

Suitbert Ertel

Healing an Ailing Model



Universitätsverlag Göttingen

Suitbert Ertel
Factor Analysis

This work is licensed under the
[Creative Commons](#) License 3.0 “by-sa”,
allowing you to download, distribute and print the
document in a few copies for private or educational
use, given that the document stays unchanged
and the creator is mentioned.



Published in 2013 by Universitätsverlag Göttingen

Suitbert Ertel

Factor Analysis

Healing an Ailing Model



Universitätsverlag Göttingen
2013

Bibliographische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

Address of the Author

Suitbert Ertel

e-mail: sertel@uni-goettingen.de

This work is protected by German Intellectual Property Right Law. It is also available as an Open Access version through the publisher's homepage and the Online Catalogue of the State and University Library of Goettingen (<http://www.sub.uni-goettingen.de>). Users of the free online version are invited to read, download and distribute it.

Set and layout: Franziska Lorenz

Cover: Matthias Bellmann

© 2013 Universitätsverlag Göttingen

<http://univerlag.uni-goettingen.de>

ISBN: 978-3-86395-133-7

**Dedicated to Elisabeth
With Gratitude**

Contents

Abbreviations7

Foreword.....9

Preface11

Chapter 1 Critique of the simple structure doctrine.....13

Introduction.....13

01. The present state of factor analytical research15

02. The doctrine of simple structure (SS).....19

03. The fallacy’s consequences.....21

04. Detailed error analysis.....23

05. Reorientation.....27

06. Where did we go wrong?32

07. Unheeded critical voices34

08. Can non-factorial procedures take us forward?36

Discussion of chapter 1 and outlook	41
Chapter 2 Finding complex structures	43
Introduction	43
Questions and Test Runs	45
Test run 1: Evaluating phoneme similarities	49
Test run 2: Similarity judgments of British coins	53
Test run 3: Differentiation of response styles at responding to questionnaires.	59
Test run 4: Semantic features of kinship terms	62
Test run 5: Intellectual development in childhood	66
Test run 6: Body size and body shape in cattle.....	69
Test run 7: Intelligence tests and performance tests.....	71
Test run 8: Psychophysiological activity indicators (Data: Köhler &).	73
Test run 9: Knowledge test with varying test methods.	75
Test run 10: Self-assessment and external assessment of children	77
Discussion of chapter 2.....	79
Chapter 3 Decathlon data under analysis	85
Introduction	85
Description of data	86
Results: Factor analyses	88
1. Interpreting Varimin factors.....	90
2. Attempt at an interpretation of Varimax factors.....	92
3. Attempt at an interpretation of initial factors.....	94
4. Expert rankings for validating Varimin factors	95
Discussion of chapter 3.....	97
Chapter 4 Intelligence data under analysis	101
Study I: Varimin analysis of IST factors	102
Data analysis.....	104
Results with comments.....	105

Summary of Study I.....	108
Study II: Validations of Varimin-rotated IST' factors.....	109
Objective.....	109
Validation 1: School performance.....	109
Validation 2: Culture-free IQ Test CFT, spelling and numeracy test.....	110
Validation 3: Culture-free IQ test FRT'	115
Summary of Study II	116
Discussion of chapter 4	117
Chapter 5 Varimin factors from Big Five personality data	119
Point of departure.....	119
Material and analysis.....	120
Towards an interpretation of Varimin factors of personality	123
More preliminaries.....	123
Varimin factor interpretations	126
Concluding remarks.....	129
The interpretations of Varimin factors on probation	130
Re-interpreting Varimax factors by profiles of Varimin factors.....	132
Discussion of chapter 5	133
Insight and Outlook.....	135
Review of chapters 1 to 5.....	137
Appendix	141
References.....	145

Abbreviations

CFA	Confirmatory factor analysis
CFT	Culture-free intelligence test
CS	Complex structure
CSM	Complex structure modelling
DV	Dependent variable
EFA	Exploratory factor analysis
ERS	Empirical relations system
ESEM	Exploratory structural equation modelling
FA	Factor analysis
FRS	Formal relations system
FRT	Figure Reasoning Test
g	General factor
IST	Intelligenz-Struktur-Test
IV	Independent variable
MDS	Multidimensional scaling
MTMM	Multitrait-multimethod (analysis)
NMDS	Nonmetric multidimensional scaling
PCA	Principal component analysis
SD	Social desirability
SEM	Structural equation modelling
SS	Simple structure
SSM	Simple structure modelling

Foreword

Paul Barrett

This book is about factor analysis. It explains what happens when you model covariances among sets of variables with a method sensitive to underlying complex relations. An idealized model that forces a structural simplicity onto such relations, is considered as leading us astray. Thus stripping away all the technicalities, it comes down to a simple question: Do you force a model onto data irrespective of reality, or do you let the data speak for itself?

Make no mistake: This question contrasts a more realistic perception of human behaviours, cognitions and attributes against a hypothetical statistical ideal. The statistical technique of factor extraction is considered healthy. The ailment arises when the investigator chooses to construct a model of factor loadings corresponding to Thurstone's Simple Structure.

The defining principle of Simple Structure is that variables should load highly on one single factor and near-zero (or zero) on all other factors. This kind of solution produces clusters of homogenous variables and interpretations of factorial meanings constructed from the content of variables within clusters. Variables that possess one or more sizeable loadings in a factor analysis are selected/rejected from solutions on the basis of their complexity. They are usually rejected completely from a solution if they possess "cross-loadings" or at least two high-loading values across two factors. The attraction of such solutions is that they appear to be easily interpretable. However, the drawback is that they may bear no relation to the reality of the underlying explanatory processes. Removing the complexity inherent in many psychological attribute interrelations is perhaps the very opposite of what the social scientist must now begin to consider.

This is the fundamental thesis of Suitbert Ertel's propositions: Do not try to force simplicity on what is (or is observed to be) complex. Instead, model the complexity itself (if present) and work with structured variables that account for that complexity. To achieve this goal, the methodology created and set out in this book is called *Varimin*. If there is no complexity in the covariance patterns among variables, a simple solution will be found. But if complexity among variable relations is present, *Varimin* will produce factors that account for the inherent complexity.

As Suitbert shows in several chapters devoted to analysing several kinds of variables (including those within the fields of cognitive ability and personality), the consequences of using *Varimin* are theoretically profound. *Varimin* factors are no longer conceived as ultimate dimensions, but as components of multifactorial variables. They thus seem to align more with what we know and observe within

other areas of psychological investigation, far more so than what is generally expected from seemingly homogeneous factors of Simple Structure.

Clearly, if you work within Simple Structure-constrained analyses (whether using exploratory or confirmatory factor analysis), you are likely to be highly sceptical of *Varimin*'s logic and approach. However, it is interesting to ask yourself the reasons for that scepticism: Are these based more on scientific considerations and observations of the phenomenon of interest, or just habit and 'status-quo' recommendations?

Thurstone created Simple Structure at a time when complexity within data relations was seen as "a problem in need of a solution", partly because there were no methods or technologies to deal computationally with complexity. Just look how that view has changed in recent years, as methods across many sciences now routinely deal with complexity as a feature of multivariable interrelations. The entire field of complex systems theory is founded on *systems* approaches to understanding phenomena . as in systems biology, business dynamics and the newer network models for epidemiological and psychological concepts, such as mental disease comorbidity and personality.

While a single book cannot, by itself, change an entire field of thinking and endeavour, it can pose and answer the big questions. Within these pages, there is enough content matter and avenues of investigation to kick-start several Masters and PhD theses, exploring the consequences of *Varimin* in areas where investigators are more familiar with representing data structures. One cannot help being curious about *Varimin* because of its impact upon how we might theorise about the nature of factors in the future. Even if you are sceptical this is a fascinating *volte face* proposition and methodology in its own right, which just might, over time, become the new established method for factor-analytic investigations.

Thurstone's Simple Structure was the 20th-century response of a pioneering psychologist to the challenge of reducing complexity within the factor analysis of questionnaire items and other kinds of variables. *Varimin* and Complex Structure is the 21st-century response from another pioneering psychologist, namely Suitbert Ertel, to the challenge of the complexities inherent in the functioning of human cognitive and other multivariable systems.

Preface

Why is factor analysis considered an ailing model in this book? I deem Simple Structure, a basic principle for factor rotation, introduced by Thurstone, as mistaken. In his foreword, Paul Barrett provides grounds for my view. I should add that two detrimental conditions have prolonged the methodical ailment: Not being able to diagnose the actual reasons for the sickly symptoms under which this model has suffered and the widespread belief that Simple Structure is an indispensable ideal. Thurstone's tenet states that factors are uninterpretable without rotating them. By rotation towards simplicity – he holds – individual variables should obtain as few factorial loadings as possible. Users of his procedure have never questioned this. Varimax rotation is the widely preferred technique, but the outcomes are misleading, which as Barrett has intimated in his preface, is the fundamental thesis of this book.

Readers who are trained by Thurstone's verbally impressive principle may be irritated that I dare reject it in the first place and may demand an alternative. This will be provided after replacing the ideal of simplicity – which cannot be achieved by Thurstone's rotation anyway – with complexity, the aim of rotating extracted factors by letting individual variables display as many factorial components as are suggested by empirical data. The alternative is called *Varimin*.

By doing the opposite of what Thurstone considered necessary for grasping factorial meanings, it may appear that we arrive at a bewildering quandary. In chapters #1 and #2 an alternative method called “minimal pair comparison” is introduced, a procedure imported from linguistics. It will be shown that the meaning of *Varimin* factors can safely be discerned. Two variables whose loadings are equal (or nearly equal) for N-1 factors, are successively paired while the loadings on only one factor are extremely different or, ideally, opposite in sign. A difference in meaning between the two paired variables must then be considered as due only to the factor whose loadings on the two variables are extremely different. *Varimin* factors eventually turn out to be latent components. Factor rotation? Yes. For improving interpretability? Yes. However, this is only achievable by *Varimin* and is impeded by Varimax, as will be demonstrated with ten test runs #2, in chapter #3 with sports data, #4 with intelligence data, and #5 with personality data. The benefit of the new paradigm of factor analysis may also be discovered with non-psychological multivariate data. I do hope that the reader will attempt to replicate findings as illustrated in my book so that he/she may also discover that the future of factor analysis has indeed a hopeful prognosis.

Acknowledgements

I was encouraged and inspired in my renegade methodical search by the advice and constructive criticism and comments of Paul Barrett, André Beauducel, Elisabeth Cott, Herbert Götzl, Gerd Lüer, Pierre Sachse, and Tatjana Schnell. Uwe Engeland implemented my ideas of a program that models latent complexity. Jürgen Guthke and his assistant, Barbara Seiwald, provided the first opportunity to test this program in their research. They had failed after applying conventional SSM methods, but using CSM they obtained the predicted solution. Two monographs in German (Ertel, 2011a,b) preceded this English translation which has been recommended by inspired readers of my German publications. The translation was initiated and encouraged by Matthias Bellmann who supported this project with valuable editorial and time-consuming technical improvements. Jürgen Hecker and Werner Rawe translated my German into English.

Chapter 1

Critique of the simple structure doctrine

There is no place for dogma in science. The scientist is free to ask any question, to seek any evidence, to correct any error. ... Dogmatism has found itself incompatible with the progress of science. J. Robert Oppenheimer. (1904–1967)

Introduction

In chapter 1 I shall examine what I deem a serious error by factor analysts adhering to Thurstone's lead. Ever since its creation in 1935/47, his simple structure (SS) principle has been as problematic as it is attractive. How did methodologists deal with this ambivalence? Were there no critics who recognised the fallacy that SS had introduced? Yes, there were a few, but they were largely ignored. Why was that?

Researchers eventually changed track, i.e., abandoned exploratory research that kept on producing questionable results. They switched to confirmatory methods, above all to structural equation modelling. Was that progress? Hardly, because even structural equation users stuck to simple structure modelling, SSM, which I believe to be the main source of errors. Why did they cling to SSM? Because no-

body dared touch SSM which had become a doctrine of statistical reasoning that nobody saw through. How could this happen?

Readers who have never doubted the validity of the SS principle may dislike my attempt to prove them wrong. I am aware that theoretical considerations alone in this first chapter will hardly change deep-rooted convictions, not least because creating truly simple structures seems to be an undeniable goal of all science, not only of factor analysis.

I hope, nevertheless, that my criticism will make sense to you as you read chapters 2 to 5. I want to show what can be achieved by factorial analysis when it is freed from the constraints of SSM and when complexity is revealed by appropriate statistical decisions.

Chapters 1, 2 and 4 were first published in 2009 and 2010 in the little known *Psychologie des Alltagshandelns* and were reworked for a German monograph Ertel (2011a). The content of chapter 3 was published in *Personality and Individual Differences* (Ertel, 2011c) and revised for this book. Chapter 5 (on personality) is based on another German monograph (Ertel, 2011b), condensing and adapting the message of that more extended work.

Chapter 1 gives a theoretical overview. Obvious weaknesses of previous methodical reasoning are discussed. They could have been avoided if analytical procedures had been applied more prudently and if common sense had also been given a say in the matter. Common sense tells us that complexity of conditions of manifest behaviour is self-evident. I do not only voice my own critique but also quote the supporting opinions of many others which unfortunately have been and are being studiously ignored by the majority of experts in this field.

In chapter 2, *complex structure modelling*, CSM, is presented as an alternative to conventional *simple structure modelling*, SSM. Varimax rotation is replaced with *Varimin* rotation. The use of *Varimin* is explained by giving ten empirical examples using data from published factor analyses.

In chapter 3, data from decathlon (ten physical events performed by Olympic athletes) are subjected to *Varimin* analysis. I considered that an interpretation of factors of physical sports events was easier compared with the interpretation of factors of mental performance which are more commonly analysed, but more liable to controversy.

In chapter 4, intelligence data obtained from a study using the well-known German test of general intelligence are analysed by *Varimin* and for comparison, also by Varimax. Does the commonly accepted SS-based distinction between fluid and crystallised intelligence remain valid if identical data are subjected to *Varimin* rotation?

Chapter 5 conveys the gist of my book by a CS-based analysis of personality data (Ertel, 2011b), where the Big Five personality data are selected as the main focus. An interpretation of factors obtained from such data is more demanding and cannot avoid first attempts at theoretical reasoning. Pertinent discussions

expounded in the German monograph (Ertel, 2011b) might be helpful and will be more comprehensively translated in a future publication.

The five chapters of this book may be read independently. There are some overlaps among chapters where basic issues are viewed from complementary perspectives. At the end of the book is an abstract for each chapter. Readers may use, by clicking URL <http://www.varimin.com>, Dr Uwe Engeland's statistical program "factor analysis" online which allows application of principal component analysis with *Varimin* and Varimax rotations. A user manual is available on this website.

01. The present state of factor analytical research

Detrimental characterisations and metaphors employed by disappointed authors are symptomatic of the chronic anomalies encountered in factor analytical research: *"uneasiness in factor analysis"* (Kallina, 1967); *"alarming lack of commitment"*; *"subjectivity in factor analysis"* (Horn, 1967); *"ambivalence of factorial research"* (Meili, 1968); *"destruction of generality"* (Davies, 1971); *"product of chance and imaginary evidence"* (Greif, 1972); *"nonsensical effort"* (Revenstorf, 1978); *"ambiguity of factorial rotation"* (Buse & Pawlik, 1978); *"dubious legacy"* (Schönemann, 1981); *"faktoranalytits"* (Jäger & Hörmann, 1981); *"myth of factor analysis"* (Lenk, 1983); *"factors are fictions"* (Revelle, 1983); *"morass of factor analysis"* (Eysenck, 1992); *"psychopathology of factor indeterminacy"* (Schönemann, 1996)¹; *"pathology of psychometrics"* (Borsboom, 2003).

The shortcomings of FA, however, are trivialised by most users and quite happily buried under optimism; they thumb their nose at critics and maintain that there exist, after all, "significant results". This is opined, for example, by Pawlik (1977) in a comprehensive German overview of the first decades of FA research.

But, say critics prepared to face the dilemma, FA research has miscarried: *"Exploratory factor analysis has never been developed to anything approaching its full promise and potential, despite the eighty-year history of its efforts ..."* (Yates, 1987, p. 325). In an overview of "fifty years of test theory", Blinkhorn (1997) concludes that neither the "considerable technical strides" made during the past decades nor the "well-known contributions of Jöreskog and McDonalds" basically changed the dilemma:

¹ Schönemann (1981) and Steiger (Steiger & Schönemann, 1975), after Guttman (1955), belong to the middle generation of methodologists critical of FA. Their criticism was harsh ("theoretical problems", "users are generally uninformed about the defects of this model", pp. 175, 188), but they did not focus on the *simple structure* principle. Instead they confined themselves to the "indeterminacy" of factorial dimensions and their "lack of identifiability". Following the re-analysis of 13 published FA studies that resulted in devastating criticism of these studies (Schönemann & Wang, 1972), Schönemann and Steiger (1976) developed an alternative method for multivariate data reduction (*Regression Component Decomposition*, RCD). It promised greater conceptual clarity and computational efficiency plus the possibility of model falsification. But this approach remained unnoticed given the success of Thurstone's "multiple factor analysis". Admittedly, the alternative approach did not offer new insights into the transformation of "components" determined by RCD. Moreover, the results of RCDs did not seem much different from those supplied by Thurstone's factor analysis.

“How curious ... that we are so little further forward in our understanding of the psychology of individual differences as a result of these advances ... Can anyone identify a single publication in the last 50 years in which the use of factor analysis has led to counter-intuitive, or surprising, or genuinely enlightening outcomes?” (Blinkhorn, 1997, p. 181). Already 50 years ago one could and should have noted what Schönemann reported retrospectively (1994) about Louis Guttman, who had delivered a “eulogy” for multiple factor analysis in 1955: *“It was left to Louis Guttman to read the eulogy (p. 209, p. 406): The era of Multiple Factor Analysis had come to an end – for knowledgeable people at any rate. ... It was logical, then, to ask: What lies ahead for Factor Analysis? (Guttman, 1958). He answered it with a vision that challenged habits of thought that had led nowhere. This vision he kept pursuing for the rest of his life.”* Schönemann and Borg (1996, p. 249) took stock: *“Today we know that the explorative factor analysis era that Thurstone heralded brought very few lasting insights.”*²

Two calamitous results of exploratory factor analysis (EFA) are particularly deplorable:

- **EFA research engendered a myriad of constructs in psychology and thus produced the opposite of what it set out to achieve**

EFA was supposed to describe the multitudes of correlating manifest variables parsimoniously and advantageously. This was thought to be achieved by extracting from them a small number of factorial variables which were assigned the role of *latent dimensions*³.

Decades of EFA research produced an inexhaustible number of latent dimensions supposedly underlying the observable variables. At the *11th European Conference on Personality* (2002) in Jena, Lee Sechrest pointed out the glut of construct variables in psychology, citing an author who had counted 7800. Many are new creations of EFAs. Is Sechrest’s number unrealistic? An internet search of article titles containing the word “scale” from the PubMed database provided 889 differ-

² Moosbrugger and Hartig (2002) and Fabrigar et al. (1999) delivered survey papers on EFA research methods with an implementation orientation – most of them without seminal criticism. Their papers were preceded by articles with similar objectives: Stevenson (1993), Tinsley & Tinsley (1987), Ford et al. (1986), Glass & Taylor (1966), Cattell (1965), Peel (1953). The text books most often quoted on EFA are: Comrey & Lee (1992), Child (2006), Gorsuch (1983), Harman (1976), Mulaik (1972), Weber (1978), Revenstorf (1976), Überla (1971). Lienert (1969) has an introduction as an appendix in a textbook, and Bortz (2005, 1977) devotes one textbook chapter (chapter15) to factor analysis (FA).

³ By applying the term “dimension”, a claim is staked for a metric that was never challenged. Just as the three dimensions of Euclidean space serve to locate objects in space, it was thought that the primary variables of psychological observation could be positioned with factorially acquired “dimensions”. This is overtaxing of the limits of the space dimension metaphor (the same goes for the phrase “semantic space” by C. E. Osgood). Thus the term “dimension”, while legitimate in mathematics, is misleading and superfluous when merely a naming of sources of variance is required.

ent scale denominations ranging from the *Abel and Becker Cognition Scale* to the *Zung Self Rating Anxiety Scale*. Most scales were compiled or adapted by FA. Thus, if one scale delivers on average, say, three or four factors, some 3000 factorial constructs were generated in clinical psychology and medicine alone.

Moreover, many scales and derived constructs have been developed in non-clinical differential psychology. Every issue of the journal *Personality and Individual Differences* offers new material, so that the guestimate made earlier seems realistic: “*The idle practice of producing new personality scales continues unabated, making it less likely that they will ever arrive in the promised land of the paradigm which alone would endow our efforts with scientific respectability*” (Eysenck, 1992, p. 672).⁴

- **SS-based constructs identified as dimensions lack theoretical connections**

Factorial constructs obtained from SSM-oriented analyses are unrelated and thus isolated from one another; i.e., they form mere aggregates. Only if they enter into relationships may constructs be conceived as components of some processual whole. As early as 1956 Stephenson wrote: “... *simple structure may have resulted in an analysis into too many unrelated, and UNRELATABLE, primaries [primary factors]*” (Stephenson, 1956, p. 6; emphasis by S.E.). Andresen’s (1998) comprehensive critical and historic overview of EFA personality research leaves behind a chaotic impression.

The Big Five factor model, developed since the 1990s in personality research, was welcomed enthusiastically and soon achieved reputation. Did it remedy the theoretical shortcomings? No, it merely showcased five middling invariant dimensions in the “chaotic plethora of personality constructs” (Funder, 2001, p. 200). The invariance of constructs, however, does not signify validity since inferior constructs may be as invariant as high quality ones.

Some proponents of the Big Five model believe that their factors were analogous to chemical elements (this idea seems to have started with Goldberg (1981)). Such optimism is out of place. The discovery of chemical elements in the nineteenth century introduced a scientific revolution⁵. Advocates of the Big Five

⁴ Rutkowski (1974) tried to capture the totality of typological constructions in differential psychology, and not only those of FA origin. Sponcel (1998) commented as follows: “*Worldwide, there are more than 1,000 personality or character typologies (Rutkowski, 1974). Most of them are probably ... contentious. Many have disappeared in cultural or scientific history. Many overlap. It seems as if a random number of constructions are possible – depending on differing goals and purposes.*” Gigerenzer and Strube (1987, p. 85) arrived at a similar conclusion: “*It is the crux of factor analytical research to have come up with so many ‘accepted’ personality factors that even simple dichotomisation of dimensions leaves us with a number for the resulting available high-order quadrants of approximately 2^{50} ... which is around four hundred thousand times the population of Earth.*”

⁵ Blinkhorn (1997, p. 180) criticises the excessive hopes held by the pioneers of FA: “*The words they use, for example ‘primary mental abilities’ (Thurstone) or ‘source traits’ (R. B. Cattell), are witness to the*

model claimed that complex differential psychological constructs, like molecules made of chemical elements, could be put together using five element-like dimensions. The notion arose that future extractions of factors in the domain of personality would only be legitimate if they correlated with the Big Five⁶.

However, *"this comparison [with chemical elements] did not hold water"* (Lukesch & Kleiter, 1974, p. 294). H, He, C, Ca, N, etc. have a functionally definable place within the periodic table. Atoms form molecules because of bonding properties caused by the number of protons in the nuclei, the density of electrons, etc. In brief, chemical elements are related by their components and compositions. The Big Five personality "dimensions", however, do not exhibit components that would allow an assessment of similarities and differences. Very few observers take exception to this general belief (for example Briggs, 1989, and Block, 1995⁷).

H. J. Eysenck was irritated by EFA research that lacked theoretical underpinnings and accused Big Five researchers of not transcending superficial taxonomic goals. In so doing, said Eysenck, they remained at the psychometric surface instead of developing biologically interpretable models of relatedness (Eysenck, 1992, 1997).

Eysenck attempted to conjoin the three dimensions of his PEN model (psychoticism, extraversion, neuroticism). He postulated differential cortical areas assigning them neuro-psychological roles that were supposed to have functional relationship. As welcome as Eysenck's aim may have been in principle, he did not achieve it⁸. He did not recognise the true cause of the lamented *"morass of factor analysis"* (1992, p. 672), it could not be found where he was looking for it.

Factor analytical data analysis has also been conducted in numerous non-psychological disciplines (cf. Rummel, 1970, Reymont & Jöreskog, 1993, and Figure 1.01), and it is not uncommon that discomfort is also voiced there. Earth scientist Davies, for example, who methodically utilised SSM-orientated EFA, comments: *"By Varimax] we may be butchering our results; cutting up the body of generality into a*

faith and trust placed in factor analysis as revealing the psychological analogue of the periodic table of elements, or the list of subatomic particles."

⁶ Ozer and Reise (1994) *"characterized the Big Five as the 'latitude and longitude' along which any new personality construct should be routinely mapped"* (Funder, 2001, p. 200).

⁷ Briggs (1989): *"... a coherent and falsifiable explanation for the five factors has yet to be put forward. There is no theoretical reason why it should be these five rather than some other five."* (p. 249). *"The structure of trait attributions may not correspond straightforwardly to the deep structure or neurophysiological basis of human tendencies."* (p. 250). *"Perhaps the critical step in elucidating these concepts [interpreting the five factors] ... is the specification of their exact nature: What are the elements or components of each factor? How are they interrelated?"* (p. 253). Block (1995) quotes Briggs and criticises more specifically: *"No functioning psychological 'system', with its rules and bounds, is designated or implied by the 'Big Five' formulation; it does not offer a sense of what goes on within the structured, motivation-processing, system-maintaining individual."* (p. 188).

⁸ *"How should the Big-5-or-6 be understood in psychological terms? Sadly, despite many years of research – especially into extraversion – the picture is still very unclear (see e.g., A. Gale & M. W. Eysenck, 1992, Handbook of Individual Differences: Biological Perspectives; G. Matthews, 1993, in A. Smith & D. Jones, Factors Affecting Human Performance.). Here are some possibilities that still look viable, yet falsifiable."*

set of unrelated fragments without ever realizing that these fragments can ever be considered as part of a larger entity” (Davies, 1971: p. 113). Davies repeatedly characterises the effect the Varimax rotation has on factorised data as “destructive”. This will be examined more closely in the following section.

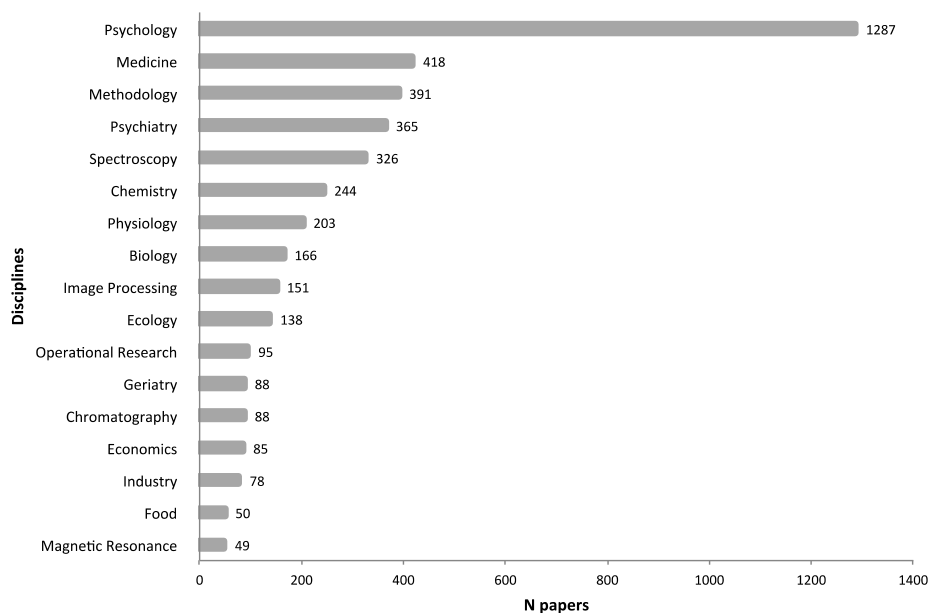


Figure 1.01: Papers on factor analysis, by discipline, identified by Kaplunowsky (2007).

02. The doctrine of simple structure (SS)

The above account of the situation of factor analytical research helps understand where the calamity comes from. An “unease in factor analysis” is generally ascribed to an arbitrariness of procedural decision taking. Arbitrariness occurs when variables for correlations are selected, when samples of individuals are formed, when the number factors to be extracted are determined, when the choice between orthogonal or oblique rotation is made, and when one rotation procedure is selected from among a large number of options (cf. Finch & West 1997, p. 464 et sqq.). To me the effect of such arbitrariness on the results of FA appears negligible compared with what caused EFA’s most serious defect.⁹

⁹ Velicer (1977) found that extraction procedures of maximum likelihood, image analysis, and principal components analysis had “extremely similar” (p.18) results when tested in nine sets of test data. Using a representative data set of trait descriptions, Goldberg (1990) was able to deliver an almost invariant reproduction of the Big Five factor model, regardless of whether he varied the method of factor extraction (principal components, principal factors, alpha-factoring,

The idea of SS is widely known and is considered “intuitively compelling” (Kaiser, 1958, p. 188). In the coordinate system of initial factors, clouds of points represent variables (cf. Figure 1.02A). Statisticians draw regression lines through such clouds so that squared distances between regression lines and variable points are reduced to a minimum. To achieve this, the system’s coordinates are rotated in such a way that they coincide with the regression vector (Figure 1.02B). Since two coordinates are always rotated simultaneously, the convergence of both coordinates will be optimised simultaneously. Hard-core SSM methodologists are even more radical than proponents of orthogonal rotation because they opt for disadvantageous oblique rotations just to come closer to SS¹⁰.

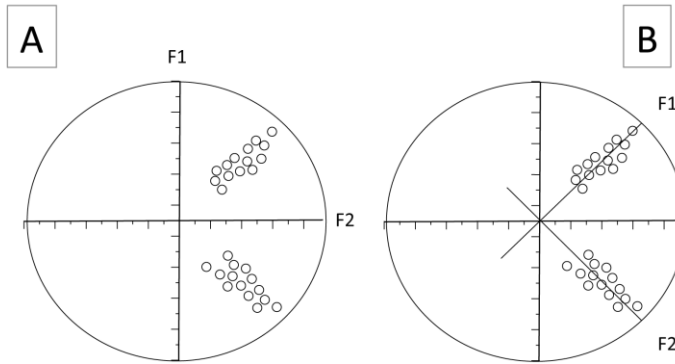


Figure 1.02: Two clusters of variables in an initial factor system F_1 , F_2 (A) and following orthogonal simple structure rotation (B).

SS was, incidentally, considered an inevitable continuation of the simplification strategy that reduced the large diversity of variables to a smaller number of fac-

image-factoring, maximum-likelihood), or whether he applied orthogonal rotation (Varimax), or oblique rotation (Oblimin). But without exception these methodical variations were conducted with SS orientation. Alternatives to SS were not considered, although at least Goldberg seemed to consider a renunciation of SS, according to an Eysenck quote (1981): “*Correlational psychology cannot in the nature of things come up with objective, universally agreed dimensions or categories; there are innumerable, mathematically equivalent ways of rotating factors, for instance, and no statistical magic key (not even simple structure) can close the door on alternative solutions* (p. 43).”

¹⁰ In an overview of factor rotation’s analytical methods, Warburton (1963) describes the logic of mathematical parsimony in simple terms: “He [Ferguson] approached the problem of parsimony by considering a single variable, represented by a point, and asking himself what was its most parsimonious description. He suggested that, intuitively, the most parsimonious description results when one of the axes passes through the point. It is seen that when the reference frame is rotated so that one of the axes approaches the point, the product of the two co-ordinates grows smaller ...” (p. 169).

tors: “In the factor problem, striving for simplicity aimed at coming up with the smallest possible number of factors ... The rotation problem ... seeks to design the correlation of variables and factors ... as simply as possible within the ... predetermined ... factor space” (Überla, 1971, p. 176). The simplification principle of “less is better”, which factor extraction rightfully employs, is carried over to subsequent factor rotations, where it is, however, no longer valid or legitimate. It is tantamount to transferring a principle that helped solve one problem to all other problems.

03. The fallacy’s consequences

The simplicity enforced on factorial data modelling is the result of formal reasoning but with damaging consequences for the representation of reality that the model should bring forth. Inadvertently, Überla (1971) provides an instructive example: 90 men had their systolic and diastolic blood pressure measured six times within half an hour. The initial results of a FA of the $2 \times 6 = 12$ variables are shown in Figure 1.03A. Figure 1.03B shows the Varimax results for these factors, orthogonally optimised by SS. Überla claims the solution depicted in 1.03B is the only sensible one, arguing that the measurement variations are “*basically determined by two parameters, the systolic and the diastolic blood pressure*” (p. 265).

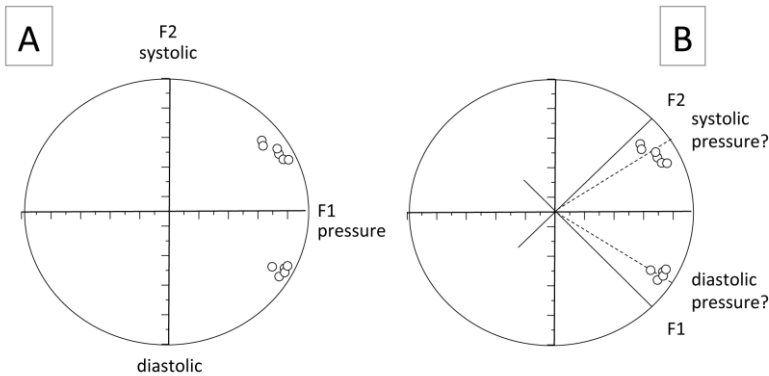


Figure 1.03: Initial factor solution of blood pressure data (Überla, 1971) (A) and rotation to simple structure (B) — orthogonal - - - oblique.

But in fact, “systolic” and “diastolic” are not “parameters”, but merely two different measuring instances reflecting varying conditions of pressure. Blood pressure is a unit that varies with location and time. At the occurrence of systole and diastole, blood pressure varies just as atmospheric pressure varies with geographical altitudes. Regulation of blood pressure by therapeutic drugs does not specify systolic or diastolic “parameters”, they aim at improving blood pressure as a systemic feature (cf. *Journal of Human Hypertension and Hypotension*)¹¹. A FA of Überla’s twelve blood pressure variables would reflect the physiological facts best if the highly correlating variance of “systolic-diastolic”, i.e., the generality of inter- and intra-individual differences of blood pressure, were represented by a general factor. The first layout of the required general factor is already provided by the initial solution, as shown in Figure 1.03A, but disappears by an SSM rotation (1.03B).

The second initial factor, however, suggests some difference between “systolic” and “diastolic” measures. Physiological reasons give rise to a second initial factor, because the two blood pressure measurements show additional variance under particular conditions. Systolic data generally react somewhat more strongly to external influences than diastolic data. With circadian measurement readings, the systolic variance is greater than the diastolic variance (Halberg, 1980, Fig. 8, p. 552). For age-related hypertension measures, systolic values are generally larger than diastolic values, etc. Obviously, this variance of the two measurements is present in the initial solution by a second factor explaining, as is to be expected, a smaller percentage than the first factor.

A similar situation arises with factorial intelligence research. For intelligence, too, a general factor g is theoretically demanded. In most instances a notable approximation to g exists already in an initial solution. SSM makes g disappear. But since a g -factor is expected, g is often reconstructed in a cumbersome way¹², for example, by orthogonalising obliquely rotated primary factors using a particular procedure designed by Schmid-Leiman (1957). Thus “second-order” factors are

¹¹ When internists employed a common SS-oriented EFA to determine the insulin resistance syndrome (ISR), of which blood pressure is a part, much to their chagrin the researchers discovered that systolic and diastolic blood pressure did not load on the same factor, although “... *systolic and diastolic blood pressure are more strongly associated with each other than they are with other components of the insulin resistance syndrome, something which most clinicians would expect*” (Lawlor et al., 2004, p. 6). They got around this problem by only leaving one of the two blood pressure variables in the set of ISR variables to avoid the methodical artefact: “*Therefore, the evidence for inclusion or exclusion of hypertension in the definition of the syndrome is based on whether one or two blood pressure measurements are included in the model rather than on any sound clinical or pathophysiological reasoning*” (Lawlor et al., 2004, p. 1016).

¹² Even in 1958 this procedure was caricatured by British authors who did not feel bound by the Thurstone doctrine: “*It has been said that what was thus thrown out of the door [the general factor] returned through the window: for correlated factors in turn give rise to a second order factor, and this is virtually the general factor of the centroid ... under another name*” (Hamilton, 1958, p. 167).

derived and a hierarchy of factors is introduced while theoretical follow-up problems and hazards of consequences are ignored¹³.

Jensen very succinctly describes the effect of an SSM rotation on the initial first factor in intelligence test analyses: *“The tests are all positively correlated and therefore all had some factor in common - a general factor, or g. The general factor that was so prominent in the analysis depicted in Figure 3.3 [showing the initial solution] seems to have disappeared from Figure 3.4 [showing the simple structure-rotated solution] as a result of rotating the factor axes. Actually, it has simply been dispersed (or redistributed) among the rotated factors ... So if you ask where g went, the answer is that it has been divided up and lies ‘hidden’ among all of the tests’ smaller loadings on all of the orthogonally rotated factors. Its variance has not disappeared; it has simply been obscured by being dispersed throughout the whole factor matrix”* (Jensen, 1998, p. 66).

Pett et al. (2003, p. 143) quote authors dealing with this problem: *“Nunnally and Bernstein (1994) ... warn against prematurely concluding that, based on a Varimax solution, a general factor is absent, because Varimax is designed to eliminate general factors (Gorsuch, 1983). Comrey and Lee (1992) suggest that the researcher avoid including too many factors in a Varimax rotation solution because it tends to overinflate the importance of lesser factors. Although the authors do not indicate how many factors are ‘too many’ they point out that trial and error is the only way to arrive at the appropriate number of factors.”*

The fact that an identified error is readily corrected by haphazard trial and error operations shows how constricting the predicament has become.

04. Detailed error analysis

- The demotion of initial factor structures

Thurstone and all those adhering to his “American school of thought” considered an initial orthogonal FA solution to be *“generally uninterpretable”* (Überla, 1971, p. 175) and unstable. Regarding the alleged instability of initial factors, Überla remarks: *“[The initial factors] ... change from sample to sample ... because new variables often shift weights significantly and, in so doing, change the position of the axes”* (p. 175).

But the claim that initial factors lack invariance is unsubstantiated. Literature on FA offers no evidence that rotated factors are more invariant than unrotated factors (Andresen (1998) bemoans this empirical deficit ¹⁴), whereas there are

¹³ “Hierarchy” suggests a classifying order with subordinate elements “contained in” superordinate units. But bio-psychological reality does not proceed in such a manner. What reality does show are different concurring influences. A more comprehensive factor, for instance general intelligence ‘g’, is not made up of smaller factorial units. Intelligence may become functionally manifest without including, say, specialised tools of expression, such as verbal tools (Revenstorff, 1976, p. 313, refers to a similar observation).

¹⁴ *“For construct exploration tasks in personality questionnaire scales, final proof still needs to be delivered of the superiority of the simple structure optimising method with regard to definitely improved cross-variable sample structure replications.”* (Andresen, 1998, p. 74).

definitely complaints about a lack of invariance in *rotated* factors (Fittkau, 1968, Butler, 1969)¹⁵.

The idea that initial factors are not interpretable must also be discounted as ill-founded. Although variables loading on only one factor, after SSM rotation, might immediately make more sense than variables loading on multiple initial factors, more comprehensible interpretations cannot indiscriminately be presumed to be more valid than less comprehensible interpretations. One might as well continue to believe that the universe revolves around the Earth, and reject the Copernican view.

The claim that initial factors have to be transformed by SS rotation was occasionally also inspired by the argument that a most prominent initial “general factor” must be a methodical artefact, because initially general factors are, as a rule, extracted from greatly differing data sets. Überla voiced this widely held notion: “*The [initial] factors ... depict an arbitrary distribution of variance. The variance distribution is not derived from data, but inherent to the method.*” (1971, p. 175). Here, a mathematical regularity is deemed “arbitrary” or coincidental. One should rather consider that a FA of sources of variance underlying some set of variables might generally reveal one dominating source of variance. Sources of variance of factors 2, 3, 4 ... on the other hand, generally reveal less variance what should already be expected from ranked eigenvalues of respective vectors¹⁶.

Despite warnings by the authors of textbooks, researchers often accept the first unrotated factor of factor analyses as an expected final result. Many intelligence researchers discover an expected general intelligence factor (explicated by Jensen, 1998, pp. 65-68) as an unrotated first factor. To conduct a Varimax rotation of intelligence factors, Maxwell (1972) limited factor rotations to factors F_2 to F_5 and did not rotate the initial factor (F_1). Dealing with questionnaire data, personality researchers often recognise hypothetically preferred factors operationalised in initial first factors (starting with Hamilton, 1960, Lumsden, 1961, Leventhal & Stedman, 1970).

By an SSM rotation, however, an initial F_1 very often disappears. To prevent a first initial factor from disappearing, some FA exponents resort to some kind of a trick: They repeat a first analysis after eliminating variables with unwelcome high loadings on a second and third factor, so as to get the second or third etc. eigenvalues falling under the Kaiser and Guttman (= 1.0) criterion of exclusion. This way, extractions of second and third factors are avoided and the ostensible obli-

¹⁵ Fittkau (1968, p. 110): “... *the results of all the analytical rotation processes [are] ... not invariant to adding other variables or replacing some variables with others.*” Butler (1969, p. 13): “*the simple structure concept does not solve one of the most crucial and fundamental problems of factor analysis, the problem of the likelihood of factorial invariance.*”

¹⁶ The assertion that the position of axes in an initial solution was arbitrary and thus had no interpretative value was Thurstone’s (1934): “*A characteristic of the multiple factor problem is that the location of the axes is arbitrary and that hence the factorial components are to that extent arbitrary and without fundamental psychological significance.*”

tion to conduct an SS rotation is circumvented without violating the convention. This is a formalistic approach more akin to exploiting legal loopholes than to facing facts. The trick as such is obvious, but rarely does anyone take offence at it¹⁷. Gibbons and Hedeker (1992) proceeded using a tidier method by utilising a so-called bi-factor model. In this model a general factor (F_1) with all variables participating is accepted, while each variable is allowed to load one additional factor. The bi-factor rotation model that conserves the general factor fitted significantly better than SSM rotations when data from knowledge tests and from depression questionnaires were analysed¹⁸.

Quite early on, special rotation techniques were developed to rescue the general factor as it disappeared as a result of SS rotations. To a certain extent, the orthogonal Quartimax rotation by Neuhaus and Wrigley (1954) manages to do this. There are, however, obvious weaknesses (Gorsuch, 1974, p. 191) that Quartimax does not eliminate.

Nonetheless, Überla and all those sharing his textbook views believe that the allegedly “incomprehensible” and “arbitrary” initial factor solutions must be transformed into allegedly more comprehensible and more stable solutions. Simple structure is taken as the only justifiable or even the only possible guide-line. Alternatives are not debated.

At the dawn of FA, Thurstone’s “American school of thought” was being opposed and criticised by the British school headed by Cyril Burt. Burt had recognised the presence of valuable information in initial factorial solutions, or in the initial bi-polar structures as he called them: “... *to the experienced factorist both the regularities and irregularities [of the pattern of signs] will yield considerable insight into the data he is analysing, even without any further rotation or analysis*” (Burt, 1954, p. 16). Burt also routinely transformed bi-polar initial factors into uni-polar solutions that he called “group factor solutions”. But he did not depart as far from initial solutions as did

¹⁷ Thalbourne (1998) is a typical example of the popular exclusion of items to sidestep the obligatory SS rotation. For reasons of interpretation, Groner and Groner (1991) even go so far as to limit themselves to just the first of 17 (!) initial factors with eigenvalues greater than 1. Gangestad and Snyder (1985) and Snyder and Gangestad (1985, 1986) present an illuminating argument for saving the concept of “self-monitoring”, which would be lost if SS rotations were applied. The dilemmas are increasingly neutralised by methodical compromises. These include the extraction of second-order factors (causing hierarchical models) and the monitoring of various model possibilities by confirmatory and structural equation procedures (Undheim and Gustafsson (1987) aim at “Restoring general intelligence”). The problem’s real sources are hereby disguised.

¹⁸ Hamilton (1958) employed the then still topical method of “simple summation” to find a general anxiety factor (F_1) in an anxiety questionnaire. He also found a weaker factor F_2 , indicating the centre of anxiety (psychic vs. somatic symptom dominance). Benactyzin, an antidepressant which was given to patients participating in the experiment, caused changes in F_1 , as expected, and not in F_2 . The author therefore criticised Thurstone’s SS procedure which did not reveal actual general anxiety reduction. A *Varimin* re-analysis corroborated Hamilton’s criticism and confirmed, using a published correlation matrix, his factor interpretation (results of an unpublished re-analysis by the author).

the American factorists. Burroughs and Miller (1961, p. 35-37) also had reservations about misleading SS rotations: “... *the subsequent rotations are apt to obscure [the objective and dichotomous classification based directly on the data]*” (p. 36). Had the competition between British and American factorists continued a little longer – unfortunately it did not – mistakes and failures caused by Thurstone’s position might have been delayed or possibly even prevented.

- The “positive manifold” is misinterpreted

Correlation coefficients of intelligence tests are usually positive (called “positive manifold”). Hence, initial factor solutions from test intercorrelations like blood pressure measurements, exhibit a unidirectional, positive general factor $F_1 (= g)$. As a rule, however, they also exhibit additional bi-polar factors F_2, F_3 , etc., i.e., quite a few variables possess negative loadings. Now, early intelligence researchers expected factors representing intellectual *abilities*, which by definition could not have negative manifestations. They were convinced that intelligence test factors could only have positive loadings or at best almost-zero loadings, meaning no ability. Therefore it seemed necessary to transform bi-polar loadings to positive-only loadings. SS transformations, bringing this about, were therefore considered even more justified (Thurstone, 1947, p. 341)¹⁹.

In his blood pressure FA, Überla (1971) had assumed, like researchers of intelligence, that only positive loadings were admissible because the smallest value of the blood pressure on a mmHg measurement scale is zero and not negative. Here the properties of a metric that is suited for manifest observations (blood pressure and intelligence measurements use ratio scales) are transferred to the metric suitable for latent conditions (showing properties of an interval scale). “... *rigorous measures, such as direct counts, latency, or duration, are excellent measures if used as descriptions of behaviour but may become arbitrary metrics if they are used to infer some psychological construct*” (Kazdin, 2006).

Following the publication of five contributions to the discussion about “arbitrary metrics” (*American Psychologist*, 61, 2006), agreement may soon be reached about the metric resulting at the latent level of factor loadings. It may be shown

¹⁹ “It is ... *natural to postulate that when a unique simple structure is found for a battery of tests of mental abilities, then the non-vanishing entries in the factorial matrix are positive*” (Thurstone, 1947, p. 341). Thurstone says that the limiting condition of the “positive manifold” was not required for applying the simple structure principle. But SS rotation would be appropriate especially when “*the factor loadings shall be positive or zero*”. (p. 23) Berneyer (1957) reports: “*The different methods of [factor] analysis [of mental aptitudes] yield factors which have negative loadings ... Such factors, so Thurstone contends, must be devoid of ‘scientific meaning’. They do not permit us to ‘interpret the various tests as functions of the mental aptitudes which those tests elicit’*” (p. 23). C. Burt also shares Thurstone’s “positive manifold” view: “... *an ability for x is by definition a dispositional property that facilitates doing x , i.e., it denotes a positive and never a negative tendency. Hence we shall be compelled to seek factors with positive saturations only.*” (Burt, 1954, p. 18).

that arbitrariness of zero-points and signs is not only permitted but indeed required. Vukovich (1967) argued at an early date for liberal scaling decisions: “*You can treat measurements any which way you like as long as you do not compromise their illustrative character. If empirical comparisons do not contain evidence about the zero-point or the size of measurement units, they can be chosen freely and they will conveniently be determined to facilitate the lucid description of larger contextual conditions*” (p. 114).

In short: An initial general factor must not be downgraded to a methodical artefact (Überla, 1971) and need not be removed at all. It may actually be the rule that one source of variance out of several sources emerges as predominant. Additional minor sources of variance (factors) might orthogonally modify the main effect with bipolar distributions.

As a rule, a general factor is extracted first which serves as a reference for variance sources of subsequently extracted factors. This is equivalent to a standardised distribution of values whose mean, a zero point, is the reference for values deviating from the mean in positive or negative direction. If negative loadings occur together with positive loadings for a second, third, etc. factor, this signifies that additional sources of variance are set in relation to the main source. Similar views were voiced by Thompson (1963), which, however, apparently have not been taken up by other factor analysts²⁰.

The present criticism of Thurstone and his followers’ assessment of positive manifolds is not meant to consider initial factor extractions as final. Initial solutions often require improvements by Complex Structure (CS) rotation. This will be elaborated in sections 05 and 08.

05. Reorientation

- Natural processes are complex

The FA of blood pressure measures showed that the variance total of variables can be split into two latent sources. One represents blood pressure independent of heart beat, the other yields proportions of systolic-diastolic variance. This result adds empirical reasons supporting my SSM criticism.

It may sound trivial, but in almost all domains of nature observed variables are dependent on multiple conditions. SSM-orientated FA ignores this phenomenon. The economic pay-off of the first step of FA, achieved by assigning large numbers of manifest variables to smaller numbers of latent sources of variance (factorial “building blocks”), is not disputed. But this achievement is gambled away by ap-

²⁰ “If they are able to choose one or the other ... psychologists tend to prefer unidirectional to bi-polar measurement, probably as a result of the prestige of such measures as the standard metre in physics. ... Thurstone expressed a preference for a positive manifold (without, in the writer’s opinion, giving a fully convincing explanation) ... Bi-polar and unidirectional measurement are both needed in psychology.” (Thompson, 1963, p.22). The latter statement is justified at length in Thompson (1962).

plying factor transformations that leave just one latent factor to each manifest variable, or as few as possible. The very opposite should be required (see additional support for this view in section 08).

It is remarkable that the SS fallacy did not lead to doubts and reflections earlier because the SS ideal was almost never fulfilled. All the agonised wrestling about SS was a Don Quixote-like tilting at windmills. Even Harry Harman, a leading author on FA, had to admit: *“An orthogonal uni-factor solution is practically impossible with empirical data and not very likely even when the factors are permitted to be oblique. Nonetheless ... it is towards that end that the simple structure principles are proposed for the multiple factor solution”* (Harman, 1968, p. 99).²¹

FA may be compared with multivariate analysis of variance (MANOVA). Factor analysts, however, generally deal with independent variables (IVs, sources of variance, “factors”, for example, abilities) that are latent or hypothetical and merely assumed to generate manifest, observed, dependent variables (DVs, for example, test measures). For conducting a FA, IVs need not be known, they are deemed to manifest themselves through factors. ANOVA researchers, however, do not only deal with manifest dependent variables (DVs), but also with manifest independent variables (IVs) that need to be manipulated experimentally. Their focus is not on underlying causes that make measurable variables become manifest.

Another difference between conventional SSM factor analysis and MANOVA may be pointed out: MANOVA does take into account “structural” conditions of the investigated DVs, i.e., their interactions. However, MANOVA is only interested in interactions among manifest variables. SSM-orientated FA excludes latent-level interactions.

Had common elementary knowledge been impartially considered, the SS principle would have appeared suspect from the outset. Variables investigated in science are generally based on components interrelated on lower levels. Atoms consist of protons, electrons, and neutrons, salt consists of sodium and chloride, and

²¹ Criticism of the SS principle was voiced on rare occasions: *“Deciding which of many possible mathematical solutions to use depended on a formal rule, the simple structure principle, which lacks any theoretical substantiation. Freed of all necessity to conduct more elementary deliberations, (researchers) now delivered bulk work ... The result of this method, as could have been predicted, was a surplus production of superficially defined factors”* (Meili, 1969, p. 278). Similarly, Revenstorf (1976) deemed SS *“generally unlikely ... [especially] in large feature compilations (questionnaires)”*, because *“features can increasingly depict an endless range of combinations of factorial measures. In this case, the features of the configuration of variables are scattered across the entire factor space, and a simple structure in a Thurstone sense ... can no longer be discerned”* (p. 321). Countless examples of doubtful SS-factor interpretations exist: *“It is obvious that the simple structure is lacking”* (Bierhoff, 2000); *“... despite the comparatively high intercorrelations we cannot assume a simple structure”* (Schaper & Baumgart, 2002); *“Generally, simple structure does not appear to be succinct”* (Beauducel, Strobel, & Brocke, 2003), *“... not a simple structure according to the Bargmann test”* (Herzberg, 2002); *“An acceptable simple structure for the factor loading matrix could not be achieved by an orthogonal or an oblique rotation of the three main axes”* (Schmitt, 2000); *“The result does not indicate a simple structure”* (Lambert et al., 2002).

genes have an effect with complex diversity only (“polygenic” effects)²². Perception of an individual colour is the result of electrical excitation of three cone pigments. To continue: clauses consist of words, words of morphemes, morphemes of phonemes. Each phoneme is based on several phonetic features. Applied to FA, it should be expected that manifest variables in a study are based on concurring variance sources.

Complexity, i.e., the concurrence of conditional components, is in accordance with Occam’s principle of economy, indeed it is an immediate consequence thereof. When available units of operation can be combined, new adaptive processes result, the cumbersome production of particular units for additional purposes becomes superfluous. Evolution did not develop additional receptors for the perception of purple, ochre, ultramarine etc.

A recent investigation into the evolution of language culminated in the conclusion that even the highest mental attainments are the result of successfully combining process-related resources. Initiating the evolution of an extra programme for establishing linguistic communication of hominids would have amounted to a waste of resources, say Bates et al. (1992)²³ and Gould et al. (2002).

Wolf Singer (2003) does not tire of describing the human cerebral cortex as having the ability “*to play a combinatory game*” (p. 84). Elementary brain structures have “*a similar format*” and are of a “*surprisingly monotonous build*” (p. 44), “*nature is very conservative ...*” (p. 46). When centralisation is lacking, brain activity produces its “*performance through constructivist bonds*” (i.e., combinatorics) (p. 75). “[*In the*] brain’s functional architecture [*it is*] crucial, who gets in touch with whom, how intensely, and whether it happens in an inhibitive or in an agitating manner” (p. 38).

Tens of thousands of dendrites facilitate contacts among brain cells. Transferring Thurstone’s principle of SS onto brain activity based on neural units as solitary pegs, would contradict the principle of free bonds. Singer regards the bond phenomenon as “*the core problem of neuroscience*” (p. 57). Apparently processes of living nature and even of inanimate nature are subject to combinatorics of elementary building blocks. The biologist Humberto Maturana (1998, pp. 158-189) refers to the universality of “structural determination” of nature’s activities²⁴.

²² Simple anthropometric characteristics like height and hair colour are based on the concurrence of multiple genes.

²³ “*Language learning appears to be based on a relatively plastic mix of neural systems that also serve other functions. I believe that this conclusion renders the mysteries of language evolution ... somewhat more tractable. That is, the continuities that we have observed between language and other cognitive systems make it easier to see how this capacity came about in the first place.*” (Bates, E. Modularity, domain specificity, and the development of language. URL: <http://www.ecs.soton.ac.uk/~harnad/Papers/Py104/bates-1994.html>, no longer accessible).

²⁴ Numerous quotable statements support this view, e.g.: “*A living system is a structural determinant system and everything in this system happens as a result of relations of its constituent parts ...*” (Maturana, 1998, p. 184).

The complexity issue deserves consideration from a more encompassing perspective, too. Malcolm Forster (1998) examines the terms “*parsimony and simplicity*” and concludes: “*The paucity of parameters is a limited notion ... There are no compelling ideas about why such properties should count in favour of one theory being closer to the truth than another [...] Whether a simple curve is preferable to some more complex alternative, or the reverse is true, has nothing to do with simplicity and everything to do with predictive accuracy*” (Forster & Sober, 1994, p. 28).

In *Complexity - a Philosophical Overview*, Nicholas Rescher states: “... In the development of knowledge – as elsewhere in the domain of human artifice – progress is always a matter of complexification. An inherent impetus towards greater complexity pervades the entire realm of human creative effort.” (Rescher, 1998, p. 58). “There really are no adequate grounds for supposing the ‘simplicity’ of the world’s make-up. Instead, the so-called ‘principle of simplicity’ is really a principle of complexity-management.” (p. 61). By the same token one finds positive resonance in Bechtel and Richardson (1993): “*It is ultimately the pattern of connections in the system, and not the jobs performed by specific units in the system ..., that is critical to the behavioural systems*”.

Gawronski (2000) qualifies the simplicity posit for epistemological reasons: Simplicity is a “vague criterion”. “*If clusters of variables are to be described by a mathematical function, there may still be a certain consensus about the ‘simplest’ graph. But if verbal scientific theories are concerned, judgments about simplicity can vary greatly ... depending on structures of knowledge. In this sense, the demand for simplicity becomes relative.*” (p. 10).

- The new guideline of factor transformation is *complex structure modelling* or CSM

Factor analytical research needs a new guideline which turns Thurstone’s *simple structure principle* upside down. This is it: *complex structure modelling*.

Initial EFA factors should be transformed in such a way so as to optimise the simultaneous presence of extracted factors with individual variables. In chapter 2 I show that this can be done, supported by maths and statistics, and that it prevails over conventional procedures.

It will also be shown that initial factors already engender complex solutions, even without factor rotation, provided the data sets are based on only small numbers of sources of variance (factors), as was the case with blood pressure measurements. When working with more than two substantial factors, however, a rotation procedure is called for to improve the complexity of their representations. The procedure should assign factorial sources of variance, established for the respective domain, to all individual variables, if possible, and then have the results tested against empirical boundaries²⁵.

²⁵ An extensive German-language criticism of factor analysis (Holz-Ebeling, 1995) takes exception to the vagueness of factor interpretations. The author is right in stating that investigated variables (DV) are generally dependent on multiple conditions (“multi-IV conditionality”). She adds

CSM aims to show actual complexity, while SSM tries to enforce non-existent simplicity. The ideal of solitary variables (each variable is allegedly explained by only one factor or by as few factors as possible) actually generates non-simplicity and non-parsimony.

Employing CSM for factor transformation is no rebuff of Newton's "*natura est simplex*" or Occam's razor. Occam's razor just needs to be applied in such a way as to reveal the strategy of parsimony that nature itself brings forth (by utilising its bonding capacities as shown earlier)²⁶.

Incidentally, the *complex structure model* does not expect to be taken as an innovation. It transfers a view generally held in science to a particular field of methodical operations where until now it has not gained a foothold. An opposite view of enforced simplification, associated with SS, has survived to this day in a protected corner of statistical methodology and has been leading research astray^{27 28}.

that factor analysis does not do justice to the multi-IV conditionality of DVs. She suggests using a procedure methodically related to variance analysis to "replace or complement" factor analysis. Multi-IVs influencing singular DVs are accessible by variance analysts. The author's impractical (her words) suggestion is interesting in as much as it attempts to correct errors caused by the anti-complexity SS doctrine. She is aiming in the right direction but does not touch the dilemma's real causes.

²⁶ Time and again, factor analysts confronted by unwieldy and complex variables have encountered the limits of the SS principle. Guilford and Zimmermann (1963) risked a liberalisation of the parsimony principle by admitting complexity: "*In general, investigators need to relax somewhat their drive to achieve parsimony. Long ago, psychology should have progressed beyond the stage in which investigators continue to look for the 'philosopher's stone'*" (p. 299). When modelling complexity, parsimony cannot be curtailed. Nature applies different economic strategies than those that Thurstone used for depicting nature's variables geometrically.

²⁷ Oblique rotation, devised by Thurstone and vehemently propagated by Cattell, is comprehensible from a CSM perspective. Oblique rotation is actually a dubious balancing act between obtaining monofactorial variables, while also having to take into account the complexity of variables as witnessed in the natural universe. "*It is unreasonable to expect that a great variety of influences operating and interacting in the same universe would be completely uncorrelated*" (Cattell & Dickman, 1962, p. 390). The factor analyst applying oblique rotation will thus allow for factor correlations in a non-orthogonal factor space. By permitting factors to correlate, a back door is kept open for some functional linkage. This compromise was criticised decades ago by Guilford and Zimmermann (1963, p. 289): "*[This amounts to] a hollow and accidental victory for oblique methods of rotation*". The pseudo-solution by oblique rotations is mentioned here to illustrate the consequences of a frustrating search for SS and the attempts of researchers to get rid of the self-created evil by questionable decisions.

²⁸ Cattell and Radcliffe's (1962) attempt to eliminate, by suppressing unwanted variance, actually existing complexity that was not removed by SS rotation is just another symptom of misguided SSM: "*If we are right in assuming that behaviour which has a large personality factor variance will generally be factorially complex, then the unwanted common factor variance will, as a rule, be far from negligible*." The authors intend developing "unifactor scales" and to this end they manufacture a process "... which will reduce the contribution of unwanted factors by suppression" (p. 125).

06. Where did we go wrong?

How did the erroneous idea of SSM come about? Human liability to cognitive errors probably contributed much. Gestalt psychologists focusing on “Gestalt laws” have repeatedly warned about detrimental side effects. *“Thought misled by Gestalt tendencies”* (Witte 1974) and *“Visual Prägnanz”* [conciseness] *as an obstacle to problem solving*” (Kanizsa, 1975) seems to affect visualisations of data points in factor space, clusters of variables are irresistibly attracted by surrounding coordinates due to the laws of proximity and grouping. Apparently, researchers adhered to “intuitively compelling” perceptions that provided an anchoring of the variables. It was easy to employ “curve fitting” by utilising freely rotatable coordinates. Gorsuch (1974, p. 164 f) talks of a “visual compellingness” that guided Thurstone in his early visualisation of extracted factors resulting in his oft-quoted five criteria of simple structure rotation²⁹.

Arkes (1991) describes another reason for judgmental errors possibly underlying the argument of “easy interpretation” of SS factors: *“Suppose a person adopts a quick and dirty strategy to solve a problem. Because it is so quick, it is easy to execute. This is a benefit. Because it is dirty, it results in more errors than a more meticulous strategy. This is a cost. Although the choice of this strategy results in [more errors] ..., this cost may be outweighed by the time and effort saved”* (p. 487). Arkes goes on to call this, not quite accurately, a “strategy-based judgment error”. He should have called it a “cost-saving judgment error”.

In a similar vein, Edmonds (2002) refers to a universal bias for simplicity. *“... the simpler theory is not more likely to be true and is not likely to be nearer the truth ... For human beings it is much easier to elaborate ... [a failing] theory, or otherwise tinker with it, than to undertake a more radical shift (for example, by scrapping the theory and starting again). This elaboration may take on many forms, including ... complicating the model with extra equations or rules ... or using more complicated functions.”* Edmonds summarises: *“... Model selection, for the sake of simplicity is either simply laziness ... [or] due to pragmatic reasons”*. He advocates, as a matter of principle, foregoing “simplicity for fear of innovation”. *“... The elaboration of [an existing] theory in order to fit a known set of data should be resisted ... The lack of success of a theory should lead to a more thorough and deeper analysis than we are usually inclined to perform.”*

Without doubt, the “power of words” has boosted SSM’s immunity. Hardly anyone dares oppose a term indicating the seductive attributes *simple* and *structure*.

²⁹ 1. Each row of the factor matrix should contain at least one zero.
 2. If there are m common factors, each column of the factor matrix should have at least m zeros
 3. For every pair of columns in the factor matrix, there should be several variables for which entries approach zero in the one column but not in the other
 4. For every pair of columns in the factor matrix, a large proportion of the variables should have entries approaching zero in both columns when there are four or more factors
 5. For every pair of columns in the factor matrix, there should be only a small number of variables with non-zero entries in both columns.

Had Thurstone named his rotational principle, say, the *Principle of Solitary Factor Contribution*, which would have been more modest and closer to the truth, its validity would have soon been doubted. Now, after a worldwide dispersion of the magic term “simple structure”, it will be hard to get rid of it.

Hard-core SSM has been sealed off by leading methodologists. They countered an apparent need for more complex modelling by more sophisticated procedures that showed flaws right from the outset. The cautious attitude of an exploratory researcher, who is prepared to be surprised by unexpected findings, has increasingly been replaced with an attitude of mere administrators of variables, who determine the units of their domains arbitrarily and assign them functions as they see fit.

The current epistemic climate in psychology places feasibility at the top of value rankings. Models that can be manufactured and imposed or inflicted on nature garner more respect than models emerging from nature itself. The intricacy of models constructed by amateur tinkerers easily conceals the fact that discoveries of true value mostly occur by careful bottom-up observations.

Using Lakatos’ (1978) historical approach, the SS principle may be taken as resulting from a “hard core” attitude accompanying long-term developments of conventional research. Core beliefs may turn hard core when a community of researchers takes them for granted and no longer questions them³⁰. Michell (2000) considers blind clinging to fundamental premises a symptom of “scientific pathology”: “*A hypothesis is accepted without a serious attempt being made to test it and this failure of critical inquiry is ignored*” (p. 648). Schönemann described such a symptom by mentioning the outrage of “traditionalists” when Bargmann wanted to obtain significance values for SS statistically: “[*They*] scorned it [*Bargmann’s Test*] as a sacrilege of their cherished belief that simple structure is a law of nature” (Schönemann, 1994, p. 293).

Long before Lakatos, Fleck (1935) presented case studies to describe sociological excesses of “thought collectives”: “*Once a fully developed, closed system of opinions has been formed, it will consistently persist against all opposition ... What does not fit the system will not be seen or is concealed ... or is declared as not opposing the system by employing huge exertion.*” (p. 35). “*The tendency of opinion systems to persist proves that they should be seen as ... stylistic structures. As harmonic entities ... they exhibit special stylistic features that determine every single cognitive function. ... [They create a ‘harmony of deceptions’], which cannot possibly be dissolved within the ambit of a certain way of thinking.*” (p. 45).

Lakatos maintains that hard-core conditions in scientific research will lead, after a while, to empirical anomalies. As the number of anomalies grows, research programmes move into a “degenerative” phase. “Protective belts” are constructed and reinforced when danger looms. For the past two and a half decades or so, one-sided orientated methodologists have been developing highly complicated

³⁰ “*It may be contended that we should cling at all costs to the conception of simple structure, because we have no satisfactory alternative*” (Reyburn & Raath, 1949, p. 127).

procedures for analysing multivariate data. These methods are clearly protective belts whose function is to immunise and shield the doctrine of SS from its disappearance (also refer to section 08).

07. Unheeded critical voices

Did no one ever take umbrage at the SSM principle? Yes, the limitations of this principle were challenged occasionally, but its basic legitimacy was almost never questioned.

W. Stephenson (1956) was one of the earliest doubters. He realised that attempts to make personality traits dependent on singular factors will clash with empirical complexity: *“The problem is to explain complex traits in terms of relatively few primaries”* (p. 7). He used more liberal rotation methods for what he rightly named “compound traits”. But the usage of his method was laborious, it was much like the Circumplex procedures of Hofstee et al. (1992) and de Raad et al. (1994) (more in section 08) and their confusing results that were also based on SSM’s rotation ideal.

Stephenson favoured the factorial Q design which aims at analysing “person variables”. He obtained them by so-called Q-sort assessments that he himself had developed. Person variables were intercorrelated and factorised unlike what is required for factorial R designs that select behavioural variables (such as test results, ratings, not persons). Stephenson thereby circumvented the unsolved question: Which are the basic features that make complex experience, behaviour and trait variables, theoretically understandable? The answer can only be found by employing exploratory R studies, since the constructs needed to interpret Q results are not provided by the Q FA itself.

Another researcher highly dissatisfied with Thurstone’s SSM was J. P. Guilford (1974, pp.498 f): *“Thurstone’s principle of simple structure ... is by no means sufficient if we want logical psychological meaning ... In numerous instances [in Thurstone’s and his students’ studies] tests of very different character are often thrown together significantly on the same factors. This does not make good psychological sense ... Rarely were varimax factors clean-cut and easy to interpret psychologically ...”*.

Guilford reacted to deficient results of FA of intelligence variables by proposing what he called a “structural model” of intelligence. He tried to improve the factorial SSM results by “logical and psychological” means. Alas, he did not detect inherent errors in Thurstone’s principle. He set out to rectify factorial deficiencies by applying non-factorial conceptual tools.

A more recent critic providing similar arguments is Allen Yates. Better than anybody before or after him he hit the nail on the head: *“The factors that result from cluster-oriented factor analysis [from blind application of simple structure] are simply an index of success an investigator has had in putting together groups of collinear variables (clusters) – regardless of how complex these same variables might be in terms of their latent determinants. In other*

words, manifest collinearity among variables is an indication only that they share the same pattern of causal determination; it does not in any way suggest that the shared pattern involves just one latent causal factor." (Yates, 1987, p. 39).

Yates' monograph delivers the most concerted attack on the model of SS that I have found. His last chapter gets started with: *"Only a radical reorientation of current perspectives will allow researchers to apply exploratory factor analysis in the manner envisioned by its originators as a powerful technique for routine discovery of the underlying bases of observed covariation"* (Yates, 1987, p. 323).

But even Yates does not touch the core of the SS calamity. According to him, Thurstone's original approach that had generally been ignored was "more liberal" (Thurstone even ignored it himself later). Yates developed new rotation algorithms. His alternatives (*Direct Geomin* and *Direct Geoplane*) are extremely complicated and researcher-dependent, much like the procedures of other researchers who tried to escape the unwelcome side effects of SS analyses. The success of Yates' innovation was meant to be a way out of the dilemma, but psychometricians did not take notice of his approach.

SSM problems have also been discerned by Rozeboom (1991), who saw no way out: *"For diagnosing the causal grain of common-factor space, rotation to simple structure is so disquietingly fallible that we would surely prefer another criterion were any plausible alternative at hand. (Churchill's aphorism on the inferiority of democracy comes to mind here, namely that democracy is the worst form of government there is – except for all the others ... Read 'simple structure' for 'democracy' and 'rotation criterion' for 'form of government'.)"* (p. 587). Rozeboom's HYBALL rotation, in which multiple coordinates (sub-spaces) instead of individual coordinates were successively rotated (*"somewhat more holistic than a simple sequence of planar rotations"*, p. 587), proved to be just as dubious a compromise as the one Yates had invented.

Schönemann and Borg (1996) were among the critics categorically doubting SSM. With much resignation they stated: *"the simple structure criterion (formulated as Varimax criterion) [is] routinely applied in factor analytical practice"*. And, more critically, they add: *"Given many systems of variables in the first place, the important question why a simple structure should be expected completely falls by the wayside"* (Steiger, 1994, p. 204). Basically, *"the hypothesis of a simple structure was not very plausible ..."* It asserted that test questions with non-zero loadings on all factors were "impossible" and, in so doing, "puts the cart before the horse" (according to Guttman, 1992, p. 186).

A more general underlying flaw of researchers was diagnosed by Gigerenzer (1978), albeit for another example of dimensional model generation. Uncritical model users in psychology tend to ignore the fact that their models are, by pre-suppositions, connected to the psychological domain of observations. These pre-suppositions are not directly tested and Gigerenzer says that what he calls "proposition of implication" is generally ignored: *"[This proposition] holds that every mathematical system (e.g. a method for dimensional analysis) implies some psychological theory about the*

respective domain of observations [...]” (p. 110).³¹ The formalised system of relations (FRS) is directly and inseparably entwined with the empirical relations system (ERS) (Gigerenzer refers to supporting views of leading methodologists Suppes and Zinnes (1963)). Should FRS and ERS diverge in essential aspects, divergence artefacts ensue, leading to “theoretically worthless results” (p. 111).

Applied to the present context this means SS (an example of FRS) assumes, without evidence, that the variables accessible at the ERS level are based on solitary sources of variance. Gigerenzer demands accordingly: *“To prevent an investigative result being interpreted as divergence artefact, the researcher has to ... explicate the observed psychological model that was implied by the mathematical model ...”* (p. 116).

This, though, is precisely what factor analysts have thus far neglected to do. They believed that their analyses, applied with mathematical precision, would more or less automatically deliver structural models of psychological reality. Utilising independent reasoning was deemed superfluous and even disreputable; this carried the blemish of “subjective” intrusions into the discourse on “objective” facts.

08. Can non-factorial procedures take us forward?

- What do Circumplex procedures achieve?

Methodologists have often reacted to unsatisfactory EFA results by inventing additional procedures. Mathematical superstructures should cure the symptoms (Revenstorf, 1980, p. 12). Elaborate calculations were conducted to cope with multiple loadings of variables (e.g., questionnaire items) deviating from the SS ideal. Hofstee et al. (1992) and their AB5C model (Abridged Big Five Dimensional Circumplex) is an example. The authors tried to improve SS results of adjectival personal descriptions by applying Wiggins’ (1979) Circumplex method. The Circumplex method allows two factorial dimensions to be associated with variables simultaneously, and thus does away with SSM restrictions, but anyway only partially.

The Circumplex model is just another hybrid compromise, the method *“does not deal adequately with those ... variables that load highly on more than two ... factors”*. Circumplex results are thus neither “definitive nor comprehensive” (Hofstee et al., p. 161). The SS principle is not shaken; the dimension problem is not stirred. The authors merely *“propose a partial liberalization of simple structure”* (p. 147), while what is required is the abolition of the SS principle for solving the complexity problem.

³¹ Gigerenzer’s implication proposition antedated Smith and Jones’ (1975) “central thesis” published in a paper on multi-dimensional scaling: *“All data analysis and all scaling involve fundamental assumptions about the psychological processes that lead to the data and the scaling solutions under consideration. In particular, current scaling methods, in our view, should be regarded with deepest suspicion, precisely because they are based on doubtful or untested psychological assumptions.”* (p. 44).

Acton and Revelle (2004) surveyed psychometric criteria for an application of the Circumplex (p. 26). They also found that this method fails when “all items are best described by more than two factors”. The requirements for Circumplex application are incidentally very discerning – the authors name ten requirements –, making the Circumplex hardly commendable even if data sets can be described by only two factors.

- Are confirmatory procedures an alternative?

Confirmatory factor analysis (CFA) and its more flexible sequel “structural equation modelling” (SEM), seem to have largely displaced, with sophisticated algorithms, exploratory (EFA) procedures. Instead of considering, as EFA researchers generally do, unpredictable latent variables in their calculations, it has become fashionable to invent latent variables willy-nilly and to use confirmatory techniques in order to find fits between the thought-up model and empirical reality. To this end vast numbers of fitting attempts are conducted, mostly via blind trial and error.

But structural equation models are still geared to the SS principle. Therefore these models also do not uncover the complexity of sources of variance³². “*The more recent structural equation models – initially as ever euphorically celebrated ... exacerbate factor analysis’ problems*” (Schönemann & Borg, 1996, p. 241). Nearly always they fall short of underlying structures³³. Critical remarks are sometimes recorded: “*The assumption of simple structure is probably a typical ... simplification bias*”, but unfortunately “*necessary*” (Beauducel & Wittmann, 2005, p. 43). “*... Simple structure models of personality are unlikely to meet conventional or even fairly relaxed goodness-of-fit criteria ... Overemphasis on simple structure ... may explain some of these problems*” (p. 44)³⁴. SEM results are sparse (“poor results”, Beauducel & Wittmann, 2005, p. 42). In a critical review, MacCallum and Austin (2000) point at problems of the confirmatory method and the “confirmation bias” of its users. Users tend to make do with *goodness-of-fit values* and arbitrarily chosen criteria, resulting in make-believe fits. Instead, results should be evaluated using factual information and not primarily formal

³² Basilevsky (1994, p. 415) describes the trend to dissect singular factors from the examined variables in CFA practice as follows: “... *we may wish to impose zero restrictions on the loadings. Values other than zeroes can also be used, but zeroes are most common in practice*”.

³³ “*The LISREL manual discreetly conceals the fact that none of the latent causes have been positively defined*” (Schönemann & Borg, 1996, p. 250). The “indeterminability problem”, a purely mathematical formal problem, “was actually compounded” in the LISREL case “because many more latent variables were postulated there than in the multiple factor analysis model ...” (p. 250).

³⁴ A recent attempt to solve the acknowledged problems with SEM research has been undertaken by Marsh et al. (2010) who try to combine the advantages of unbiased exploratory analyses with confirmatory procedures (the approach is called *exploratory structural equation modelling*, ESEM). Marsh et al. address symptoms of the malaise which they are eager to correct, but the underlying source of the symptoms (SS bias) is not recognised. The Big Five factors are not questioned, they are fully replicated, the ostensible advantage are improved statistical properties in detail.

measures³⁵. When models do not make the fit criterion, this is often ignored (“working with imperfect models” is the title of an article by MacCallum (2003)). Or it remains unconsidered that the selected model’s fit could easily be surpassed by other, non-tested models. Even Kaiser did not think much of CFA: *“I cannot resist saying that, for me at least, the earlier exploratory thrashing about was much more fun – and perhaps even represented more progress – than the forthcoming confirmatory prettying-up”* (Kaiser, 1970, p. 406).

Sobering results from other CFA and conventional EFA comparisons can be found in Church and Burke (1994) and Ferrando and Lorenzo-Seva (2000). Criticism by Cliff (1983) beats the same drum. A recent comprehensive CFA study done with simulated data finds as follows: *“... trait models [of personality] assuming simple structure tend to be rejected with CFA ...”. “... there will always be some small distortion of simple structure”* (Beauducel & Wittmann, 2005, p. 72). When applying SSM to data sets (especially personality data) that are subjected to CFA, *“a gap [is created] between the large body of results based on exploratory factor analysis and CFA in personality psychology”* (p. 73).

Conventional EFA is orientated toward SSM but, as a rule, for many variables unwelcome secondary loadings are noticeable. These cannot even be manipulated when factor loadings are blindly distributed beforehand (*“there is no [prior] knowledge of secondary loadings”*) (Beauducel & Wittmann, 2005, p. 43). Vittadini (1989) considers LISREL results indeterminate because latent model variables are made dependent on manifest variables: *“one may actually be confirming the model because the manifest variables are determined by other variables than those hypothesized, which happen to have the same pattern of relationship to the manifest variables as given by one’s hypothesis ... One can never regard structural hypothesis as true as opposed to ‘confirmed’”* (p. 428).

Only where latent variables have already been verified by exploratory analyses, may confirmatory procedures be applied. Velicer and Jackson’s (1990) reasoning points that way: *“Exploratory analytic approaches ... should be preferred except for those cases where a well-defined theory exists. Exploratory approaches avoid a confirmation bias, do not force a theory-oriented approach prematurely, and represent a conservative strategy.”* (p. 21). But it does not make much sense to conduct CFA calculations for obtaining a result that has already been discovered by means of EFA³⁶.

³⁵ *“The LISREL model’s practically boundless plasticity ... not only undermines its claim to statistical inference but also gets close to soliciting abuse, because if you have sufficient patience you are bound to discover some kind of causal model that does not have to be declined for the currently available data.”* (Schönemann & Borg, 1996, p. 250).

³⁶ Rost (2002): *“The model specialist [specialist in modeling with structural equations, log-linear models, item-response models, etc.] ... cannot use data at all if he is not told which variable is supposed to interact with which other variable ..., which latent variables should be there, ... etc.”* My comment: The present-day “model specialist” can gain an approximate insight into the latent parameters of human thinking, feeling, and behaviour right from the outset. The exploratory factor analyst from days gone by, however, was unable to gain this knowledge, not even with the benefit of hindsight, despite having

- How should mathematical tools of research be generally evaluated?

A growing tendency in current science is to trust mathematical techniques blindly and to let statistical tool makers dominate research. Papers of some self-critical methodologists contain warnings: *“Those who firmly believe that rigorous science must consist largely of mathematics and statistics have something to unlearn. Such a belief implies the emasculation of the basic substantive nature of science. Mathematics is content-less, and hence not – in itself – empirical science ... rigorous treatment of content or subject matter is needed before some mathematics can be thought of as a possibly useful (but limited) partner of empirical science”* (Guttman, 1971, p. 42).³⁷ Schönemann quotes the sceptical Guttman: *“There remains the danger of seeking data merely to fit axioms”*, and comments: *“In hindsight, these warnings sound positively prophetic in anticipating the present malaise in mathematical psychology some 20 years before Cliff (1992) noticed it ...”*. Schönemann (1994, p. 294) also mentions Narens and Luce (1993) as critics of the malaise.

As early as 1975, a pioneer of mathematical psychology, William K. Estes, complained about the shortcomings of his field: *“... it is clear that many investigators in our field are not entirely happy with their current situation”* (Estes, 1975, p. 263). He laments the chasm between mathematical and content orientated psychology and, to support his own review, quotes Leont’ev and Dzhamfarov (1973, p. 20): *“An analysis of the present situation shows that contemporary psychology and contemporary mathematical instruments are still not compatible enough with one another to allow mathematization to assume a central place in the development of psychological knowledge; the reason for this is not only the low level of sophistication of the latter ... What is required is a continual interaction between mathematics and psychology, an interaction that ... would lead to a revision of existing mathematical methods into forms more amenable to the proposed mathematized conceptual systems.”*

Access to psychological content is not primarily achieved through formal models but through the totality of experience in psychological domains. Knowledge and belief acquired by previous experience are only specified and tested by research. In this process, accounts that can be communicated in everyday language may play a significant role. The fit of formal modelling of psychological data should be evaluated by taking into account non-formalised information. Critical writers have repeatedly commented on this:

masses of data to analyse. Today it is often hypothesised that these parameters can be dreamt up and that they only need to be shaken out by trial and error through model fitting.

³⁷ This quote was taken from an article by Barrett (2003), who criticised conventional psychometrics because its measurement operations rely upon untested assumptions of quantitative structure for psychological attributes (intelligence, personality). He pleads, alternatively as it were, for openness toward application issues. Barrett propagates so called “applied numeric” that would liberate research from the constraints of questionable theoretical expectations and be pragmatically much more useful (a rebuff of theoretical claims).

- “Neither algorithmic sophistication, nor axiomatic rigor alone are apt to advance our knowledge much if they are cultivated in an empirical vacuum” (Schönemann, 1981, p. 412).
- “May I ... insist once again on the absurdity of divorcing the mathematical or statistical evidence from evidence procured by other means? ... The sole claim of mathematical analysis should be to verify, by appropriate calculation, the hypotheses commonly advanced on the basis of much broader and more general lines of evidence.” (Burt, 1949, p. 107).
- “Let us try to be free of ... a priori mathematical and statistical considerations and prescriptions – especially codes of permission. Instead, let us try to think substantively ... and focus directly on the specific universe of observations with which we wish to do our business.” (Guttman, 1971, p. 346).
- “Factorists with more mathematical training than the rest of us have been addressing themselves to problems ... on a technical rather than upon a fundamental level... In most cases [their models are] irrelevant ... to the problems ... of those for whom factor analysis is a research tool...” (Butler, 1969, 252-3).
- “These techniques [“for rotating factors into ‘psychologically meaningful positions’”] ... became the stock in trade of practicing factor analysts ... It is probable that future historians will be severely critical of them and of their users; critical of the techniques because of ... their users’ extravagant claims on their behalf.” (Maxwell, 1959, p. 228).

Given the growing formalistic alignment of psychological research, we should go out of our way to nurse and nourish non-mathematical methods of knowledge acquisition. These sources of information are often pushed aside as merely “phenomenological” or “hermeneutical”. But if they were allowed to operate “on a fundamental level” (Butler) and if they were required to critically review the formalists’ “extravagant claims” (Maxwell), one should keep them alive. In the universe of our knowledge, mathematical resources can fulfil only part of our desires³⁸. Precision in detail and ingenious operations with numbers are useful, but not always essential. Numerical tools can be harmful if wrongly designed or excessively applied – as the SSM debacle shows – while more basal, comprehensive, holistic, albeit possibly at times less focused methods for acquiring knowledge are needlessly forced to stay outside³⁹. Gawronski (2000) also pleads for a holistic approach in research in order to examine whether methodically carved out observations are in fact reconcilable with our comprehensive background knowledge.

³⁸ This is also the primary concern of dissident Sigmund Koch (1999) and other lone voices criticising science that is continually drifting off toward one-sided objectivism (Bridgman, 1959).

³⁹ S. Jevons (1873): “I tend to grumble about mathematical writers because so often they cheer all the things they can do without pointing out that what they do is just the very minutest part of what could actually be done. They exhibit the general tendency of not even mentioning the existence of stubborn or intractable problems ...” (quoted in Rescher, 1985, p. 124).

Discussion of chapter 1 and outlook

This chapter does not downgrade FA, but just its habitualised wrong application of today. Gigerenzer and Strube (1978) emphasised that applying research methods in psychology should always go along with “*critically scrutinizing their assumptions*”. By adhering to this recommendation, Yates’s ideal could have been reached earlier. Yates considered a paradigm change inevitable,⁴⁰ because without fundamental changes the “pathology” of factor analytical research could not be cured while the “morass” of its previous results would endure.

Ultimately, the ritual called Little Jiffy, decried by Gigerenzer and Strube (p. 81), should come to an end. (Little Jiffy is a recipe-like application of FA culminating in a Varimax rotation). Also, the spirit of criticism that was revitalised at the seminal “Munich Symposium” and withered away again needs to be revived. Its proponents were Kallina (1967), Kalveram (1970), Gigerenzer and Strube (1978), and Revenstorf (1980). By employing FA guided by CSM (see introduction to Varimin in chapter 2) the Thurstone doctrine will certainly be shaken up. After further refining this approach, perhaps debugging it, an optimal understanding of variance sources underlying manifest observables may eventually be achieved.

Even Henry F. Kaiser (1927–1992), who created the preconditions for Little Jiffy, might readily have supported a radical new orientation of factor transformation – unfortunately he passed away too soon. In closing a talk on “Second generation Little Jiffy” to the Psychometric Society he explicitly offered, should such a situation arise: “*For the future, I can assure you of one thing: if any of you folk ... come up with some Big Breakthroughs I shall be waiting in the wings ready and eager to paste them together to produce the next generation Little Jiffy*” (Kaiser, 1970, p. 414). In his eulogy, Gene Glass (1991) characterises Kaiser as “disrespectful”, a trait that seemed to suit Kaiser well, because, Glass continued: “*Irreverence must be a necessary ingredient in the recipe for creativity. Whoever worships received wisdom too ardently will never see beyond it.*” (p. 159-171).

Complex structure modelling might appear “disrespectful”, because it turns Kaiser’s Varimax criterion upside down. But Kaiser might have welcomed such insubordination, since he practised it himself and expected his students – and probably those of coming generations – to do likewise.

⁴⁰ Some current formal modelling experts seem categorically to exclude any “paradigm change”. Rost (2003), for instance, talks about “laws cast in iron that will survive fashion tides ... and will even emerge from them stronger”. Rost believes that psychology’s *zeitgeist* can affect the use of methods only marginally, but would not lead to “*the ... arsenal of research methods proving to be wrong or unusable. Rather [fashion] will result in the method arsenal being expanded and broadened by important aspects.*” To complete his generalising review, Rost would have to add that flawed methods contained in this arsenal can cause huge damage, and must be fundamentally redesigned.

Chapter 2

Finding complex structures

Introduction

Empirical observations never assert themselves more vigorously than when they thwart our expectations and disabuse us. When I began contemplating *Varimin* rotation as a possibly better alternative to Varimax, I was concerned that my expectations might be frustrated. I could merely hope that the new method would surprise me by revealing complex structures (CS) of analysed variables and that they would model examined domains more appropriately than conventional simple structure (SS) procedures.

My uncertainty was justified. *Varimin* structures of FA could not differ too much from initial structures which are complex to start with. If complexity of factorial structures were actually as veridical as I kept hoping it, why then had the partial advantage of complex initial solutions not been recognised during the decades of practice with this method? I feared that *Varimin* rotation might increase a model's complexity to such an extent that valid models of psychological reality might be missed. Without knowing exactly where *complex structure modelling* (CSM) was headed, this could not be ruled out.

Are factors of complex solutions interpretable? Thurstone and his followers maintain that an interpretation of extracted factors requires a simple structure (SS) transformation. Apparently, they had not considered or were not aware of a sim-

ple method, useful for feature and componential interpretations, developed by linguistics, by phonologists in the first place (1952), called “minimal pair comparison”. Two variables A and B may both be complex because each bears multiple latent features. These cannot be captured in their entirety. But if (n-1) features of A and B are equally pronounced and if only one feature has opposing characteristics (e.g., positive vs. negative factorial loadings), then it is possible, by comparing the meanings of A and B, to attribute a conspicuous difference between A and B to the contrasting feature only (cf. Table 2.01).

Table 2.01: Clarifying the minimal pair procedure: Variables A and B display contrasting values of feature d only. Similarly pronounced characteristics of A and B (a, b, c, and e) need not be considered if the characteristic, causing contrast (d), is to be identified.

Variables	Factors				
	a	b	c	d	e
A	-2	+1	-1	+3	+2
B	-2	+1	-1	-2	+2

A chair and a stool differ by the presence or absence of a backrest, a mare and a stallion by their gender, a mountain and a hill by their height. It is more reliable to identify an individual contrasting characteristic by minimal pair comparison, as manifested by *Varimin* CSM, than to try to embrace communalities in a cluster of unanalysed SSM variables. The issue of interpretability of factors will be examined further under “Question II”.

In this chapter, the advantages of complexity-oriented factor analyses will be discussed by using ten empirical test runs. For this purpose, data sets are preferably used whose characteristics are largely transparent and almost self-evident even without factor analyses. Why did methodologists hardly ever make use of this obvious testing strategy when the efficiency of their procedures needed proof?

Questions and Test Runs

Chapter 1 found fault with conventional EFA. It was argued that correlations among manifest variables of an empirical domain were generally based on *multiple* variance and covariance sources. Individual variables do not reveal simple structures; but almost exclusively complex structures. Thurstone’s SS principle presuming a preponderance of monofactorial conditions of manifest variables was discarded as misconception.

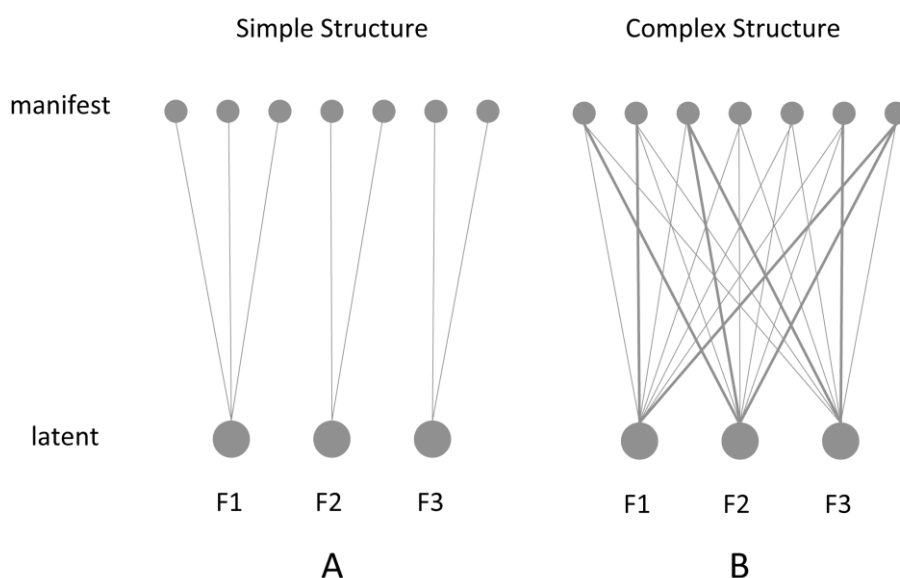


Figure 2.01: Simple structure (A) and complex structure (B) of relationships between sources of covariance (possibly latent) and manifest variables.

Thurstone’s simple structure model (SSM) differs from the complex structure model (CSM) as follows: SS rotation (Figure 2.01A) assigns only one extracted factor to individual manifest variables, or as few factors as possible. On the other hand, CSM rotation (Figure 2.01B) aims to link individual variables with as many extracted factors as empirically possible. How is this goal achieved? By replacing SSM rotation with its inverse. The most frequently used SSM procedure, Varimax (see Figure 2.02), is replaced with what I call *Varimin*, a term denoting that what is maximised by Varimax is minimised by *Varimin*. Varimax *increases* the variance of the squared factor loadings per factor by pairwise rotation of the factor coordinates. This procedure is repeated until the sum of loading variances for all factors

cannot be increased any further. Criterion V (cf. equation 1), which is the Varimax criterion, is maximised. *Varimin* coordinates are rotated with the aim of iteratively reducing the variance until the sum of the squared loadings, i.e., criterion V, cannot be reduced any further.

$$V = n \sum_{p=1}^m \sum_{j=1}^n \left(\frac{b_{jp}}{h_j} \right)^4 - \sum_{p=1}^m \left(\sum_{j=1}^n \frac{b_{jp}^2}{h_j^2} \right)^2$$

Equation 1

b = factor communality

p = running index for factors 1 to m

b = factor loading

j = running index for variables 1 to n

In transformations towards SS, orthogonal procedures are the method of choice, as it also is for Kaiser's Varimax-rotation (cf. Figure 2.01). But unlike SSM, CSM does not also consider oblique rotation, oblique rotations had been introduced to come closer to the SSM utopia which runs counter to a CSM realism.

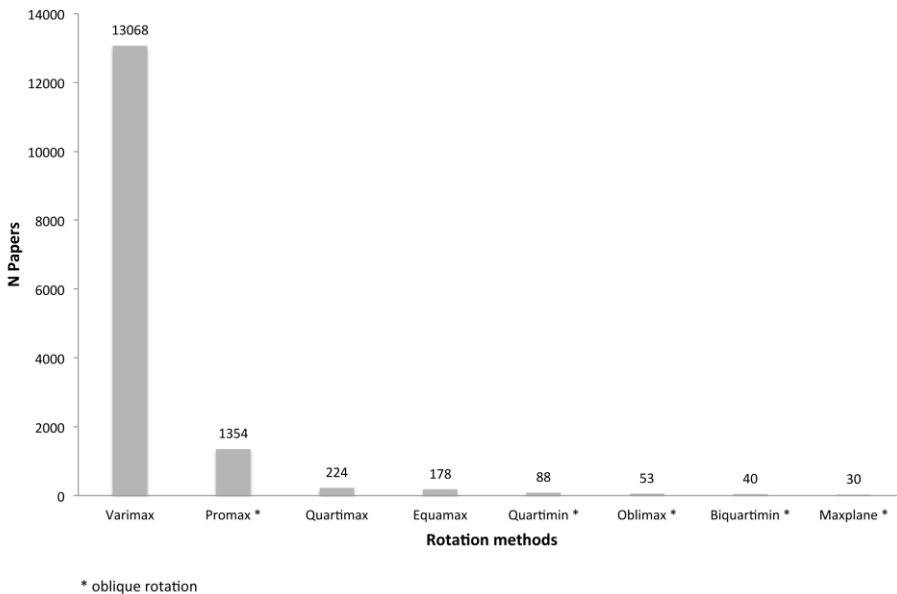


Figure 2.02: Number of articles with reference to factor rotation aiming at simple structure (SS). Result of an Internet search with keywords indicating rotation procedures (Science Direct, Elsevier, 2008).

Assuming an initial structure of Figure 2.03A is available, applying Varimax transforms this structure into structure 2.03C. In 2.03C, the coordinates intersect clusters of variables, and for individual factors the sum of their squared weights is maximised. Applying *Varimin* rotation transforms 2.03A into 2.03B. The distance between the clusters of variables and the coordinates is increased as much as possible. The sum of the squared weights for the factors is minimised.

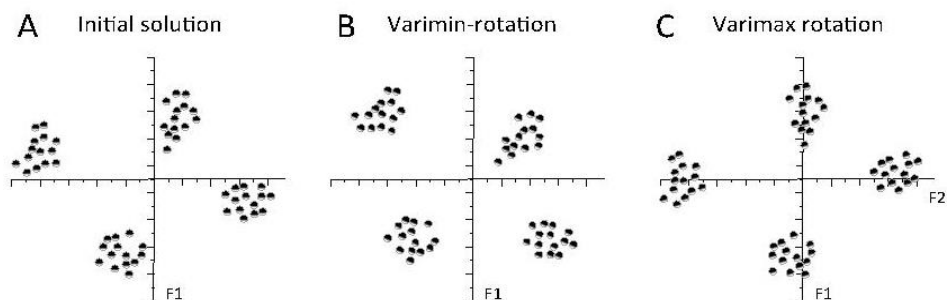


Figure 2.03: Two-factorial initial solution with fictitious variables (A) after *Varimin* rotation (B) and *Varimax* rotation (C).

Introducing *Varimin* as a procedure for factor transformation engenders new questions. Six of the most urgent ones are dealt with in the following, including, wherever possible and advisable, support by empirical checks.

Question I:

Why use complex structure rotation at all?

Are initial solutions not complex enough?

A general result obtained by numerous *Varimin* rotations allows the conclusion that initial factor solutions tend to reliably announce the result obtained from Varimin rotations, but only where not more than two factors are validly interpretable. With three and, above all, more factors the results of unrotated and *Varimin*-rotated factors may considerably differ. Three and more varimin-rotated factors of an analysis are generally more interpretable compared with factors from initial solutions. Since they are more interpretable in case of $k > 2$ substantive extracted factors, I recommend that *Varimin* rotations should always be applied to initial factors disregarding that – with only two interpretable factors – numerical differences of loadings between initial and Varimin-rotated factors are generally negligible.

Why are solutions with three or more *Varimin*-interpretable factors less safely interpretable on initial extraction levels or not at all interpretable? The reason is

that once a factor has been extracted, the variance contained in a correlation matrix is exploited for that factor more than optimally. After initial (PCA) extractions, factors are uncorrelated (they show zero intercorrelations or intercongruences). Factors of initial solutions therefore are exactly orthogonal to each other. The mathematical constraint towards zero congruence among extracted factors ignores empirical deviations from an imposed model of complete factorial independence. To some extent, *Varimin* considers not quite model-fitting underlying correlations among factors. Therefore, after applying *Varimin* rotations some small factor intercongruences are generally noticeable. For comparison, after applying Varimax, intercongruences are considerably larger. Thus, given orthogonal rotation in both cases, Varimax does not achieve orthogonality as perfectly as *Varimin* does and minor deviations from perfect independency of factors by Varimin rotation is a reasonable compromise with empirical conditions.

• **Question II:**
**How can *Varimin*-transformed factors
 be interpreted?**

SS rotation of factor solutions is traditionally deemed necessary because initial factor solutions are complex and factors are regarded as not interpretable if manifest variables have more than one substantial factor loading (Guilford, 1952, p. 27, Burroughs & Miller, 1961, p. 37, Überla, 1971, p. 175, Gorsuch, 1974, p. 162, Comrey, 1978, p. 653 f., Reise et al., 2000, p. 292).

The argument that factorial complexity of variables impedes or even prevents interpretation loses its weight if the concepts ‘distinctive features’ and ‘minimal pair comparison’ are considered methodically. These concepts have been developed in phonology by Jacobson and Halle (1956). They have been widely acknowledged as a significant methodological improvement in linguistics. In a generalised form, the logic of this procedure can be transferred to and used for other disciplines.

The following example provides details: Phonemes are distinguishable by “features”, provided they have a distinctive function within a specific language. By individual observation and objective research, three categories of features were found to be relevant for every phoneme specifying their articulation: Duration of articulation (*short* vs. *long*)⁴¹, sonority (*voiced*, *unvoiced*), and the location of articula-

⁴¹ Other distinguishing features generally used in determining the articulation mode (plosive, fricative, nasal, etc.) are more specific and cannot be determined with the phoneme sample chosen here. [m] and [f] are phonemes with longer articulation (as opposed to short articulation of the plosive sounds [b], [p], [t], [k]). The articulation of [m] is categorised as a nasal sound, whereas [f] is classified as a fricative sound. As all plosive phonemes are short, it would have been possible to create a bipolar category of “plosive” vs. “non-plosive” phonemes and to expect these to emerge as a factor in this assessment.

tion (*bilabial*, *labial-dental*, etc. all the way to *uvular*). Here, the terminology of FA may be applied, because it is safe to say that phonemes are manifest units. Underlying (“latent”) sources of variance, distinctive features, are conceivable as factors. Every phoneme can be defined by three sources of variance. [b], for example, is a phoneme defined by short duration of articulation (plosive), is voiced, and is bilabial (formed with both lips).

Now, how can the three features of phoneme [b] be detected? *Minimal pair comparison* is required. To form a pair for [b], another similar unit from the same domain is needed. In German, the phoneme [p] shares two features with [b] (same length and location of articulation), [b] and [p] differ only regarding sonority. Using phonemes as variables including [b] and [p], their differences/similarities can be rated, yielding a matrix of intercorrelations. A factor analysis should extract three distinctive features as factors, and a subsequent *Varimin* rotation should identify the three factors F_1 , F_2 , and F_3 .

Phonemes [b] and [p] should show similar loadings for, say, factors F_1 and F_2 , but then F_3 should show contrasting loadings. A factor analyst would only need to interpret the difference between [b] and [p] in F_3 : In this case, F_3 would exhibit the difference regarding sonority. This interpretation could be verified if more minimal pairs were put together taken from this data set, for instance, the pairs [d] vs. [t] as well as [g] vs. [k] which all differ regarding sonority. They would exhibit the same F_3 difference⁴².

Test run 1: Evaluating phoneme similarities

(Data new: unpublished)

The above considerations were used in an experiment with two German-speaking students, one a psychology student with an obvious gift for languages and the other a student of advanced linguistics with phonology as sub-discipline. Both participants were asked to assess similarities among 10 German phonemes ([b], [d], [f], [g], [k], [m], [n], [p], [t], [v]) on bipolar seven-point Likert-scales. These 10 phonemes under investigation were used in pairs with all combinations (=45 combinations) in random succession.

⁴² The factor analytical research strategy favoured here can be directly tied to theoretical and empirical approaches, where concepts and other cognitively represented objects may be perceived as bundles or as structures of more or less latent components (“component model” of objects with common settings (Feger, 1979), “Feature Pattern Analysis” (Feger & Brehm, 2001). Also, the strategy is in line with the efforts to explain the phenomena of “experiencing similarities”. Shepard (1974) analysed them as “hidden structures” and Tversky (1977) as “collections of features”.

The phonemes had to be rated according to their similarities on scales such as

[d] 3---2---1---0---1---2---3 [m]
 [b] 3---2---1---0---1---2---3 [d] etc.

and on 43 other scales. If a phoneme represented a scale polarity, e.g., if [d] was to be rated on a scale like [d] 3—2—1—0—1—2—3 [m], the participants were instructed to mark level 3 on the left (maximal ‘resemblance’ or identity).

The ten resulting profiles comprising 45 judgments each were intercorrelated for each participant and subjected to independent principal component analyses (PCA). The first three extracted factors (=components) were *Varimin* rotated, since three interpretable factors were expected (eigenvalues of the first five factors: 2.23, 1.58, 1.40, 1.12, 1.05 for participant 1, and 4.86, 1.68, 1.09, 1.03, 0.59 for participant 2).

The validity of *Varimin* rotations is estimated by how well the factorial loadings comply with the expected classifications. According to phonological classification, the following phonemes have short duration of articulation: [b], [d], [g], [k], [p], [t] (called *plosives*). The following phonemes [f], [m], [n], [v] are articulated longer with applying continuous air flow. Phonemes [b], [d], [g], [m], [n], [v] are voiced, whereas [f], [k], [p], [t] are unvoiced.

While *articulation duration* and *sonority* result in bipolar classifications without further differentiation, the *location of articulation* has four alternatives with linear order. In the following examples, the *location of articulation* ranges from “fully in front” to “far back”: (1) bilabial phonemes [b], [m], [p], (2) labiodental phonemes [f], [v], (3) alveolar phonemes [t], [d], [n], and (4) velar phonemes [k], [g]. As a next step, the loadings of the *Varimin* rotated factors, based on participant data, are related by point-biserial correlations to *length of articulation* (short=1 and long=2) and *sonority* (unvoiced=1 and voiced=2). The four ordinal levels of the *location of articulation* (1=fully in front, to 4=far back) are product-moment correlated with the obtained factor loadings.

This correlation is shown in Table 2.02. It can be seen that the correlations are almost all larger than .90. The only exception is the psychology participant’s correlation for articulation location ($r = .671$). F_1 represents articulation duration while F_2 represents sonority. Apparently, because of her training, the linguistics student had acquired a finer ability to perceive locations of articulation (the Tucker F_3 congruence for both participants amounts to only .574). But as the loadings for F_3 are much larger for the linguistics student than for the psychology student, averaging both sets of data with Fishers Z transformation results in an F_3 correlation of $r = .922$ with articulation location. It is thus safe to display unified sets of data for all three factors.

Table 2.02: Correlations of factor loadings with objective rankings and Tucker congruencies of factors for two female participants.

Factors		Correlations of factor loadings with objective rankings			Congruence of factors of participants
		Participant Psychology	Participant Linguistics	Both Students	
F ₁	Short-long	.972	.990	.996	.943
F ₂	Voiced- voiceless	.917	.957	.966	.936
F ₃	Place of articulation	.671	.976	.922	.574

Figure 2.04 shows the *Varimin* results of the two students’ combined data. Positive factor loadings are represented by dark circles, negative ones by lighter circles. The different sizes of the circles represent absolute loadings. Zero loading would be represented by a point without dimension. If these three factors were not already correlated with expert judgments, the minimal pairs [t] vs. [f] and [d] vs. [n] might have been formed for F₁, articulation duration as a distinctive feature. In case of F₂ the minimal pairs of “sonority yes” vs. “sonority no2 would have stood out: [b] vs. [p], [d] vs. [t], [g] vs. [k] and [v] vs. [f]. In F₃, the minimal pairs [p] vs. [k] and [b] vs. [g] would have manifested the “front – back” contrast of articulation locations

Varimax results of the two students differ considerably. The linguistic student’s results are as follows: Varimax F₁ with positive sign clusters *long, voiced* phonemes [m], [n], and [v], and with negative sign the *short, unvoiced* phonemes [t] and [k]. The *short, voiced* phonemes [g] and [d] with positive sign are clustered by F₂. The *long, voiced* phoneme [f] has a negative F₂ loading. Unipolar factor F₃ clusters the bilabial plosive phonemes [p] and [b]. It is clear that Varimax rotation clusters similar phonemes. But these clusters have multiple phonetic features that do not stand out as phonetic features. Phonetic features underlying similarities and differences among phonemes are *not* factorially discovered, rather they are disguised. (The results of the psychology student would not change this conclusion). Thus, CSM-transformed factors (by *Varimin*) can be interpreted as latent sources of variance without difficulty, as long as minimal pairs are formed. The interpretation of factors of SSM-oriented factor analyses is considerably more difficult, since *global similarities* among factorially clustered variables, based on underlying multiple features, are unsuitable in principle to identify these features. Seemingly paradoxically, the CSM result is simpler than the SSM result.

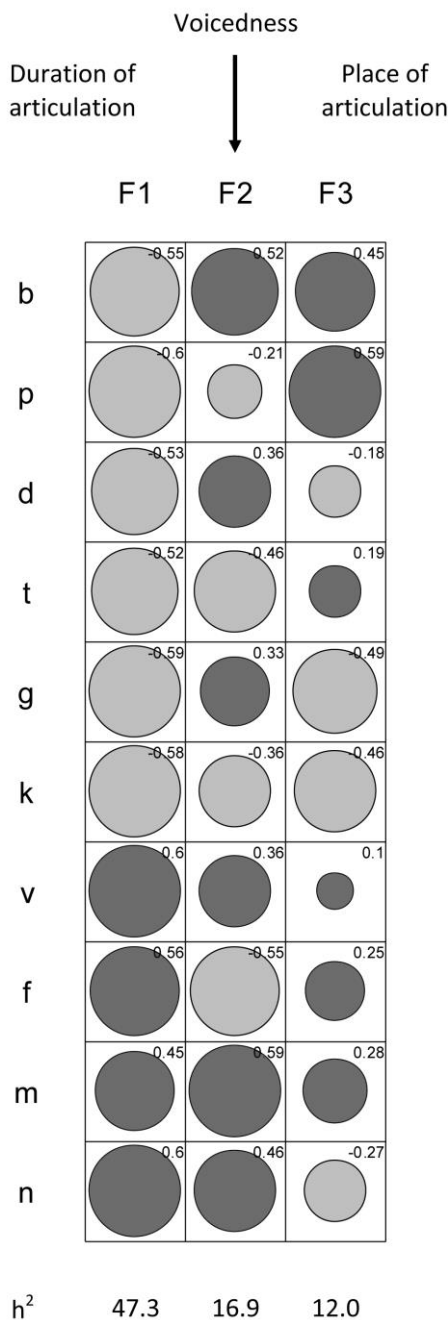


Figure 2.04: Varimin-transformed factor loadings of ten German phonemes (based on similarity judgments).

Differential results from *Varimin* and Varimax factor rotation are just as convincing in a study using similarity judgments of British coins.

Test run 2: Similarity judgments of British coins

(Novel data: unpublished)

British coins were chosen because similarity ratings of coins, even more than of phonemes, are based on perceivable objective features. i.e., with British coins on size, shape, and colour (see Figure 2.05). Coins of adjacent values differ in size (e.g., *1 pence* small, *2 pence* large, *5 pence* small, *10 pence* large, etc.).

Natural pairs of coins are thus formed by their size, additional pairs or groups suggest themselves with colour of the metal (e.g. *5* and *10 pence* are silver, *1* and *2 pence* are not silver) or with their shape (e.g. *5* and *10 pence* are round, *20* and *50 pence* are heptagonal). These features are expected to influence the similarity ratings among these coins. The diameters and weights of the coins are given in Table 2.03.

Coins	Colour	Size	Form
 1 Pence	not silver	smaller	round
 2 Pence	not silver	bigger	round
 5 Pence	silver	smaller	round
 10 Pence	silver	bigger	round
 20 Pence	silver	smaller	not round
 50 Pence	silver	bigger	not round
 1 Pound	not silver	smaller	round
 2 Pound	not silver	bigger	round

Figure 2.05: British coins with the attributes of colour, size, and form.

The main experiment was conducted using a German student (IS) attending a college in Cambridge, UK. She was asked to rate all eight current British currency coins (1, 2, 5, 10, 20, and 50 pence and 1 and 2 pound) according to similarity. The coins were stuck on cardboard and presented in pairs: Each coin was paired with every other coin, beginning with the pairs 1 penny⁴³ – 2 pence, 1 pence – 5 pence etc. up to the pair 1 pound – 2 pound, with 28 pairs in all.

The student was asked to hold one of the eight coins and compare it with all pairs of coins on the cardboard. On a seven-point bipolar Likert scale she had to indicate whether the coin in her hand resembled the coin on the left or the coin on the right on the cardboard. For example, she may think the 50 pence coin resembles the 5 pence coin better than the pound coin. In this case she should mark the 5 pence – 1 pound scale at a point close to the 5 pence side. An all embracing, holistic judgment was requested. All features influencing similarity and difference were to be considered concurrently, but the face value of the coins was to be ignored. The eight coins obtained 28 ratings each.

Table 2.03: Diameter and weight of British coins.

	1 Penny	2 Pence	5 Pence	10 Pence	20 Pence	50 Pence	1 Pound	2 Pound
Weight (g)	3.28	7.10	5.52	6.50	5.00	8.20	9.48	11.92
Diameter (mm)	20	25	18	24	20	28	22	28

In this manner, an assessment profile for each coin, consisting of 28 individual similarity ratings, was created. The profiles for the eight coins were intercorrelated, the correlation matrix was subjected to Principal Component Analysis (PCA), and the extracted factors were subjected to *Varimin* and Varimax rotations. I expected that the features of the British coins would assert themselves factorially by a *Varimin* rotation, but not by a Varimax rotation, and that the *Varimin* rotation would show the features of coins better than the initial solution.

The results (Figure 2.06) met all expectations⁴⁴. The most prominent factor (F₁), distinguishes with positive and negative signs markedly silver coins from non-silver ones. The second factor reveals, with bipolar loadings, the larger and smaller coins. The third factor distinguishes between the two shapes of coins⁴⁵. An interpretation by meticulous minimal pair comparisons is not necessary, given the transparency of this “latent” set of conditions.

⁴³ Colloquially, *one pence* is often preferred to *one penny*, which is the correct expression.

⁴⁴ Circle diameters indicate loading levels. The lighter circles represent positive loadings, the dark ones negative loadings. Numerical values are displayed in small print in the cells of the matrix. Percentages of factor communality appear below the columns.

⁴⁵ The slight variations in loading levels of F₂ (especially the 2 pence coin) may possibly be attributed to the subjective feature evaluations of the coins. Actually, the 2 pence coin seems very large considering its low value. Size seems to make a greater impression than shape.

Which information does the initial solution provide? (see Figure 2.06): Among the initial factors, F_2 may be interpreted as a manifestation of coin size. F_1 seems to be the colour factor, but the loadings of the 5 and 10 pence coins do not tie in with colour. F_3 cannot be interpreted as a shape attribute and remains a mystery in the initial solution. The initial solution is thus less satisfactory as are many other initial results of other such studies. This suggests one practical conclusion: PCA factors should always be rotated with *Varimin*, even if the result does not differ significantly from an initial solution.

The Varimax solution: Only in half the coins is the ideal of a monofactorial loading achieved. Both the 10 pence and the 20 pence coins deviate significantly from a solitary ideal loading. While multiple factor loadings have always been common in simple structure practice, they were tolerated as an exasperating nuisance. More importantly, Varimax does not reach descriptive simplicity at a contextual level. Neither colour, nor form, nor size is represented by Varimax factors – they are not even hinted at.

To test whether the results, obtained with participant T.S., may be generalised, the experiment was repeated with eight students who had no experience with British coins. The results were averaged and analysed in the same manner as were the results of T.S. The eigenvalues 2.07, 1.90, 1.30, 1.20 allowed an extraction of four factors. In the case of T.S., only three factors were extractable (eigenvalues: 2.87, 1.92, 1.24, 0.97).

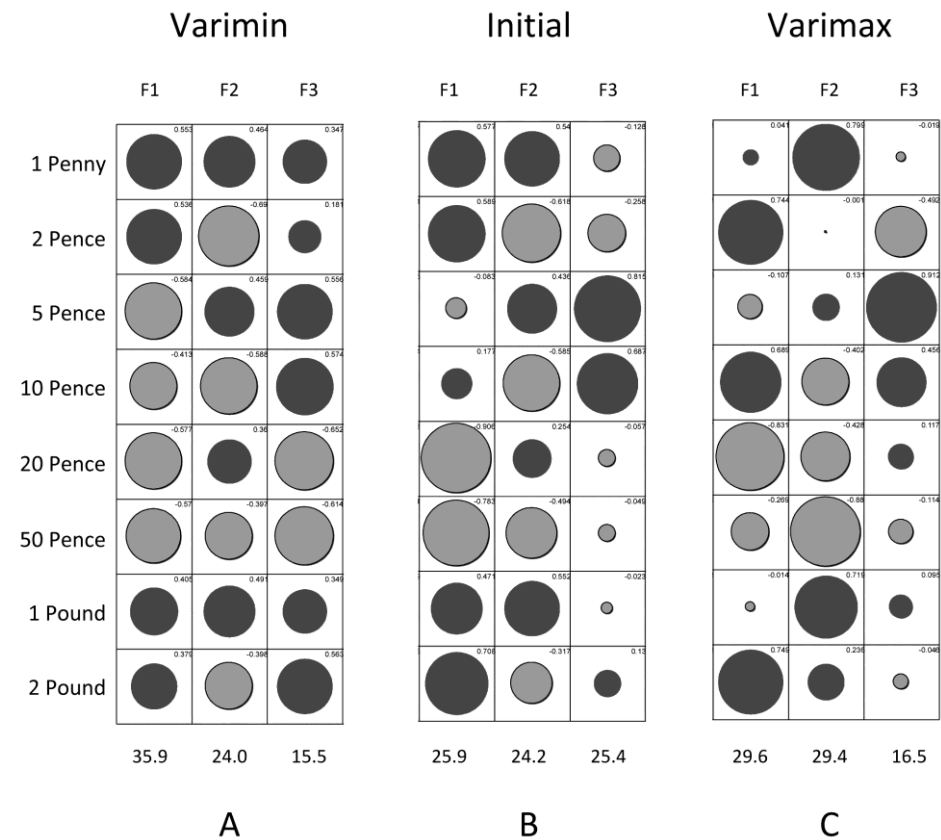


Figure 2.06: *Varimin* (A), *Initial* (B) and *Varimax-transformed* (C) solutions. Similarity judgments of British coins.

The factorial congruences of the solutions to be compared, i.e. of T.S. and the students, were significant for F₁ (factor of colour) and for F₂ (factor of size)⁴⁶, amounting to .981 and .973 respectively. However, there was not even a hint of congruence between F₃ of T.S. and either F₃ or F₄ of the students (.312 and .224, respectively). However, the students' data showed a high correlation between the eight F₃ loadings and the eight face values of the coins, while the largest correlation T.S. achieved with face values was only .14 (for F₁).

It seems that the similarity judgments of the students were influenced not only by the colour and size of the coins but also by their monetary value. They ignored the shape of the coins which was clearly manifest for T.S. as F₃. It had been men-

⁴⁶ The interpretation of *Varimin* F₂ as a factor of coin size is supported by correlating the F₂ loadings with the measured size of a coin: $r = .81$ (T. S.), $r = .85$ (students). The weight of the coins in grams only marginally correlates with the factor loadings, although it correlates with their size up to $r = .80$.

tioned in the written instructions to the students that the monetary value of the coins should be disregarded, but this was less emphasised compared with the oral instructions given to T.S.

• **Question III:**

**Interpretations of factorial simple structure-solutions
must have been fairly satisfactory in the past,
why else would they have been constantly used?**

The variables in the studies referred to so far were physical objects whose perceived features had been assessed and compared by test participants. Such data are rarely subjected to SSM factor analysis. Similarity judgements and ratings of complex givens are unsuited for SSM analyses. Alternative multidimensional procedures are used for this purpose, e.g. multidimensional scaling (MDS).

In psychology, FA is extensively used for rating people, for self-and external assessments, assessing personality traits, behavioural dispositions, attitudes etc. Traits and behaviours are rated by verbal items without objective references. The validity of factors extracted from semantic material cannot be appraised by objective means. Factors extracted from such material are prone to noncommittal interpretations with considerable subjective freedom.

I think the lack of rigorous methods for validating factors obtained from verbal material is one of the reasons why FA of psychological data has often also been regarded as fairly satisfactory. The reason – I guess – is that the modelling quality of SSM factors cannot be based on evident criteria, as is the case with phonemes or coins. Factors based on semantics can easily somehow “make sense”, they can almost always be interpreted in one way or another. Using verbal variables, denoting mental or otherwise non-evident matter, deficiencies of SSM cannot readily be identified. Although occasional reports exist about significant correlations between factorial self- or external assessments and objective behavioural data, factor analysts have to construe meanings conceptually anyway for determining latent sources of factorial variance.

In the following study, correlations between rating preferences are factor-analysed. At this stage, semantic problems are avoided because the variables used are adverbs without content indicating mere degrees of Likert-scale judgments. They are taken from a study by Carl (1968) about response-style behaviour. With the help of expected *Varimin* results it can be shown, nevertheless, how meanings of SS factors are construed.

Test run 3: Differentiation of response styles at responding to questionnaires

(Data: Carl, 1968).

Carl (1968) aimed to determine response sets of participants responding to Likert scales with five response alternatives. He collected from $N = 580$ participants Likert ratings for 580 items of the *Minnesota Multiphasic Personality Inventory* (MMPI) from 100 persons. Across items he summed the ratings 1 to 5 separately for each rating point and for each participant. The scale ranged from “strong approval” to “strong disapproval”, intermediate steps of the scales were not verbalised). Carl excluded artefact correlations by separating five subsamples of items, rendered parallel with regard to content, one subsample of items served for one rating point.

For the present purpose, a 5×5 intercorrelation matrix was picked from Carl’s paper, each correlation was based on frequencies of usage of respective rating points. For this matrix, the initial PCA factor structure was determined that was rotated using *Varimin* and for comparison, also *Varimax*.

An expert in response set research would expect certain results: Acquiescence, the tendency to respond affirmatively, is well known and should generate a factor with positive loadings for the two affirmative response alternatives and negative loadings for the two disapproving response alternatives. Equally well known is the tendency to give extreme responses, and this should produce a factor in which the two extreme alternatives, one for extreme affirmation and one for extreme disapproval, should have positive loadings and the intermediate alternatives between the extremes negative loadings. At least these two factors should come to light by *Varimin* rotation.

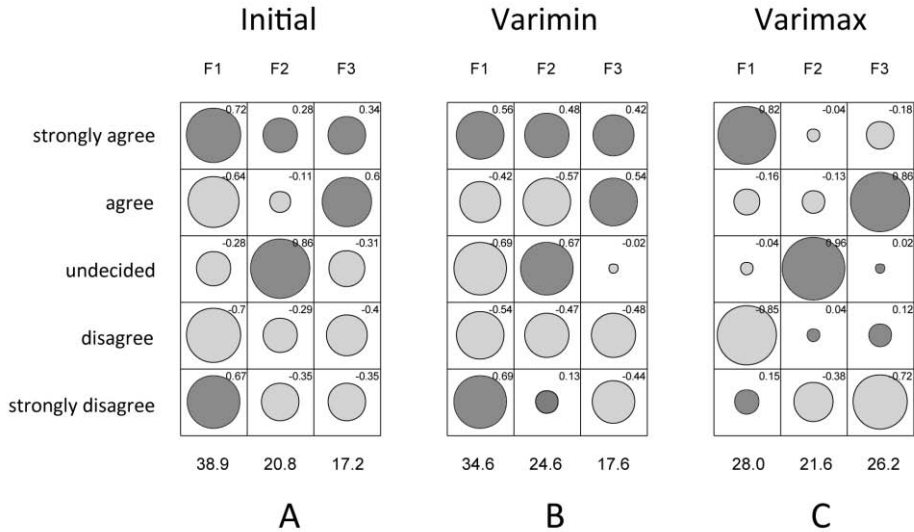


Figure 2.07: Initial (A), Varimin- (B), and Varimax: structure (C) of response set data (source: Carl, 1968).

The *Varimin* result is shown in Figure 2.07B. The expected acquiescence factor is F_3 and the extreme response factor F_1 . However, a substantial unexpected F_2 is evident and demands interpretation. The matter can be followed up with Herrmann (1965), who identified what he called “distinctiveness of judgment” (*Urteilsnuanciertheit*) as another response tendency. Accordingly, in scales with response alternatives respondents are not only distinguished by approving and disapproving a statement, but also by the extent to which they differentiate between “approve” and “greatly approve”, “disapprove” and “greatly disapprove”, as well as between “approve /disapprove” on the one hand and “undecided” or “don’t know” on the other. In pertinent literature, this response variance is rarely discussed, presumably because it does not distort results to the extent that the other response tendencies do (Hinz et al., 2003). In my opinion, factor F_2 of the present *Varimin* analysis may safely be interpreted as “distinctiveness of judgment”.

Are the three response sets already recognisable in the *initial* factor structure (cf. Figure 2.07A)? The initial factor structure resembles the *Varimin* factor structure. In both solutions, the F_1 factors are virtually identical. Hence, the initial F_1 factor may *also* be interpreted as indicating an extreme response set. The initial solution, however, is not satisfactory with F_3 , i.e., because the middle rating alternative, used for refraining from judgment, features a considerable negative loading. This should not be the case, because judgment abstention is supposed to be dependent on factors other than negative judgment. Furthermore, the *Varimin*

solution for F_2 is more distinct. In the initial solution “disapproval” and “strong disapproval” in F_2 are not distinguishable. Also, the numerical difference between “approval” and “strong approval” in the initial solution is considerably weaker than in the *Varimin* solution. The initial F_2 factor can thus hardly represent distinctiveness of judgment.

Our main concern here is the Varimax result (Figure 2.07C). How is this to be interpreted? Comparing Varimax and *Varimin* solutions (2.07C and 2.07B) sheds light on this. The Varimax rotation led to a bipolar clustering of “strong approval” (positive loading in F_1) and “disapproval” (negative loading in F_1). Why? It is noticeable that the “strong approval” and “disapproval” loading signs have opposite signs of the three features as differentiated by *Varimin*. Contrasts in the *Varimin* feature profiles are also found for Varimax factor F_3 , a bipolar factor as well. In the *Varimin* result, there is no polar opposite for the “undecided” category of judgment, hence “undecided” remains fairly isolated in Varimax F_2 .

These findings conform to the results of the *Varimin*-Varimax coin factor comparison. In short, variables with similar *Varimin* profiles of features tend to be clustered by Varimax transformations⁴⁷. In the case of bipolar factors, Varimax also clusters variables with opposite profiles (with their signs reversed). Varimax does not analyse features. Varimax clusters variables, using *Varimin* features, feature differentiation is thus prevented⁴⁸.

What might a conventional factor analyst publish after a Varimax analysis of these response data? He might argue that acquiescence is not a monofactorial construct, as has been traditionally assumed. Rather, he would go on, a distinction should be made between acquiescence I (F_1) associated with extreme response tendency and acquiescence II (F_3) *without* such tendency. Moreover, he might interpret F_2 as an “undecided” factor and simply ignore that F_2 shows considerable negative loading even for “strong disapproval”. He might rely on his SSM data processing and believe he had discovered three new psychological constructs (acquiescence I, acquiescence II and ‘undecided’. Since the terminology does not appear senseless, nobody might notice that these factors represent a rather useless collection of variables, as they are constantly put forth by SSM. The following - analysis of *verbal* data from an MDS study further exposes this non-committal practice of interpretation.

⁴⁷ Zimmermann (1953) remarked that “a test which actually contains variance on two or more factors may appear with all of that variance confined on a single factor.” This he calls the “composite factor”. The author thus sticks to the literal meaning of the initial factors (centroid factors), which he deems composed in the rotated factor. “It is my feeling that the failure to give composite factors the attention they merit must be considered either a serious oversight or a serious error or omission” (Zimmermann, 1953, p. 389).

⁴⁸ Overall (1964) explicitly claims: “Rotation to simple structure can be understood as an elaborate approach to cluster analysis. It identifies clusters of tests which measure the same things, but there is no assurance that these ‘same things’ are simple and primary dimensions.” (p. 271), and “there is no need to assume that simple structure factors will correspond to any particular set of fundamental dimensions of the objects ...” (p. 276).

Test run 4: Semantic features of kinship terms

(Data: Marx & Hejj, 1989)

The participants were asked to sort cards hierarchically with 16 kinship terms. The results were used by Marx and Hejj (1989) to determine a matrix of similarities (cf. Tables 2.03–2.5, p. 112 in the original). The authors fairly successfully applied an NMDS process to obtain the semantic terms of kinship from the resulting stacks of cards⁴⁹. To this end, the original matrix of sort sequences serving as indicators for distance or dissimilarity (DS) in the 16 terms was mirrored diagonally and extended into a square matrix. Then, the DS-measures were transformed into measures of similarity (S) by $S = 1 - DS/1000$ ⁵⁰. The highest similarity value of each relevant column/row was inserted into the diagonal. The columns of the similarity matrix were subsequently correlated amongst themselves and the inter-correlation matrix was subjected to PCA. Five substantial factors were extracted⁵¹ and transformed by *Varimin* and Varimax.

Figure 2.08A shows the *Varimin* solution. The variance of loadings of the general factor F_1 is only minimal. Apparently, F_1 is a methodical product and may therefore be ignored⁵². Factors F_2 up to F_5 are bipolar. With contrasting signs and by applying minimal pair comparison, the expected semantic features stand out: Lineality (F_2), nuclear family (F_3), gender (F_4), and age or generation (F_5)⁵³. For example, the minimal pair of *brother* and *sister* shows differing loadings only with factor F_4 . It should thus be interpreted as the “gender” factor (male vs. female). Other minimal pairs are easily identifiable, e.g., *father* and *son* or *mother* and *daughter*, whose loading directions contrast only with F_5 , the “generation” factor. The factor structure does not show subtler differences, e.g., distinctions between the youngest, the middle and the oldest generations (as for son, father, grandfather). Also, differences due to the first-person perspective cannot be recognised by factors, e.g. the distinction between (*my*) *brother* and (*my father's*) *son* – denoting the same person. The result shows the main kinship characteristics only.

Figure 2.08B shows the result of a Varimax rotation of the same factors that were rotated by *Varimin* for Figure 2.08A. Of the five rotated factors, F_1 and F_2 cannot be interpreted at all and F_3 only by taking considerable liberties. F_3 combines *grandparents* and *grandchildren*. That may make some sense, since the genera-

⁵⁰ These linear transformations were made to help interpret the individual values in the table.

⁵¹ Eigenvalues: 4.31, 2.34, 1.54, 1.09, 1.05, 0.90, 0.79, 0.75 ...

⁵² F_1 seems to be a result of the hierarchical clustering procedure.

⁵³ These interpretations of *Varimin* factors F_2 , F_4 and F_5 are also found in a similar MDS study by Romney and d'Andrade (1964) (gender, generation, consanguinity). In F_2 (lineality), the present study also differentiates between “nuclear family” and non-nuclear family (by F_3). Moreover, the study by Romney and d'Andrade offers an excellent introduction to the terminological and methodological principles for componential analysis of concepts. It also gives insight into the conclusions based on the existence of discriminative stimuli (sememes) in search of definitions.

tional extremes in the same lineage are here grouped. This cluster might be termed “very young or very old in the dominant lineage”.

Regarding factors F_2 and F_4 , a certain similarity is noticeable between the Varimax and *Varimin* results. Varimax F_2 loads the kinship terms of the *non-lineal* lineage (“extended family”). Unlike the *Varimin* solution, Varimax does not indicate *lineal* kinship (“close relatives”) with opposite signs ([+] “feature present” vs. [-] “feature absent”). Instead, all Varimax factors are unipolar. In F_4 , the members of the nuclear family are grouped, but non-members again remain without a sign (no negative sign). Also, the F_4 loadings for *granddaughter*, *niece*, and *female cousin* are quite high, which, however, has no semantic reason.

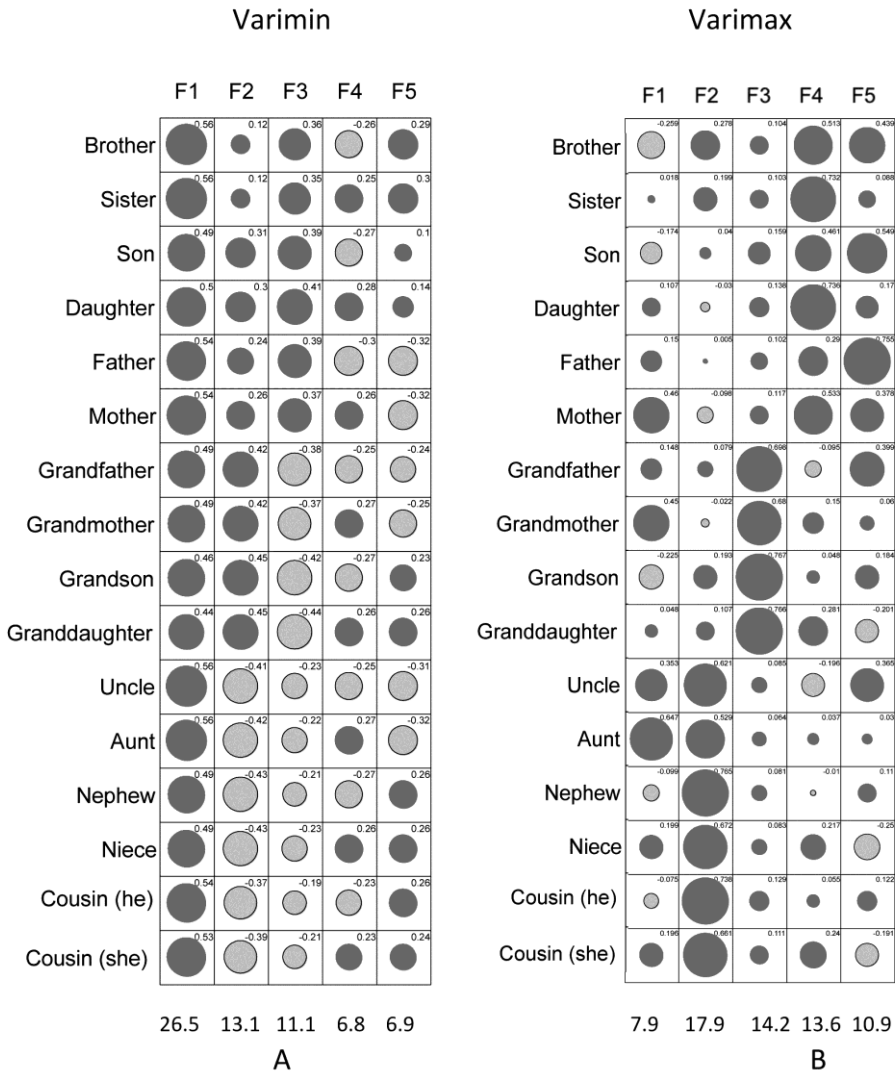


Figure 2.08: Varimin- (A) and Varimax-rotated factors (B) of similarities among kinship terms (data source: Marx & Hejj, 1989).

Nonetheless, while Varimax factors are not entirely uninterpretable, semantic overlaps in this solution render most interpretations unsatisfactory. This becomes evident by using the semantics of kinship terms. The sources of semantic similarity and dissimilarity (generation, gender, etc.), that should be revealed by FA, remain hidden in the Varimax solution. Grandparents and grandchildren are clustered according to lineality and generation, but Varimax does not distinguish the features.

Table 2.04: Data of hierarchical sorting (source: Marx & Heij, 1989).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 brother	1															
2 cousin (he)	.759	1														
3 cousin (she)	.858	.451	1													
4 grandson	.777	.770	.866	1												
5 granddaughter	.842	.865	.792	.454	1											
6 grandmother	.860	.882	.830	.702	.570	1										
7 grandfather	.783	.822	.877	.564	.695	.444	1									
8 mother	.747	.942	.895	.836	.803	.679	.834	1								
9 nephew	.787	.538	.674	.802	.882	.925	.865	.970	1							
10 niece	.900	.685	.542	.875	.783	.869	.926	.927	.465	1						
11 uncle	.819	.594	.711	.804	.875	.856	.772	.913	.592	.719	1					
12 sister	.432	.857	.764	.841	.781	.774	.853	.673	.884	.797	.892	1				
13 son	.632	.892	.953	.721	.869	.864	.764	.710	.900	.969	.912	.743	1			
14 aunt	.905	.717	.602	.896	.821	.766	.843	.840	.725	.576	.446	.810	.960	1		
15 daughter	.728	.872	.890	.865	.735	.775	.869	.561	.971	.902	.972	.625	.476	.906	1	
16 father	.614	.896	.947	.717	.894	.823	.693	.477	.915	.981	.831	.744	.571	.924	.718	1

• Question IV:

Can any substantial information be gained from the bipolarity of *Varimin* factors?

The bipolarity of *Varimin* factor loadings deserves special attention. Bipolarity is quite common in *Varimin* solutions, but less common in Varimax or other simple structure solutions, if it occurs there at all. The occurrence of negative *Varimin* loadings may mean that these variables (as opposed to positively loaded variables) might have a detrimental effect on the respective co-variance source. For example, in the response set data of Carl (1968) it may be assumed that variables with negative factor loadings might have an inhibitive effect: A subject giving many extreme yes-no answers would obviously not give many moderate answers, and vice versa. This is due to his/her inclination to either avoid or favour moderate judgments⁵⁴. To bipolar factors of acquiescence or distinguishability, motivational and functional interpretation of sign differences in factor loadings are applicable.

Bipolarity in kinship data has to be interpreted differently. For example, in the factor for gender, bipolarity cannot be regarded as functional. The *male*-feature is present not because the *female*-feature is absent. Rather, it is an organismic condition, precluding (as a rule) the feature *female* if *male* is present⁵⁵. Lineality, however, only needs a *yes*- or *no*-answer about the lineage position of a relative in the genealogical tree. In such cases, a positive or negative sign indicates the presence or absence, respectively, of a feature. This interpretation of plus-minus signs is more common in fields such as linguistics than in psychology.

⁵⁴ Although it is possible that participants submitting many negations equally tick “undecided” or give moderate affirmations, this does not seem likely.

⁵⁵ This is not meant to comment on the spiritual polarity in “Animus vs. Anima” by Carl Jung, which maintains that male and female tendencies occur in one and the same biological gender.

If a negative loading of a *Varimin* factor indicates the absence of an effect or function, it may sometimes be concluded that this factor is irrelevant for the respective variables. If a positively loaded bipolar factor is present, this may also depend on other variables of the sample⁵⁶. I therefore suggest one should always interpret negative signs initially to mean “feature not present” only then, – and with contextual knowledge – a functional interpretation (constraint or polarised characteristics) should be made, if appropriate^{57 58}.

The following transparent example should aid a better understanding of the determinants of bipolarity in factors.

Test run 5: Intellectual development in childhood

(Data: Humphreys & Davey, 1988)

In a longitudinal study, Humphreys and Davey (1988) used four age-adjusted intelligence tests on children aged three months to nine years. The authors wanted to determine the consistency of the children’s intellectual development in this period. I copied their resulting intercorrelation matrix for 14 test repetitions in the longitudinal section and applied PCA and *Varimin*- and Varimax-rotation. Figure 2.09 shows the results (the initial solution is not listed; it is virtually identical with the *Varimin* solution). Only two factors are substantial (eigenvalues: 5.705, 1.473, 0.827, 0.788 ...).

The first *Varimin* factor (communality of 40.7%) represents the stable percentage of general intelligence. In the longitudinal test period, the individual averages of factor loadings roughly remain the same. The second factor is bipolar and shows 10.6% communality. It represents the time-dependent variance in intelligence development. This variance is due to specific favourable or unfavourable

⁵⁶ That is why multivariate causal models which were introduced by confirmatory FA and structural equation models are questionable. Modelling sources of covariance – which is the main objective of multivariate analysis – does not necessarily reveal causal sources.

⁵⁷ Because the rotation method allots positive or negative signs irrespective of the assigned content of the measurements, it additionally has to be decided whether the signs, resulting from the statistical program, should be kept or be reversed to improve their comprehension. To facilitate comprehension, variables with a factor indicating functionally restricting influence or where the factorially expressed material has to be regarded as not present, may be given, by sign reversal, a negative sign if it is not present originally. It has become apparent (a rule of thumb so far) that absence of a feature or negative influence of a variance source within a factorial vector tends to appear specifically in those variables that have (1) negative loadings or (2) predominantly either plus or minus signs.

⁵⁸ A neglected hint from the beginnings of centroid factor analysis which grants informational value to unrotated bipolar factors can be found in a note by Zimmermann (1953): “*It is well-known in dealing with intellectual variables that the first centroid loadings are usually all positive and the second centroid, as well as those that follow, divide positive and negative variables equally. What is apparently overlooked is the tendency for the second centroid to split the most obvious dichotomy, the third centroid to split the next most obvious dichotomy and so on. For example, if the battery contains both linguistic and quantitative tests, the second centroid will most likely separate these two major groups ...*” (p. 389).

individual influences exerted on the participants: education, varying psychological or physical condition or illness etc.

The bipolarity of F_2 and the monotonous succession of the factor loadings are to be interpreted as follows: The amount of covariance not exhausted by F_1 is spread evenly across all ages. Changes within a single test interval are smaller than across two or more test intervals. They are largest between the first and the last measuring point.

Accordingly, test results of adjacent test intervals correlate more closely than those spaced further apart. Test results in the middle range of the longitudinal section are equidistant regarding their difference from results at the beginning and at the end, as indicated by correlation coefficients. In the *Varimin* model, the testing occasion located in the middle between the first and the last testing is regarded as the zero point of the F_2 loading-vector. Factor F_2 thus represents the degree of individual instability in intelligence test performance with respect to the average value of its stable level.

This example is informative in as much as it shows that the F_2 factor loadings are related to those in F_1 . The particular meaning of this relationship has to be specified with the help of appropriate contextual knowledge. In the present case, negative F_2 loadings do not indicate missing or contrary effects or logical exclusion, but instead differences in temporal change of individual intelligence test performance which may have, on an individual level, sometimes a positive and sometimes a negative direction. As the F_2 factor loadings are polarised positively or negatively, a scale emerges showing changes in the test results (degree of fluctuation) steadily increasing in the course of this longitudinal study.

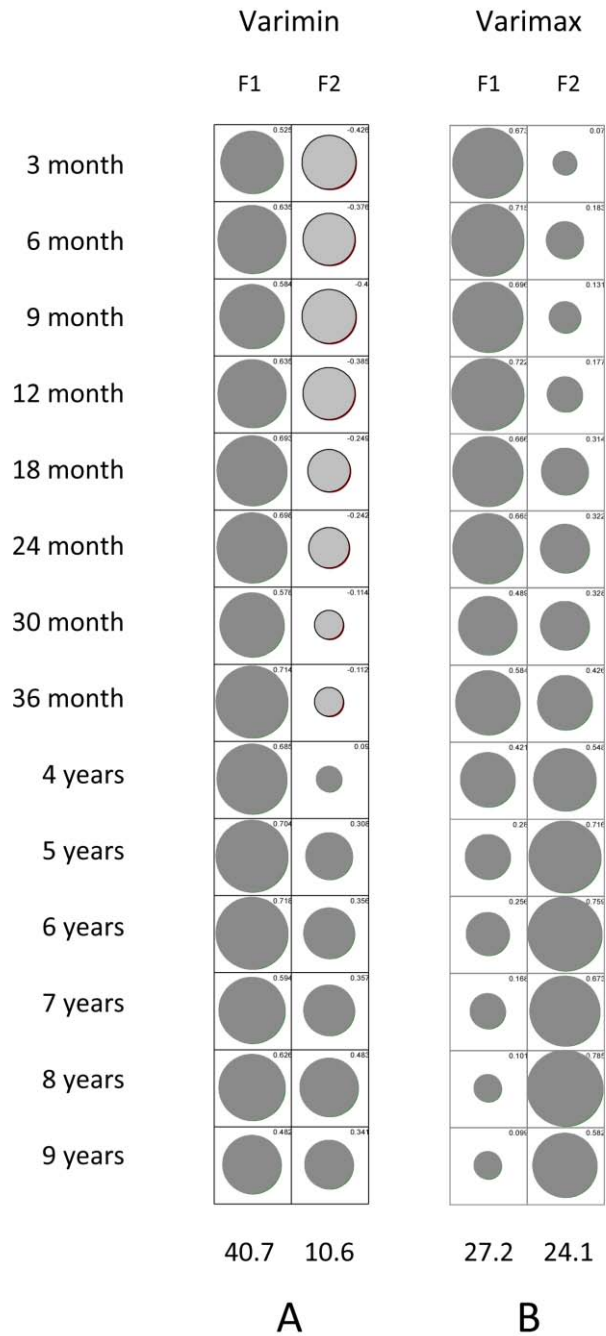


Figure 2.09: Varimin (A) and Varimax solution (B) of longitudinal data of intelligence development (data source: Humphreys and Davy, 1988).

The Varimax solution eliminates the general factor and thus ignores the fact that interindividual variance of intelligence test performance remains fairly constant in this longitudinal study. Instead the Varimax solution comes up with two factors which might be wrongly interpreted as two independent kinds of intelligence. One of them (F_1) might be conceived as affecting early childhood and eventually being superseded by the functioning of the second sort of intelligence (F_2) – an absurd notion that is unlikely to be entertained even by SSM factorists. Given repeated longitudinal section data, these researchers might simply regard their SS method as unsuitable and would probably not use it.

A different example is shown in the following study: Here, the negative characteristics of a factor not only reveal the absence of a feature, but also indicate the presence of a scalometrically independent feature.

Test run 6: Body size and body shape in cattle

(Data: Rasch/Weber, 1962)

Rasch (1962) gathered twelve measures of the body bulk of 107 female cattle (heifers). Height, width, and length were measured. Our usual factorial processing according to E. Weber's intercorrelation matrix produces the bifactorial solution (eigenvalues: 7.69, 1.20, 0.74 ...) in Figure 2.10. The initial structure is not shown, it is virtually identical to the *Varimin* structure.

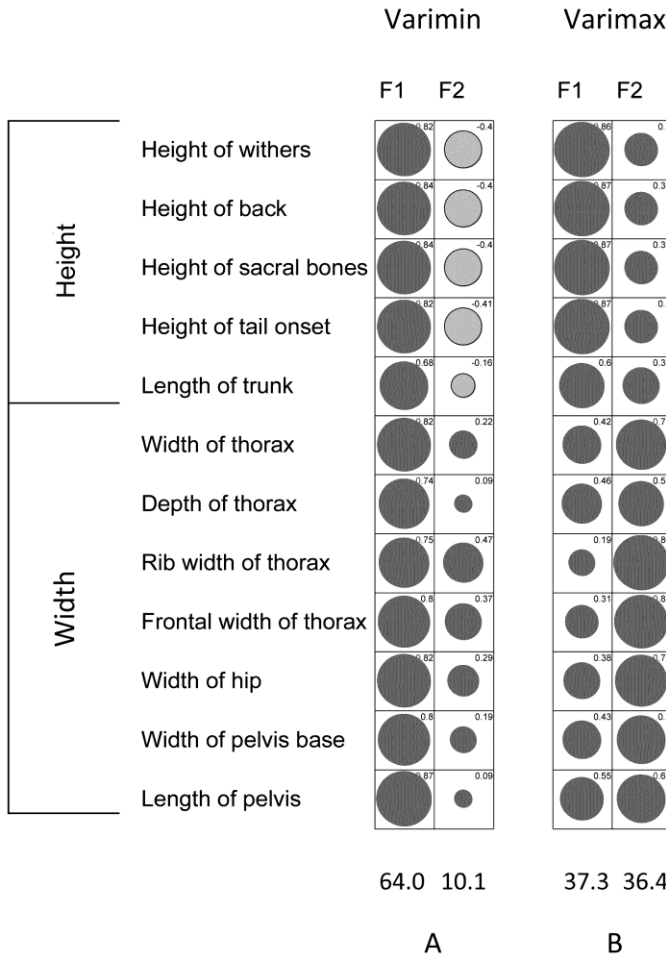


Figure 2.10: *Varimin (A) and Varimax (B) solution of body size and body shape measurements in cattle (data source: Rasch, 1962).*

The first factor (communality 64.0%) represents the body bulk (body size, or mass) without specifying spatial dimensions. It shows that cows varying above or below average in height also tend to vary above or below average in length and in width. Here, F_1 can be interpreted as the general factor of morphometric measurements. The bipolar second factor (communality 10.1%) represents the variance in body shape. Among large and small cattle relatively delicate (slender) or bulky (fat) animals exist. The polarity indicates that growth in height occurs ‘at the expense of width’, as it were, and growth in width takes place ‘at the expense of height’.

In the Varimax solution the general factor of body bulk is lost. The variables slenderness and fatness are evident as well, but they are represented as non-polar orthogonal dimensions. The interconnected and somewhat opposing manner of expansion in height and width does not take effect in the SSM result. However, when using *Varimin*, this manifests itself as an influential bipolar factor for body shape, aside from the dominant volume or mass factor⁵⁹.

Even in psychological data, interpreting negative factor loadings as bipolar traits can seem obvious, as demonstrated in the following test run.

Test run 7: Intelligence tests and performance tests

(Data: Holzinger & Swineford/Jöreskog & Sörbom, 2003)

Holzinger and Swineford's data are used occasionally in educational textbooks to demonstrate model calculations, for instance by Jöreskog and Sörbom (2003). The table of intercorrelations was taken from their website (Google search: <LISREL 8.52 Jöreskog>).

Holzinger and Swineford used three tests each for visual, verbal and for speed performances (speed tests), respectively. Figure 2.11 contains the *Varimin* and Varimax standard results. Unipolar *Varimin* F_1 (Figure 2.11A) apparently is general factor g , general intelligence, which is generally expectable from intelligence test series. Bipolar factor F_2 contrasts, by positive sign, three speed tests against six other tests (negative signs), where the emphasis is clearly on “concentration” and “power” rather than on speed.

This is a case of bipolarity in the domain of intelligence. Knowledge of psychological context allows negative loadings to be interpreted as due to functional opposition. A heightened ability and inclination of a participant for speed should have a positive impact on speed tests. In tests requiring concentration and deeper problem solving, tendencies and abilities to the advantage of speed will probably have some negative effect. A corresponding counter effect is possible regarding concentration skills and aptitude. These tendencies may be counter-productive for tests requiring skills and aptitude for speed. This differential speed effect seems to have given rise to the bipolarity of F_2 .

⁵⁹ Bipolarity of factorial constructs has often vanished when treated with SSM: By SSM rotation, responses to emotion items with negative valence turn out to be statistically independent from responses to emotion items with positive valence. Against all common sense, SSM psychologists believe that positive and negative emotions are functionally unrelated (Diener & Emmons, 1983, Watson & Clark, 1988, late correction after model comparisons by Crawford & Henry, 2004). Optimism and pessimism appear similarly independent in factorial SSM results. According to common experience, these two attitudes are functionally bipolar (opposite) expectations towards the future (Marshall et al., 1992). The polarity in the gender typology is also lost by SSM methods. Due to factorial orthogonality, generated by SSM, the traits femininity and masculinity are deemed unrelated (Bem, 1981).

The bipolarity of bipolar *Varimin* F_3 needs another comment. It is noteworthy that the bipolarity of F_3 is not limited to verbal vs. visual performance, both belonging to the “power” category. The contrast here is found as verbal vs. nonverbal performance. It is possible – but mere speculation so far – that an inclination and ability preferring verbal performance might be somewhat disadvantageous (inhibitive) for nonverbal intellectual performance or, vice versa, that nonverbal inclinations and abilities are somewhat disadvantageous for verbal performance. This, however, cannot safely be concluded⁶⁰.

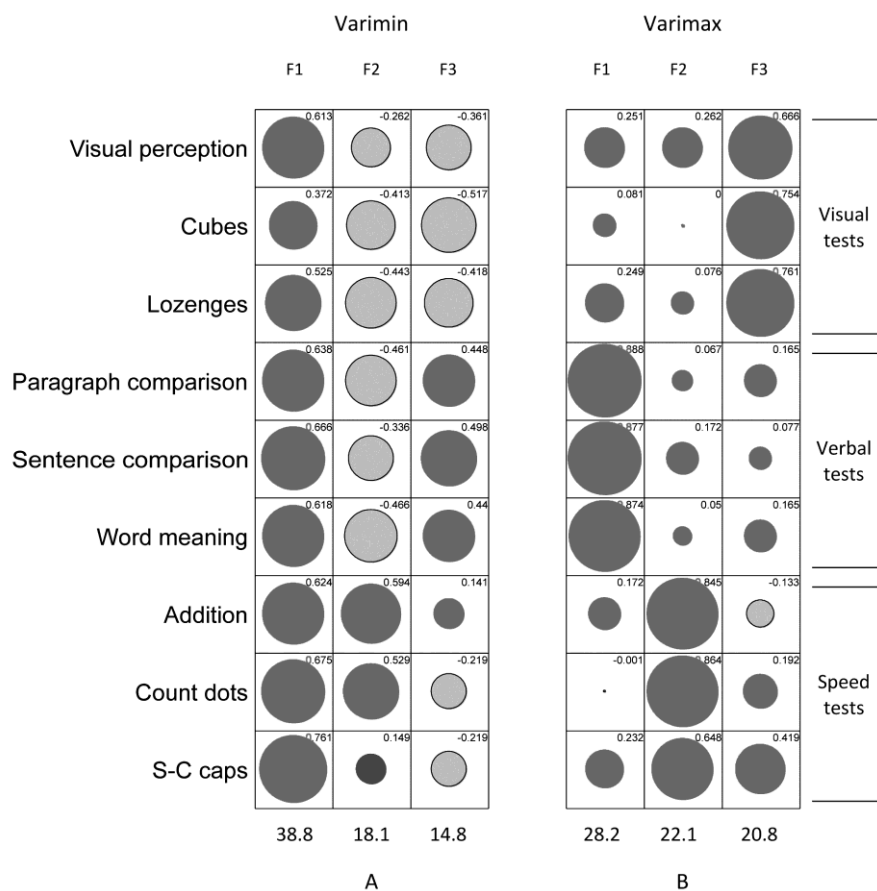


Figure 2.11: *Varimin* (A) and *Varimax* solution (B) of intelligence tests (data: Holzinger and Swineford)

⁶⁰ These tests should also be run with two additional instructions, with a speed instruction to encourage the participants to solve the tasks fast and with a power instruction to emphasise quiet concentration and deliberation. Depending on the test, a differential increase or decrease in performance will probably result.

The Varimax solution (Figure 2.11B) causes g to vanish, and differentiates the three types of tests. In this way, no antagonism is implied between these tests and their factorial conditions. Jöreskog and Sörbom's LISREL solution draw the same wrong conclusion as is drawn from the Varimax solution of 2.11B, however by exerting considerably more statistical effort. The authors thus verified the SSM model and assumed three latent sources of variance that they deem independent. Jöreskog and Sörbom did not find the more straightforward and theoretically more plausible solution, something that *Varimin* has revealed without mathematical extra investment.

Bipolar solutions do not always have variables with positive or negative loadings, sometimes they may have near-zero loadings. This may actually be informative as shown by the following example.

Test run 8: Psychophysiological activity indicators

(Data: Köhler & Troester, 1991).

Köhler and Troester (1991) tried to validate palmar sweat (PSI, palmar sweat index) as an indicator of psycho-physiological activation. They tested 50 subjects for three states of rest and for one state of strain (Strain: Participants were to successively subtract 7 from the number 2007). On these four test occasions the authors collected 16 psychophysiological values per person: PSI values taken at the middle and at the index fingers (PSI-M, PSI-F), spontaneous fluctuations of sweat production (SF), skin conductance level (SCL), and heart rate (HR). These five values were taken 16 times per person and they were correlated intra-individually. The correlations were averaged across all 50 participants.



Figure 2.12: *Varimin* (A) and *Varimax* solution (B) of human physical measures (data: Köhler and Troester, 1991).

Figure 2.12 shows the results of a factorial standard analysis of the intercorrelations. Because of their loading pattern and the order of eigenvalues 3.50, 0.76, 0.33, 0.20 ... two factors were deemed interpretable. The initial solution is not shown, as it is very similar to the *Varimin*-rotated solution.

Results: *Varimin* F₁ is a general factor showing that individual differences of the five measurements apparently have the same indicator value, i.e. activation. *Varimin* F₂ represents additional variance from skin conductance and heart rate. The sweat variables seem to have nothing to do with this source of variance. This is indicated by their near-zero loadings. The F₂ loadings with opposite signs in SCL and in HR seem to indicate that these variables have an antagonistic relationship: There are participants whose skin conductance reacts more strongly to activation than their heart rate. In others, the heart rate reacts more strongly than skin conductance⁶¹.

⁶¹ Troester, the first author, briefly expressed approval of my *Varimin* interpretation of his results in an email correspondence.

Although both variables generally indicate the same function of activation in F_1 , a small residue of variance remains. This can be explained by slight preferences for either SCL or HR. One might speak here of a “forked effect”: The greater an effect takes a direction X, the smaller it is for direction Y, and vice versa. Sweat production is not influenced by these coordinated preferential effects.

Varimax solution presents two insoluble puzzles. The activation effect is split into two independent components – puzzle number one. Physiologically, two independent sources of activation are hardly imaginable. The second conundrum has SRL belonging to one of the activation branches and HR to the other. Thus Varimax succeeds in fouling up the sources of variance beyond recognition.

The above examples show the following: FAs of variables allowing for hypothetical complexities of their sources of variance engender results that can be more easily interpreted than results that have been forced into the straitjacket of SS. The question of latent conditions will be further dealt with in detail below.

Question V **Can CSM-orientated factor analysis** **capture method-dependent variance sources?**

In 1959, Campbell and Fiske introduced a new methodological technique, multi-trait-multimethod analysis (MTMM). The authors tackled a previously largely neglected phenomenon. Personality researchers are not only confronted with a multitude of latent traits, they also have to expect variance whenever they use different testing methods to assess people’s behaviour. Soon, other determinants of variance such as by changing samples of informants, were included (self-assessment vs. external assessment). Situational factors, possibly influencing the covariances of measures (test rerun effects etc.), were included. In subsequent decades numerous MTMM data sets were analysed. As multiple sources of variance were successfully revealed by using this “complexity friendly” MTMM procedure, it seems reasonable to expect corresponding outputs from CSM factor analyses of MTMM data.

Test run 9: Knowledge test with varying test methods (Data: Campbell & Fiske, 1959).

Campbell and Fiske (1959) developed a somewhat subjective procedure for revealing from MTMM data methodical sources of variance. Effectively, this amounted to systematic inspections of tables of intercorrelations. The authors selected correlation data from published papers to demonstrate how their procedure worked. For test run 9, one of Campbell and Fiske’s correlational data sets is taken, the data had been used before by Cronbach and Vernon for other purposes – without indicating who had collected the data. Apparently, students as participants had

been submitted to some physics knowledge test that included subject matters of mechanics and electricity, the tasks were presented either by verbal questions or with the aid of visual displays.

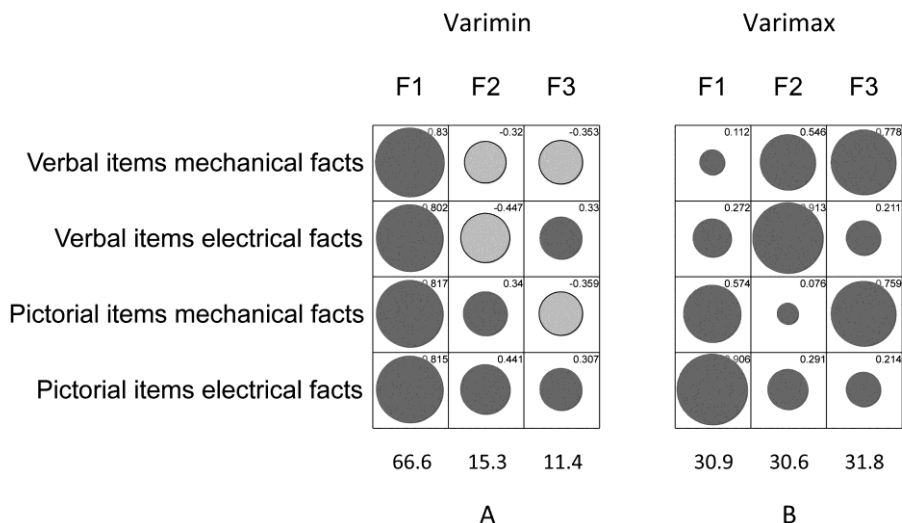


Figure 2.13: *Varimin* (A) and *Varimax* (B) solution of knowledge tests (source of data: Campbell & Fiske, 1959).

Figure 2.13 shows the results from applying *Varimin* and *Varimax* analyses. Considering the order of eigenvalues (2.67, 0.62, 0.45, and 0.27), traditional interpretation rules would call for interpreting F₁, at most F₂ in addition. By inspection of these correlations, without FA, Campbell and Fiske identified effects of two sources of variance⁶².

Results: *Varimin* F₁ represents the general factor associated with a high proportion of communality. Intelligence differences amongst the participants and differences in diligence for physics may have contributed to this, as did, albeit to a far lesser extent, F₂ as a method factor (speech vs. image presentation) and F₃ as the factor for subject matter (electricity vs. mechanics).

Apparently, some participants found it easier to deal with verbal questions while others preferred to solve problems with the aid of pictures (F₂). Some seem

⁶² The unreliability of the criteria of eigenvalue ≥ 1 or the bend of eigenvalues in a Scree Plot, has been noted before (Fabrigar, 1999, p. 287). The “parallel test” procedure also suggested by Fabrigar cannot be conducted with correlation data only. My own experience with re-analysing factorial data shows that the percentages of the summed communalities held by initial factors can be helpful. A factor less than 10% of summed h^2 is usually not interpretable. In Campbell and Fiske’s data, the percentages of communality of the four initial factors amount to 66.7, 15.4, 11.3, and 6.7. If a minimum of 10% communality is set as a criterion, F₂ and F₃ still have to be taken into consideration.

to have had more interest and training for mechanics, others for electricity (F_3). Again, bipolarity turns out to be an indicator for competing conditions, not only for the presence or absence of one single condition. Opposing effects in verbal vs. visual test material are lost in a Varimax solution. The factorial content for electricity is not represented in a Varimax solution either. Only mechanics is deemed to have been considered by Varimax (F_3).

The last test run tries to establish whether *Varimin* is suited to expose a source of variance that emerges when samples of informants are swapped.

Test run 10: Self-assessment and external assessment of children

(Data: Matson & Nieminen, 1987)

In a questionnaire survey on behaviour disorders, depression, and fear in children, Matson and Nieminen (1987) used six scales with items to be rated by children. They were shown to the children themselves and to their teachers. Figure 2.14 gives the result of their ratings. The eigenvalues are 3.98, 2.14, 1.41, 0.99, 0.86 ...; the percentages of explained variance in the initial factors are 33.1, 17.8, 11.8, 8.2 ..., meaning that three factors are probably substantial.

The result of the *Varimin* solution is given in Figure 2.14A: The communality for the source of variance in children and teachers is evident in the general factor F_1 . A tentative interpretation suggests that the children's actual dysfunctional symptoms and their inter-individual variance may be expressed by the children's and teachers' judgments. Caution is required, however, as in most questionnaire surveys; because acquiescence may give rise to or boost a first factor. The argument that in this survey the children's and teachers' judgment may be different is not valid because a disposition for acquiescence (tendency to say yes) in questionnaires may be equally distributed among teachers and children. Additionally, F_1 might have had influence by *social desirability* (SD) in children *and* in teachers equally. The SD component is not identifiable, since all scales include negatively rated experiences and behaviours.

The bipolar factor F_2 seems to reflect differences due to variance of informants. As such, this is not in doubt, but its origin is not clear. The children might tend to judge themselves as less maladjusted than their teachers did, or the other way around, or they may be more or less prone to acquiescence or SD than their teachers. These influences may add up in F_1 , much in the same way partial influences do.

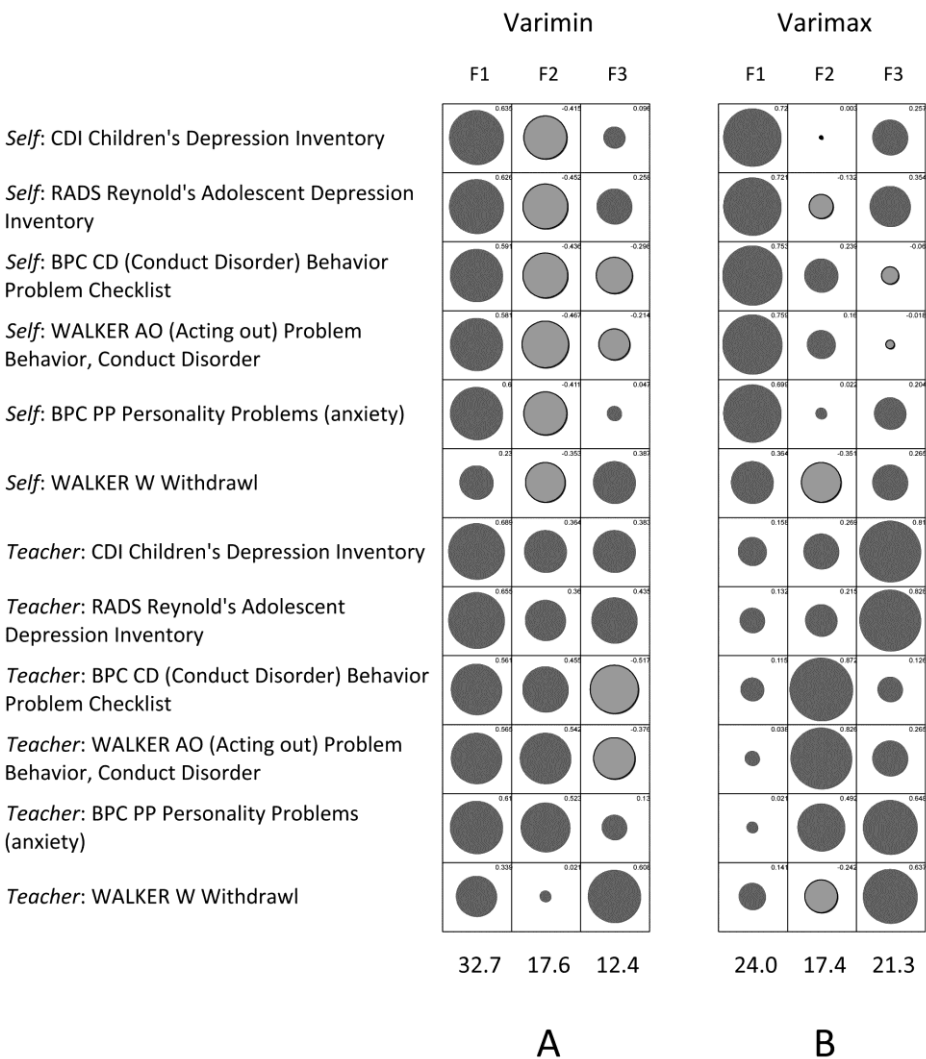


Figure 2.14: Varimin (A) and Varimax (B) solution of dysfunctional statements (source of data: Matson & Nieminen).

The bipolar factor F₃ apparently differentiates dysfunctional symptoms by opposite signs, both for children and teachers alike. On one side (by variables with positive signs), disorders with disinhibiting effects (acting out, conduct disorder) are clustered, and on the other side (by positive signs) disorders with inhibitive effects (withdrawal, depression, anxiety). Here too, an antagonistic relationship seems to be present. Dysfunctional dispositions may come into effect either with or without inhibition. Disinhibition can lead to an unrestrained “acting out” (e.g.,

aggressions). Psychoanalytically speaking, the manifestation of fear occurs under dissipating conditions.

The Varimax solution (Figure 2.14B) brings forth three clusters of scales for pupils and teachers. One cluster combines the six scales assessed by pupils, the second cluster combines the three scales for disinhibition assessed by teachers, and the third combines the two scales for inhibition, also assessed by the teachers.

However, the Varimax result is marred by a number of runaway values and multi-factorial loadings. Conceptually, these clusters are useless. A re-analysis of Matson and Nieminen's data again makes it clear that the term "latent" (invisible, hidden) is not always an appropriate term for factors that are causal for manifest variables. Their possible influence must be revealed by an analysis designed to indicate the existence or non-existence of sources of variance. Whether they are known or unknown, manifest or latent, does not play any role.

In the present case, *Varimin* F₂ (informants) is explained by transparent conditions. Interpretation of F₃ (dysfunctional inhibition vs. dysfunctional disinhibition) had to refer to some less transparent psychological abstractions.

Discussion of chapter 2

Five methodological questions raised by *Varimin*, were addressed. Ten empirical tests of its usage were conducted to help answer them.

Should the tests have covered more psychological data?

In my view, the efficiency of new research procedures should initially be put to the test with data promising positive results, provided the procedure works. If reasonably expected results are not achieved, it stands to reason that the procedure is faulty. This testing strategy, I believe, is too rarely used by psychologists. Psychological data, above all descriptive verbal material from this domain, is not suitable for testing FA methods. Even though language provides an abundance of words and expressions denoting psychological phenomena, the semantics of mental experience and human behaviour is much less transparent than the semantics of, say, kinship terms.

Obviously, FA, once it has passed its methodological checks, should be employed for solving problems in our discipline. Based on available methodological research, a significant conclusion may be predicted even now: FA with *Varimin*-rotation of manifest verbal variables referring to psychological content is prone to reveal essential components. However, these are more difficult to identify and to

separate from each other than features we find in the perceivable and tangible world we live in⁶³.

Thus for interpreting *Varimin* factors of a psychological domain, factor designations based on manifest verbal variables (e.g., questionnaire items, lexical units) cannot suffice, while users of SSM procedures took this for granted. Terms of everyday language, e.g., trait concepts like *conscientious*, *agreeable*, *open*, *sociable*, etc., are not appropriate because it is not these terms in the first place, that we are interested in, but their sources of variance. An optimal approach should resemble what we did for analysing components of the kinship domain, where the descent to a “latent” level was required. Manifest terms like *mother*, *sister*, *aunt* etc. had to be ignored as “merely manifest”, because their *distinctive features* were sought, not these terms themselves nor their Varimax clusters⁶⁴.

The empirical examples of this chapter using relatively transparent domains, proved that *Varimin* factor solutions are indeed interpretable and, what is more, could be interpreted more satisfactorily than Varimax factor solutions. Minimal pair comparisons, permissible for complex *Varimin* interpretations, reveal differences between paired variables of only one feature. This is easier and more reliable than looking for similarities in unanalysed feature aggregates that Varimax usually clusters.

To be fair, three limitations have to be mentioned.

Firstly, using minimal pair comparisons requires suitable pairs of variables in the data set. Reliable interpretations of factors may presuppose the presence of multiple minimal pairs displaying not only opposite loadings of focal factors but also near-identities of non-focal factors. This requirement cannot always be considered beforehand when variables are selected; so some factors may at times not be reliably interpretable because minimal pairs in the sample of variables are missing. Additional pairs may be found only in extended samples of data.

The second limitation is due to the fact that the sources of variance found with *Varimin* are always more abstract than those factors obtained from SS rotations. This may be made clear with the kinship terms: The feature *generation*, revealed by *Varimin*, is more abstract than *the youngest and the oldest in the same lineage*, which is required, as shown, in a Varimax analysis which had to factorially combine *grandparents* and *grandchildren*. It will be more difficult, in CS analyses of data from the psychological domain, to identify abstract features than, in SS analyses,

⁶³ “Characterizing concepts in terms of features works very well for certain types of words ... [those] that are very well structured by physical, social, or biological dimensions, which then serve as the basis of semantic features. Other types of words ... which do not come from well-structured domains, are more difficult to characterize as a set of features. ... No methodology directly reveals the meaning components of a word. Despite the difficulty in confirming what is or is not a feature of a word, it is generally accepted that ... words are comprehended by accessing various features or meaning components.” (Just & Carpenter, 1987, p. 63 f).

⁶⁴ This will be expounded and exemplified in chapter 5.

to make sense with more tangible feature compounds. By the same token it is harder to conceptualise *furniture* than *wardrobe*, *armchair* or *cupboard*.

The third limitation is due to the fact that minimal pairs can hardly be formed for general factors. General-factor loadings of variables of almost any domain are generally consistently positive and show little variance. On the other hand, a general factor normally represents, to a considerable extent, the dominant source of variance in which researchers are usually most interested.

It needs to be considered, in addition, that general factors may represent more than one source of variance. A general factor of intelligence tests may be based on intelligence *plus ambition*. In questionnaires, a general factor may represent trait dispositions as indicated by the items used *plus an inclination toward acquiescence*. Information about relative proportions of such additional sources of variance of factors cannot be retrieved from the data itself. *Varimin* will display the complexity of latent conditions only to the extent that sources of variance affect the variance of manifest variables with *different* effect contributions.

A number of examples in this chapter have shown that *Varimin* rotations facilitate interpretations of initial complex structures provided they contain more than two substantial factors. While conducting numerous factorial re-analyses, I did not come across a single case where the result of a *Varimin* transformation was more difficult to interpret than the result of an initial solution⁶⁵.

For certain data types, the interpretability of SS results left much to be desired. In practice, SS-oriented factor analyses are not generally used for such data. Alternative multidimensional procedures, such as MDS (or NMDS), MTMM, or Circumplex are preferred. The significant advantage of CS modelling is that it may be used to analyse a large variety of data sets, while SS modelling procedures exclude applications of that sort.

FA with SS rotation may remain useful, however, for particular purposes, not for discovering “dimensions” in domains being examined, as was believed thus far. It may be useful for clustering domain variables if this is what is desired. However, in that case a discovery of cluster-producing features will not occur. Varimax and especially oblique rotation procedures would prove useful only for discovering covarying variables as clusters without regarding the sources of their clustering.

Varimax may also be useful if variables have to be analysed that are known to be based on only one *single* source of variance. That would be the case if underlying features revealed by *Varimin* were used as manifest variables, e.g., as items in questionnaires, and if a correlation matrix of these variables is then factor analysed.

Varimin analyses may run into complications if it is not sufficiently known right from the start whether the selected variables are based, truly without excep-

⁶⁵ The *Varimin* program was tested on more than 500 data sets.

tion, on a larger number of variance sources. So far I have assumed that variables are generally multifactorially determined. This does not exclude exceptions, and it is unclear how many exceptions are tolerable in particular cases and how they can be recognised as exceptions.

This uncertainty as to how to treat *Varimin* rotated factors was discussed earlier when the question was raised: Should *negative* factor loadings be interpreted, and if so, how? The analysis itself does not answer the question whether there are functionally antagonistic determinants present in bipolar structures tagged with plus or minus signs, or whether the signs merely indicate the presence or absence of some condition. These questions can only be answered with contextual knowledge of the selected domain and by considering relevant assumptions.

As was indicated previously, *Varimin* research requires samples of variables to be representative of the analysed domain. From the start, the variables should be suspected to reveal multiple latent conditions. Variables providing only little information about sources of variance should not be included. For example, if age-dependent variables were introduced into a personality questionnaire (“I am becoming more forgetful”, “my health is deteriorating”), they would possibly produce an additional factor (“signs of aging”). With positive loadings, they would contribute to such variance – if present –, but they would also yield negative loadings for variables unrelated to age. With meaningless negative loadings, communality would be “wasted” by such variables⁶⁶.

Re-analyses with *Varimin* rotations do not always yield satisfactory results. Unsatisfactory results are mostly attributable to an inadequate sampling of variables. Applying measures of sampling adequacy (Kaiser, 1970) may help to determine routinely whether data sets are adequate for exploratory factor analyses.

The agenda also requires us to attend to the question how the problem of invariance can be solved. Do *Varimin* results deliver stable solutions or do they considerably change if additions to or deletions from a sample of variables are made, or when samples of participants are changed? An unpublished study with intelligence data showed significantly more stability for *Varimin* than for Varimax results when artificial test variables were added to a previously analysed sample of naturalistic variables⁶⁷.

The method of optimal factor extraction is another issue still to be looked at. In the present project so far, principal component analysis (PCA) was used exclusively, as has become standard in FA research. It is possible that initial results might be slightly distorted by entering – as PCA requires – the number 1 into the

⁶⁶ Negative factor loadings of a source of variance, which, for most variables, denotes “effect not present in this variable” are incorporated into the variable concerned, thus adding to the communality of such factors. The communality in such cases would not be informative.

⁶⁷ A lack of invariance in SS solutions was already criticised by Butler (1969): “*the simple structure concept does not solve one of the most crucial and fundamental problems of factor analysis, the problem of the likelihood of factorial invariance*” (p.13). “*Normal varimax factors cannot be regarded as factorially invariant ...*” (p. 24).

diagonal fields of correlation matrices, especially if intercorrelations are predominantly low. Methodological research should also test and compare the modelling quality of the principal-axes and the maximum-likelihood method. Further improvements might ensue if other extraction methods were to yield better models⁶⁸.

The present study limited itself to the development of *Varimin* as a CSM procedure for exploratory factorial research. The method stood numerous tests with varying sets of data, only some were related to psychological issues. A more systematic treatment of psychological issues is a requisite. Two *Varimin* studies on intelligence will be reported in chapter 4 and one on personality in chapter 5.

Lastly, there may be objections to the general strategy of this approach that uses issues of formal data processing (in this case, factor rotation) to reach conclusions by considering much non-formal knowledge. Are non-quantitative arguments admissible for evaluating mathematical-statistical tools?

Such an objection has been discussed in chapter 1. SS and CS principles are formal principles. Yet these were introduced with non-formal motives because the value of formal operations (rotations) was made dependent on results suggesting non-formal interpretations⁶⁹. Such interpretations require psychological understanding, i.e., knowledge exceeding formalistic techniques. This is the reason why the present study up to now has preferred examples where conclusions required predominantly non-formal knowledge. By entertaining the non-formalised notion of “understanding”, Thurstonean formalists would move closer to researchers, who feel committed to “understanding” without much restraint. Where the positions held by model formalists overlap with those of less narrowly-focused researchers, the latter should be allowed to demand a say. They should be allowed to point out the mistakes of formalists, provided they understand their language. They should be allowed to test their mistakes and to rethink them if they cannot be explained away. If their skills permit it, they should also be allowed to help correct their errors.

⁶⁸ This can hardly be expected in the light of the comparisons of extraction methods available so far: *“The major conclusion of this article is that there is little basis to prefer component analysis or factor analysis. For practical purposes the choice of method is not a decision that will greatly affect empirical results or substantive conclusions.”* (Velicer & Jackson, 1990, p. 19).

⁶⁹ *“... consistent psychological meaning is by far the most important criterion for the success of a factor analysis that is designed to illuminate psychological phenomena. The simple structure criterion was designed only as a means to that end.”* (Guilford & Hoepfner, 1969, p. 6). *“The arguments in favour of rotation are not mathematical; and in each research the investigator has to decide its merits on non-mathematical grounds”* (Burroughs & Miller, 1961, p. 35).

Chapter 3

Decathlon data under analysis

Introduction

The historical beginnings of FA raise a perplexing question: Why was the new method predominantly applied to analyse psychological variables such as mental abilities, attitudes and personality traits? Such variables make use of words and sentences as this is required for depicting psychological functioning. But verbal communication about mental experience and performance tends to be imprecise, semantically fuzzy and often ambiguous. If data sets with unambiguous variables had been submitted to FA at the outset, for instance variables of physical performance, a consensus about the efficacy of FA might have been attained more readily.

Such deliberations inspired me to subject sports results to FA. For decathlon sports, athletes deliver performances of ten events. Fortunately, valuable data sets of Olympic decathlon performances are available, in print and downloadable via internet. I do not understand why FA results of these data have almost never been published. I suspect the reason for this omission is that the results of factor analysing decathlon data by procedures committed to SSM are rarely interpretable, if at all. This is a challenge for CSM. An analysis of decathlon data (10 events) with CS rotation of its factors (by *Varimin*) should generate an easily comprehensible factor structure provided the *Varimin* procedure is valid. An interpretation of

decathlon factors (related to physical conditions) should be much easier than an interpretation of intelligence factors (related to mental processes). Researchers in sports psychology have generally complained about inefficiencies of FA in their field (Bös, 1987, Teipel, 1988, p. 341 et sqq., Büsch et al. 2001)⁷⁰.

Description of data

Table 3.01a: Intercorrelations of decathlon performance (sources: Linden (1977) and Kunz (1980)).

	1	2	3	4	5	6	7	8	9	10
	100 m	Long jump	Shot put	High jump	400 m	110 m hurdles	Discus throw	Javelin throw	Pole vault	1500 m
1	-	.59	.35	.34	.63	.40	.28	.20	.11	-.07
2	.50	-	.42	.51	.49	.52	.31	.36	.21	.09
3	.15	.38	-	.38	.19	.36	.73	.24	.44	-.08
4	.17	.20	.11	-	.29	.46	.27	.39	.17	.18
5	.68	.33	-.01	.27	-	.34	.17	.23	.13	.39
6	.55	.50	.28	.36	.55	-	.32	.33	.18	.00
7	.15	.39	.54	.16	.11	.31	-	.24	.34	-.02
8	.16	.37	.16	.28	.39	.35	.28	-	.24	.17
9	-.19	.23	.27	.01	-.13	.02	.33	.06	-	.00
10	.25	.19	-.13	.10	.62	.34	.07	.33	.19	-

Notes:

Upper triangle matrix: Source Linden (1977). Performance at Olympics 1948-1976

Lower triangle matrix: Source Kunz (1980). Performance of Swiss top athletes.

The signs of correlations between race disciplines (higher achievement is obtained with less time measures) and all other disciplines (higher achievement is obtained with more length or height) have been reversed. Original correlations were negative.

⁷⁰ Discussing the *factorial validity* of motor activity tests in sports, presented in a textbook by Bös (1987), he states (p. 141): “A detailed analysis of the quoted [factor analytical] findings shows an ‘alarming non-commitment’ (Orlik, 1967, p. 87) of the factor analytical dimension analyses.” The results often caused “contradictions”; they did not permit “definitive interpretations ... the validity of the statements was over-rated”. “Factor analysis (was) not practical.” There were “doubts about the viability of factor analytical findings ... Factor structures could only rarely be replicated.” (p. 461).

Table 3.01b: Intercorrelations of decathlon performance (sources: website and Zarnowski).

	1	2	3	4	5	6	7	8	9	10
	100 m	Long jump	Shot put	High jump	400 m	110 m hurdles	Discus throw	Javelin throw	Pole vault	1500 m
1	-	.58	.31	.15	.64	.52	.25	.14	.12	.11
2	.66	-	.38	.40	.39	.55	.34	.49	.18	.15
3	.51	.56	-	.20	.18	.46	.79	.49	.60	-.03
4	.45	.64	.50	-	.17	.34	.16	.18	.26	.09
5	.66	.60	.39	.54	-	.43	.15	.24	.11	.52
6	.62	.67	.55	.63	.58	-	.48	.43	.13	.07
7	.43	.48	.80	.44	.37	.49	-	.42	.47	-.05
8	.48	.61	.56	.70	.59	.63	.52	-	.31	.15
9	.34	.46	.59	.40	.44	.41	.51	.52	-	-.04
10	.09	.21	.09	.35	.53	.21	.12	.37	.24	-

Notes:
Upper triangle matrix: Source website: Austrian top athletes (N = 2,400).
Lower triangle matrix: Source Zarnowski (Olympics 1989); Performance at Olympics 1948-1988.
Signs of correlations were partially reversed, see note in Table 3.01a

Linden (1977), who has so far conducted the only FA of decathlon variables (as described by Basilevsky (1994)), obtained four factors. This author transformed them using Varimax and interpreted them as follows: F₁: *short-distance run*, F₂: *explosive arm power*, F₃: *running endurance*, F₄: *explosive leg power*. Linden’s Varimax solution yields no *g* factor that would indicate general athletic competence. The specificity of his labels suggests that he did not reach the level of components characterising the events. Why should arm power and leg power, for example, manifest themselves separately and only “explosively”? Why should the ability to run fast become evident only in short-distance races? Linden does not mention pole vault in his categories; apparently it does not fit into his factorial system. I used three further decathlon data sets to back up the results: An intercorrelation matrix by Kunz (1980), original data on individual athletes by Zarnowski (1989), and original data from an internet source (2004). Tables 3.01a and 3.01b show mean intercorrelations of performances in 10 sports disciplines, obtained from these four sources.

1. Linden (1977)⁷¹
Linden’s correlation matrix is based on Olympic decathlon performances of n = 106 athletes at eight Olympic Games (1948–1976). Linden subjected this data to FA. In the present study Linden’s data were used (as reported by Basilevsky (1994)).

⁷¹ Because of their transparency, Linden’s data are occasionally used for exercises in statistics courses. On the internet, they are currently available at <http://math.usask.ca/~miket/f-03.pdf>

2. Kunz (1980)

Kunz's correlation matrix (p. 161) is based on decathlon performances of $n = 27$ Swiss top decathletes (the database contained results of 90 competitions, many athletes participated in more than one event). While Kunz interprets $n(n-1)/2$ intercorrelations by inspection, he does not use FA. In an appendix (p. 212/13), Kunz provides the original individual data.

3. Zarnowski (1998)

This source supplied the raw data of 233 decathletes whose scores were obtained at eleven Olympic Games (1948–1988). The performances of 75 athletes who did not participate in all 10 disciplines were ignored. Linden's data are included in Zarnowski's data.

4. http://www.10k.at/site/zk_oesterreich/_home.html

From this Austrian website, raw data of $N = 108$ Austrian *all-time decathlon* athletes were taken, including $N = 64$ *all-time decathlon juniors* (aged < 20).

These data can be found at werthner.casc.at/bin/results/alltime.php.

Table 3.02: Eigenvalues 1 to 6 for four principal component analyses with data from Linden, Kunz, Zarnowski, and website.

	1	2	3	4	5	6
Linden	3.79	1.52	1.11	0.91	0.72	0.60
Kunz	4.43	1.85	1.18	0.96	0.71	0.59
Zarnowski	5.44	1.24	0.92	0.63	0.50	0.33
Website	3.88	1.76	1.08	0.96	0.80	0.53
Mean congruence	4.39	1.59	1.07	0.87	0.68	0.51

Results: Factor analyses

Table 3.02 shows the six first eigenvalues of factors extracted from the four data sets (by PCA). Applying the Kaiser-Guttman factor-extraction criterion, three factors should be extracted for rotation in the datasets except for Zarnowski's. In this latter case, the criterion suggested just two factors. But Zarnowski's data would, with interpretability as a criterion, also suggest a three-factor solution.

The four initial PCA solutions were first transformed by *Varimin* and then by Varimax for comparison. Firstly, similarities of the four factorial solutions were obtained separately for *Varimin* and Varimax results. Table 3.03 shows congruence coefficients of Tucker-Wrigley and Neuhaus (cf. Harman, 1968, p. 270).

Table 3.03: Congruences of *Varimin* and Varimax-rotated factors F₁, F₂, and F₃, separately for four data sources. Low congruences in columns F₁, F₂, and F₃ are bold.

	F ₁		F ₂		F ₃	
	Varimin	Varimax	Varimin	Varimax	Varimin	Varimax
Linden : Kunz	.98	.93	.94	.66	.94	.79
Linden : Zarnowski	1.00	.99	.97	.97	.98	.97
Linden : Website	.98	.99	.98	.98	.92	.88
Kunz : Zarnowski	.95	.94	.93	.53	.87	.89
Kunz : Website	.98	.90	.90	.63	.79	.82
Zarnowski : Website	.98	.98	.95	.97	.96	.89
Mean congruence	.978	.955	.945	.790	.910	.873

In Varimax solutions, congruences of factors are somewhat smaller for factors (F₁, F₂ and F₃) than for *Varimin* solutions indicating that *Varimin* solutions are more invariant when data sources are changed. (cf. Table 3.03).

As congruences in *Varimin* rotated factors of different data sets are large, it suffices to use the results of only one data set for factor interpretation. Zarnowski’s dataset seems to be the most suitable source as it has the largest N of decathletes from 49 nations. Zarnowski includes data from the less comprehensive Linden source. Zarnowski’s initial data are original performances, 10 per athlete per Olympiad. By way of example, the seven best Olympic performances are shown in Table 3.04.

Table 3.04: Seven top decathlon records of Olympic participants based on official record counting rules (source: Zarnowski).

Year	Venue	Rank	Nation	Name of athlete	100 m	Long jump	Shot put	High jump	400 m	110 m hurdles	Discus throw	Javelin throw	Pole vault	1500 m	Total score
1984	Los Angeles	1	GB	Daley Thompson	10.44	8.01	15.72	2.03	46,97	14,33	46.56	65.24	5.00	4.35,00	8847
1984	Los Angeles	2	D	Jürgen Hingsen	10.91	7.80	15.87	2.12	47,69	14,29	50.82	60.44	4.50	4.22,60	8695
1976	Montreal	1	USA	Bruce Jenner	10.94	7.22	15.35	2.03	47,51	14,84	50.04	68.52	4.80	4.12,61	8634
1988	Seoul	1	D	Christian Schenk	11.25	7.43	15.48	2.27	48,90	15,13	49.28	61.32	4.70	4.28,95	8488
1972	Munich	1	SU	Nikolay Avilov	11.00	7.68	14.36	2.12	48,45	14,31	46.98	61.66	4.55	4.22,82	8466
1976	Montreal	2	D	Guido Kratschmer	10.66	7.39	14.74	2.03	48,19	14,58	45.70	66.32	4.60	4.29,09	8407
1984	Los Angeles	3	D	Siegfried Wentz	10.99	7.11	15.87	2.09	47,78	14,35	46.60	67.68	4.50	4.33,96	8416

Some athletes participated in more than one Olympiad. Table 3.05 shows the *Varimin* factor loadings (data by Zarnowski) including Varimax factor loadings and initial solution.

Table 3.05: *Varimin*, Varimax, and unrotated factor loadings for ten decathlon disciplines.

Discipline	Varimin-solution				Varimax-solution				Initial solution			
	F ₁	F ₂	F ₃	h ²	F ₁	F ₂	F ₃	h ²	F ₁	F ₂	F ₃	h ²
1 100 m	.49	.61	.42	.79	.22	-.03	.86	.79	.73	-.07	.5	.79
2 Long jump	.66	.47	.3	.75	.36	.15	.77	.75	.82	-.01	.27	.75
3 Shot put	.69	.51	-.37	.86	.86	-.01	.35	.86	.77	-.46	-.22	.86
4 High jump	.74	.18	.23	.64	.37	.40	.59	.64	.78	.17	.03	.64
5 400 m	.75	.06	.46	.78	.17	.54	.67	.78	.77	.41	.13	.78
6 110 m hurdles	.65	.47	.30	.75	.36	.15	.76	.75	.81	-.01	.27	.75
7 Discus throw	.66	.43	-.42	.79	.85	.02	.26	.79	.71	-.45	-.30	.79
8 Pole vault	.82	.16	.10	.67	.50	.44	.52	.67	.82	.12	-.11	.67
9 Javelin throw	.73	.11	-.35	.71	.74	.31	.15	.71	.66	-.15	-.45	.71
10 1500 m	.65	-.65	.23	.89	.02	.94	.06	.89	.39	.76	-.40	.89
% Variance	47.2	17.2	11.2	76.0	27.2	16.7	32.0	76.0	54.4	12.4	9.2	76.0

Note:

Results of four race disciplines are printed bold to improve clarity and comparability.

1. Interpreting *Varimin* factors

F₁ is the expected general factor *g* of decathlon events. Individual F₁ factor scores of the 233 athletes correlate highly with their overall decathlon performance (= .95). Individual total scores, as officially calculated according to the rules of the decathlon sports association, are weighted sums of the performances in the 10 track and field events⁷². To avoid redundancy, the official points total, provided by Zarnowski, was not included as a variable in the sample for FA. Based on their high correlation with the total score, F₁ factor loadings are thus to be considered as independent indicators of overall achievement in the 10 decathlon events.

Some events show higher F₁ loadings than others (e.g., pole vault) and, accordingly, somewhat lower F₂ and/or F₃ loadings than events with lower F₁ loadings (e.g., 100 m race). Kunz (1980) already noticed conspicuous correlations between pole vault and various more specialised events, such as race, jumping and throwing. He concludes that pole vault is an “*exceptionally many-sided event*” (Kunz, 1980, p. 166). Many-sidedness implies dependence on multiple athletic abilities whose joint effects may be revealed by *g* factor loadings. Another interpretation suggests that high *g* loadings indicate particular contributions through training and that lesser *g* loadings indicate more contributions through congenital advantage (for example, bones and muscles in the case of the 100 m race). *Varimin* F₂ is a bipolar

⁷² The decathlon scoring system was determined in 1985 by an IAAF committee. No mathematical or statistical explanation of the individual rating was published.

factor, with its highest loading, negative in sign, for the 1500 m race ($r = -.65$). No other event shows a negative F_2 loading. For interpreting *Varimin* factors, minimal pair comparisons are often indispensable (cf. chapter 2). We find a minimal pair with the 1500 m and 100 m races, as shown in Table 3.06.

Table 3.06
A minimal pair of variables.
Large difference of loadings on one factor only

	F ₁	F ₂	F ₃
1500 m race	.65	-.65	.23
100 m race	.49	.61	.42

The loadings of F_1 and F_3 for the two race events are similar, but loadings of F_2 have opposite directions, thus giving rise to an optimal minimal pair. *Endurance* is the descriptive term (or *demand of endurance*) characterising the 1500 m race and *speed* or *demand of speed* as a descriptive term for the 100 m. The 1500 m race requires continuous power expenditure with longer duration while the 100 m race requires an explosive expenditure of power for a short time period. Plausibly, other race events also require additional constrained power, e.g., the 110 m hurdles ($F_2: .47$), while the 400 m race, likewise plausibly, is positioned between the 400 m and the 1500 m race with an F_2 score of .06 on the bipolar F_2 scale.

Distinctions between more extended expenditure of effort vs. concentrated expenditure can also be found among throw events. Shot put requires an “explosion” of strength ($F_2 = .51$), as does the discus throw ($F_2 = .43$), while for the javelin throw energy expenditure (.11) is less tight. Kunz called the javelin throw a “many-sided event”, “probably extremely demanding” (Kunz, 1980, p. 167). The javelin throw differs from typical power disciplines, i.e., shot put and discus, among the throwing events.

The proposed F_2 interpretation (speed) also applies to differences among jumping events. It makes sense to expect concentrated effort expenditure for the long jump ($F_2 = .47$). The high jump ($F_2 = .18$) requires more skilful and coordinated body movements, not merely peaks of energy expenditure. The same holds true for pole vault ($F_2 = .16$). In sum, F_2 seems to indicate the speed demands of energy expenditure⁷³.

⁷³ For power-endurance (F_2) in athletic performances there is an analogous polarity (speed-power) in mental performances, which has been debated and empirically examined (Jensen, 1993, pp. 492-509) in intelligence research.

F_3 is another bipolar factor which, by signs of loadings, distinguishes events requiring predominant energy expenditure either of the upper or lower extremities. Muscle power of the arms is required for throwing events (discus throw, $F_3 = -.42$, javelin throw, $F_3 = -.35$, and shot put, $F_3 = -.37$). Muscle power of the legs is required for the race events of 400 m ($F_3 = .46$) and 100m ($F_3 = .42$) and the 110 m hurdles ($F_3 = .30$). The 1500 m race does not seem to require particular leg power ($F_3 = .23$), endurance of energy expenditure (F_2) seems to be more important.

In sum, an interpretation of *Varimin* factors of decathlon scores is straightforward. Each event is characterised by g , (1) a general disposition, by (2) the predominant source of demanded muscle power (upper vs. lower body parts) and (3) by the temporal pacing of energy expenditure (speedy vs. enduring).

The knowledge of conditions obtained from analysing the correlations of decathlon events appears to be applicable also for research in sports physiology. Final conclusions can be attempted for generalising concomitant effects of training. Does discipline B benefit from training discipline A? At present it suffices to state that an immediately comprehensible result has been achieved, as soon as the way was paved for analysing co-functioning conditions.

For example, the 100 m race is not viewed as dependent on a single condition, as Linden would have it (his one-factor label: *short distance run*). Instead, the manifestations of three factors are incorporated. A superior performance in 100 m races requires a general physical potential (congenital physical condition plus generalised benefits by training). In addition, leg power is required as well as a concerted input of physical energy (power). The other decathlon disciplines can also be characterised by variable, but conjunct contributions of the same factors, whose functioning depends on linked aspects of the same process⁷⁴.

2. Attempt at an interpretation of Varimax factors

The three decathlon factors are unipolar, after *Varimin* rotation they have only positive loadings just like factors obtained from intelligence tests. No general factor shows up. Factor scores of the 233 Olympic athletes correlate with official total scores, i.e. with external criteria for a general factor, as follows: $r = .60$ (for F_1), $r = -.45$ (for F_2) and $r = -.64$ (for F_3). For none of the three factors does a conspicuous correlation between factor scores and Olympic total record appear. This is to be expected since Varimax distributes communality of a general factor, present to a large extent in an initial F_1 factor, among additional factors F_2 , F_3

⁷⁴ It needs to be remembered that when evaluating differences in athletes' achievements, only the performance-related parameters responsible for these differences can be identified. The latent parameters vary considerably between individuals. They remain latent and can only be researched with other methodological procedures.

etc.⁷⁵. The Varimax result is flawed even by its own criteria, because single loadings are found for only three of ten sports variables. According to SS, seven disciplines are thus unwelcome hybrids.

Are Varimax factors useful, though?

Varimax F₁:

F₁ shows highest loadings for shot put (.86), discus (.85), and javelin throw (.74). The F₁ cluster of variables might be termed “throw events”. However, pole vault is not a throw event despite its high F₁ loading (.51). Also, long jump and high jump with considerable F₁ loadings do not have anything in common with throw events.

Varimax F₂:

Varimax F₂, with its highest loading for the 1500 m race (.94), appears to indicate endurance. The 400 m race (F₂ = .54) still fits this interpretation. But how to explain high F₂ loadings for pole vault (.44) and high jump (.40)? One would be hard pressed to claim similarities between the 1500 m race and two high jump disciplines.

Varimax F₃:

It is difficult to make sense out of Varimax F₃. Varimax F₃ shows highest loadings in the 100 m race (.86), in long jump (.77), in 100 m hurdles (.76), in the 400 m race, and in high jump (.59). In F₃, race and jumping are shuffled as if they were related, which can hardly be explained⁷⁶.

Conclusion: Varimax factors are not useful.

⁷⁵ Negative signs present with two loadings need not be taken as unexpected since signs of factor loadings are generally arbitrary. Highmore and Taylor (1954) also criticise the falsifying effect of an SSM rotation in sports test data: “... the ‘basic’ factor, representing general athletic ability (in which we are primarily interested), necessarily disappears, and the group factors [of simple structure rotation] show little relation to the classification indicated by the [initial] bipolar matrix.” (p. 4).

⁷⁶ After completing this manuscript, I discovered a more comprehensive Austrian data source containing 4586 individual scores of 2674 decathletes: www.werther.at/zehnkampf in “Alltime-Liste Wien, 1993-2004”, a non-Olympic selection. The above-reported Olympic results were taken from the source “Wien 1993-2004”: For F₁ (general factor) the congruence between the Olympic and non-Olympic selection was .99, for F₂ (the limb factor) .95 and for F₃ (pacing of energy factor) .94. Such high congruence in factor structures despite widely differing proportions of variance in the factors is interesting. The proportions of variance for the Viennese (non-Olympic) athletes were 63.8%, 11.1%, and 10.1% for F₁ to F₃ respectively (total 85%), and for the Olympic athletes 47.5%, 17.5%, and 11.2% (total 76.2%). The scores of Viennese athletes are certainly larger than those of the Olympic athletes due to different selection yardsticks. This is indicated by a larger *g* factor proportion (F₁) for the Viennese athletes. The sources of variance for F₂ and F₃, however, are not much different in the two samples. This seems to indicate the fact that anatomical and physiological performance conditions represented by F₂ and F₃ are as valid for top athletes as for less successful ones.

3. Attempt at an interpretation of initial factors

It may be suspected that *Varimin* factors do not differ, or hardly differ, from initial factors so that differing initial solutions may compete with *Varimin* solutions. The case at hand contradicts this assumption. Even though there is a very high congruence (Tucker (1951), Wrigley & Neuhaus (1955)) between *Varimin* F_1 and Initial F_1 (cf. Table 3.07b), for *Varimin* F_2 and initial F_2 , the congruence shrinks to .62 and for F_3 to .69. Thus factors F_2 and F_3 for initial and *Varimin* solutions are not congruent with one another.

Regarding an interpretation of initial F_2 , Table 3.04 also shows an extremely high positive loading (.76) in the 1500 m race. Interpreting the score as an indication of stamina seems appropriate here, as shot put and discus throw also have positive loadings (.46 and .45) as they did in the *Varimin* solution, representing the opposite of stamina, i.e., power. But the initial F_2 lacks a high (power) loading for the 100 m race. Because of their sustained or concentrated energy input, the 1500 m and 100 m race should facilitate an ideal minimal pair comparison. But these two running disciplines do not represent an ideal pair in the initial solution. Only after *Varimin* rotation do they become useable. The initial F_2 likewise lacks evidence of concerted race performances in the 110 m hurdle while *Varimin* can extract this from the correlations.

Initial F_3 would not even be remotely interpretable as a factor differentiating between arm power and leg power. The javelin throw ($F_3 = -.45$) and the 1500 m race (-.40) should have lower arm power loading in the initial F_3 than shot put (-.22), as would be expected if F_3 was interpreted to mean polarity of arm power and leg power. By the same token, an interpretation of the initial F_3 cannot (as might have been hoped) be interpreted by adding other traits. The initial factors F_2 and F_3 are thus difficult to interpret and much less plausible than the *Varimin* factors F_2 and F_3 .

Conclusion: Initial factor F_1 is an approximation to *Varimin* F_1 and may thus be considered as usable. Initial factors F_2 and F_3 , however, lack clarity and consistency, compared with *Varimin*-rotated factors. Initial factorial solutions, except for F_1 , need not and should not be considered when final conclusions about FA results of any data category are drawn.

Table 3.07: Congruences among Varimin, Varimax and initial factors from Zarnowski's data set (Olympic performances, N = 233). Note: Congruence values > .70 are bold.

		7a			7b			7c						
		Varimin			Varimin			Varimax						
		F ₁	F ₂	F ₃				F ₁	F ₂	F ₃				
Varimax	F ₁	.86	.68	-.24	Initial	F ₁	.98	.66	.28	Initial	F ₁	.87	.62	.93
	F ₂	.74	-.17	.41		F ₂	.11	-.62	.67		F ₂	-.38	.70	.06
	F ₃	.86	.73	.55		F ₃	-.15	.41	.69		F ₃	-.32	-.34	.37

Table 3.07 shows congruence values of the factorial solutions that elicit the following commentary:

1. Table 3.07a: The first *Varimin* factor F₁ shows relatively high congruences with the three Varimax factors (F₁, F₂, F₃). This is apparently caused by SS transformations distributing the loadings of initial factor F₁ evenly amongst the three Varimax factors. With *Varimin*, initial F₁ is maintained or even optimised (cf. Table 3.07b).
2. Table 3.07a: *Varimin* factor F₂ seems to be related to Varimax F₁ (.68) and to Varimax F₃ (.73), as if underlying parameters, revealed by *Varimin* F₂ in its bipolarity, were distributed amongst Varimax F₁ and F₃.
3. Table 3.07b: Apart from a close resemblance of *Varimin* F₁ and initial F₁, there is no significant configural congruence between the *Varimin* and initial solution. This has already been noted when attempting to interpret initial factors.
4. Table 3.07c: By comparing Varimax and initial factors it is noted that Varimax F₃ is surprisingly close (.93) to initial F₁. It is hard to explain this similarity, but this does not seem worth worrying about.

4. Expert rankings for validating *Varimin* factors

Twelve decathletes were asked via an internet survey to rank the 10 track and field events of their sports. The instruction was:

1st ranking: *Give rank 1 to the event requiring most arm activity and least leg activity. Give rank 10 to the event requiring most leg activity and least arm activity. Ranks 2 to 9 should be distributed according to relative arm-leg activity between the extremes.*

2nd ranking: *Give rank 1 to the event requiring an exertion of bodily energy with highest concentration within a short time period and give rank 10 to the event requiring an exertion of energy spreading over a longer time period (stamina). Ranks 2 to 9 should be distributed between the extremes according to relative contributions of concentration and stamina.*

The aim of this survey was to find out if the experience of decathletes corresponds with our interpretations of *Varimin* factors F_2 and F_3 . In Figure 3.01, the Y-axis shows averages of the first rankings. On the X-axis, the *Varimin* factor loadings are indicated. The correlation is $r = .80$ ($p = .003$).

It stands to reason that in pole vault the athletes probably slightly overrated the use of their arms, and the use of their legs was probably overrated for the 1500 m race. According to Kunz (1980, p. 166), as noted earlier, for pole vault the functions of the entire body are required conjointly (see above, “versatility of pole vault”).

For the 1500 m race, a continued optimal dosage of energy exertion seems to be more important than leg power. This is shown plausibly in Figure 3.02 which depicts the results of the second ranking (F_2). The correlation between rankings and *Varimin* F_2 is lower ($r = .70$, $p = .01$) probably because the athletes overrated the stamina for the 400 m race and underrated the stamina needed for the javelin throw. It is also possible, though less probable, that *Varimin* factor loadings, and not the rankings of athletes, slightly distorted the empirical givens. At this point it suffices to state as a result that the majority of rankings of the athletes confirmed the interpretation of *Varimin* factors.

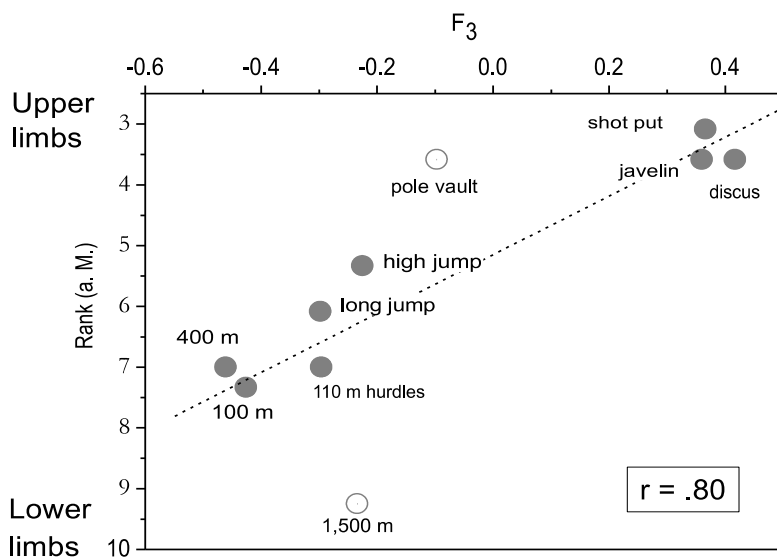


Figure 3.01: Ten decathlon disciplines plotted for *Varimin* factor loadings (X-axis) and mean ranks obtained on a scale (Y-axis) on which decathlon athletes rated relative amounts of arm and leg energy required for good results in these disciplines.

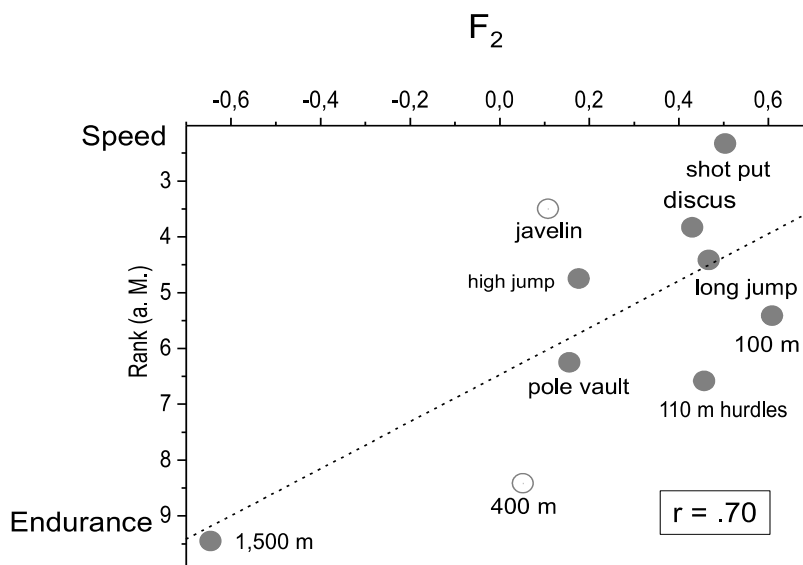


Figure 3.02: Ten decathlon disciplines plotted for Varimin factor loadings (X-axis) and mean ranks obtained on a scale (Y-axis) on which decathlon athletes rated speed of energy expenditure relative to endurance of energy expenditure required to obtain good results in these disciplines.

Discussion of chapter 3

Factorial analyses of athletic performance show that applying the conventional EFA paradigm based on SS structure modelling (SSM) leaves much to be desired. An interpretation of Varimax factors is either very difficult or even impossible. Complex structure modelling (CSM), on the other hand, yields transparent structures that are easily interpretable. In addition, it is easier to interpret *Varimin* factors than initial factors which may also escape comprehension. It has also been noted that a *Varimin* solution is more invariant than a Varimax solution against changes of participating samples. Moreover, an interpretation of *Varimin* factors of decathlon events was supported by self-assessments provided by experienced decathletes.

Varimin results of this study correspond with the results of other sports psychological studies, i.e. with the results of Szopa et al. (1998), who reported, without using FA, multifunctional relationships among various sport activities. The researchers used 42 tests of motor performance; the participants were 143 men and 91 women. The authors summarised their results by distinguishing five main motor abilities; the first three match perfectly with *Varimin* factors. Szopa et al.'s "ability to develop global strength" manifests itself in *Varimin* F_1 , an "ability to develop local strength (of lower or upper extremities)" manifests itself in *Varimin*

F₃, and an “ability of muscular endurance” matches one polar characteristic of the bipolar *Varimin* F₂ factor.

The bipolarity of energy pacing (explosive vs. enduring), as indicated by *Varimin* F₂, is also related to physiological observations: “*The energy at muscular activity can be supplied either a) anaerobically, as it is during short bursts of activity of high intensity with an accumulation of lactic acid as a result, or b) aerobically, as during more prolonged work, when oxygen intake balances the oxygen demand ... In aerobic work, respiration and circulation will play a dominating role ...*” (Åstrand, 1956, p. 307). Both physiological processes (a) and (b) may be used relatively independently despite bi-polar amounts of contributions: “*The development of the anaerobic and aerobic processes are not parallel. It cannot, therefore, be expected that a test procedure where the capacity of the aerobic processes is determined will give accurate information about the capacity of the man for aerobic work*” (Åstrand, 1956, p. 307). According to Milhorn (1982), the basic physiological requirement for physical stamina is cardiovascular fitness.

The demands of physical activity on muscles for different decathlon events are rarely equal. This explains *Varimin* factor F₂, showing that priority is given either to the upper or the lower extremities, depending on the activity.

Experts will hardly be surprised by this result. Athletes competing in various different disciplines “*take up the challenge of competing across the whole range of athletic disciplines ... [but] it is not possible for a driving mechanism to exist which is equally optimal or which can provide an identical maximum neural activation for all disciplines*” (Tidow, 2000, p. 245). Certain “movement affinities” occur, different activities require the use of different muscle sub-systems.

The results of Varimax transformations are not supported by physiological findings. SSM transformations of factors of decathlon variables were disappointing. Manning et al. (1988) gets right to the point in his recap of poor factor analytical results gained in anaerobic tests by conventional methods: “*Results showed no single factor emerged and that unrelated aspects existed among these tests and that they were not measuring similar qualities. It is suggested that anaerobic tests that are used to evaluate anaerobic power be performed as specifically as the skill being tested.*” – in other words: Conventional FA is not suited for analysing athletic activities. It should not be used, the authors conclude. This begs the question: Why should one refrain from FA (with SS-orientated rotation) of human physical performance, but not from FA of intellectual and other mental performance? The answer seems to be that in the realm of ambiguous intellectual and mental processing expectancies are vague, if there are expectancies at all, and hence deviations from expectation cannot be recognised and their underlying causes not identified.

The goal of earlier sports psychologists (Guilford, 1958, Pöhlmann et al., 1979, Bös & Mechling, 1984 etc.) is being approached. Thurstone’s obstructive SS principle should not be used when an investigation of components of physical fitness is required (Ismail et al., 1965). The harsh critique by these authors depicts the present situation of conventional EFA research that Ismail et al. call totally “*hope*

less” (see also Lykken, 1991, Michell, 1997, Koch, 1999, Breiman, 2001, Gigerenzer, 2004, Barrett, 2005). But this does not have to be the case, provided there is a willingness to model complexity ⁷⁷

⁷⁷ An example of serious criticism of the failure of conventional multivariate research is found in Barrett (2005, p. 45): *“The pressure for change is building - and it looks like a paradigm change – for example, not merely a transition from say Classical Test Theory to IRT – but the entire loss ... of psychometric test theory altogether over time.”* Borsboom’s *“attack of the psychometricians”*, to which Barrett refers, is considered an attempt to bridge the “home-grown rift between psychometrics and psychology”. The modelling of psychometrics, which is deemed modelling without substance, has to be replaced by modelling with substance (Borsboom, 2006).

Chapter 4

Intelligence data under analysis

A complex structure analysis of IST data
(Intelligence Structure Test)

Point of departure and objectives

Chapter 4 is devoted to a factor analysis (FA) of intelligence data. Intelligence was the first domain on which the pioneers of FA, Spearman, and his early followers, Cattell, Thurstone, Vernon, and Burt, applied the new method. Textbooks today distinguish between two main intelligence factors that were introduced by Cattell and Horn (1963), namely “fluid” intelligence (G_f) (supposedly based on congenital conditions) and “crystallised” intelligence (G_c) (based on long-term learning).

This distinction was the result of analyses that considered, without further ado, SS structure modelling (SSM) to be right and safe. Today the distinction between (G_f) and (G_c) is regarded worldwide as an established fact. In the present study, 18 samples of participants in the Intelligenz-Struktur-Test (IST), which is very popular in Germany, will be re-analysed in order to find out whether the alleged subdivision of intelligence is replicable if data formerly analysed by Varimax are re-analysed using *Varimin*, the Complex Structure Modelling (CSM) device. It seems likely that CSM will engender different factors.

In the following, intelligence factors are subjected to *Varimin* and, for comparison, to Varimax rotations. The interpretation of intelligence factors rotated to SS and CS, provided in Study I, will be extended by collecting external validation data in Study II.

Study I: *Varimin* analysis of IST factors

Objective

The first question we have to answer is whether *Varimin* rotation yields meaningful IST factors.

Data

In order to increase the stability of results, 18 IST intercorrelation tables, found in psychology journals, were factorised. Factor loadings of eight subtest variables were averaged across these 18 studies. In Table 4.01 the sources are provided, with additional information. Test versions IST-55 and IST-70 were considered, version IST-2000 did not provide sufficient correlation tables.

Expectation

Amthauer et al. (1999) obtained a two-dimensional factorial structure using the IST-2000 test version. Following an SSM-orientated rotation, these authors interpreted their findings as manifestations of fluid and crystallised intelligence. I expect that after *Varimin* rotation IST factors must be interpreted differently and that *Varimin* interpretations will be more satisfactory than Varimax interpretations.

Table 4.01: Sources of intelligence test data. IST-55 was applied in test studies 1–8, IST-70 was applied in test studies 9–18.

#	Year	Authors	Subjects	N	Factors	Additional variables
01	1958	Fischer, H.	Secondary school pupils (14-16 yrs.)	122	3 *	4 Variables of Primary Mental Abilities
02	1964	Höger, D.	Gymnasium pupils (12-19 yrs)	519	6 *	Grades of 11 school subjects
03	1966	Bäumler, G. & Weiß, R.	Pupils	200	6 *	10 Variables: CFT Pauli-Test
04	1967	Haenschke, B. & Mehl, J.	Ordinary pupils (13.2 yrs)	77	6 *	4 Tests productivity Grades for 5 school subjects
05	1967	Haenschke, B. & Mehl, J.	Special pupils (13.0 yrs)	67	6 *	4 Tests productivity Grades for 5 school subjects
06	1969	Seitz, W. & Löser, G.	Gymnasium pupils (17.6 yrs)	124	3 *	4 CFT, grades for 7 school subjects 12 HSPQ (Personality)
07	1970	Amthauer, R.	Sample heterogeneous	799	X	No additional tests used
08	1970	Bäumler, G. & Breitenbach, W.	Psychology students	55	7 *	Pauli-Test 5 motivation variables
09	1970	Sassenscheidt, H. & Buggle, F.	Craftsmen, technicians	132	2	4 PTV-variables Technical ability 16 PF
10	1972	Langner, E. & Olbrich, M.	Pilot applicants	397	3 #	5 Variables EVT (Intelligence test)
11	1972	Langner, E. & Olbrich, M.	High school military applicants	190	3 #	5 Variables EVT (Intelligence test)
12	1972	Langner, E. & Olbrich, M.	University applicants	210	3 #	5 Variables EVT (Intelligence test)
13	1972	Langner, E. & Olbrich, M.	Primary school pupils	236	3 #	5 Variables EVT (Intelligence test)
14	1972	Langner, E. & Olbrich, M.	Conscripts	150	3 #	5 Variables EVT (Intelligence test)
15	1993	Schmidt-Atzert, L. et al.	Hauptschul-pupils (17.5 yrs)	196	2	CFT Dictation. Basic arithmetic
16	1993	Schmidt-Atzert, L. et al.	Realschul-pupils (18.1 yrs)	394	2	CFT. Dictation. Basic arithmetic
17	1993	Schmidt-Atzert, L. et al.	Gymnasium pupils (21 yrs)	397	2	CFT. Dictation. Basic arithmetic
18	1998	Brocke, B. et al.	Aptitude test-subjects (38.8 yrs)	279	2	Dictation. Basic arithmetic. FRT (Intelligence test)

Notes:

* Factor analyses I-S-T conducted with additional variables.

Three factors were considered as appropriate, even though 2-5 factor solutions had been tried.

X No factor analysis conducted.

Table 4.02: Mean factor loadings (averaged across 18 analyses) of two-factor and three-factor solutions after *Varimin* and Varimax rotations of individual initial solutions.

Solution	A. Varimin					B. Varimax				
	Two factor solution		Three factor solution			Two factor solution		Three factor solution		
Interpretation	g	L	g	L		c	f	c	f	A
	F ₁	F ₂	F ₁	F ₂	F ₃	F ₁	F ₂	F ₁	F ₂	F ₃
Subtests										
1 SC	.63	.37	.63	.35	-.12	.70	.24	.72	.16	.24
2 WS	.62	.32	.63	.33	.06	.67	.23	.65	.16	.43
3 AN	.70	.29	.68	.31	-.06	.69	.32	.59	.23	.42
4 CO	.65	.35	.63	.38	.17	.68	.31	.67	.18	.35
5 NU	.71	-.04	.69	-.22	-.19	.53	.53	.25	.45	.64
6 NS	.68	-.21	.66	-.20	-.08	.39	.62	.25	.52	.50
7 FS	.65	-.38	.67	-.35	.11	.22	.72	.12	.72	.32
8 DI	.57	-.51	.59	-.38	.09	.11	.74	.10	.74	.29
Variance %	79	21	79	19	02	54	46	38	34	28

Notes

g = General or basic intelligence L = Learning capital (stock)

f = Fluid c = Crystallized intelligence A = Artefact (uninterpretable factor)

SC Sentence completion, WS word selections, AN analogies,

GE commonalities, NU numeracy, NS number series,

FS figure selection, DI dice

Data analysis

A memory test (ME) was eliminated from the nine subtests of IST, leaving eight for analysis. Reviewers of IST-70 (Schmidt-Atzert & Hommers, 1996, Schmidt-Atzert, 1997, Brocke et al., 1998) likewise did not regard memory as sufficiently represented by ME alone. The use of ME might also interfere with the factorial structure of the whole test. When non-IST variables (for instance, on personality) had been used in addition, the correlations of these variables were also discarded.

The 18 correlation matrices of eight remaining subtest variables were PCA-factorised in each database. For each factor matrix a two-factor and a three-factor solution was rotated, using *Varimin* as well as Varimax. The obtained factor loadings were averaged across the 18 analyses using Fisher's Z transformation. Before determining averages, the sequences of factors for the 18 analyses were synchronised, because identical or almost identical factors did not always share the same position within a sequence of extractions.

The signs of loadings varying between analyses were also coordinated in order to equalise structures and to align them. The loading patterns in the various analyses were adjusted by visual comparison. Ambiguity occurred in just two or three cases for three-factor solutions. Uncertainty appeared occasionally when two fac-

tors with very similar loading profiles had to be coordinated. Unnoticed incorrect decisions at such matchings are minimal and have been neglected.

Results with comments

Table 4.02 depicts aggregates of factor structure of *Varimin* and Varimax rotations.

1. Comparing two-factor and three-factor solutions

To begin with, we shall look at differences between two-factor and three-factor solutions. In 18 original IST analyses, following SSM rotations, the authors often kept a third extracted factor for rotation if the factor seemed to account for non-negligible and interpretable variance. But the *Varimin* aggregates for F_3 explained only 2% variance. Obviously, *Varimin* F_3 is irrelevant and is therefore discarded; only *Varimin* F_1 and F_2 are apparently substantial⁷⁸.

2. Comparing communalities

Compared with the *Varimin*-based factor aggregate, the Varimax aggregate shows entirely different communalities. The two-factor solution has 54% (F_1) vs. 46% (F_2) communality for Varimax, as opposed to 79% (F_1) vs. 21% (F_2) for *Varimin*. The three-factor solution yields loadings of 38%, 34%, and 28% for Varimax factors F_1 , F_2 , and F_3 , respectively, as opposed to 79%, 19%, and 2% in the case of *Varimin* factors.

The contrast regarding communality for F_3 loading percentages is striking: Varimax F_3 explains 28%, *Varimin* F_3 explains 2%. The relatively high percentage for Varimax F_3 would be a notable result if F_3 were interpretable. But it is not, because the variance of F_3 factor loadings across those eight variables is very small. Accordingly, they provide no information about conditions of performance. Varimax F_3 in Table 4.02B is thus an artefact and will be ignored.

⁷⁸ My experience with *Varimin* rotations, still to be specified, holds that to be recognised as making a substantial contribution to the total variance, a factor should explain >10% of variance.

3. Interpreting Varimax F1 and F2 (Table 4.02B)

F₁ and F₂ profiles of earlier IST FAs, based on Varimax, are obviously similar to the Varimax profiles at hand. Amthauer et al. (2001) used Oblimin-rotated factors. Oblimin rotation is another SS procedure, and the loadings of Oblimin are generally similar to Varimax loadings. Varimax F₁ of the present study is conventionally interpreted as fluid intelligence (G_f or g_f)⁷⁹; Varimax F₂ as crystallised intelligence (G_c or g_c). Fluid intelligence is ostensibly based on congenital conditions. Cattell (1963) and Horn (1976) found that what they called “crystallised intelligence” indicated intelligence transformed by educational and cultural learning. This idea resulted from second-order FAs and SSM rotations.

The idea is plausible. Performance in intelligence tests depends, to some extent, on antecedent educational and cultural learning. Even some researchers who did not find crystallised intelligence in their data when they employed conventional methods (Johnson & Bouchard, 2005) assumed that learning will modify intelligence test results (*“learning processes preceding the test could cause a noticeable variance of performance”*) (p. 410). A FA of IST data should therefore disclose such additional sources of variance of intellectual performance.

SSM, however, can hardly fulfil this expectation in the first place. By Varimax rotation the main source of variance, general intelligence, loses its unity. Much of F₁ variance is passed on to F₂ so as to give rise to two allegedly different types of intelligence. An influence of learning on performance that is actually small is thus blown up and falsified by scrambling F₂ with g. General intelligence thus disappears by SS rotation and is recovered in a cumbersome way by second-order FA procedures.

4. Jensen (1998) describes the result of simple structure rotation: *“... So if you ask where g went, the answer is that it has been divided up and lies ‘hidden’ among all of the tests’ smaller loadings on all of the orthogonally rotated factors. Its variance has not disappeared, it has simply been obscured by being dispersed throughout the whole factor matrix.”* (Jensen, 1998, p. 66).⁸⁰ Interpreting Varimin F₁: Basic intelligence (g).

Obviously Varimin F₁ represents g, the base of intellectual performance. The term *basic intelligence* emphasises the importance of Varimin F₁ as a fundamental condition for intellectual performance. This interpretation will be tested in Study II by examining correlations of Varimin F₁ with IST-70 performance on the one hand and with culture-free tests on the other (see below). The abbreviation g will henceforth be used to denote general intelligence factors obtained by applying Varimin in order to facilitate the distinction from general intelligence g obtained through Varimax and second-order procedures.

⁷⁹ Alternative notations have been used for fluid and crystallised intelligence: G_f vs. G_e, g_f vs. g_c, g_f vs. g_c.

⁸⁰ The correlations of Varimax F₁ and F₂ with the IST-70 original totals are .74 and .63, respectively.

5. Interpreting *Varimin* F₂: Learning assets (L)

Varimin F₂ is a bipolar factor. Traditionally, negative intelligence factor loadings were not tolerated, SS rotation largely removed negative signs⁸¹. In chapter 1, detailed reasons were given for interpreting bipolarity straightforwardly. *Varimin* F₂ has significant positive loadings on subtests requiring verbal operations (SC, WS AN, CO). F₂ loadings of figural tasks FS and CT bear negative signs. This seems to indicate that F₂ emphasises verbal operations playing a predominant role in education. Only rarely do schools require students to master pictorial tasks and to perform language-free formal operations. A relative deficit of learning advantage for FS and CT performance might result. Predominant educational practice requires linguistic operations. Number tasks, showing near-zero *Varimin* F₂ loadings, may require less school training, and almost no educational practice seems to be required for pure figural operations.

“Learning assets”, the label for *Varimin* F₂, may be taken as a metaphor. Learning requires storage. For many intelligent operations stored knowledge gives helpful returns, much like interest on capital accumulation⁸². The distinctions made here correspond to Weinert’s “triangle” of *aptitude*, *knowledge* and *learning* (Weinert, 1996). In the first instance, *aptitude* or *talent*, according to Weinert, requires basic intelligence. *Knowledge* may be seen as accumulated learning assets. In our present analysis, *learning* processes are presupposed, not directly investigated. Learning generates an increase of knowledge and abilities during an individual’s lifetime and education (Waldmann et al., 2003). Research clarifying the interpretation of *Varimin* F₂ is taken up in Study II of this chapter.

Interpreting *Varimin* F₂ as learning assets need not be the final word. It might be argued that a bipolar factor F₂ might indicate a preference for one of two polar cognitive styles, either for holistic operations as are required for language-demanding tasks (F₂ positive), or for more analytical and detailed operations required for solving figural tasks (F₂ negative). Numerical tests might require the two operations in balanced proportions. Likewise, educational “learning assets” are possibly more easily acquired by students who prefer holistic cognitive operation. Differences of cognitive style may be based, just like basic intelligence, on genetic endowment. A final decision about this issue is neither possible here nor required.

⁸¹ Berneyer (1957) comments along these lines: “*The different methods of [factor] analysis [of mental aptitudes] yield factors which have negative loadings ... Such factors, Thurstone contends, must be devoid of ‘scientific meaning’ They do not permit us to ‘interpret’ the various tests as functions of the mental aptitudes which those tests elicit.*” (p. 23).

⁸² Cattell (1971) uses a similar metaphor. With his “investment theory” he attempts to connect fluid and crystallised intelligence: Fluid intelligence is the “investment” made by learning throughout a person’s life, says Cattell (Holling et al., 2004, p. 21). But in his SS-based analyses his learning “investments” are confounded with intelligence: the capital and resulting interest are mixed.

Summary of Study I

Mean factor loadings of *Varimin* F_3 for IST subtests in three-factor solutions are small while Varimax F_3 subtests attract significant loadings. However, since F_3 loadings are unstable across 18 studies, F_3 appears to be an artefact and is furthermore ignored.

IST's conventional F_1 and F_2 interpretations of Varimax F_1 as crystallised and Varimax F_2 as fluid intelligence cannot also apply to *Varimin* solutions. *Varimin* F_1 apparently represents basic intelligence, g. *Varimin* F_2 is best understood as an additional source of performance, i.e., as learning assets (L) (deemed independent of congenital basic intelligence). Participants of IST draw more or less advantage from L depending on the amount of learning assets they have acquired. This might hinge on antecedent availability and practice with tasks demanding cognitive processes akin to those that verbal IST subtests require.

Study II: Validations of *Varimin*-rotated IST factors

Objective

In what follows, the validity of *Varimin* factors F_1 and F_2 , which have been interpreted as g (basic intelligence) and L (learning assets), shall be tested. Also, comparisons of correlations by *Varimin* factors g and L with correlations by Varimax factors G_f and G_e will be made. Apart from IST data, the following data will be considered: Firstly, school performances collected by Höger (1964) and Cronmeyer (1983); secondly, spelling and numerical performances plus results from a culture-free intelligence test by Schmidt-Atzert et al. (1995); thirdly, data of two culture-free intelligence tests collected by Brocke et al. (1998).

Validation 1: School performance

Data 1: Höger and Cronmeyer

Höger (1964, p. 435) and Cronmeyer (1983, p. 172) independently collected IST test data and the participants' school grades (cf. Table 4.03). Höger's school grades were obtained from 519 gymnasium school pupils (age 10–13), Cronmeyer's grades were those of 656 high school graduates.

Table 4.03: Correlations between F_2 (*Varimin*) of IST subtests (row 1) and school grades. Sources: Höger (row 2) and Cronmeyer (row 3).

	SC	WS	AN	CO	NU	NS	FS	DI		
1 F_2 <i>Varimin</i>	.37	.32	.29	.35	-.04	-.21	-.38	-.51	total F_2	p
2 r Höger	.09	.12	.25	.17	.12	.09	.08	.06	r Höger :	.62 .05
3 r Cronmeyer	.17	.21	.32	.25	.3	.23	.08	.07	r Cronmeyer :	.63 .05
4 r mean	.14	.17	.29	.22	.22	.17	.08	.07	r mean :	.67 .03

Notes:

SC sentence completions, WS word selections, AN analogies, CO communalities, NU numeracy, NS number series, FS figure selection, DI dice

Expectation 1

If learning assets is the correct interpretation for *Varimin* F_2 , then this should be revealed by differential F_2 loadings of IST subtests and by differential correlations between F_2 loadings and school grades.

Assessment 1 and comments

Varimin factor loadings of the eight IST subtests, obtained from a factor aggregate (study I data), and correlations of the eight subtests with school grades (data provided by Höger and Cronemeyer), need to be focused.

Clarifying Table 4.03 entries. In Höger's study, the correlation of pupils' SC subtest scores (sentence completion) with their school grade average is .09 (line 2). Correlations in line 2 and line 3 and their means in line 4 (lines 2 and 3 correlate with each other with $r = .77$) and were subsequently correlated with *Varimin* F_2 factor loadings of line 1 (cf. Table 4.02). The three correlations in the second to last column of Table 4.03 (.62, .63, .67) are significant. They indicate that school learning is related to F_2 loadings of IST subtests. The interpretation of F_2 in terms of learning assets has found support.

Varimin F_2 (learning assets) is factorially independent of *Varimin* F_1 (basic intelligence). School performance of pupils is expected to be influenced, in the first place, by basic intelligence. But applying *Varimin* F_1 (basic intelligence) in conjunction with F_2 (learning assets) as independent variables for multiple regression, while using school performance as dependent variable, a correlation of $r = .67$ between F_1 and school performance alone is increased, by adding F_2 as predictor, to $r = .80$.

The original student data used by Höger and Cronemeyer are not available, so the correlation of school achievement with F_1 and F_2 *factor scores* cannot be calculated on an individual level. This would have been preferable.

Validation 2: Culture-free IQ Test CFT, spelling and numeracy test

Data 2

A database provided by Lothar Schmidt-Atzert (and partially utilised in Study I (see Table 4.01, nos. 15–17) with anonymised test results of 908 participants included original IST-70 data plus individual IQ values from the language-free and culture-free intelligence test CFT IQ. Results of a spelling test (dictation) and a typical school numeracy test were also available. The sample was made up of 397 *gymnasium* students, 394 junior-high school students, and 196 secondary school students. The sample of pupils has missing data for some variables and thus has a somewhat lower N in the table.

Expectation 2.1 (regarding basic intelligence):

The data provide an opportunity for validating *Varimin* F₁ as a measure of basic intelligence and for validating *Varimin* F₂ as a measure of learning assets.

For each participant, two measures of general intelligence g, i.e. factor scores from IST *Varimin* F₁ and IQs from CFT, are available. Also available, for each participant, is one factorial measure of L (learning assets, factor scores of *Varimin* F₂ from IST) as well as two derived measures of learning assets (L_d and L_p) based on relative performance with IST subtests with positive F₂ loadings and performance with IST subtests with negative F₂ loadings (see note to Table 4.04). It is expected that the two g measures are highly correlated and that g indicators do not correlate appreciably with the two L indicators⁸³. The aim was to find out whether original raw scores of IST would validate the construct L without using factorial F₂ loadings.

Table 4.04: Intercorrelations among indicators of basic intelligence and learning assets.

		Indicators of "Learning assets" (L)			g-Indicators	
		1	2	3	4	5
		% Learning capital	Difference	Varimin-F ₂	Varimin-F ₁	CFT-IQ
		(L _p) *	(L _d) **			
1	% Learning capital (L _p)	1	.94	.93	-.03	.09
2	Difference (L _d)	.94	1	.99	-.03	.09
3	Varimin-F ₂	.93	.99	1	.00	-.10
4	Varimin-F ₁	-.03	-.03	.00	1	.73
5	CFT-IQ	.09	.09	-.10	.73	1

Note:

* L_p = (SE+GE)/ SE+GE+FA+WU)*100

* L_d = (SE+GE)-(FA+WU)

⁸³ Two measures are introduced, L_p and L_d, as estimates of learning assets. IST subtests SC and CO represent the positive pole of *Varimin* F₂, best indicators of learning assets. Subtests FS and CT represent the negative pole, best indicators of basic intelligence (see Table 4.03, first row). All subtest variables had been transformed into standard values (Schmidt-Atzert had done so, getting an average of 100). Next, two indicators for a learning component were set up, L_p and L_d, by relating SC+CO and FS+CT:

$L_p = (SC+CO)/(SE+CO+FS+CT)*100$ (idealised mean = 50)

$L_d = (SC+CO)-(FS+CT)$ (idealised mean = 0)

(L = learning indicator, p = proportion, d = difference)

Assessment 2.1 and comments

It can be seen that

- The three intercorrelations among measures of learning assets are large, as they should be.
- Correlations between measures or estimates of learning assets and measures of general intelligence (*Varimin* F_1 and CFT-IQ) are low, as they should be.
- *Varimin* F_1 (obtained from IST data) proves to be a measure of general intelligence since its correlation with a culture-free intelligence test (CFT-IQ) is considerable ($r = .73$).

Expectation 2.2 (regarding learning assets):

If *Varimin* F_2 obtained from IST data is a valid indicator of learning assets then *Varimin* factor scores F_2 should correlate significantly with typical school tests like spelling and numeracy.

Assessment 2.2 and comments

Figure 4.01 visualises the correlation between learning assets L , obtained from F_2 of IST, and one of the two variables of external validation (spelling performance). Spelling performance should correlate significantly with L because good spelling is possible only after sufficient learning and practice. In Figure 4.01 spelling data of 945 participants in the Schmidt-Atzert sample (Y-scale) are spread across the L_p scale (i.e., the X scale). The correlation between L_p indicating learning assets and spelling requiring much of those assets is not large, but highly significant ($r = .30$). Without considering basic intelligence (indicated by F_1 of IST), not shown in this graph), it may be assumed that learning assets contributes significantly to spelling test results.

Another meaningful result from Figure 4.01 is worth mentioning, but details must be renounced. Basic intelligence of participants with best spelling results (those in the upper horizontal section D_5) is not evenly distributed across sections L_1 to L_5 , instead L_1 participants in D_5 turn out to be more intelligent, L_2 less intelligent and L_5 participants least intelligent (intelligence measured by CFT IQ). Correlations with intelligence vary from D_5 participants down to D_1 participants, but the tendency of change is progressively reverse: Participants whose spelling scores are low and who possess scant learning assets tend to be more intelligent than those whose spelling scores are low and who dispose of more learning assets. In other words, good spelling scores can be obtained by participants with above-average learning assets without extraordinary intelligence, or with above-average intelligence without extraordinary learning assets.

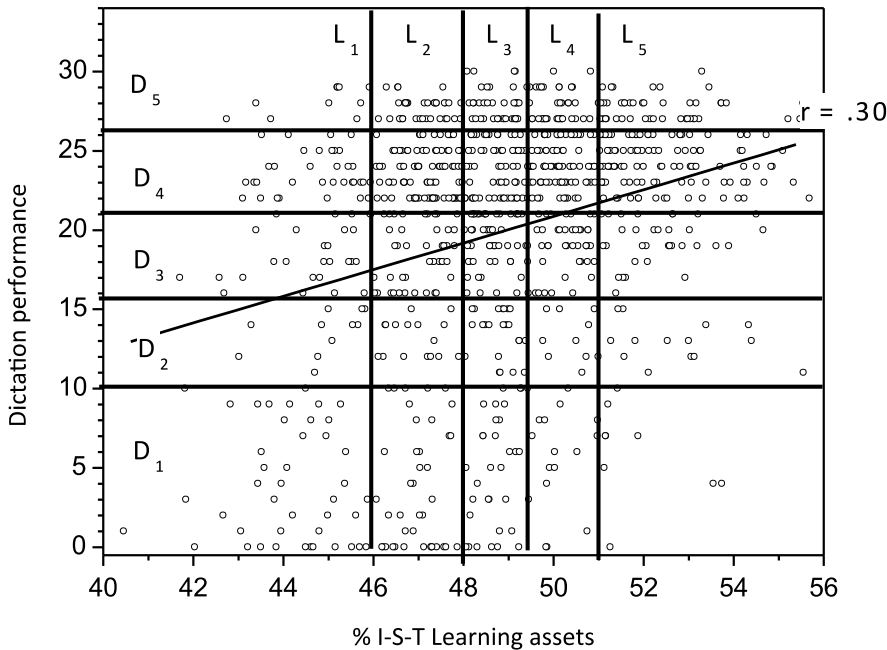


Figure 4.01: Distribution of spelling performance (Y-axis) across learning assets (X-axis) for five subsections for Y (spelling D₁–D₅) and five subsections for X (learning assets L₁–L₅) facilitate the visual impression of the correlation.

More information about the *Varimin* and Varimax results can be gained from Table 4.05.

Varimin results (Table 4.05)

Table 4.05 section A:

CFT IQ values indicating basic intelligence correlate .73 with *Varimin* F₁ (basic intelligence), but only marginally with *Varimin* F₂ (.11). Why is the latter correlation so marginal? Apparently because F₂, based primarily on IST subtests for which learning assets are beneficial, does not contribute to scores of tests like CFT, for which basic intelligence is mainly responsible.

Table 4.05 section B:

Spelling performance that is largely dependent on verbal practice in school correlates more highly with *Varimin* F₂ factor scores (learning assets), $r = .31$. Spelling performance is thus also correlated with basic intelligence (*Varimin* F₁ factor scores), which is not surprising.

Table 4.05 section C:

The numeracy results meet expectations only partially. The correlation between numeracy and *Varimin* F₁, factor scores of basic intelligence, is high as expected ($r = .64$). An expected additional, although somewhat lower, correlation with *Varimin* F₂ factor scores (learning assets) did not show up. The correlation was $r = .07$, only barely significant.

Varimax results (Table 4.05)

Correlations between numeracy and Varimax factors are not elucidating. As expected, the correlation of numeracy is somewhat higher for fluid intelligence (with factor scores of F₂, $r = .59$) than for crystallised intelligence (with factor scores of F₁, $r = .44$).

As expected, spelling and numeracy performances correlate better with crystallised intelligence than with fluid intelligence (numeracy .64 with crystallised intelligence vs. .21 with fluid intelligence, numeracy .50 with crystallised intelligence vs. .40 with fluid intelligence). Comparisons of *Varimin* vs. Varimax rotational validities thus are favourable for *Varimin* once again.

Table 4.05: Pearson correlations between *Varimin* and Varimax factor scores F₁ and F₂ obtained from IST and CFT IQ, a dictation, and a test of numeracy.

Rotation: Factor interpretation:		A. CFT IQ				B. Dictation				C. Basic arithmetic			
		Varimin		Varimax		Varimin		Varimax		Varimin		Varimax	
		g	L	c	f	g	L	c	f	g	L	c	f
School types	N	F ₁	F ₂	F ₁	F ₂	F ₁	F ₂	F ₁	F ₂	F ₁	F ₂	F ₁	F ₂
<i>Gymnasium</i>	381	.66	-.10	.40	.53	.44	.18	.50	.12	.55	-.14	.29	.48
<i>Realschule</i>	198	.60	.12	.50	.34	.46	.26	.51	.13	.42	-.01	.29	.31
<i>Hauptschule</i>	381	.60	.02	.44	.41	.47	.27	.52	.14	.39	-.05	.24	.31
Total	970	.73	.11	.59	.44	.60	.31	.64	.21	.64	.07	.50	.40

Notes:

g = General or basic intelligence

L = Learning capital

f = Fluid c = Crystallized intelligence

“Gymnasium” (a 9-year selective high school for gifted students, with an academic focus).

“Realschule” (a 6-year selective high school for middle-tier students, providing an allrounded education).

“Hauptschule” (a 5-year high school open to all students, focused on developing hands-on, practical skills).

Validation 3: Culture-free IQ test FRT

Data 3

Brocke et al. (1998) published an intercorrelation table that has already been referred to in Study I (cf. Table 4.01, line 18). Apart from using the eight IST subtests, $N = 279$ for IST), they also applied correlations with the Figure Reasoning Test (FRT) ($N = 241$ for FRT). The authors hoped “*to come up with evidence for IST-70’s internal validity*” (p. 94), above all by correlations of “fluid” IST-70 subtests with FRT. FRT’s correlation with Raven’s figural SPM, widely used for fluid intelligence, was $r = .93$.

Expectation 3

FA of intercorrelations of FRT variables should show a high loading on *Varimin* factor F_1 , provided *Varimin* F_1 is correctly interpreted as indicating general intelligence. FRT should show a low loading of *Varimin* factor F_2 , provided *Varimin* F_2 is correctly interpreted to be the learning assets factor. Varimax factors F_1 and F_2 , allegedly indicating fluid and crystallised intelligence, should divide intelligence into two sub-branches, without indicating which factor, F_1 or F_2 , would represent fluid and which crystallised intelligence.

Assessment 3 and comments

Table 4.06 shows the results of *Varimin* and Varimax analyses of the eight IST variables (in rows 1 to 8, ME subtest excluded). The FRT variable was added in row 9. It turned out:

Varimin result: FRT reveals the expected high *Varimin* F_1 loading (.70, basic intelligence) as well as an expected low F_2 loading (-.15, learning assets). The reason that FRT is unrelated to F_2 is that advantages by schooling and cultural experience are not supposed to help solving FRT tasks just as such experience is not helpful at solving IST Figure selection FS (-.12) or, even more so (-.59), for solving IST task dice (DI).

The main outcome is that the factorial distinction between “fluid” and “crystallised” intelligence (via Varimax) is much less marked than the distinction between “basic intelligence” g and “learning assets” (via *Varimin*), L . The empirical and conceptual separation of “intelligence” and “learning assets” is thus considerably more marked when these two constructs, intermingled by Varimax rotation, are untangled by *Varimin* rotation.

The authors unfortunately did not apply FA to IST subtests together with FRT subtests or with FRT total. They merely provided an alternative way of relating test results from the two sources. They employed multiple regressions using IST variables as independent variables (IVs) and the FRT total as the dependent

variable (DV). The calculations were complicated and the result was, in their own view, unconvincing.

Table 4.06: *Varimin* and Varimax rotated factor loadings of FRT data (row 9) and loadings of IST 70 subtest variables (rows 1 – 8).

Interpretation Subtests		Varimin		Varimax		h^2
		g F ₁	L _p F ₂	c F ₁	f F ₂	
1	SC	.63	.51	.80	.09	.65
2	WS	.63	.43	.75	.14	.59
3	AN	.77	.37	.81	.28	.73
4	CO	.75	.16	.65	.41	.58
5	NU	.76	.25	.72	.37	.66
6	NS	.81	-.01	.56	.58	.65
7	FS	.76	-.12	.45	.62	.59
8	DI	.66	-.58	.05	.88	.77
9	FRT	.76	-.15	.43	.65	.60
% Variance		53	12	39	26	65

Notes:

g = General intelligence

L_p = Learning capital

f = Fluid c = Crystallized intelligence

SC Sentence completion, WS word selections, AN analogies,

GE communalities, NU numeracy, NS number series,

FS figure selection, DI dice

Summary of Study II

After validating *Varimin* and Varimax factors, measured by tests that facilitate their validation, *Varimin* comes out on top. Performances in two culture-free tests (CFT and FRT) are better predicted by *Varimin* F₁, interpretable as basic or general intelligence, than by Varimax F₂, which since Cattell and Horn has been considered as indicating “fluid”, i.e., general intelligence. In other words, Varimax-based general intelligence is outdone by *Varimin*-based general intelligence which – in order to account for this property – I prefer to call “*basic* intelligence”.

The performances with two tests as external criteria (spelling and numeracy), supposed to benefit from preceding cultural learning, are better predicted by *Varimin* F₂, the learning assets factor, than by Varimax F₁ interpreted as crystal-lised intelligence, which allegedly manifests itself by effects of school and cultural

learning. Varimax F_1 should have shown its contribution to scores in spelling and numeracy. But it did so less convincingly than *Varimin* F_2 .

Discussion of chapter 4

Studies I and II presented in this chapter were devoted to the question whether CSM rotation (by *Varimin*) of intelligence test data would be more valid than conventional SSM rotation by Varimax. The answer was in the affirmative, the results are convincing. *Varimin* separated the two main sources of variance of IST performance; intelligence and learning. The proposed interpretation of the two factors proved valid.

Proposing new research strategies with claims that they outdo conventional measuring tools should go along with exposing weaknesses in past research. I believe and repeat that the main weakness of past research lies in Thurstone's SS model, on which past research was based. SSM ignores what Jäger terms a "core assumption" that belongs at the beginning of all intelligence research: "*Each intelligence performance involves (along with other conditions) all intellectual abilities, albeit with significantly differing weights. Every performance's variance can be dismantled into its respective components.*" (Jäger, 1997, p. 4).⁸⁴ But Jäger's own ways to attain this goal did not question the SS principle.

Varimin analyses of intelligence data showed that, to attain good test scores, increased learning efforts in schools and beyond can compensate, to a certain degree, for lack of intelligence, and vice versa: Educational psychology (Weinert, 1996) has been aware of this for quite some time, as has actually anybody with common sense. What appears to be new, though, is that the concept of *learning influence* has been freed, methodically and conceptually, from the concept of intelligence. The two concepts had been and still are generally bonded together, without justification, by an SSM artefact called "crystallised" intelligence. CSM separates the two conditions, nature and nurture, facilitating their comparison and theoretical evaluation.

Factor rotation aiming at SS, which assigns only one factor (one source of variance) to test variables, destroys a general factor g which is preformed by initial factor solutions, instead of improving its representation. SSM spreads variance of the main factor's contribution to subsequently extracted factors attributing to them an unjustified premium of communality.

To make sense out of *Varimin* F_2 , no new "ability" (such as "crystallised intelligence") needs to be invented. F_2 can be understood as "earned interest",

⁸⁴ In his data analyses, Jäger did not break away from the SS principle. Instead, he had to employ alternative methods to prove empirically the concurrence of latent functional contributions to intelligence test performance. This he tried to achieve in his BIS model. (*Berliner Intelligenz-Strukturmodell*).

summed by performance throughout life-long learning, from which test participants benefit differently, according to varying amounts of preceding practice. *Varimin* F_2 is a fertiliser, while F_1 is a measure of an innate potential for growth of its output.

A practical consequence of these results would have us recommending that the IST authors revise the utilisation of their test. They should focus the user's attention on basic intelligence g and put it centre stage. Thus far, the IST manual does not even make mention of basic intelligence although the test can capably provide information about basic g .

Information about learning assets accumulated by educational and school influence might also be obtained from IST. One might take raw test results from IST, transform them by employing *Varimin* F_2 factor loadings and by using suitable conversion tables. Two separate scores might be obtained to compare summed F_1 and F_2 test values (e.g., by using L_d or L_p). The two scores can be related, e.g., after standardisation. This strategy would allow for an assessment of relative contributions of past learning and practice to levels of achievement. One might eventually even be able to distinguish between “diligent” and “lazy” test candidates, by relating individual effort expenditure to individual giftedness.

Chapter 5

Varimin factors from Big Five personality data

Point of departure

In this chapter I apply *Varimin* rotation to factorial personality data. *Varimin* rotation is expected to suspend the Big Five personality factors of extraversion-introversion, stability-neuroticism, openness, conscientiousness, and agreeableness that presently make up the core of prominent personality dimensions. Theoretically, the loss of these dimensions will hardly be harmful since in fact there is no theory requiring the Big Five as necessary building blocks. An analysis aiming at complexity might lead to first steps towards theoretical re-orientation. Details cannot be predicted, the course a complexity analysis of personality descriptions will take is open to bottom-up surprises. The issues examined in chapter 5 have been presented in my German-language publication, *Basiskomponenten der Persönlichkeit*.⁸⁵ Non-German readers are invited to contact me regarding any matters that have been dealt with insufficiently in the present monograph⁸⁶.

⁸⁵ Ertel (2011b).

⁸⁶ Contact:: sertel@uni-goettingen.de

Material and analysis

The data used here ($N = 11,724$) was obtained from the standardised German-language self-assessment form, NEO-PI-R (Ostendorf & Angleitner, 2004)⁸⁷. Intercorrelations of six facet variables for each of five factors were available, i.e., a total of 30 variables (cf. Ostendorf & Angleitner, 2004, Table 1). A PCA of the correlations and a subsequent *Varimin* rotation yielded the factor structure shown in Table 5.01 and in Figure 5.01. The factorial results obtained by Varimax rotations of such data are well-known. In Figure 5.01 they are indicated by assigning them the letters N, E, O, A, C as labels. The Varimax factors will be referred to by using these labels. .

⁸⁷ The correlation matrix was kindly provided by the first author, Fritz Ostendorf.

Table 5.01: *Varimin*-rotated factors of the 30 NEO-PI-R facet variables
Data source: Ostendorf & Angleitner (2004). N = 11,724.

	F ₁	F ₂	F ₃	F ₄	F ₅	h ²	
1:	-0.401	0.350	-0.323	-0.434	0.398	0.734	N1 Anxiety
2:	-0.375	0.521	-0.058	-0.101	0.533	0.710	N2 Hostility
3:	-0.510	0.400	-0.340	-0.402	0.230	0.750	N3 Depression
4:	-0.488	0.229	-0.310	-0.457	0.145	0.616	N4 Self consciousness
5:	0.029	0.407	-0.426	0.250	0.392	0.565	N5 Impulsivity
6:	-0.548	0.329	-0.444	-0.297	0.216	0.741	N6 Vulnerability
7:	0.573	-0.310	-0.448	0.103	0.276	0.712	E1 Warmth
8:	0.335	-0.185	-0.340	0.426	0.334	0.555	E2 Gregariousness
9:	0.492	0.009	0.323	0.426	0.285	0.609	E3 Assertiveness
10:	0.419	-0.025	0.150	0.201	0.534	0.525	E4 Activity (open to)
11:	0.192	0.193	-0.123	0.478	0.273	0.393	E5 Excitement seeking.
12:	0.620	-0.078	-0.294	0.311	0.260	0.642	E6 Positive emotions
13:	0.382	0.553	-0.352	0.035	0.023	0.577	O1 Fantasy
14:	0.503	0.373	-0.323	-0.341	0.029	0.613	O2 Aesthetics (open to)
15:	0.522	0.375	-0.383	-0.195	0.249	0.660	O3 Feelings (open to)
16:	0.534	0.247	-0.200	0.245	-0.092	0.454	O4 Ideas (open to)
17:	0.595	0.407	0.113	-0.168	-0.179	0.593	O5 Actions (open to)
18:	0.432	0.310	-0.236	-0.016	-0.287	0.422	O6 Values (open to)
19:	0.399	-0.429	-0.382	0.004	-0.111	0.502	A1 Trust
20:	-0.077	-0.450	-0.274	-0.438	-0.191	0.512	A2 Straightforwardness
21:	0.348	-0.478	-0.473	-0.265	0.058	0.646	A3 Altruism
22:	0.006	-0.480	-0.370	-0.311	-0.368	0.600	A4 Compliance
23:	-0.290	-0.354	-0.373	-0.390	-0.123	0.516	A5 Modesty
24:	0.218	-0.136	-0.571	-0.312	0.062	0.494	A6 Tender-mindedness
25:	0.622	-0.239	0.441	-0.052	0.131	0.658	C1 Competence
26:	0.135	-0.276	0.480	-0.407	0.298	0.580	C2 Order
27:	0.236	-0.459	0.345	-0.470	0.218	0.653	C3 Dutifulness
28:	0.438	-0.054	0.459	-0.281	0.415	0.657	C4 Achievement striving
29:	0.355	-0.446	0.515	-0.267	0.213	0.706	C5 Self discipline
30:	-0.011	-0.255	0.420	-0.531	-0.123	0.539	C6 Deliberation
R ²	16.9	11.9	13.2	10.4	7.3	59.7	
λ	5.71	4.65	3.47	2.50	1.61		

Note:
h² = Communality
R² = Explained variance
λ = Eigenvalues

Signs of loadings of factors F₁ and F₃ were reversed after extraction in order to match this arbitrary feature with the positive and negative meanings of these bipolar factors.

Varimin factors	Varimax factors
F ₁ = Functionality	F ₁ = Neuroticism
F ₂ = Representation	F ₂ = Extraversion
F ₃ = Regulation	F ₃ = Openness
F ₄ = Activation trend	F ₄ = Agreeableness
F ₅ = Activation level	F ₅ = Conscientiousness

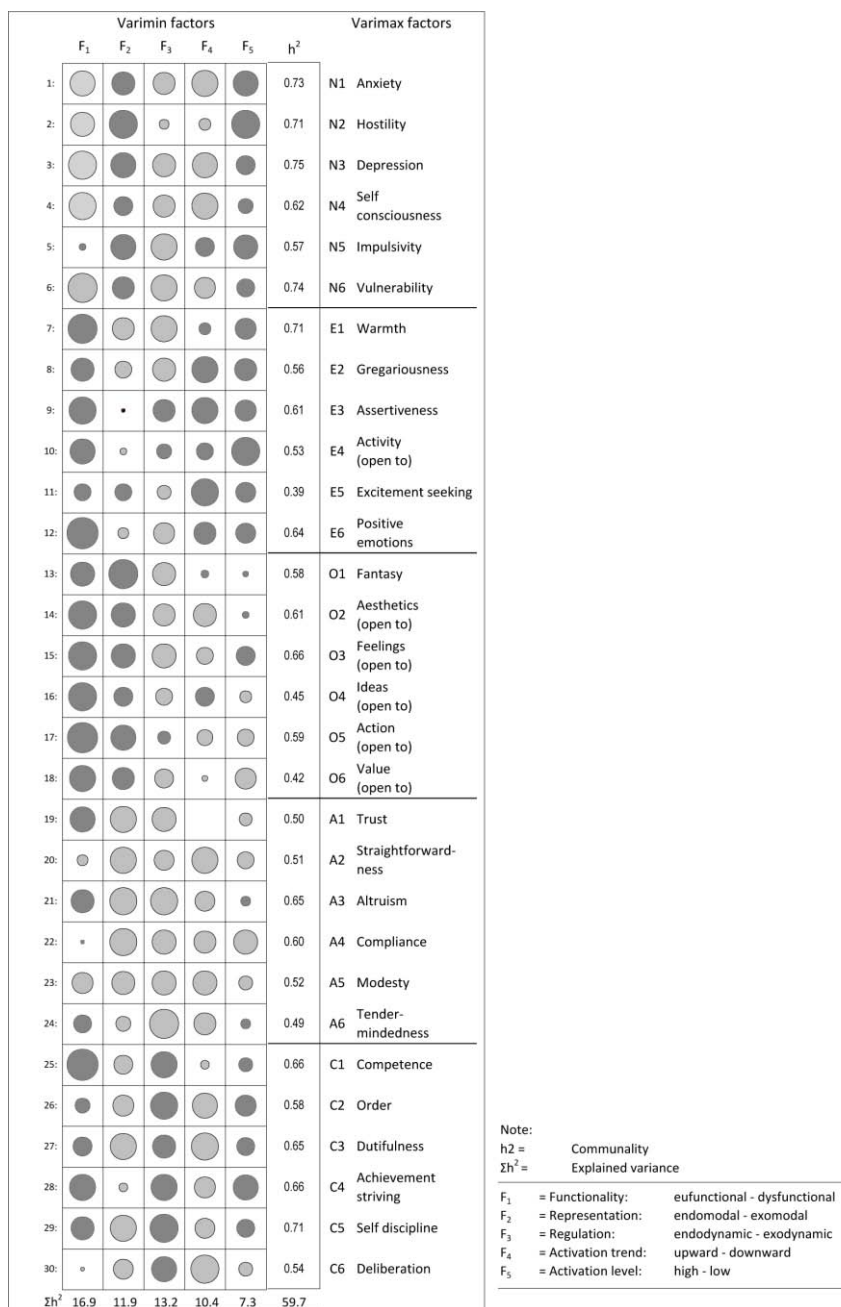


Figure 5.01: Varimin-rotated factors of the 30 NEO-PI-R facet variables
Data source: Ostendorf & Angleitner (2004) N = 11,724.

Towards an interpretation of *Varimin* factors of personality

The distribution of *Varimin* factor loadings engenders two preliminary comments (cf. Figure 5.01):

1. *Varimin* factors, represented by dark and light circles (for positive or negative loadings, respectively) is not entirely unrelated to the Varimax Big Five clusters (N, E, O, A, C). Facet variables belonging to Varimax clusters have *Varimin* loading patterns that are similar to each other. This is not surprising, because the Varimax clustering of the Big Five facets reveals similarity among these facets in the first place. Any orthogonal transformation of variables preserves similarities among them. The question remaining is what causes facet variables to have profiles of *Varimin* factor loadings that are similar to each other (more about this issue below).
2. If one compares *Varimin* factor loadings column by column, it is clear that *Varimin* factor loadings with the same loading signs are not located exactly within the limits of Varimax dimensions, especially in the case of F_2 , F_4 , and F_5 (not quite as pronounced in the case of F_3). This indicates that variables of the five Varimax dimensions, which have thus far been conceived as being independent of each other, may indeed have something in common. Variables of different Varimax dimensions may have loadings on equal *Varimin* factors, albeit possibly of different amount and with different loading signs.

More preliminaries

Interpreting *Varimin* personality factors is more elaborate and more involved than the general practice of ad-hoc labelling of factors. Findings must be embedded semantically into contextual knowledge. It turned out, as a surprise, that *Varimin* factors are conceptually interrelated. I shall pay attention to their relationships with delineations at an unusually abstract level. The naming of *Varimin* factors can hardly make use of familiar trait vocabularies (where one might find, e.g., *openness*, *conscientiousness*, *agreeableness* or habitual technical terms such as *extraversion*, *neuroticism*).

In chapter 2, minimal pair comparisons of variables were recommended for carving out the meanings of *Varimin* factors. In view of the larger number of variables of the present data set the description of detailed pair comparisons would require a great deal of space. Pairs of facets will only occasionally be contrasted. Since *Varimin* factor loadings generally have bipolar directions, a short-cut method similar to contrasting minimal pairs has been applied in Table 5.02. Five variables with extreme positive loadings on one factor have been grouped and contrasted to five variables with extreme negative loadings on that factor. This procedure does not produce distinct trait opposites such as *assertive vs. yielding*, *altruistic vs. egoistic*,

warm vs. cold. In Table 5.02 *Varimin* factor terms are chosen to denote the respective polarities, not their positive or negative manifestations. This procedure requires sensitivity to semantic nuances. The size of loadings of a *Varimin* factor on variables does not determine their meaning in the first place, as it has always been taken for granted for interpreting Varimax factors.

In addition, the importance of factors is not determined by the amount of statistical variance they explain, but by their contribution to an emerging nomological network. In the case at hand, for example, factors extracted later (with less communality) appeared to be semantically more basic than factors extracted earlier. Factor F₅, which was extracted last (*level of activation*), contributed most to understanding factors F₂ to F₄. According to Cronbach and Meehl (1955), one should always aim for nomological networks when conducting factor-analytical research⁸⁸. In this regard, Fischer (1967, p. 125) also demanded: “*Actually, all factors of any structure should be interpreted simultaneously, because they depend on each other to a significant degree.*”

⁸⁸ “A necessary condition for a construct to be scientifically admissible is that it occurs in a nomological net...” (Cronbach & Meehl, 1955, p. 290). “To validate a claim that a test measures a construct, a nomological net surrounding the concept must exist” (p. 291). “As research proceeds, the construct sends out roots in many directions, which attach it to more and more facts or other constructs” (p. 291). Constructs can enter a nomological net of theoretical rank, if they do justice to the nature of what is being described. They should also be suited to being successfully applied in other areas of psychological research belonging to the same system. Fiske (1976, p. 877) provides support by saying: “Construct validation requires the investigation of construct-operation units in an explicit conceptual framework”.

Table 5.02: Facet variables with extreme positive and negative loadings of five Varimin factors from NEO-PI-R data.

Positive loadings		Negative loadings	
F ₁ : State of functionality			
	Eufunctional		Dysfunctional
.622	Competence (C ₁)	-.548	Vulnerability (N ₆)
.620	Positive emotions (E ₆)	-.510	Depression (N ₃)
.595	Actions (open to)(O ₅)	-.488	Self consciousness (N ₄)
.573	Warmth (E ₁)	-.403	Anxiety (N ₁)
.522	Feelings (open to)(O ₃)	-.375	Hostility (N ₂)
F ₂ : Mode of representation			
	Eufunctional		Exomodal
.553	Fantasy (O ₁)	-.480	Compliance (A ₄)
.521	Hostility (N ₂)	-.478	Altruism (A ₃)
.407	Impulsivity (N ₅)	-.459	Dutifulness (C ₃)
.375	Feelings (open to) (O ₃)	-.450	Straightforwardness (A ₂)
.373	Aesthetics (open to) (O ₂)	-.446	Self discipline (C ₅)
F ₃ : Source of regulation			
	Endodynamic		Exodynamic
.515	Self discipline (C ₅)	-.571	Tender-mindedness (A ₆)
.480	Order (C ₂)	-.473	Altruism (A ₃)
.459	Achievement striving (C ₄)	-.448	Warmth(E ₁)
.441	Competence (C ₁)	-.444	Vulnerability (N ₆)
.345	Dutifulness (C ₃)	-.426	Impulsivity (N ₅)
F ₄ : Trend of activation			
	Upward		Downward
.478	Excitement seeking(E ₅)	-.531	Deliberation (C ₆)
.426	Gragariousness (E ₂)	-.470	Dutifulness (C ₃)
.426	Assertiveness (E ₃)	-.457	Self consciousness (N ₄)
.250	Impulsivity (N ₅)	-.434	Anxiety (N ₁)
.245	Ideas (open to) (O ₄)	-.390	Modesty (A ₆)
F ₅ : Level of activation			
	High		Low
.534	Activity (E ₄)	-.368	Compliance (A ₄)
.533	Vulnerability (N ₂)	-.287	Values (open to) (O ₆)
.415	Achievement striving (C ₄)	-.191	Straightforwardness (A ₂)
.398	Anxiety (N ₁)	-.179	Ideas (open to) (O ₅)
.392	Impulsivity (N ₅)	-.123	Modesty (A ₅)
Facet designations			
A = Agreeableness			
C = Conscientiousness			
E = Extraversion			
N = Neuroticism			

Varimin factor interpretations

1. Level of activation (Varimin F₅)

When comparing the five most positively loaded F₅ facets with the five most negatively loaded F₅ facets (cf. Table 5.02), it is hard to avoid the interpretation that the difference is basically engendered by amounts of psychophysical energy (activation) required for manifesting the respective behaviours. Initially we can disregard the fact that releasing psychophysical energy might either be beneficial or inhibitory for the system as a whole (cf. in this context, *State of functionality* (factor F₁) as an extra feature).

Activation as a term was coined by general psychological research (Lindsley, 1951, Duffy, 1962)⁸⁹. Factor F₅ indicates the level of *psychophysical energy* associated with behavioural and experiential phenomena. The facets of *Deliberation* (C₆) vs. *Dutifulness* (C₃) constitute a minimal pair (cf. Table 5.02) regarding the expenditure of energy in the *Conscientiousness* cluster. *Dutifulness* (positive F₅ loading) normally requires more energy-demanding actions than *Sobriety* and *Thoughtfulness* (negative F₅ loading). The data reveal more minimal pairs of variables for F₅ with contrasting levels of apparent energy expenditure (cf. Table 5.01). F₅ shows higher values across greatly differing Varimax facets: *Activity* (E₄), *Hostility* (N₂), *Achievement striving* (C₄), and lower values for different A and O facets: A₄ *Compliance*, A₂ *Straightforwardness*, A₅ *Modesty*, and O₅ *Ideas* (open to)⁹⁰.

2. Trend of activation (Varimin F₄)

Trend of activation refers to a tendency to change the amount of energy expenditure. While energy is being expended a person may want to generate even more energy (*ascending* or *positive trend of activation*). On the other hand, too much energy might be used up (depending on conditions of homeostasis), which is generally associated with the need to lower its level (*descending* or *negative trend*). Even in cases of low energy expenditure an upward or downward trend of energy expenditure may follow depending on discrepancies between actual and optimal activation (cf. Eysenck, 1973, on the differential phenomena of activation).

F₄ has its strongest positive loading with *Excitement seeking* (E₆), a prototype for striving to heightened activation. Extraversion (E₂ *Gregariousness* and E₃ *Assertiveness*) are also associated with upward trends of psychophysical energy release (effort).

⁸⁹ “Characteristic individual differences in activation, or responsiveness, are suggested as the basis from which certain other differences in behaviour may be derived” (Duffy, 1962, 322): Duffy alternatively uses “energy level”, “energy mobilization”, and “degree of excitation”.

⁹⁰ “Active or controlled processes are energy or resource dependent, while passive or automatic processes are not...” (Sanders, 1983, p. 74).

The downward trend of activation is strongest with the C facets of *Deliberation* (C₆) and *Dutifulness* (C₃). People who like to deliberate do not act hastily, they want to consider issues in peace, to remain thoughtful, composed, and serene. The dutiful person acts in predetermined ways and avoids breaking the rules, is obedient, meticulous, loyal. *Self-consciousness* (N₄) and *Anxiety* (N₁) facets that load the Varimax factor N (*Neuroticism*) also have negative *Varimin* F₄ loadings. With these facets optimal mental energy expenditure (F₅) is exceeded, a need to reduce energy consumption is prevalent. This is consistent with clinical research: As a rule, clients suffering from mental instability are in want of a reduction of tension. This is one of the main goals of all variants of psychotherapy. The extraversion facets are likewise associated with an increased level of energy production (cf. F₅), as was shown. Extraverted and neurotic persons are thus similar in this respect. But extraversion is associated with some heightened need more activation (F₄), a desire for increased energy release, while neuroticism is associated with a need for diminished energy release.

3. Source of regulation (*Varimin* F₃)

The *Varimin* factor F₃ has pronounced positive loadings on the facets of *Conscientiousness* [*Self-discipline* (C₅), *Order* (C₂), *Dutifulness* (C₃), *Achievement striving* (C₄), and *Competence* (C₁)]. An inner source of regulation, an ego or will centre, plays a determinant role. This feature denotes a system-internal causation of behaviour and may be called *endodynamic regulation* or *endoregulation*.

The opposite polarity is called *exodynamic regulation*, or *exoregulation*, which denotes regulation by non-ego determinants, i.e., by environmental stimulation, excitation, enticement, temptation, commands, etc. One should also distinguish between positively and negatively evaluated endo- and exoregulation. Forms of behaviour triggered by exoregulation, if positively valued, are frequent among facets of *Tender-mindedness* (A₆) and *Altruism* (A₃). These forms of behaviour are responsive, not agentive, generally initiated by others without strong obligations. A *modest* (A₅) and *compliant* (A₄) person avoids imposing his will on other people, so as to maintain *for them* unobtrusive conditions of endoregulation.

Exodynamic regulation is predominantly found with facets of *neuroticism*. This makes sense because a *self-conscious* (N₄), *vulnerable* (N₆), *depressive* (N₃) and *fearful* (*anxiety*) (N₁) person feels helpless and delivered to his own mental and emotional urges. Given such lack of will power, his troubled *endoregulation* is turned into *exoregulation*. The person is overwhelmed by his own emotions. Such lines of thought were advanced by de Charms (1968), Heider (1958), and Deci and Ryan (1985). Boekaerts, Pintrich, and Zeidner (2000), and Baumeister and Vohs (2004) provide an overview of research focusing on self-regulation (unfortunately without sufficiently considering exoregulation). In Ertel (2011b) more theoretical support for this view is provided, including references to Sigmund Freud's *ego-id* polarity, Hen-

ry Murray's (1938) concept of *press*, and James Gibson's (1963) *endo-exo* perspectives in his revolutionary theory of perception.

4. Mode of representation (Varimin F₂)

The endo-exo distinction is abstract, but vital. A bipolar endo-exo concept is also suggested to underlie major distinctions of mental representations. Representation means that behaviour, regulated by internal or external determinants, is embedded in and surrounded by a myriad of mental units bearing endo or exo characteristics. Features of an *inner* world may be described as vague, unarticulated, holistic, feeling-like and subjective. These characteristics shall henceforth be called *endomodal*. Features of mental representation belonging to an *outer* world are articulated, distinct, delimited, marked-off, detailed, and often called "objective". These characteristics will henceforth be called *exomodal*. This is suggested by the negative polarity of factor F₂ which is termed *mode of representation*.

With positive F₂ loadings (indicating *endomodal quality*) facets of *openness* and *neuroticism* appear comparable. This finding may initially come as a surprise because *openness* and *neuroticism* appear to be quite different. But within the context construed here, they are plausibly comparable. *Hostile* (N₂), *impulsive* (N₅), and *depressive* (N₃) people with a much disordered ego are out of touch with reality, and extremely subjective because of concomitant emotional traits. The frequently found correlation between *neuroticism* and *introversion*, which thus far was hardly understood, suddenly seems plausible. By the same token it becomes comprehensible that endomodal experiences are predominant in humans who are *fantasy prone* (O₁) and *open to feeling* (O₃), showing much subjective supplements when dealing with their surroundings, like those with an interest in *aesthetics* (O₂), or people entertaining *ideas* (O₄) and who are *open to values* (O₆). In these cases subjective information processing (endomodal qualities) develops spontaneously, perhaps by hereditary disposition, while for people with an acquired neurotic dysfunctioning internal tensions engender an inflation of endomodal ways of dealing with self and others. An overview of pertinent research supporting the factor interpretation presented here is contained in a meta-analysis by Mor and Winquist (2002) about self-focused attention and negative affect.

The concept of *mode of representation* cannot easily be built into further contextual knowledge. Many earlier attempts to grasp variance and distinctions in this field are referred to in my more-extended study (Ertel, 2011b). They include references to William Stern (1930) (polarity of *subjective vs. factual matters*); to Carl Jung (1930) (*feeling vs. thinking*); to P. Lewicki (2005) (*internal vs. external encoding*); to B. Shanon (1993) (*presentational vs. representational*); to S. Epstein (2003) (*experiential-intuitive vs. rational*); to P. S. Holzman and G. S. Klein (1954) (*levelling vs. sharpening*); and to H. A. Witkin (1959) (*field-dependence vs. field independence*).

5. State of functionality (F_1)

Factor F_1 manifests the personal system's actual or enduring *state of functionality*. The relationship of system components with each other and vis-à-vis their surroundings is more or less positive, balanced and undisturbed or negative, unbalanced and disturbed – with intermediate degrees separating the extremes. A positive *state of functionality* displays *eufunctional* system processes that are appropriate and undisturbed, even if they may be tense and strained. Excessive or prolonged stress and an unexpected trauma will cause *dysfunctional* symptoms indicating that the system is unstable.

All *neuroticism* facets are negatively loaded with *Varimin* F_1 , which was to be expected. Of the remaining *Varimax* facets *Modesty* (A_5) has slightly negative F_1 loading (cf. Table 5.02), probably because this facet contains numerous items of an implied ego weakness (*Self-consciousness*). Other variables have positive F_1 loadings, expressing *eufunctional* conditions in behaviour and experience, the strongest being *Competence* (G_1), *Positive emotions* (E_6), and *Open to actions* (O_5). The positive signs that have been affected by the sign reversal of F_1 and F_5 loadings are informative and just as fit as the negative ones, because an absence of dysfunctionality and the presence of mental health are preconditions for approaching the world and life with curiosity, competence and pleasure. Dysfunctionality requires expenditure of energy to eliminating disturbance in the system, thus energy will be lacking which must be freely mobilised for the person to become and remain open to their environment.

Concluding remarks

Having characterised the five *Varimin* factors as aspects of a functional whole, they shall henceforth be called *basic components of personality functioning*⁹¹. Personality is considered as a system of constituents cofunctioning and cooperating with each other with greater or lesser degrees of involvement for the system as a whole (Cronbach & Meehl, 1955). The task of differential psychology is to identify general and individual parameters of these components.

⁹¹ The term facets would not, unlike components or constituents, indicate the contributory function that basic traits have in shaping the whole. Endoregulation, for instance, is not only a facet of personality, but also a component of its manifestation. An addition of basic as in basic traits is recommended in order to indicate the latent level of its functioning.

The interpretations of *Varimin* factors on probation

In order to test whether my interpretation of *Varimin* factors remains stable across different judges, 26 students were presented with descriptions of the 30 Varimax facets. The description of each individual facet was written on a card. The descriptions used are exemplified for *gregariousness* and *aesthetics*.

Gregariousness: “I need to have people around me, I dislike being alone, I think parties and get-togethers are stimulating, I do not want to work alone in my profession.” (a facet of extraversion).

Aesthetics: “I have a keen sense for beauty in nature and in art and can be totally absorbed by music. Ballet, dance, poetry fascinate me.” (a facet of openness).

Participants were asked to rank order the 30 facet descriptions five times, once for every one of the five *Varimin* factor descriptions. The appendix contains *Varimin* factor descriptions for F₁ to F₄. The description for activation (F₅) e.g., was as follows:

Increased energy expenditure. Please give top ranks (1, 2, 3 etc.) to facets associated with highest energy expenditure and give further ranks to facets requiring less energy expenditure.

High energy expenditure means: *The individual uses a great deal of energy. He/she may want to achieve ambitious goals that require much energy consumption. It could also be that his/her energy is absorbed spontaneously by strong impulses, urges, and emotional experience. Or a great deal of energy is expended because of internal or external impediments that the individual has to deal with. Energy does not only manifest itself as unhindered power, but may be consumed by continued tension and blockage.*

Reduced energy expenditure. Please rank lowest (30, 29, 28 etc.) facets with lowest apparent energy expenditure.

Reduced energy expenditure means: *The individual uses only small amounts of energy. He/she may be pursuing less ambitious goals that can be reached with little effort. Maybe, by nature, he/she does not seek motivational or emotional stimulation. Possibly, he/she only wants to maintain his/her energy reserves. He/she may also need to overcome fewer internal and external impediments which, if present, would call for an increased energy input. At any rate, increased effort expenditure or continued tension and blockages occur less often in his/her case.*

The participants could, for instance, have assigned ranks 1 and 2 to the facets *Activity* and *Vulnerability* (these facets are highly positively F₅ loaded), and the facets *Compliance* and *Open to values* might have obtained the lowest ranks 29 and 30 (these facets are highly negatively F₅ loaded). Facets not yet ranked for increased

or decreased energy expenditure were eventually to be given intermediate ranks by the participants.

For each *Varimin* factor, ranks of the 30 facets were averaged across the 26 participants (possible range of means 1–30). For *Varimin* factor interpretations (F₁ to F₅), five mean rankings ensued. They were then ordered alphabetically by facet descriptions. By the same token, the facets’ *Varimin* factor loadings were also ordered alphabetically by facet descriptions, eventually providing five vectors of mean rankings and five vectors of factor loadings that were intercorrelated.

Figure 5.02 shows the result. As expected, the highest positive correlations of mean ranks with *Varimin* factor loadings are mainly found in the diagonal fields. Only the *source of regulation* factor deviates by showing a smaller correlation in the associated diagonal field. Post hoc the suspicion arises that the participants were not properly instructed about possible sources of regulation and that it was not made sufficiently clear that exodynamic regulation could indicate either weak will power (lack of ego regulation) or heightened susceptibility to external impulses. In four out of five cases, however, the factor interpretations proved to be replicable by independent judges.

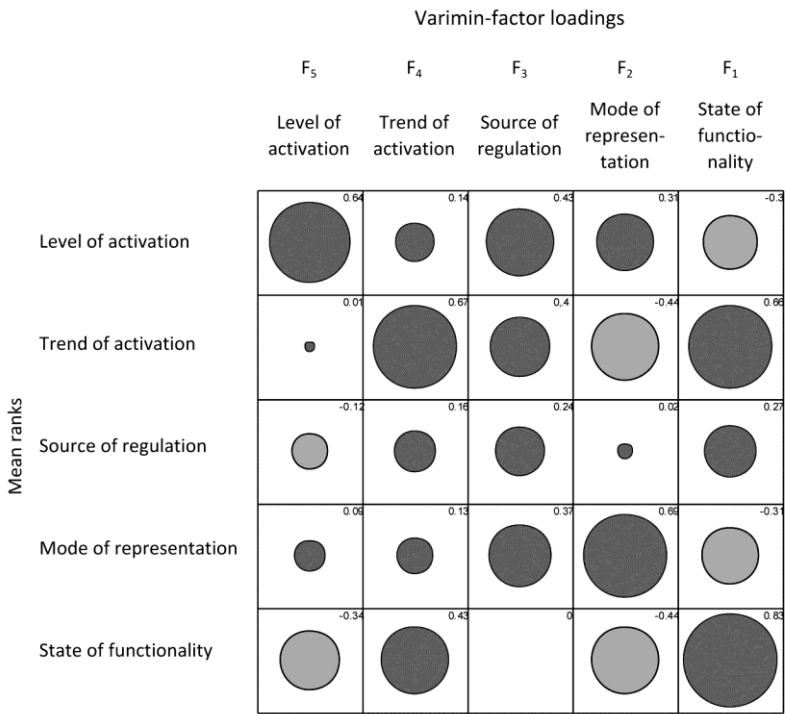


Figure 5.02: Correlations between *Varimin* factor loadings of 30 facet variables with average ranks of judged meanings.

Re-interpreting Varimax factors by profiles of *Varimin* factors

Varimax factors can be conceived as clusters of *Varimin* components. This can be exemplified by reinterpreting *neuroticism* and *extraversion* in terms of *Varimin* factors (for the remainder, refer to Ertel, 2011b).

1. Neuroticism: The most striking characteristic of people contributing to the *neuroticism* cluster is *disturbed functionality*. *Dysfunctionality* goes along with an increased level of activated energy and with the desire to tone down energy expenditure. The energy expenditure of neurotically disturbed persons is dependent on environmental impediments and on uncontrolled endodynamic impulses (*quasi-exoregulated*). The readiness for and/or ability to endoregulation is missing. Excessive exoregulation is often called *ego weakness* by psychoanalysts. A *disordered functionality* also leads to unbalanced *modes of representation*. Information is chiefly processed *endomodally* (subjectively with excessive feeling quality). Increased *self-consciousness* is accompanied by reduced assessments of reality. An increased energy expenditure in disordered systems makes sense, because solutions are sought for reducing painful activation (tension), often with support requested from experts (therapists). Activated energy can be employed differently, depending on whether the system is balanced or imbalanced. Many authors have noted this. Thayer (1985) uses bipolar attributes to indicate undisturbed vs. disturbed energy expenditure within person systems: *energetic arousal* (in the case of extraversion, for instance) and *tense arousal* (in the case of *neuroticism*). This terminological distinction helps to compare the use of mental energy in *undisturbed* and *disturbed* systems.

2. Extraversion-Introversion: Most pronounced in individuals with high *extraversion* in Big Five questionnaires is an increased *level of activation* associated with an *upward trend*⁹². Representations of internal and external experiences are preferably processed in an exomodal manner, i.e., focused on external objectives. Endomodal concomitants are often neglected. The regulation of energy for extraverts is, however, not uniform. While most extraversion facets are associated with exoregulation (behaviour is guided by external stimulation, social incitements are prominent), facets clustered as extravert do not totally lack endoregulation. Extraverts manage their actions, they know what they want. The functional status of extraverts is balanced and has positive signs.

⁹² SS-orientated factor analyses make an activation/energy trait disappear. Zuckerman (1994) comments: "...Energy Level is considered a major trait ... and regarded as a basic aspect of temperament in neo-Pavlovian models ..., but it tends to get lost or subsumed under Extraversion in the Big Three and the standard Big-Five" (p. 66).

For the opposite pole, for introverts, the statements characterising extroverts must be reversed. The behaviour of introverts is based on a sub-average energy level. Moreover, introverts tend to lower their activation levels. Introverts tend to have endomodal representations even when other people generally have exo-modal experiences. Introverts also prefer endoregulated activities. They do not like to be guided by other people. Their functionality status often indicates some lack of adjustment.

Concluding from these two examples: neuroticism, extraversion-introversion and the other Varimax “dimensions” are not elementary building blocks of some personality theory, instead they are clusters based on definable *Varimin* profiles (cf. Table 5.07). Watson and Clark (1997, p. 780) likewise emphasise that positive affectivity is a “*central feature of the construct [extraversion]*” which combines a number of sub-constructs (“*positive emotional experience forms the core of the higher order construct*”). The question remains open why *Varimin* components of activation, regulation, representation etc. combine as profiles called extra- and introversion. But this question is of subordinate importance here.

Discussion of chapter 5

Objections to the present interpretation of *Varimin* factors are fairly predictable. It might be argued that the suggested interpretations are not compelling, different researchers might suggest different factorial meanings. It might likewise be objected that these interpretations are primarily based on semantics, not on objective data. However:

1. Writers of questionnaire items and subjects responding to them operate with the semantics of verbal units. Any production and comprehension of language is dependent, in the first place, on interpretations, not on quantitative units.
2. With the aid of words and terms references are established to trans-linguistic reality. The components of kinship terms (e.g. generation, gender etc., cf. chapter 2) are undoubtedly anchored in non-linguistic realms. In the semantic field of personality, components obtained by the same methodology should no less be considered as referring to some essentially trans-linguistic reality.
3. Doubts are merely legitimate as to whether terminological decisions are optimally appropriate and useful for more comprehensive contexts demanding an overall understanding.
4. Terminological decisions made in this study seem to take directions towards a developing conceptual network. *Varimin* factors emerging here reveal relationships

to more encompassing themes and the theories of scholars ranging from Pavlov to Freud to Murray to Shanon etc.

5. One of the most urgent theoretical desiderata in psychology is that a bridge be built joining differential psychology with general psychology. SSM work in differential psychology has not yet built and apparently cannot build connections with general psychology⁹³. Surprisingly, CSM factors obtained from interindividual covariances, using a bottom-up strategy, proved to be useful for describing processes of general psychological functioning.

6. It is possible that the componential concepts obtained from *Varimin* analyses of personality data are liable to be used with bio-psychological and neuro-psychological aims (with the help of constructs like energy, activation, regulation, representation, functionality).

7. Objections to terminological decisions: I should nevertheless always consider them if objections are embedded in theoretical contexts replacing and significantly improving the beginnings that have been developed here.

⁹³ “It would be advantageous, if not magnificent, if a between-subjects five-factor model would imply ... exchangeability [with a within-subjects model]... However, the required equivalence has not been shown, and [we] expect that it will not, in general, be a tenable assumption.” (Borsboom, 2013, p. 213). Apparently, the basic features gathered from inter-individual correlations of the NEO-PI-R facets by using *Varimin* rotation (activation, regulation, representation, etc.), meets the demands formulated by the authors: “If it is shown that a given set of ... processes [within individuals] leads to a particular latent variable structure [of personalities], we could therefore say that this set of processes realizes the latent [personality] variables in question” (p. 215 f).

Insight and Outlook

My dispute with alleged misconceptions of factor analytical research may appear “rebellious”. Could I have prevented giving this impression? When I went into the lateral thinking mode and during subsequent years of empirical testing - not all results have been presented here - did I perhaps not abide by the rules of proper scientific conduct? I used verifiable data trying to prove the theory that SS modelling was an error. At the same time, I supplied evidence for the benefits of the CS modelling strategy. Even at the risk of exaggeration I dare forecast that adjusting psychological research to CSM might open new doors to latent domains of mental processing. Componential analyses might diminish the chaos of unanalysed constructs that has long inhibited psychological progress. New perspectives might become visible and initiate prolific developments.

Or should I state in mitigation that *Varimin* is just another rotation method to be added, with pluralistic tolerance, to the long list of earlier such techniques aiming to solve the continuing problem of indeterminacy of extracted factors? Should I maintain that conventional SSM data analyses, including SEM, might have the promising future that most methodologists of today, endorsing these approaches, are expecting.

What could prove me wrong claiming that simple structure as advocated by Thurstone and his followers had gravely detrimental consequences to this day (2013) and that SSM was and sadly remains the main source of the misery afflict-

ing multivariate analyses? What could stop me believing that by taking a 180-degree turn the past mistakes are rectified? Unrecognized logical errors in my reasoning? Errors in my empirical work? I have braced myself for criticism and would enjoy being confronted with surprising new facts. I would be happy if sceptical conventionalists were up in arms, if they would deal with the challenge by undertaking empirical attempts to defend their approach against my claim that they have failed. For me it is inconceivable that the majority of my results can be brushed away. But I can wait and see whether this book is just an instance of an aberrant scientist's need for adventure or the beginning of a turn in multivariate methodology towards ensuing conceptual and empirical innovations.

Review of chapters 1 to 5

Chapter 1: Critique of the simple structure doctrine

In chapter 1 the aim of the book is outlined: A methodological doctrine is revealed and critiqued that has prejudiced FA since its beginning. Simple structure (SS), the guiding principle for factor rotation (Thurstone, 1935/1947) is unveiled as questionable because it generally distorts latent sources of variance of manifest empirical variables instead of revealing them. The critique is based on theoretical considerations and is supported by many verbatim quotations from critical authors. The present calamity of factorial research is deemed to be due to these methodological flaws. One-sided mathematical formalisation in the discipline has lost its objectives by unjustifiably ignoring ordinary sources of gaining knowledge, including common sense. The problem of SS cannot be solved by circumplex and structural equation modelling (SEM) which suffer no less from SS errors. An alternative factor transformation leading to complex structures is demanded. A paradigm change is overdue.

Chapter 2: Finding complex structures

This chapter resumes the preceding criticism. The rotation procedure Varimax which is commonly used to generate SS is replaced with *Varimin* which aims to manifest latent complex structures (CS). *Varimin* optimises the model of complexity which – being already announced by initial unrotated structures – still needs

improvement. The new method raises various questions, of which five are discussed. How can *Varimin* factors be interpreted? Do latent sources of covariance not already appear sufficiently complex with initial solutions? Are SS solutions not fairly interpretable, how else could they have been routinely used? How to interpret the commonly encountered bipolarity of *Varimin* factor loadings? Is FA with complex structure transformation applicable to data affected by method factors? Ten empirical applications of *Varimin* transformation serve as exemplary tests. Particular features of transformation to CS, revealing latent sources of covariance (by *Varimin*), are elucidated by comparing pertinent results with those obtained from transformations to SS (by Varimax). Varimax will remain useful for merely clustering objectives. Attention is also drawn to limitations of the methodical innovation.

Chapter 3: Decathlon data under analysis

The results of a factorial study are reported using transparent sports data: Decathlon record scores covering 10 sporting events. The aim is to compare *Varimin* and Varimax results regarding factorial stability and interpretability. It is shown that *Varimin* factors reveal latent components of sports activity in interaction, while Varimax factors yield obscure clusters of features. In addition, factor structures obtained by *Varimin* rotation are more robust to changing data sources than those obtained by Varimax rotation. The results of this study are consistent with pertinent non-factorial results of sports physiology.

Chapter 4: Intelligence data under analysis

Eighteen matrices of intercorrelations of eight subtest variables of the intelligence test IST are subjected to principal component analysis, the resulting factors are rotated by *Varimin* to model complex structure (CS). The 18 *Varimin* solutions are aggregated, two factors result: *Varimin*-F₁ represents a general factor g (“basic intelligence”), *Varimin*-F₂ represent a performance-modifying factor, apparently based on previous educational training and learning effects (to be termed “learning assets”, l). The validity of *Varimin*-F₁, basic intelligence, is ascertained by high correlations between g and test scores of general intelligence, operationalised by culture-free CFT and FRT. The interpretation of *Varimin*-F₂ as learning assets finds support by significant correlations with school grades and scores in orthography and arithmetic. The 18 PCA-factors are also transformed by Varimax to SS. This transformation causes the splitting up of initial g into two seemingly separate factors, called “fluid” and “crystallised” intelligence by convention. In addition, differences between Varimax F₁ and F₂ of correlations with external criteria of general intelligence versus school grades and training scores in orthography and arithmetic that should emerge are missing. Apparently, SS modelling of intelligence test data amalgamates general intelligence with learning effects. Rotation of

intelligence data to SS does not reveal independent contributions of latent functional components, but hinders their detection.

Chapter 5: Varimin factors from Big Five personality data

Varimin rotation is applied to five PCA factors obtained from 30 facet variables of NEO-PI-R (Ostendorf & Angleitner, 2004). As expected, *Varimin*-rotated factors do not replicate the Big Five (*neuroticism*, *extraversion*, etc.), but instead reveal five distinctive bipolar factorial components: *level of activation* (high-low), *trend of activation* (ascending-descending), *source of regulation* (endodynamic-exodynamic), *mode of presentation* (endomodal-exomodal), and *status of functionality* (eufunctional vs. dysfunctional). The well-known Big Five factors turn out to be clusters of *Varimin* components rather than unique dimensions. The validity of the five features obtained by *Varimin* has largely been confirmed by independent rankings of the 30 NEO-PI-R facets using *Varimin* features as ranking criteria. Replacing SS analysis on a broader scale through CS procedures might lead to building blocks for conceptualising personality and individual differences as future theoretical goals.

Appendix

Bipolar meanings of *Varimin* factors provided to participants as criteria for rank ordering the 30 NEO-PI-R facets.

For factor F₅, level of activation, refer to section *Varimin factor interpretations* p.126. *The interpretation of Varimin factors on probation* for F₄, F₃, F₂, and F₁ are as follows:

F₄: **Trend of activation**

Ascending trend

The individual often feels a need to expend more or greater energy than he or she currently expends. Even though an energy release may be high, the person may want to increase it. Energy expenditure can also be too low and thus cause discomfort. The person feels insufficiently stimulated and challenged and therefore strives to make use of more energy. The manners in which persons attempt to increase energy expenditure may differ – myriad forms of stimulation may be sought to increase the intensity level of experience. The individual may also feel he or she has an increased influence on his/her environment and other people. None of this is crucial, though. Likewise it is not important whether the individual actually manages to increase the desired energy expenditure. It is only essential that a need exists for increased activation of the individual's energy reserves.

Descending trend

The individual feels a need to reduce energy expenditure. The current level of energy being expended is too high. The individual may seek to reduce an unpleasant excess of stress and tenseness. Possibly, and although the person may be expending little energy anyway, he/she may want to reach an even better state of rest that would further improve the energetic condition. It is not critical in which way the individual attempts to achieve a reduction of energy expenditure and whether these attempts are met by success. The only aspect of importance is the presence of a desire to reduce energy manifestations.

F₃: Sources of regulation

Endodynamic regulation (self or internal regulation)

The behaviour of an individual follows a volitional programme. The person makes decisions and seeks to reach various goals, wanting to influence and change external conditions. It is not important which activities the person initiates and whether they are focused reflexively on the person him- or herself or on external objectives. The actions should merely be initiated and executed by the person's ego and not by others or circumstances. The acting individual should also not be urged by their own uncontrolled emotional drives and should remain free from internal constraints and compulsions.

Exodynamic regulation (external or uncontrollable internal regulation)

The individual's behaviour is initiated by other people or environmental factors. The person reacts to external stimulations, triggers, or seductions. Self-determined decisions are rare. The person is unwilling or unable to manage his or her environment. Instead, his or her behaviour predominantly consists of reactions to external challenges or to inner impulses and constraints. Possibly, the individual cannot control him- or herself well despite being eager to do so. At any rate, the individual is considerably dependent on external or internal will-restricting conditions.

F₂: Mode of representation

Endomodal representation

The individual tends to prefer subjective views. Thinking about things and experiencing them is more important than the things themselves, feelings are more effective than what is transmitted by sensory channels. This kind of subjectivity may have positive or negative effects. An openness to feeling qualities might result as well as the neglect of objective and sensually discernible realities.

Exomodal representation

The individual tends to prefer objective views. Experiencing things is less important than things as they are, feelings are less important than what is transmitted via sensory channels. This kind of objectivity may have positive or negative effects. A precaution against subjective errors might result as well as a neglect of beneficial subjective contributions to reality as a whole. The individual's inclination to cling to what is factual might be inflated at the cost of fruitful subjectivity.

F₁: Status of functionality

Eufunctional or balanced status

The individual is consistent with him- or herself; varying directions of thinking, experience, and aspirations are in harmony. Energy expenditures, however strong they may be, are aligned with one another having balanced relationships. Tensions may be present, but they do not strain the overall system. Should major conflicts occur, they do not survive for long and are soon resolved.

Dysfunctional

The individual is at odds with him- or herself. The person's thinking, experience, and behaviour are insufficiently aligned and not well adjusted; they are not in harmony and interfere with each other. The relationship of activated energies is unbalanced. Ambivalences may cause tensions and conflicts which tend to endure and are only overcome slowly, if at all.

References

- Acton, G. S. & Revelle, W. (2004). Evaluation of ten psychometric criteria for circumplex structure. *Methods of Psychological Research Online*, 9 (1), 1-29.
- Amthauer, R. (1970). IST70. *Intelligenz-Struktur-Test 70*. Göttingen: Hogrefe.
- Amthauer, R., Brocke, B., Liepmann, D., & Beauducel, A. (2001). *I-S-T-2000-R. Intelligenz-Struktur-Test 2000-R*. Göttingen: Hogrefe.
- Amthauer, R., Brocke, B., Liepmann, D., & Beauducel, A. (1999). *I-S-T-2000. Intelligenz-Struktur-Test 2000*. Göttingen: Hogrefe.
- Andresen, B. (1998). *Persönlichkeitsstruktur und Psychosetendenz, Bd. 1: Grundlagen, Methoden, Basisfaktoren*. Westerau: Psychometrie- & Psychodiagnostik-Verlag.
- Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin*, 110, 486-498.
- Asendorpf, J. B. & van Aken, M. A. (1999). Resilient, overcontrolled, and undercontrolled personality prototypes in childhood. Replicability, predictive power, and the trait-type issue. *Journal of Personality and Social Psychology*, 77, 815-832.
- Åstrand, P.-O. (1956). Human physical fitness with special reference to sex and age. *Physiological Reviews*, 36, 307-335.
- Barrett, P. T. (2003). Beyond psychometrics: Measurement, non-quantitative structure, and applied numerics. *Journal of Managerial Psychology*, 18, 421-439.

- Barrett, P. T. (2005). *Validity and utility in I/O psychology*. Auckland I/O and HRINZ SIG meeting at Auckland Town Hall, New Zealand, 10 May 2005.
- Basilevsky, A. (1994). *Statistical factor analysis and related methods. Theory and applications*. New York: Wiley.
- Bates, E., Thal, D., & Janowsky, J. (1992). Early language development and its neural correlates. In I. Rapin & S. Segalowitz (Eds.), *Handbook of neuropsychology* (Vol. 7, pp. 69-110). New York: Elsevier Science Publishers.
- Baumeister, R. F. & Vohs, K. D. (2004). *Handbook of self-regulation*. New York: Guilford.
- Beauducel, A., Strobel, A., & Brocke, B. (2003). Psychometrische Eigenschaften und Normen einer deutschsprachigen Fassung der Sensation Seeking-Skalen, Form V. *Diagnostica*, 49, 61-72.
- Beauducel, A. & Wittmann, W.W. (2005). Simulation study on fit indexes in CFA based on data with slightly distorted simple structure. *Structural Equation Modeling*, 12, 41-75.
- Bechtel, W. & Richardson, R. C. (1993). *Discovering complexity: Decomposition and localization as scientific research strategies*. Princeton NJ: Princeton University Press.
- Bem, S.L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review*, 88, 354-364.
- Berneyer, G. (1957). Psychological factors. Their number, nature, and identification. *The British Journal of Statistical Psychology*, 10, 17-27.
- Bierhoff, H. W. (2000). Skala der sozialen Verantwortung nach Berkowitz und Daniels: Entwicklung und Validierung. *Diagnostica*, 46, 18-28.
- Blinkhorn, S. F. (1997). Past imperfect, future conditional: Fifty years of test theory. *British Journal of Mathematical and Statistical Psychology*, 50, 175-185.
- Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, 117, 187-215.
- Boekaerts, M., Pintrich, P. R., & Zeidner, M. (Eds), (2000). *Handbook of self-regulation*. San Diego: Academic Press.
- Borsboom, B., Mellenbergh, G.J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, 110, 203-219.
- Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika*, 71, 425-440.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. Heidelberg: Springer, 2005 (1st edition 1977).
- Bös, K. (1987). *Handbuch sportmotorischer Tests*. Göttingen: Hogrefe.
- Bös, K. & Mechling, H. (1984). Dimensionen sportmotorischer Leistungen als Grundlage für einen fähigkeitsorientierten Ansatz. In G. Hagedorn, H. Karl, & K. Bös (Red.). *Handeln im Sport*. Sechster Sportwissenschaftlicher Hochschultag. Clausthal-Zellerfeld.

- Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science*, 16, 199-231.
- Briggs, S. R. (1989). The optimal level of measurement for personality constructs. In D. M. Buss & N. Cantor (Eds.), *Personality Psychology. Recent trends and emerging directions* (pp. 246-260). New York: Springer.
- Brocke, B., Beauducel, A., & Tasche, K. (1998). Der Intelligenz-Struktur-Test: Analysen zur theoretischen Grundlage und technischen Güte. *Diagnostica*, 44, 84-99.
- Burroughs, G. E. R. & Miller, H. W. L. (1961). The rotation of principal components. *The British Journal of Statistical Psychology*, 14, 35-50.
- Burt, C. (1949). The structure of the mind: A review of the results of factor analysis. *The British Journal of Educational Psychology*, 19, 100-111, 176-199.
- Burt, C. (1954). The sign pattern of factor-matrices. *The British Journal of Statistical Psychology*, 7, 15-29.
- Büsch, D., Hagemann, N., & Thielke, S. (2001). Gibt es eine Ballfähigkeit. Interpretation eines faktorenanalytischen Ergebnisses. *Psychologie und Sport*, 8 (2), 57-66.
- Buse, L. & Pawlik, K. (1978). Stichprobenfehler und Rotationsambiguität von Faktorladungen. *Zeitschrift für experimentelle und angewandte Psychologie*, 25, 10-34.
- Butler, J. M. (1969). Simple structure reconsidered: Distinguish ability and invariance in factor analysis. *Multivariate Behavioral Research*, 4, 5-28.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Carl, W. (1968). Eine Untersuchung zur Faktorenstruktur von Antworttendenzen ('response sets') bei Antwortskalen verschiedener Stufenzahl. *Zeitschrift für experimentelle und angewandte Psychologie*, 15, 419-434.
- Carroll, J. B. (1993). *Human cognitive abilities. A survey of factor-analytic studies*. Cambridge: Cambridge University Press.
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology* 54, 1-22.
- Cattell, R. B. (1965). Factor analysis: An introduction to essentials. II: The role of factor analysis in research. *Biometrics*, 21, 405-435.
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. New York: Houghton Mifflin.
- Cattell, R. B. & Dickman, K. (1962). A dynamic model of physical influences demonstrating the necessity of oblique simple structure. *Psychological Bulletin*, 59, 389-400.
- Cattell, R. B. & Horn, J. L. (1963). A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *Journal of Educational Measurement*, 15, 139-164.

- Cattell, R. B. & Radcliffe, J. A. (1962). Reliabilities and validities of simple and extended weighted and buffered unifactor scales. *The British Journal of Statistical Psychology*, 15, 113-128.
- Child, D. (2006). *The essentials of factor analysis*. 3rd edition. London: Cassell, 2006 (1st edition: 1971).
- Church, A. T. & Burke, P. J. (1994). Exploratory and confirmatory tests of the big five and Tellegen's three- and four-dimensional models. *Journal of Personality and Social Psychology*, 66, 93-114.
- Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research*, 18, 115-126.
- Cliff, N. (1992). Abstract measurement theory and the revolution that never happened. *Psychological Science*, 3, 186-190.
- Comrey, A. L. (1978). Common methodological problems in factor analytic studies. *Journal of Consulting and Clinical Psychology*, 46, 648-659.
- Comrey, A. L. & Lee, H. B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Lawrence Erlbaum, 2nd edition. (1st edition Comrey alone: 1973, New York: Academic Press).
- Crawford, J. R. & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 43, 245-265.
- Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
- Cronmeyer, I. (1983). Analyse der Intelligenz- und Persönlichkeitsstruktur von Abiturienten der neu gestalteten Gymnasialen Oberstufe in Nordrhein-Westfalen. Eine Untersuchung über die Beziehungen zwischen der Abitur-durchschnittsnote und dem „Intelligenz-Struktur-Test“ von Amthauer sowie dem „Freiburger-Persönlichkeits-Inventar“. Dissertation. Erziehungswissenschaftliche Fakultät Universität Köln.
- Davies, W. K. D. (1971). Varimax and the destruction of generality: A methodological note. *Area*, 3, 112-118.
- De Charms, R. (1968). *Personal causation: The internal affective determinants of behavior*. New York: Academic Press.
- De Raad, B., Hendriks, A. A. J., & Hofstee, W. K. B. (1994). The Big Five: A tip of the iceberg of individual differences. In C. J. Halverson, Jr., G. A. Kohnstamm, and R. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 91-109). Hillsdale, NJ: Erlbaum.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Diener, E. & Emmons, R. A. (1983). The independence of positive and negative affect. *Journal of Personality and Social Psychology*, 47, 1105-1117.

- Duffy, E. (1962). *Activation and behavior*. New York: Wiley.
- Edmonds, B. (2002). *Simplicity is not truth-indicative*. In C. Gershenson, D. Aerts, & B. Edmonds (Eds.). *Philosophy and complexity* (pp. 65-80). Singapore: World Scientific.
- Epstein, S. (2003). Cognitive-experiential self theory of personality. In Millon, T. & Lerner, M. J. (Eds.). *Comprehensive handbook of psychology*. Vol. 5. *Personality and Social Psychology* (pp. 159-184). Hoboken, NJ: Wiley & Sons.
- Ertel, S. (2009a). Kritik am Dogma „Simple Structure“. *Psychologie des Alltagshandelns* 2 (1), 36-63.
- Ertel, S. (2009b). Latente Strukturen aufspüren. *Faktorenanalysen mit Varimaxrotation*. *Psychologie des Alltagshandelns* 2 (2), 33-60.
- Ertel, S. (2010). Grundintelligenz (*g*) und Lernkapital (*h*). Komplexstrukturelle Faktorenanalyse des I-S-T. *Psychologie des Alltagshandelns* 3 (2), 49-61.
- Ertel, S. (2011a). *Komplexität modellieren. Faktorenanalyse am Scheideweg*. Universitätsverlag Göttingen.
- Ertel, S. (2011b). *Basiskomponenten der Persönlichkeit*. Universitätsverlag Göttingen.
- Ertel, S. (2011c). Exploratory factor analysis revealing complex structure. *Personality and Individual Differences*, 50, 196-200.
- Estes, W. K. (1975). Some targets for Mathematical Psychology. *Journal of Mathematical Psychology*, 12, 263-282.
- Eysenck, H. J. (1973). Personality and the law of effect. In D. E. Berlyne & K. B. Madsen (eds.), *Pleasure, reward, and preference* (pp. 133-168). New York: Academic Press.
- Eysenck, H. J. (1990). *Rebel with a Cause: The Autobiography of Hans Eysenck*. New Brunswick: Transaction Publishers, 1997.
- Eysenck, H. J. (1992). Four ways five factors are *not* basic. *Personality and Individual Differences*, 13, 667-673.
- Eysenck, H. J. (1997). Personality and experimental psychology: The unification of psychology and the possibility of a paradigm. *Journal of Personality and Social Psychology*, 73, 1224 - 1237.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods* 4, 272-299.
- Feger, H. (1979). Einstellungsstruktur und Einstellungsänderung: Ergebnisse, Probleme und ein Komponentenmodell der Einstellungsobjekte. *Zeitschrift für Sozialpsychologie*, 10, 331-349.
- Feger, H. & Brehm, M. (2001). *New developments in feature pattern analysis*. Lengerich: Pabst Science Publishers.

- Ferrando, P. J. & Lorenzo-Seva, U. (2000). Unrestricted versus restricted factor analysis of multidimensional test items: Some aspects of the problem and some suggestions. *Psicológica*, 2000, 21, 301-323.
- Finch, J. F. & West, S. G. (1997). The investigation of personality structure: Statistical models. *Journal of Research in Personality*, 31, 439-485.
- Fischer, G. H. (1967). Zum Problem der Interpretation faktorenanalytischer Ergebnisse. *Psychologische Beiträge*, 10, 122-135.
- Fischer, H. (1958). Ein Vergleich zwischen dem IST von Amthauer und dem PMA von Thurstone. *Diagnostica*, 4, 25-32.
- Fittkau, B. (1968). Die Mehrdeutigkeit der Varimax-Rotation und einige sich daraus ergebende Folgerungen. *Archiv für die gesamte Psychologie*, 120, 106-114.
- Fleck, L. (1935). *Entstehung und Entwicklung einer wissenschaftlichen Tatsache. Einführung in die Lehre vom Denkstil und Denkkollektiv*. Basel: Schwabe.
- Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, 39, 291-314.
- Forster, M. (1998). Parsimony and simplicity. Available from: <http://philosophy.wisc.edu/forster/220/simplicity.html>. Supply date 2011
- Forster, M. & Sober, E. (1994). How to tell when simple, more unified, or less ad hoc theories will provide more accurate predictions. *The British Journal for the Philosophy of Science*, 45, 1-35.
- Funder, D. C. (2001). Personality. *Annual Review of Psychology*, 52, 197-121.
- Gangestad, S. & Snyder, M. (1985). "To carve nature at its joints": On the existence of discrete classes in personality. *Psychological Review*, 92, 317-349.
- Gawronski, B. (2000). Falsifikationismus und Holismus in der experimentellen Psychologie: Logische Grundlagen und methodologische Konsequenzen. *Zeitschrift für Sozialpsychologie*, 31, 3-17.
- Gibbons, R. D. & Hedeker, D. R. (1992). Full-information item bi-factor analysis. *Psychometrika*, 57, 423-436.
- Gibson, J. J. (1963). The useful dimensions of sensitivity. *American Psychologist*, 18, 1-15.
- Gigerenzer, G. (1978). Artefakte in der dimensionsanalytischen Erfassung von Urteilsstrukturen. *Zeitschrift für Sozialpsychologie*, 9, 110-116.
- Gigerenzer, G. (2004). Mindless statistics. *Journal of Socio-Economics*, 33, 587-606.
- Gigerenzer, G. & Strube, G. (1978). Zur Revision der üblichen Anwendung dimensionsanalytischer Verfahren. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 10, 75-86.
- Glass, G. V. (1991). Eulogy for Henry Kaiser (1927-1992). *Multivariate Behavioral Research Monographs*, 27, 159-171.

- Glass, G. V. & Taylor, P. A. (1966). Factor analytic methodology. *Review of Educational Research*, 36, 566-587.
- Goldberg, L. R. (1981). Developing a taxonomy of trait-descriptive terms. In D. W. Fiske (Ed.), *Problems with language imprecision: New directions for methodology of social and behavioral science*, No. 9 (pp. 43-65). San Francisco: Jossey-Bass.
- Goldberg, L. R. (1990). An alternative "Description of Personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59, 1216-1229.
- Gorsuch, R. L. (1983). *Factor analysis*, 2nd edition. Hillsdale NJ: Lawrence Erlbaum. (1st edition, 1974).
- Gould, S. J. (2002). *The structure of evolutionary theory*. Cambridge MA: Harvard University Press-Belknap Press.
- Greif, S. (1972). *Gruppenintelligenztests – Untersuchungen am WTT, IST, LPS und AIT*. Bern, Frankfurt: Peter Lang.
- Groner, R. & Groner, M. T. (1991). Heuristische versus algorithmische Orientierung als Dimension des individuellen kognitiven Stils. In K. Grawe, R. Hänni, N. Semmer & F. Tschann (Eds.), *Über die richtige Art, Psychologie zu betreiben*. Göttingen: Hogrefe.
- Guilford, J. P. (1952). When not to factor analyze. *Psychological Bulletin*, 49, 26-37.
- Guilford, J. P. (1958). A system of the psychomotor abilities. *American Journal of Psychology*, 71, 164-174.
- Guilford, J. P. & Hoepfner, R. (1969). Comparison of varimax rotations with rotations to theoretical targets. *Educational and Psychological Measurement*, 29, 3-23.
- Guilford, J. P. & Zimmerman, W. S. (1963). Some variable-sampling problems in the rotation of axes in factor analysis. *Psychological Bulletin*, 60, 289-301.
- Guilford, J. P. (1974). Rotation problems in factor analysis. *Psychological Bulletin*, 81, 498-501.
- Guttman, L. (1955). The determinacy of factor score matrices with implications for five other basic problems of common-factor theory. *British Journal of Statistical Psychology*, 8, 65-81.
- Guttman, L. (1971). Measurement as structural theory. *Psychometrika*, 36, 329-347.
- Guttman, L. (1992). The irrelevance of factor analysis for the study of group differences. *Multivariate Behavioral Research*, 27, 175-204.
- Halberg, F. (1980). Closing remarks: Chronobiology and health. In L. E. Scheving & F. Halberg (Eds.), *Chronobiology: Principles and applications to shifts in schedules* (pp. 541-562). Alphen aan den Rijn: Sijthoff and Noordhoff.
- Hamilton, M. (1958). An experimental approach to the identification of factors. *The British Journal of Statistical Psychology*, 11, 161-169.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry*, 23, 56-62.

- Harman, H. H. (1967). *Modern factor analysis*. Chicago: The University of Chicago Press.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Herrmann, T. (1965). Psychologie der kognitiven Ordnung. Series: C. F. Graumann & J. Linschoten (Eds.), *Phänomenologisch-Psychologische Forschungen*. (Vol. 6). Berlin: Walter de Gruyter & Co.
- Herzberg, P. Y. (2002). Zur psychometrischen Optimierung einer Reaktanzskala mittels klassischer und IRT-basierter Analysemethoden. *Diagnostica*, 48, 163-171.
- Highmore, G. & Taylor, W. R. (1954). A factorial analysis of athletic ability. *British Journal of Statistical Psychology*, 7, 1-8.
- Hinz, A., Brähler, E., Geyer, M., & Körner, A. (2003). Urteileffekte beim NEO-FFI. *Diagnostica*, 49, 157-163.
- Hofstee, W. K. B., de Raad, B., & Goldberg, L. R. (1992). Integration of the Big Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology*, 63, 146-163.
- Höger, D. (1964). Analyse der Intelligenzstruktur bei männlichen Gymnasiasten der Klassen 6-9 (Untersekunda-Oberprima). *Psychologische Forschung*, 27, 419-474.
- Holling, H., Preckel, F., & Vock, M. (2004). *Intelligenzdiagnostik. Kompendien Psychologische Diagnostik*. Band 6. Göttingen: Hogrefe.
- Holz-Ebeling, F. (1995). Faktorenanalysen und was dann? Zur Frage der Validität von Dimensionsinterpretationen. *Psychologische Rundschau*, 46, 18-35.
- Holzinger, K. J. & Swineford, F. (1939). A study of factor analysis: The stability of a bi-factor solution. *Supplementary Educational Monographs* 48. Chicago: University of Chicago, Department of Education.
- Holzman, P. S. & Klein, G. S. (1954). Cognitive system-principles of leveling and sharpening: Individual differences in assimilation effects in visual time-error. *The Journal of Psychology: Interdisciplinary and Applied*, 37, 105-122.
- Horn, J. L. (1967). On subjectivity in factor analysis. *Educational and Psychological Measurement*, 27, 811-820.
- Horn, J. L. (1976). Human abilities: A review of research and theory in the early 1970s. *Annual Review of Psychology*, 27, 437-485.
- Humphreys, L. G. & Davey, T. C. (1988). Continuity in intellectual growth from 12 months to 9 years. *Intelligence*, 12, 183-197.
- Ismail, A. H., Falls, H. B., & MacLeod, D. F. (1965). Development of a criterion for physical fitness tests from factor analysis results. *Journal of Applied Physiology*, 20, 991-999.

- Jäger, A. O. & Hörmann, H.-J. (1981). Demonstrationen von „g“ (der Allgemeinen Intelligenz) und zur Bedeutung des Variablekontextes bei exploratorischen Faktorenanalysen. *Psychologische Beiträge*, 23, 408-420.
- Jakobson, R. & Halle, M. (1956). *Fundamentals of language*. The Hague: Mouton.
- Jensen, A. R. (1998). *The g factor. The science of mental ability*. Westport CT: Praeger.
- Johnson, W. & Bouchard, T. J. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, 33, 393-416.
- Jöreskog, K. G. & Sörbom, D. (2003) LISREL 8.52. *Website*: www.ssicentral.com.
- Jung, C. G. (1930). *Psychologische Typen*. Zürich: Rascher.
- Just M. A. & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Boston: Allyn and Bacon.
- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.
- Kaiser, H. F. (1970). A second generation Little Jiffy. *Psychometrika*, 35, 401-415.
- Kallina, H. (1967). Das Unbehagen in der Faktorenanalyse. *Psychologische Beiträge*, 10, 81-86.
- Kalveram, K. T. (1970). Über Faktorenanalyse. Kritik eines theoretischen Konzepts und seine mathematische Neuformulierung. *Archiv für Psychologie*, 122, 92-118.
- Kanizsa, G. (1975). „Prägnanz“ as an obstacle to problem solving. *Giornale Italiano di Psicologia*, 2, 417-425.
- Kaplunovsky, A. S. (2007). Why using factor analysis? (dedicated to the centenary of factor analysis). HAIT Journal of Science and Engineering. Prepublication draft: <http://www.magniel.com/fa/kaplunovsky.html>.
- Kazdin, A. E. (2006). Arbitrary metrics: Implications for identifying evidence-based treatments. *American Psychologist*, 61, 42-49.
- Koch, S. (1999). *Psychology in human context. Essays in dissidence and reconstruction*. Chicago: The University of Chicago Press.
- Köhler, T. & Troester, U. (1991). Changes in the palmar sweat index during mental arithmetic. *Biological Psychology*, 32, 143-154.
- Kunz, H. (1980). *Leistungsbestimmende Faktoren im Zehnkampf. Eine Längsschnittstudie an Schweizer Spitzenathleten*. Dissertation an der Sozial- und verhaltenswissenschaftlichen Fakultät der Universität Heidelberg.
- Lakatos, I. (1978). *The methodology of scientific research programmes*. Philosophical Papers Volume 1. Cambridge: Cambridge University Press.

- Lambert, M. J., Hannöver, W., Nisslmüller, K., Richard, M., & Kordy, H. (2002). Fragebogen zum Ergebnis von Psychotherapie: Zur Reliabilität und Validität der deutschen Übersetzung des Outcome Questionnaire 45.2 (OQ-45.2). *Zeitschrift für Klinische Psychologie und Psychotherapie*, 31, 40-46.
- Langner, E. & Olbrich, M. (1972). Untersuchungen zur faktoriellen Struktur des I-S-T von Amthauer – mit einem Vergleich der Ergebnisse der Intelligenztests der Eignungs- und Verwendungsprüfung. *Wehrpsychologische Untersuchungen*, 3, 7-73.
- Lawlor, D. A., Ebrahim, S., May, M., & Smith, G. D. (2004). (Mis)use of factor analysis in the study on Insulin Resistance Syndrome. *American Journal of Epidemiology*, 159, 1013-1018.
- Lenk, W. (1983) *Faktorenanalyse: Ein Mythos? Historische und konzeptionelle Untersuchungen zur Faktorenanalyse und Intelligenzforschung*. Weinheim: Beltz.
- Leventhal, D. S. & Stedman, D. J. (1970). A factor analytic study of the Illinois test of psycholinguistic abilities. *Journal of Clinical Psychology*, 26, 473-477.
- Lewicki, P. (2005). Internal and external encoding style and social motivation. In J. P. Forgas, K. D. Williams, & S. M. Laham (Eds.). *Social motivation: Conscious and unconscious processes* (pp. 194-209). Cambridge: Cambridge University Press.
- Lienert, G. A. (1969). *Testaufbau und Testanalyse*. Dritte, durch einen Anhang über Faktorenanalyse ergänzte Auflage. Weinheim: Beltz.
- Linden, M. (1977). Factor analytical study of Olympic decathlon data. *Research Quarterly*, 48, 562-568.
- Lindsley, D. B. (1951). Emotion. In S. S. Stevens (Ed.), *Handbook of Experimental Psychology*. New York: Wiley, 473-516.
- Lukesch, H. & Kleiter, G. D. (1974). Die Anwendung der Faktorenanalyse. Darstellung und Kritik der Praxis einer Methode. *Archiv für Psychologie*, 126, 265-307.
- Lumsden, J. (1961). The construction of unidimensional tests. *Psychological Bulletin*, 58, 122-131.
- Lykken, D. T. (1991). What's wrong with psychology anyway? In D. Cicchetti & M. Grove (Eds.). *Thinking clearly about psychology. Volume 1: Matters of public interest* (pp. 3-39). Minneapolis: University of Minnesota Press.
- MacCallum, R. C. (2003). Working with imperfect models. *Multivariate Behavioral Research*, 38, 113-139.
- MacCallum, R. C. & Austin, J. T. (2000). Application of structural equation modeling in psychological research. *Annual Review of Psychology*, 51, 201-226.
- Manning, J. M., Dooly-Manning, C., & Perrin, D. H. (1988). Factor analysis of various anaerobic power tests. *Journal of Sports Medicine and Physical Fitness*, 28, 138-144.

- Marshall, G. N., Wortman, C. B., Kusulas, J. W., Hervig, L. K., & Vickers, R. R. (1992). Distinguishing optimism from pessimism: Relations to fundamental dimensions of mood and personality. *Journal of Personality and Social Psychology*, 62, 1067-1074.
- Marx, A. & Hejj, A. (1989). *Subjektive Strukturen. Ergebnisse aus der Gedächtnis-, Sprach- und Einstellungsforschung*. Göttingen: Hogrefe.
- Matson, J. L. & Nieminen, G. S. (1987). Validity of measures of conduct disorder, depression, and anxiety. *Journal of Clinical Child Psychology*, 16, 151-157.
- Maturana, H. R. (1998). *Biologie der Realität*. Frankfurt: Suhrkamp.
- Maxwell, A. E. (1959). Statistical methods in factor analysis. *Psychological Bulletin*, 56, 228-235.
- Maxwell, A. E. (1972). Factor analysis: Thomson's sampling theory recalled. *British Journal of Statistical Psychology*, 25, 1-21.
- Meili, R. (1969). Die Zwiespältigkeit faktorieller Untersuchungen der Intelligenzstruktur. *Schweizerische Zeitschrift für Psychologie*, 16, 278-294.
- Metzger, W. (1941). *Psychologie. Die Entwicklung ihrer Grundannahmen seit der Einführung des Experiments*. Darmstadt: Steinkopf (Letzte Auflage 2002, Wien: Krammer).
- Michell, J. (1997). Quantitative science and the definition of measurement in psychology. *British Journal of Psychology*, 88, 355-383.
- Michell, J. (2000). Normal science, pathological science and psychometrics. *Theory and Psychology*, 10, 639-667.
- Milhorn, H. T. (1982). Cardiovascular fitness. *American Family Physician*, 26, 163-169.
- Moosbrugger, H. & Hartig, J. (2002). Factor analysis in personality research: Some artefacts and their consequences for psychological assessment. *Psychologische Beiträge*, 44, 136-158.
- Mor, N. & Winquist, J. (2002). Self-focused attention and negative affect: A meta-analysis. *Psychological Bulletin*, 128, 638-662.
- Mulaik, S. (1972). *The foundations of factor analysis*. New York: McGraw Hill.
- Murray, H. A. (1938). *Explorations in personality*. New York: Oxford University Press.
- Narens, L. & Luce, R. D. (1993). Further comments on the "nonrevolution" arising from axiomatic measurement theory. *Psychological Science*, 4, 127-130.
- Neuhaus, J. O. & Wrigley, C. (1954). The quartimax method: An analytic approach to orthogonal simple structure. *British Journal of Statistical Psychology*, 7, 81-91.
- Nunnally, J. C. & Bernstein, I. H. (1994). *Psychometric Theory*. (3rd edition). New York: Mc Graw Hill.

- Orlik, P. (1967). Das Dilemma der Faktorenanalyse - Zeichen einer Aufbaukrise in der modernen Psychologie. Sonderheft: Münchener Symposium über Faktorenanalyse. *Psychologische Beiträge*, 10, 87-98.
- Ostendorf, F. & Angleitner, A. (2004). *NEO-PI-R. NEO-Persönlichkeitsinventar nach Costa und McCrae, Form S*: Revidierte Fassung. Göttingen: Hogrefe.
- Overall, J. E. (1964). Note on the scientific status of factors. *Psychological Bulletin* 61(4), 270-276.
- Ozer, D. J. & Reise, S. P. (1994). Personality assessment. *Annual Review of Psychology*, 45, 357-388.
- Pawlik, K. (1977). Faktorenanalytische Persönlichkeitsforschung. In G. Strube (Ed.), *Die Psychologie des 20. Jahrhunderts. Binet und die Folgen* (Vol. 5, pp. 617-712). Zürich: Kindler.
- Peel, E. A. (1953). Factorial analysis as a psychological technique. *Nordic Psychology Monographs*, Vol. 3, 7-22.
- Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Beverly Hills, CA: Sage.
- Pöhlmann, R., Kirchner, G., & Wohlgefahr, K. (1979). Der psychomotorische Fähigkeitskomplex – seine Kennzeichnung und seine Vervollkommenung. *Theorie und Praxis der Körperkultur*, 28, 898-907.
- Rasch, D. (1962). Die Faktorenanalyse und ihre Anwendung in der Tierzucht. *Biometrische Zeitschrift*, 4, 15-39.
- Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, 12, 287-297.
- Rescher, N. (1985). *Die Grenzen der Wissenschaft*. Stuttgart: Reclam.
- Rescher, N. (1998). *Complexity. A philosophical overview*. New Brunswick: Transaction Publishers.
- Revelle, W. (1983). Factors are fictions, and other comments on individuality theory. *Journal of Personality*, 51, 707-714.
- Revenstorf, D. (1976). *Lehrbuch der Faktorenanalyse*. Stuttgart: Kohlhammer.
- Revenstorf, D. (1980). *Vom unsinnigen Aufwand in der Faktorenanalyse*. *Archiv für Psychologie*, 13, 1-36.
- Reyburn, H. A. & Raath, M. J. (1949). Simple structure: A critical examination. *British Journal of Statistical Psychology*, 2, 125-133.
- Reyment, R. A. & Jöreskog, K. G. (1993). *Applied factor analysis in the natural sciences*. 2nd edition. Cambridge: Cambridge University Press.
- Romney, A. K. & d'Andrade, R. G. (1964). Cognitive aspects of English kin terms. *American Anthropologist* 66 (3), 146-170.

- Rost, J. (2002). Zeitgeist und Moden empirischer Analysemethoden. In J. van Deth (Ed.), *Von Generation zu Generation*. ZUMA Spezial Bd. 8 (pp. 21-31). Mannheim: ZUMA.
- Rozeboom, W. W. (1991). Theory and practice of analytic hyperplane optimization. *Multivariate Behavioral Research*, 26, 179-197.
- Rummel, R. J. (1970). *Applied factor analysis*. Evanston IL: Northwestern University Press.
- Rutkowski, W. (1974). *Typologien und Schichtenlehren. Bibliographie des internationalen Schrifttums bis 1970*. Amsterdam: Rodopi.
- Saunders, D. R. (1962). Trans-varimax: Some properties of the ratiomax and equamax criteria for blind orthogonal rotation. *American Psychologist*, 17, 395-396.
- Schaper, N. & Baumgart, C. (2002). TKI-Teamklima-Inventar. *Zeitschrift für Personalpsychologie*, 1, 154-158.
- Schmid, J. & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, 22, 53-61.
- Schmidt-Atzert, L. (1997). Replik zur Rezension des IST 70. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 18, 109-112.
- Schmidt-Atzert, L. & Hommers, W. (1996). IST 70: Weitere Daten – weitere Erkenntnisse? Anmerkungen zu Gerd Heydes Beitrag „Weitere Daten zum IST-70“. *Report Psychologie*, 21, 526-528.
- Schmidt-Atzert, L., Hommers, W., & Hess, M. (1995). Der I-S-T 70: Eine Analyse und Neubewertung. *Diagnostica*, 41, 108-130.
- Schmitt, M. (2000). Vorschlag zur Vereinfachung des Beck-Depressions-Inventars (BDI). *Diagnostica*, 46, 38-46.
- Schönemann, P. H. (1994). Book review: Louis Guttman on Theory and Methodology: Selected writings. *Applied Psychological Measurement*, 18, 293-297.
- Schönemann, P. H. (1996). The psychopathology of factor indeterminacy. *Multivariate Behavioral Research*, 31, 571-577.
- Schönemann, P. H. & Borg, I. (1981). Measurement, scaling, and factor analysis. In I. Borg (Ed.), *Multidimensional data representations: When and why* (pp. 380-419). Ann Arbor MI: Mathesis Press.
- Schönemann, P. H. & Borg, I. (1996). Von der Faktorenanalyse zu den Strukturgleichungsmodellen. In E. Erdfelder, R. Mausfeld, T. Meiser, & G. Rudinger (Eds.): *Quantitative Psychologische Methoden – Ein Handbuch in Schlüsselbegriffen* (pp. 241-252). Weinheim: Psychologie Verlags Union.
- Schönemann, P. H. & Steiger, J. H. (1976). Regression component analysis. *British Journal of Mathematical and Statistical Psychology*, 29, 175-189.
- Schönemann, P. H. & Wang, M. M. (1972). Some new results on factor indeterminacy. *Psychometrika* 37(1), 61-91.

- Sechrest, L. B. (2002). Paper without available title presented at the 11th European Conference on Personality. Universität Jena, 21.7.-25.7.2002.
- Shanon, B. (1993). *The representational and the presentational. An essay on cognition and the study of mind*. New York: Harvester.
- Singer, W. (2003). *Ein neues Menschenbild? Gespräche über Hirnforschung*. Frankfurt a. M.: Suhrkamp
- Smith, P. T. & Jones, K. F. (1975). Some hierarchical scaling methods for confusion matrix analysis. II. Applications to large matrices. *British Journal of Mathematical and Statistical Psychology*, 28, 30-45.
- Snyder, M. & Gangestad, S. (1986). On the nature of self-monitoring: Matters of assessment, matters of validity. *Journal of Personality and Social Psychology*, 51, 125-139.
- Sponsel, R. (1998). Axiome und Konstruktionsprinzipien Differentieller Psychologie der Persönlichkeit in der Allgemeinen und Integrativen Psychologie, Psychodiagnostik, Psychopathologie und Psychotherapie. *Internet Publikation für Allgemeine und Integrative Psychotherapie IP-GIPT*. Erlangen: <http://www.sgipr.org/gipt/diffpsy/giptp10.htm>.
- Steiger, J. H. (1994). Factor analysis in the 1980s and the 1990s: Some old debates and some new developments. In I. Borg & P. P. Mohler (Eds.) *Trends and perspectives in empirical social research* (pp. 201-223). New York: De Gruyter.
- Steiger, J. H. & Schönemann, P.H. (1975). A history of factor indeterminacy. Unpublished manuscript. Purdue University, West Lafayette, Indiana
- Stephenson, W. (1956). Methodology of trait analysis. *The British Journal of Psychology*, 47, 5-18.
- Suppes, P. & Zinnes, J. L. (1963). Basic measurement theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.) *Handbook of Mathematical Psychology*. Vol. 1 (pp. 1-76). New York: Wiley.
- Szopa, J., Chwała, W., & Ruchlewicz, T. (1998). Investigations on structure of "energetic" motor abilities and validity of their testing. *Antropomotoryka*, 17 (Retrieved from the web).
- Teipel, D. (1988). *Diagnostik koordinativer Fähigkeiten: Eine Studie zur Struktur und querschnittlich betrachteten Entwicklung fein- und grobmotorischer Leistungen*. München: Profil-Verlag.
- Thalbourne, M. A. (1998). Transliminality: Further correlates and a short measure. *The Journal of the American Society for Psychical Research*, 92, 402-429.
- Thayer, R. E. (1985). Activation (arousal): The shift from a single to a multidimensional perspective. In J. Strelau, F. H. Farley & A. Gale (Eds.), *The biological results on factor indeterminacy*. *Psychometrika*, 37, 61-91.
- Thompson, J. W. (1962). Meaningful and unmeaningful rotation of factors. *Psychological Bulletin*, 59, 211-223.

- Thompson, J. W. (1963). Bi-polar and unidirectional scales. *British Journal of Psychology*, 54, 15-24.
- Thurstone, L. L. (1934). The vectors of the mind. *Psychological Review*, 41, 1-32.
- Thurstone, L. L. (1935). *The vectors of the mind*. Chicago: The University of Chicago Press.
- Thurstone, L. L. (1947). *Multiple factor analysis. A development and expansion of The Vectors of the Mind*. Chicago: The University of Chicago Press.
- Tidow, G. (2000). *Challenge decathlon – barriers on the way to becoming the “king of athletes”*. Part I. NSA, 15, 243-256.
- Tinsley, H. E. A. & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of Counseling Psychology*, 34, 414-424.
- Tucker, L.R. (1951). A method for synthesis of factor analysis studies (Personnel Research Section Report No. 984). Washington, DC: U.S. Department of the Army.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327-352.
- Überla, K. (1971). *Faktorenanalyse. Eine systematische Einführung für Psychologen, Mediziner, Wirtschafts- und Sozialwissenschaftler*. 2. Auflage, Berlin: Springer. 1st edition 1968, last edition 1977.
- Undheim, J. O. & Gustafsson, J.-E. (1987). The hierarchical organization of cognitive abilities. Restoring general intelligence through the use of linear structural relations (LISREL). *Multivariate Behavioral Research*, 22, 149-171.
- Velicer, W. F. (1977). An empirical comparison of the similarity of principal component, image, and factor patterns. *Multivariate Behavioral Research*, 12, 3-22.
- Velicer, W. F. & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, 25, 1-28.
- Vittadini, G., 1989. Indeterminacy problems in the LISREL model. *Multivariate Behavioral Research*, 24, 397-414.
- Vukovich, A. (1967). Faktorielle Typenbestimmung. Sonderheft: Münchener Symposium über Faktorenanalyse. *Psychologische Beiträge*, 10, 112-121.
- Waldmann, M. R., Renkl, A., & Gruber, H. (2003). Das Dreieck von Begabung, Wissen und Lernen. In W. Schneider & M. Knopf. (Eds.), *Entwicklung, Lehren und Lernen* (pp. 219-23). Göttingen: Hogrefe.
- Warburton, F. W. (1963). Analytic methods of factor rotation. *British Journal of Statistical Psychology*, 16, 165-174.
- Watson, D. & Clark, L. A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology*, 54, 1063-1070.

- Watson, D. & Clark, L. A. (1997). Extraversion and its positive emotional core. In R. Hogan, J. Johnson & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 767-793). San Diego: Academic Press.
- Weber, E. (1978). Einführung in die Faktorenanalyse: mit einem Anhang zur Matrizenrechnung. Jena: Fischer (1st edition, 1974).
- Weinert, F. E. (Ed.) (1996). *Psychologie des Lernens und der Instruktion*. Göttingen: Hogrefe.
- Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology*, 37, 395-412.
- Witkin, H. A. & Goodenough, D. R. (1977). Field dependence and interpersonal behavior. *Psychological Bulletin*, 84, 651-689.
- Witte, W. (1974). Untersuchungen zur Behinderung des Denkens durch Anschauung. *Psychologische Beiträge*, 16, 277-287.
- Wrigley, C.S., & Neuhaus, J.O. (1955). The matching of two sets of factors. *American Psychologist*, 10, 418-419.
- Yates, A. (1987). Multivariate exploratory data analysis: A perspective on exploratory factor analysis. New York: State University of New York Press.
- Zarnowski, F. (1989). The decathlon. A colorful history of track and field's most challenging event. Champaign IL, Leisure Press.
- Zimmermann, W. S. (1953). A note on the recognition and interpretation of composite factors. *Psychological Bulletin*, 50, 387-389.
- Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking*. Cambridge: Cambridge University Press.

Exploratory factor analysis (EFA) is a statistical tool for digging out hidden factors which give rise to the diversity of manifest objectives in psychology, medicine and other sciences. EFA had its heyday as psychologist Leon Thurstone (1935 and 1948) based EFA on what he called the “principle of simple structure” (SS). This principle, however, was erroneous from the beginning what remained unrecognized despite subsequent inventions of more sophisticated statistical tools such as confirmatory analysis and structural equation modeling. These methods are highly recommended today as tolerable routes to model complexities of observation. But they did not remove the harmful errors that SS had left behind. Five chapters in this book demonstrate and explain the trouble. In chapter 2 the ailment of SS is healed by introducing an unconventional factor rotation, called Varimin. Varimin gives variables of an analysis an optimal opportunity to manifest functional interrelations underlying correlational observations. Ten applications of Varimin (in chapter 2) show that its results are superior to results obtained by the conventional Varimax procedure. Further applications are presented for sports achievements (chapter 3), intelligence (chapter 4), and personality (chapter 5). If Varimin keeps on standing the tests new theoretical building blocks will arise together with conceptual networks promoting a better understanding of the domains under study. Readers may check this prognosis by themselves using the statistical tool (Varimin) which is provided by open access in the internet.

