych::fa(R, nfactors = 2)
plot(NULL, xlab = \u201chonor\u201d,  ylab = "pride",
     ylim = c(-1,1), xlim = c(-1,1))
 points(fa$loadings[,1], fa$loadings[,2], pch=20, cex=.7)
```

```
text(fa$loadings[,1], fa$loadings[,2]+.1,
     paste0(\u201cI\u201d, 2:(N/2)+1), cex = .7)
l1 <- .5 ; l2 <- .6
points(l1, l2, pch=20, cex=.7)
text(l1, l2+.1, \u201cI:(0.5,0.6)\u201d, cex = .6, col="red")
lines(c(l1,l1), c(0,l2) ,lty=2, lwd=.5)
lines(c(0,l1), c(l2,l2) ,lty=2, lwd=.5)
abline(h = 0, v = 0)
```

Graphic 7

```
# Random number generator
set.seed(213)
# Determine number of items
N <- 20
loadings <- sample(seq(-.8,.8,0.01), N, replace=TRUE)
# Define a measurement model for x
# Note: values ~ loadings
fx <-matrix(loadings, ncol = 2)
# Define the structure matrix
 # On-diagonals: 1
 phi <- diag(rep(1, 2))
 # Between factor correlations
 phi[1, 2] <- phi[2, 1] <- 0.6
 # Visualize the measurement model
 # psych::structure.diagram(fx, phi,
                    main = ‘‘Measurement model for X’’)
 # Produce the correlation matrix
 R <- psych::sim.structure(fx, phi)$model
 # Scree plot
 psych::scree(R, factors = FALSE)
```

Grafik 8

```
# Random number generator
set.seed(213)
# Determine the number of items
N <- 20
# Specify the loading pattern
loadings <- sample(seq(-.8,.8,0.01), N, replace=TRUE)
```

```
# Define a measurement model for x
# Note: values ~ loadings
fx <-matrix(loadings, ncol = 2)
 # Define the structure matrix
 # On-diagonals: 1
 phi <- diag(rep(1, 2))
 # Between factor correlations
 phi[1, 2] <- phi[2, 1] <- 0.2
 # Visualize the measurement model
 # psych::structure.diagram(fx, phi,
                     main = ‘‘Measurement model for X’’)
 # Produce the correlation matrix
 X <- psych::sim.structure(fx, phi, n=50)$observed
 # Parallel analysis
 paran::paran(X, quietly = TRUE, graph = TRUE, cfa = FALSE)
```

Grafik 9

```
# Random number generator
set.seed(213)
# Determine the number of items
N <- 20
# Specify a loading pattern
loadings <- sample(seq(-.8,.8,0.01), N, replace=TRUE)
# Define a measurement model for x
# Note: values ~ loadings
fx <-matrix(loadings, ncol = 2)
 # Define the structure matrix
 # On-diagonals: 1
 phi <- diag(rep(1, 2))
 # Between factor correlations
 phi[1, 2] <- phi[2, 1] <- 0.6
 # Visualize the measurement model
 # psych::structure.diagram(fx, phi,
                     main = ‘‘Measurement model for X’’)
 # Produce the correlation matrix
 X <- psych::sim.structure(fx, phi, n=100)$observed
 # Scree plot
 EFA.MRFA::hullEFA(X, extr=’’ML",display = FALSE, graph = TRUE)
```

Bibliography

Adelson, J. L., Osborne, J. W., & Crawford, B. F. (2019). Correlation and Other Measures of Association. *The reviewer's guide to quantitative methods in the social sciences* (pp. 55–71). Routledge. https://doi.org/10.4324/9781315755649-5

Andersson, G., & Yang-Wallentin, F. (2021). Generalized Linear Factor Score Regression: A Comparison of Four Methods. *Educational and Psychological Measurement*, *81*(4), 617–643. https://doi.org/10.1177/0013164420975149

Asún, R. A., Rdz-Navarro, K., & Alvarado, J. M. (2016). Developing Multidimensional Likert Scales Using Item Factor Analysis: The Case of Four-point Items. *Sociological Methods and Research*, *45*(1), 109–133. https://doi.org/10.1177/0049124114566716

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

Bandalos, D. L. (2018). *Measurement theory and applications for the social sciences*. The Guilford Press.

Barbaranelli, C., Lee, C. S., Vellone, E., & Riegel, B. (2015). The problem with Cronbach's alpha: Comment on Sijtsma and van der Ark (2015). *Nursing Research*, *64*(2), 140–145. https://doi.org/10.1097/NNR.0000000000000079

Beck, E. M. (1996). Culture of Honor: The Psychology of Violence in the South New Directions in Social Psychology Series. *The Georgia historical quarterly*, *80*(4).

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

Bißantz, S. (2021). *Elisr: Exploratory likert scaling* [R package version 0.1.1]. https://github.com/sbissantz/elisr

Blair, G., Cooper, J., Coppock, A., Humphreys, M., Rudkin, A., & Fultz, N. (2021). *Fabricatr: Imagine your data before you collect it* [R package version 0.14.0]. https://CRAN.R-project.org/package=fabricatr

Bornstein, M. H. (2018). Growth Curve Modeling and Longitudinal Factor Analysis. *The sage encyclopedia of lifespan human development*. Thousand Oaks. https://doi.org/10.4135/9781506307633.n373

Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*. Cambridge University Press. https://doi.org/10.1017/CBO9780511490026

Bortz, J., & Schuster, C. (2016). *Statistik für Human- und Sozialwissenschaftler*. Springer.

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

Brillinger, D. R., Preisler, H. K., Ager, A. A., & Kie, J. G. (2004). An exploratory data analysis (EDA) of the paths of moving animals. *Journal of Statistical Planning and Inference*, *122*(1-2), 43–63. https://doi.org/10.1016/j.jspi.2003.06.016

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford Press.

Browne, M. W. (1968). A note on lower bounds for the number of common factors. *Psychometrika*, *33*(2), 233–236. https://doi.org/10.1007/BF02290155

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

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Ceulemans, E., & Kiers, H. A. (2006). Selecting among three-mode principal component models of different types and complexities: A numerical convex hull based method. *British Journal of Mathematical and Statistical Psychology*, *59*(1), 133–150. https://doi.org/10.1348/000711005X64817

Cho, E., & Kim, S. (2015). Cronbach's Coefficient Alpha: Well Known but Poorly Understood. *Organizational Research Methods*, *18*(2), 207–230. https://doi.org/10.1177/1094428114555994

Cole, D. A., Ciesla, J. A., & Steiger, J. H. (2007). The Insidious Effects of Failing to Include Design-Driven Correlated Residuals in Latent-Variable Covariance Structure Analysis. *Psychological Methods*, *12*(4), 381–398. https://doi.org/10.1037/1082-989X.12.4.381

Conti, G., Frühwirth-Schnatter, S., Heckman, J. J., & Piatek, R. (2014). Bayesian exploratory factor analysis. *Journal of Econometrics*, *183*(1). https://doi.org/10.1016/j.jeconom.2014.06.008

Corballis, M. C., & Traub, R. E. (1970). Longitudinal factor analysis. *Psychometrika*, *35*(1), 79–98. https://doi.org/10.1007/BF02290595

Costa, M. (2003). *A comparison between unidimensional and multidimensional approaches to the measurement of poverty* (WorkingPaper No. 2003-02).

Costa, V. M. (2018). Patriotism and Nationalism. In P. Smeyers (Ed.), *International handbook of philosophy of education* (pp. 1389–1400). Springer, Cham. https://doi.org/10.1007/978-3-319-72761-5_96

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

Cox, J., & Dale, B. G. (2001). Service quality and e-commerce: An exploratory analysis. *Managing Service Quality: An International Journal*, *11*(2), 121–131. https://doi.org/10.1108/09604520110387257

Crawford, C. B., & Koopman, P. (1979). Note: Inter-Rater Reliability of Scree Test and Mean Square Ratio Test of Number of Factors. *Perceptual and Motor Skills*, *49*(1), 223–226. https://doi.org/10.2466/pms.1979.49.1.223

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

Cudeck, R., & MacCallum, R. C. (2012). *Factor analysis at 100: Historical developments and future directions*. Taylor; Francis. https://doi.org/10.4324/9780203936764

de Leeuw, J. (2005). Unidimensional Scaling. *Encyclopedia of statistics in behavioral science*. John Wiley & Sons, Ltd. https://doi.org/10.1002/0470013192.bsa700

DeVellis, R. F. (2006). Classical test theory. *Medical Care*, *44*(11 SUPPL. 3). https://doi.org/10.1097/01.mlr.0000245426.10853.30

DeVellis, R. F. (2017). *Scale Development: Theory and Applications*. Sage Publications. https://doi.org/10.2307/2075704

de Winter, J. C., & Dodou, D. (2012). Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. *Journal of Applied Statistics*, *39*(4), 695–710. https://doi.org/10.1080/02664763.2011.610445

Dinno, A. (2009). Exploring the sensitivity of Horn's parallel Analysis to the distributional form of random data. *Multivariate Behavioral Research*, *44*(3), 362–388. https://doi.org/10.1080/00273170902938969

DiStefano, C., Zhu, M., & Mîndrilă, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research and Evaluation*, *14*(20).

Dwivedi, Y. K., Choudrie, J., & Brinkman, W. P. (2006). Development of a survey instrument to examine consumer adoption of broadband. *Industrial Management and Data Systems*, *106*(5), 700–718. https://doi.org/10.1108/02635570610666458

Ergüven, M. (2014). An empirical evaluation and comparison of Classical Test Theory and Rasch Model. *Journal of Education*, *3*(1), 33–38.

Fabrigar, L. R., MacCallum, R. C., Wegener, D. T., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. https://doi.org/10.1037/1082-989X.4.3.272

Fan, X. (1998). Item response theory and classical test theory: An empirical comparison of their item/person statistics. *Educational and Psychological Measurement*, *58*(3), 357–381. https://doi.org/10.1177/0013164498058003001

Fava, J. L., & Velicer, W. F. (1996). The effects of underextraction in factor and component analyses. *Educational and Psychological Measurement*, *56*(6), 907–929. https://doi.org/10.1177/0013164496056006001

Fischer, G. H. (1974). *Einführung in die Theorie psychologischer Tests: Grundlagen und Anwendungen*. H. Huber.

Fleishman, J., & Benson, J. (1987). Using Lisrel to Evaluate Measurement Models and Scale Reliability. *Educational and Psychological Measurement*, *47*(4), 925–939. https://doi.org/10.1177/0013164487474008

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

Gelman, A., & Rubin, D. B. (1995). Avoiding Model Selection in Bayesian Social Research. *Sociological Methodology*, *25*, 1–7. https://doi.org/10.2307/271064

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*(3), 377–393. https://doi.org/10.1177/0013164495055003002

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

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., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality1. *Educational and Psychological Measurement*, *37*(4), 827–838. https://doi.org/10.1177/001316447703700403

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

Green, S. B., & Yang, Y. (2009). Reliability of summed item scores using structural equation modeling: An alternative to coefficient alpha. *Psychometrika*, *74*(1), 155–167. https://doi.org/10.1007/s11336-008-9099-3

Gregory, R. J. (2014). *Psychological Testing History, Principles, and Applications seventh edition*. Pearson.

Grice, J. W. (2001a). A comparison of factor scores under conditions of factor obliquity. *Psychological Methods*, *6*(1), 67–83. https://doi.org/10.1037/1082-989x.6.1.67

Grice, J. W. (2001b). Computing and evaluating factor scores. *Psychological Methods*, *6*(3), 430–450. https://doi.org/10.1037/1082-989x.6.4.430

Gründl, J. (2021). *Jogru: Johann gründl's collection of r functions* [R package version 2.0].

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

Guttman, L. (1944). A Basis for Scaling Qualitative Data. *American Sociological Review*, *9*(2), 139. https://doi.org/10.2307/2086306

Hakstian, A. R., & Cattell, R. B. (1982). The behavior of number-of-factors rules with simulated data. *Multivariate Behavioral Research*, *17*(2), 193–219. https://doi.org/10.1207/s15327906mbr1702_3

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

Harris, C. W. (1964). Some recent developments in factor analysis. *Educational and Psychological Measurement*, *24*(2), 193–206. https://doi.org/10.1177/001316446402400202

Hattie, J. (1985). Methodology Review: Assessing Unidimensionality of Tests and Items. https://doi.org/10.1177/014662168500900204

Hayes, A. F., & Coutts, J. J. (2020). Use Omega Rather than Cronbach's Alpha for Estimating Reliability. But... *Communication Methods and Measures*. https://doi.org/10.1080/19312458.2020.1718629

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor Retention Decisions in Exploratory Factor Analysis: A Tutorial on Parallel Analysis. https://doi.org/10.1177/1094428104263675

Heise, D. R. (1973). Some Issues in Sociological Measurement. *Sociological Methodology*, *5*. https://doi.org/10.2307/270830

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

Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Perron, J. M., & Mumford, K. R. (2005). The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination. *Educational and Psychological Measurement*, *65*(2), 202–226. https://doi.org/10.1177/0013164404267287

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

Horn, J. L., & Engstrom, R. (1979). Cattell's scree test in relation to Bartlett's Chi-square test and other observations on the number of factors problem. *Multivariate Behavioral Research*, *14*(3), 283–300. https://doi.org/10.1207/s15327906mbr1403_1

Hwang, D.-Y. (2002). Classical test theory and item response theory: Analytical and empirical comparisons. *Annual Meeting of Southwest Educational Research Association*.

Hwang, I. H. (1970). The Usability of Item-Total Correlation as the Index of Item Discrimination. *Korean Journal of Medical Education*, *12*(1), 45–51. https://doi.org/10.3946/kjme.2000.12.1.45

Jackson, D. L. (2001). Sample size and number of parameter estimates in maximum likelihood confirmatory factor analysis: A Monte Carlo investigation. *Structural Equation Modeling*, *8*(2), 205–223. https://doi.org/10.1207/S15328007SEM0802_3

Jackson, H. H., Nisbett, R. E., & Cohen, D. (1997). Culture of Honor: The Psychology of Violence in the South. *The Journal of Southern History*, *63*(3), 661. https://doi.org/10.2307/2211676

Jackson, P. H., & Agunwamba, C. C. (1977). Lower bounds for the reliability of the total score on a test composed of non-homogeneous items: I: Algebraic lower bounds. *Psychometrika*, *42*(4), 567–578. https://doi.org/10.1007/BF02295979

Jacoby, W. G. (1991). Data theory and dimensional analysis.

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

Jöreskog, K. G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika*, *43*(4), 443–477. https://doi.org/10.1007/BF02293808

Jung, S., & Lee, S. (2011). Exploratory factor analysis for small samples. *Behavior Research Methods*, *43*(3), 701–709. https://doi.org/10.3758/s13428-011-0077-9

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. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, *20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31–36. https://doi.org/10.1007/BF02291575

Kaiser, H. F., & 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

Klein, H. J., Cooper, J. T., Molloy, J. C., & Swanson, J. A. (2014). The assessment of commitment: Advantages of a unidimensional, target-free approach. *Journal of Applied Psychology*, *99*(2), 222–238. https://doi.org/10.1037/a0034751

Kline, P. (1994). *An Easy Guide to Factor Analysis*. Routledge. https://doi.org/10.4324/9781315788135

Kohli, N., Koran, J., & Henn, L. (2015). Relationships Among Classical Test Theory and Item Response Theory Frameworks via Factor Analytic Models. *Educational and Psychological Measurement*, *75*(3), 389–405. https://doi.org/10.1177/0013164414559071

Kruskal, J. B., & Wish, M. (1978). *Multidimensional Scaling (Quantitative Applications in the Social Sciences)*. SAGE Publications Inc.

Ledesma, R. D., & Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out Parallel Analysis. *Practical Assessment, Research and Evaluation*, *12*(2). https://doi.org/10.7275/WJNC-NM63

Lee, H. B., & Comrey, A. L. (1979). Distortions in a commonly used factor analytic procedure. *Multivariate Behavioral Research*, *14*(3), 301–321. https://doi.org/10.1207/s15327906mbr1403_2

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *140*, 44–53.

Linn, R. L. (1968). A monte carlo approach to the number of factors problem. *Psychometrika*, *33*(1), 37–71. https://doi.org/10.1007/BF02289675

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

Lord, F. M., Tukey, J. W., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Addison-Wesley Pub. Co.

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

Lorenzo-Seva, U., Timmerman, M. E., & Kiers, H. A. (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

MacCallum, R. C., Widaman, K. F., Preacher, K. J., & Hong, S. (2001). Sample size in factor analysis: The role of model error. *Multivariate Behavioral Research*, *36*(4), 611–637. https://doi.org/10.1207/S15327906MBR3604_06

MacDonald, P., & Paunonen, S. V. (2002). A Monte Carlo comparison of item and person statistics based on item response theory versus classical test theory. *Educational and Psychological Measurement*, *62*(6), 921–943. https://doi.org/10.1177/0013164402238082

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

Mair, P., & Leeuw, J. D. (2015). Unidimensional Scaling. *Wiley statsref: Statistics reference online* (pp. 1–3). Wiley. https://doi.org/10.1002/9781118445112.stat06462.pub2

Majors, M. S., & Sedlacek, W. E. (2001). Using factor analysis to organize student services. *Journal of College Student Development*, *42*(3), 272–277.

McDonald, R. P. (1978). Generalizability in factorable domains: "Domain validity and generalizability". *Educational and Psychological Measurement*, *38*(1), 75–79. https://doi.org/10.1177/001316447803800111

McDonald, R. P. (1999). *Test theory: A Unified Treatment*. Erlbaum.

McElreath, R. (2020). *Statistical Rethinking*. CRC Press.

McIver, J. P., & Carmines, E. G. (1981). *Unidimensional scaling*. Sage Publications.

McNeish, D. (2018). Thanks coefficient alpha, We'll take it from here. *Psychological Methods*, *23*(3), 412–433. https://doi.org/10.1037/met0000144

Mokken, R. J. (1971). *A Theory and Procedure of Scale Analysis*. De Gruyter. https://doi.org/10.1515/9783110813203

Mosquera, P. M., Fischer, A., Manstead, A., & Zaalberg, R. (2008). Attack, disapproval, or withdrawal? The role of honour in anger and shame responses to being insulted. *Cognition and Emotion*, *22*(8), 1471–1498. https://doi.org/10.1080/02699930701822272

Müller-Schneider, T. (2001). Multiple Skalierung nach dem Kristallisationsprinzip / Multiple Scaling According to the Principle of Crystallization. *Zeitschrift für Soziologie*, *30*(4), 305–315. https://doi.org/10.1515/zfsoz-2001-0404

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

Navarro-Gonzalez, D., & Lorenzo-Seva, U. (2021). *Efa.mrfa: Dimensionality assessment using minimum rank factor analysis* [R package version 1.1.2]. https://CRAN.R-project.org/package=EFA.MRFA

Nesselroade, J. R. (1972). Note on the "longitudinal factor analysis" model. *Psychometrika*, *37*(2), 187–191. https://doi.org/10.1007/BF02306776

Nisbett, R. E. (2018). *Culture Of Honor: the Psychology Of Violence In The South*. Taylor; Francis.

O'Connor, B. P. (2021). *Efa.dimensions: Exploratory factor analysis functions for assessing dimensionality* [R package version 0.1.7.2]. https://CRAN.R-project.org/package=EFA.dimensions

Patil, V. H., Singh, S. N., Mishra, S., & Todd Donavan, D. (2008). Efficient theory development and factor retention criteria: Abandon the 'eigenvalue greater than one' criterion. *Journal of Business Research*, *61*(2), 162–170. https://doi.org/10.1016/j.jbusres.2007.05.008

Pearson, R. H., & Mundfrom, D. J. (2010). Recommended sample size for conducting exploratory factor analysis on dichotomous data. *Journal of*

Modern Applied Statistical Methods, *9*(2), 359–368. https://doi.org/10.22237/jmasm/1288584240

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 and Data Analysis*, *49*(4), 974–997. https://doi.org/10.1016/j.csda.2004.06.015

Pitombo, C. S., Kawamoto, E., & Sousa, A. J. (2011). An exploratory analysis of relationships between socioeconomic, land use, activity participation variables and travel patterns. *Transport Policy*, *18*(2), 347–357. https://doi.org/10.1016/j.tranpol.2010.10.010

Preacher, K. J., & MacCallum, R. C. (2003). Repairing Tom Swift's Electric Factor Analysis Machine. *Understanding Statistics*, *2*(1), 13–43. https://doi.org/10.1207/s15328031us0201_02

Ram, N., & Grimm, K. J. (2015). Growth Curve Modeling and Longitudinal Factor Analysis. In R. M. Lerner (Ed.), *Handbook of child psychology and developmental science* (pp. 1–31). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118963418.childpsy120

Rasch, G. (1960). *Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests.* Nielsen; Lydiche.

Revelle, W. (2021a). 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/

Revelle, W. (2021b). *Psych: Procedures for psychological, psychometric, and personality research* [R package version 2.1.6]. https://personality-project.org/r/psych/%20https://personality-project.org/r/psych-manual.pdf

Revelle, W., & Rocklin, T. (1979). Very simple structure: An alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research*, *14*(4), 403–414. https://doi.org/10.1207/s15327906mbr1404_2

Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on sijtsma. *Psychometrika*, *74*(1), 145–154. https://doi.org/10.1007/s11336-008-9102-z

Rummel, R. J. (1967). Understanding factor analysis. *Journal of Conflict Resolution*, *11*(4), 444–480. https://doi.org/10.1177/002200276701100405

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

Saccenti, E., & Timmerman, M. E. (2017). Considering Horn's Parallel Analysis from a Random Matrix Theory Point of View. *Psychometrika*, *82*(1), 186–209. https://doi.org/10.1007/s11336-016-9515-z

Saltelli, A. (2002). Sensitivity analysis for importance assessment. *Risk Analysis*, *22*(3), 579–590. https://doi.org/10.1111/0272-4332.00040

Santos Silva, J. M., & Tenreyro, S. (2011). Poisson: Some convergence issues. *Stata Journal*, *11*(2), 207–212. https://doi.org/10.1177/1536867x1101100203

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

Saucier, D. A., Stanford, A. J., Miller, S. S., Martens, A. L., Miller, A. K., Jones, T. L., McManus, J. L., & Burns, M. D. (2016). Masculine honor beliefs: Measurement and correlates. *Personality and Individual Differences*, *94*, 7–15. https://doi.org/10.1016/j.paid.2015.12.049

Schönemann, P. H., & Steiger, J. H. (1978). On the validity of indeterminate factor scores. *Bulletin of the Psychonomic Society*, *12*(4), 287–290. https://doi.org/10.3758/BF03329685

Schönemann, P. H., & Wang, M. M. (1972). Some new results on factor indeterminacy. *Psychometrika*, *37*(1), 61–91. https://doi.org/10.1007/BF02291413

Shackelford, T. K. (2005). An Evolutionary Psychological Perspective on Cultures of Honor. *Evolutionary Psychology*, *3*(1). https://doi.org/10.1177/147470490500300126

Shapiro, A., & Ten Berge, J. M. (2002). Statistical inference of minimum rank factor analysis. *Psychometrika*, *67*(1), 79–94. https://doi.org/10.1007/BF02294710

Shively, W. P. (2017). *The craft of political research*. Taylor; Francis. https://doi.org/10.4324/9781315269559

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of cronbach's alpha. *Psychometrika*, *74*(1), 107–120. https://doi.org/10.1007/s11336-008-9101-0

Sijtsma, K., & Van Der Ark, L. A. (2015). Conceptions of reliability revisited and practical recommendations. *Nursing Research*, *64*(2), 128–136. https://doi.org/10.1097/NNR.0000000000000077

Souza, M. G., Souza, B. C., Roazzi, A., & da Silva, E. S. (2017). Psychocultural mechanisms of the propensity toward criminal homicide: A multidimensional view of the Culture of Honor. *Frontiers in Psychology*, *8*(NOV). https://doi.org/10.3389/fpsyg.2017.01872

Steger, M. F. (2006). An illustration of issues in factor extraction and identification of dimensionality in psychological assessment data. *Journal of Personality Assessment*, *86*(3), 263–272. https://doi.org/10.1207/s15327752jpa8603_03

Stone, M. H., & Wright, B. D. (1999). Measurement Essentials. *Measurement*, 205.

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

Taherdoost, H., Sahibuddin, S., & Jalaliyoon, N. (2014). Exploratory factor analysis: Concepts and theory. *2nd International Conference on Mathematical, Computational and Statistical Sciences*, 375–382.

Tang, S. M. (2000). Impact factor model of Intranet adoption: An exploratory and empirical research. *Journal of Systems and Software*, *51*(3), 157–173. https://doi.org/10.1016/S0164-1212(99)00121-1

Ten Berge, J. M., & Hofstee, W. K. (1999). Coefficients alpha and reliabilities of unrotated and rotated components. *Psychometrika*, *64*(1), 83–90. https://doi.org/10.1007/BF02294321

Ten Berge, J. M., & Sočan, G. (2004). The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika*, *69*(4), 613–625. https://doi.org/10.1007/BF02289858

Thissen, D., Steinberg, L., & Mooney, J. A. (1989). Trace Lines for Testlets: A Use of Multiple-Categorical-Response Models. *Journal of Educational Measurement*, *26*(3), 247–260. https://doi.org/10.1111/j.1745-3984.1989.tb00331.x

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

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. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, *16*(2), 209–220. https://doi.org/10.1037/a0023353

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). https://doi.org/10.1002/9781118489772.ch11

Tisak, J., & Meredith, W. (1989). Exploratory longitudinal factor analysis in multiple populations. *Psychometrika*, *54*(2), 261–281. https://doi.org/10.1007/BF02294520

Tisak, J., & Meredith, W. (1990). Longitudinal Factor Analysis. In A. von Eye (Ed.), *Statistical methods in longitudinal research* (pp. 125–149). Academic Press. https://doi.org/10.1016/b978-0-12-724960-5.50009-3

Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: Congeneric and asymmetrical measurements. *Frontiers in Psychology*, *7*(MAY). https://doi.org/10.3389/fpsyg.2016.00769

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

Tucker, L. R., Koopman, R. F., & Linn, R. L. (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrika*, *34*(4), 421–459. https://doi.org/10.1007/BF02290601

Turner, N. E. (1998). The effect of common variance and structure pattern on random data eigenvalues: Implications for the accuracy of parallel analysis. *Educational and Psychological Measurement*, *58*(4), 541–568. https://doi.org/10.1177/0013164498058004001

Überla, K. (1971). *Faktorenanalyse* (2. Aufl). Springer.

van Alphen, A., Halfens, R., Hasman, A., & Imbos, T. (1994). Likert or Rasch? Nothing is more applicable than good theory. *Journal of Advanced Nursing*, *20*(1), 196–201. https://doi.org/10.1046/j.1365-2648.1994.20010196.x

van Osch, Y., Breugelmans, S. M., Zeelenberg, M., & Bölük, P. (2013). A different kind of honor culture: Family honor and aggression in Turks. *Group Processes and Intergroup Relations*, *16*(3), 334–344. https://doi.org/10.1177/1368430212467475

van Schuur, W. H. (1989). Unfolding the German Political Parties: A Description and Application of Multiple Unidimensional Unfolding. *Advances in Psychology*, *60*(100), 259–290. https://doi.org/10.1016/S0166-4115(08)60240-X

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct Explication through Factor or Component Analysis: A Review and Evaluation of Alternative Procedures for Determining the Number of Factors or Components. *Problems and solutions in human assessment* (pp. 41–71). https://doi.org/10.1007/978-1-4615-4397-8_3

Velicer, W. F., & Jackson, D. N. (1990). Component Analysis versus Common Factor Analysis: Some Issues in Selecting an Appropriate Procedure. *Multivariate Behavioral Research*, *25*(1), 1–28. https://doi.org/10.1207/s15327906mbr2501_1

Wardrop, J. L., & Loehlin, J. C. (1987). Latent Variable Models: An Introduction to Factor, Path, and Structural Analysis. *Journal of Educational Statistics*, *12*(4), 410. https://doi.org/10.2307/1165058

Wei, T., & Simko, V. (2021). *R package 'corrplot': Visualization of a correlation matrix* [(Version 0.90)]. https://github.com/taiyun/corrplot

Wickrama, K., Lee, T. K., O'Neal, C. W., & Lorenz, F. (2020). Longitudinal Confirmatory Factor Analysis and Curve-of-Factors Growth Curve Models. In K. A. S. Wickrama, T. K. Lee, C. O'Neal, & F. Lorenz (Eds.), *Higher-order growth curves and mixture modeling with mplus: A practical guide* (pp. 67–78). Routledge. https://doi.org/10.4324/9781315642741-11

Williams, B., Onsman, A., & Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Journal of Emergency Primary Health Care*, *8*(3), 1–13. https://doi.org/10.33151/ajp.8.3.93

Wood, J. M., Tataryn, D. J., & Gorsuch, R. L. (1996). Effects of under- and overextraction on principal axis factor analysis with varimax rotation. *Psychological Methods*, *1*(4), 354–365. https://doi.org/10.1037/1082-989X.1.4.354

Xitao, F. (1998). Item response theory and classical test theory : an empirical comparison of their item/peson statistics. *Educational and Psychological Measurement*, *58*(June).

Yang, Z., Peterson, R. T., & Cai, S. (2003). Services quality dimensions of Internet retailing: An exploratory analysis. *Journal of Services Marketing*, *17*(7), 685–700. https://doi.org/10.1108/08876040310501241

Yeomans, K. A., & Golder, P. A. (1982). The Guttman-Kaiser Criterion as a Predictor of the Number of Common Factors. *The Statistician*, *31*(3), 221. https://doi.org/10.2307/2987988

Yong, A. G., & 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. https://doi.org/10.20982/tqmp.09.2.p079

Zeller, R. A., & Carmines, E. G. (2009). *Measurement in the social sciences : the link between theory and data*. Cambridge University Press.

Zhao, N. (2009). The minimum sample size in factor analysis. *Wiki of Encorelab Toronto*, *250*(1992), 1–7.

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for likert rating scales. *Journal of Modern Applied Statistical Methods*, *6*(1), 21–29. https://doi.org/10.22237/jmasm/1177992180

Zwick, W. R., & Velicer, W. F. (1982). Factors influencing four rules for determining the number of components to retain. *Multivariate Behavioral Research*, *17*(2). https://doi.org/10.1207/s15327906mbr1702_5

Zwick, W. R., & Velicer, W. F. (1986a). Comparison of Five Rules for Determining the Number of Components to Retain. *Psychological Bulletin*, *99*(3), 432–442. https://doi.org/10.1037/0033-2909.99.3.432

Zwick, W. R., & Velicer, W. F. (1986b). Comparison of Five Rules for Determining the Number of Components to Retain. *Psychological Bulletin*, *99*(3), 432–442. https://doi.org/10.1037/0033-2909.99.3.432