

Komplexe Strukturen aufspüren: Faktorenanalyse mit Variminrotation

Suithbert Ertel

Georg-August-Universität Göttingen / Georg-Elias-Müller-Institut für Psychologie

ZUSAMMENFASSUNG

Der vorliegende Artikel knüpft an einer Kritik am ‚Dogma Simple Structure‘ an (Ertel, 2009). Ein neues Rotationsverfahren für extrahierte Faktoren wird vorgestellt (Varimin), welches die latenten Kovarianzquellen ermitteln will, die den manifesten kovariierenden Variablen einer Domäne zugrunde liegen. Varimin optimiert das durch die Initialstruktur einer Faktorenanalyse schon entworfene, aber oft noch korrekturbedürftige Modell multifunktionaler Bedingungsverhältnisse. Die konventionell verwendeten Rotationsverfahren, mit denen die latenten Kovarianzquellen als Einfachstruktur modelliert werden sollen, voran das Standardverfahren Varimax, werden bemängelt, weil sie die zugrunde liegende reale Komplexität der latenten Bedingungsverhältnisse verdecken, anstatt sie aufzuklären. Die neue methodologische Ausrichtung wirft Fragen auf, von denen fünf diskutiert werden. Gefragt wird: (1) Wie lassen sich Varimin-transformierte Faktoren deuten? (2) Wird die Komplexität der Kovarianzquellen nicht zumeist mit der Initiallösung schon hinreichend repräsentiert? (3) Lassen sich faktorielle Simple Structure-Lösungen nicht doch befriedigend interpretieren, wie könnte man sie sonst ständig verwenden? (4) Enthält die regelmäßig anzutreffende Bipolarität von Varimin-Faktoren substantielle Information? (5) Sind mithilfe einer Komplexstruktur-orientierten Faktorenanalyse auch methodenbedingte Einflüsse erfassbar? Die Klärung dieser Fragen wird durch Anwendung der Varimin-Transformation auf acht publizierte und zwei neu generierte Datensätze unterstützt. Die Transformation zur Komplexstruktur, mit welcher latente Kovarianzquellen der manifesten Variablen aufgespürt werden (durch Varimin), wird der Transformation zur Einfachstruktur (durch Varimax) gegenüber gestellt. Es zeigt sich, dass Varimax zur Aufklärung latenter Bedingungen keinen nennenswerten Beitrag leistet. Lediglich zum Clustern manifester Variablen sind Simple Structure-Transformationen geeignet. Auf die Grenzen der methodologischen Innovation wird hingewiesen.

Schlüsselwörter

Faktorenanalyse – Rotation – Varimax – Varimin – Einfachstruktur – Komplexstruktur

ABSTRACT

Revealing complex structures. Factor analyses with Varimin rotation.

This paper examines a continuation of the critique of „simple structure“ (Ertel, 2009) in which the rotation procedure „Varimax“ which is commonly used to generate simple structure is replaced with „Varimin“. Varimin aims at manifesting latent complex structures. It optimises the model of complexity which, although delineated by initial unrotated structures, is further improved. The new method raises a number of issues of which five are discussed at length. (1) How can varimin factors be interpreted? (2) Do latent sources of covariance not appear sufficiently complex with initial solutions? (3) Are simple structure solutions in common practice not adequately interpretable? (4) How to interpret the commonly encountered bipolarity of Varimin factor loadings? (5) Is factor analysis with complex structure transformation applicable to data affected by method factors? Ten empirical applications of Varimin transformation serve as practical examples. Particular features of transformation to complex structure, revealing latent sources of covariance (by Varimin), are elucidated by comparing pertinent results with those obtained from transformations to simple structure (by Varimax). Varimax will remain useful for clustering objectives. Attention is also drawn to limitations of the new methodical innovation.

Keywords

Factor analysis – factor rotation – varimax – varimin – simple structure – complex structure

1 Einleitung

Der vorliegende Artikel ist die Fortsetzung eines früheren (Ertel, 2009), in dem eine Kritik an der exploratorischen Faktorenanalyse begründet wurde. Es wurde dargelegt, dass das Hauptproblem der Faktorenanalyse, die geometrische Rotation bzw. algebraische Transformation extrahierter Faktoren, durch Thurstones Modell der Einfachstruktur nicht gelöst, sondern verschärft wurde. Die manifesten Variablen einer empirischen Domäne, so wurde argumentiert, resultieren generell aus *multiplen* Varianz- bzw. Kovarianzquellen. Latente Bedingungen, die die einzelnen Variablen hervorrufen, bilden keine Einfachstruktur, sie bilden fast ausnahmslos multifaktorielle Strukturen. Die monofaktorielle Bedingtheit manifesten individueller Variablen, die mit dem Prinzip der Einfachstruktur vorausgesetzt wird, darf nicht generell vorausgesetzt werden. Das Prinzip wurde als fehlerhaft verworfen.

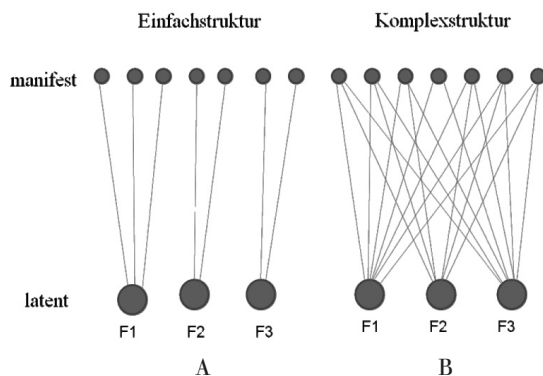


Abbildung 1: Einfachstruktur (A) und Komplexstruktur (B) des Verhältnisses zwischen latenten Kovarianzquellen und manifesten Variablen

Die Abbildungen 1A und 1B veranschaulichen den Unterschied zwischen Einfach- und Komplexstruktur. In beiden Fällen ist auszugehen von der Ebene manifesten Variablen, die interkorreliert werden mit der Absicht, durch Faktorisierung der Korrelationsmatrix die zugrunde liegenden Faktoren aufzuklären. Faktoren sind die Quellen der Kovarianz, die bei den untersuchten manifesten Variablen ermittelt wird. Das einfachstrukturelle Modell Thurstones unterscheidet sich vom hier befürworteten komplexstrukturellen Modell in Folgendem: Eine Rotation der Faktoren zur Einfachstruktur (Abbildung 1A) ordnet einzelnen manifesten Variablen jeweils nur einen extrahierten Faktor zu bzw. so wenig Faktoren wie möglich. Eine Rotation zur Komplexstruktur dagegen (Abbildung 1B) ist darauf aus, individuelle Variablen mit möglichst vielen extrahierten Faktoren hypothetisch zu verknüpfen.

Nach dieser Konzeption wird angenommen, dass die meisten Faktoren in der Regel zur Kovarianz einzelner Variablen einen – je nach Variable verschieden großen – Beitrag leisten.

Unter den Rotationsverfahren, die zur Einfachstruktur führen sollen, werden im Forschungsalltag meist die orthogonalen gewählt, oblique Verfahren nur selten, und von den orthogonalen Verfahren wird fast ausschließlich Kaisers Varimax-Rotation eingesetzt (s. Abbildung 2).

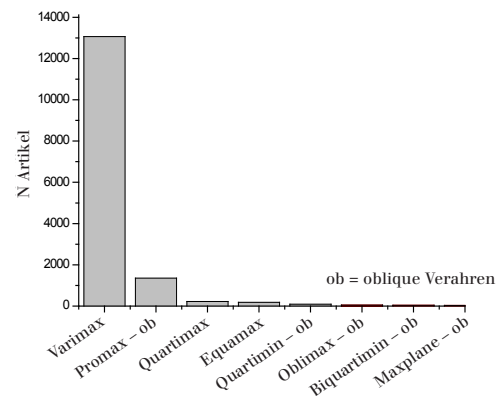


Abbildung 2: Anzahl von Artikeln, in denen auf Faktorenrotation der Einfachstruktur verwiesen wird. Ergebnis einer Internet-Suche nach Rotationsbezeichnungen in SCIENCE DIRECT (Elsevier, 2008)

Nach den Überlegungen des vorangehenden Artikels sind die Rotationsverfahren, die eine Einfachstruktur erreichen wollen (Abbildung 2), durchweg fehlerhaft. Das Modell der Einfachstruktur verhüllt die in der Initiallösung oft schon erkennbare Komplexität latenter Bedingungen, anstatt diese herauszuarbeiten. Dabei ist anzumerken, dass in der Initiallösung ein befriedigendes Komplexitätsmodell nur beim Sonderfall zweifaktorieller Lösungen zu finden ist. Für Datensätze mit mehr als zwei Faktoren wird mit der Initiallösung die Komplexität der Kovarianzquellen individueller Variablen in der Regel nicht optimal manifestiert. Der Grund ist, dass bei der Extraktion eines Faktors die Varianz, die die Matrix der Korrelationen – bzw. nach einer Faktorenextraktion die jeweilige Restmatrix – enthält, für den betreffenden Faktor überoptimal ausgeschöpft wird. Dies führt zu einem mehr oder weniger großen Überhang an Kommunalität bei den früher extrahierten Faktoren und zu einer entsprechenden Benachteiligung der später extrahierten Faktoren (Gorsuch, 1974, p. 161).¹ Zur optimalen Manifestation der Komplexität der Kovarianzquellen wurde deshalb eine Faktorenrotation auch für notwendig gehalten. Doch soll diese das Gegenteil von dem leisten, was eine Rotation zur Einfachstruktur zum Ergebnis hat.

Aus diesem Grunde wurde *Varimin* konzipiert, ein neues Rotationsverfahren des Verfassers. Varimin ist die Umkehrform von Varimax. Mit Varimax wird durch paarweises Drehen der Faktor-Koordinaten die Varianz der quadrierten Faktorladungen pro Faktor *gesteigert*. Der Vorgang wird durch Iterationen so lange wiederholt, bis sich die Summe der Ladungsvarianzen für die Gesamtheit der Faktoren nicht mehr steigern lässt. Das Kriterium V (s. Gleichung 1), das Varimax-Kriterium, wird maximiert.

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

h = Kommunalität eines Faktors

b = Ladungshöhe eines Faktors

p = Laufindex für Faktoren von 1 bis m

j = Laufindex für Variablen von 1 bis n

Für das *Varimin*-Verfahren werden die Koordinaten-Drehungen mit dem Ziel einer Varianz-*Verminderung* vorgenommen, iterativ so lange, bis die Summe der Ladungsquadrate, das Kriterium V, sich nicht mehr vermindern lässt.

Angenommen eine initiale Struktur der Abbildung 3A liege vor. Die Anwendung von Varimax transformiert die Struktur 3A in die Struktur 3C. In 3C laufen die Koordinaten durch die Punktwolken, für die einzelnen Faktoren wird die Summe ihrer Ladungsquadrate maximiert. Die Anwendung von Varimin transformiert die Struktur 3A in die Struktur 3B. Der Abstand der Variablenwolken von den Koordinaten wird so weit wie möglich vergrößert. Die Summe der Ladungsquadrate wird damit für die Faktoren minimiert.

Mit der Einführung von Varimin zur Faktorentransformation stellen sich neue Verfahrensfragen. Im Folgen-

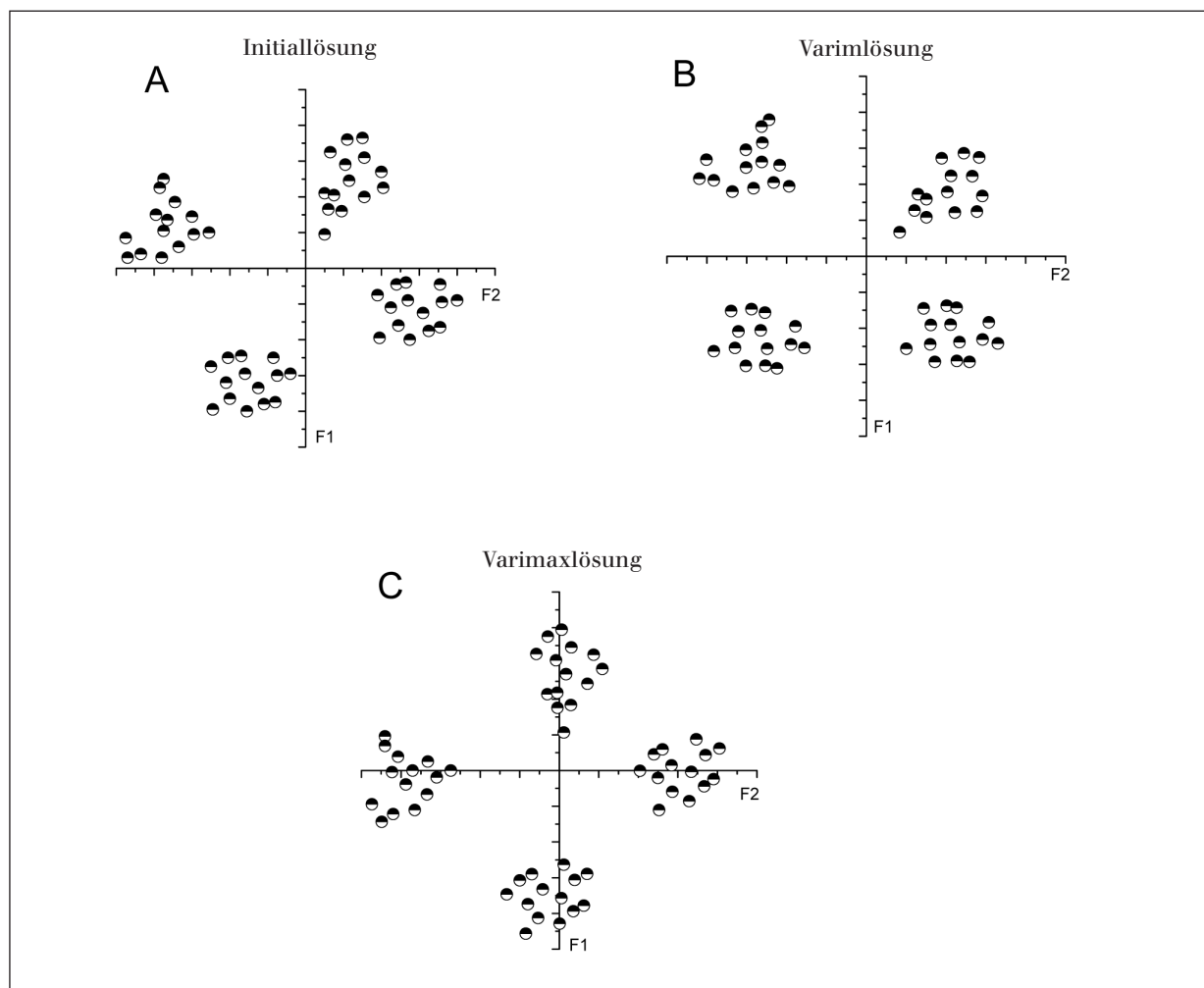


Abbildung 3: Zweifaktorielle Initiallösung mit fingierten Variablen (A), die Varimin-rotiert (B) und Varimax-rotiert wurden (C)

den werden die fünf dringendsten aufgegriffen, wobei empirische Lösungsbeispiele herangezogen werden, die zur Klärung beitragen sollen.

2 Fragen und empirische Belege

Frage I

Wie lassen sich Varimin-transformierte Faktoren deuten?

Die Rotation einer Faktorenlösung zur Einfachstruktur wird herkömmlicherweise deshalb für notwendig gehalten, weil die initiale Faktorenlösung komplex ist und Faktoren angeblich nicht gedeutet werden können, wenn die manifesten Variablen, auf die man sich bei der Faktorendeutung stützen muss, mehrfaktoriell geladen sind (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).

Die Behauptung, dass faktorielle Komplexität der Variablen eine Deutung der Faktoren verhindere oder erschwere, lässt sich jedoch nicht aufrecht erhalten, wenn man das Konzept der ‚distinctive features‘ methodisch nutzt. Das Konzept wurde in der Phonologie von Jacobson & Halle (1956) entworfen und fand in dieser Disziplin als ein gewichtiger methodischer Fortschritt weithin Anerkennung. In generalisierter Form ist seine Anwendung auf andere als phonologische und linguistische Inhalte übertragbar.

Die Details: Phoneme lassen sich mithilfe von ‚features‘ (Merkmale) unterscheiden, sofern diese eine distinktive Funktion für die untersuchte Sprache besitzen. Für jedes Phonem hat man auf dem Wege der Selbstbeobachtung und der objektiven Messung drei Klassen von Merkmalen als funktional relevant erkannt, welche die Artikulation der Phoneme spezifizieren: die Artikulationsdauer (kurz vs. lang)², die Sonorität oder Stimmbeitragsung (stimmhaft, stimmlos) und den Artikulationsort (bilabial, labio-dental usw. bis uvular). Hier lässt sich die Begrifflichkeit der Faktorenanalyse anwenden, denn man kann sagen: Phoneme sind manifeste Einheiten, distinktive Merkmale sind die zugrunde liegenden (‚latenten‘) Varianzquellen. Jedes Phonem ist durch drei Varianzquellen zu beschreiben, zum Beispiel ist [b] ein kurz artikulierendes (plosives), ein stimmhaftes und bilabiales (mit den Lippen gebildetes) Phonem.

Zu fragen bleibt, wie z. B. beim Phonem [b] die drei Merkmale methodisch aufgespürt werden können. Dies geschieht durch den *Minimalpaar-Vergleich* (*minimal pair comparison*). Zur Paarbildung für [b] ist eine ähnliche andere Einheit der gleichen Domäne mit heranzuziehen. Unter den Phonemen der deutschen Sprache findet sich auch das [p], das mit dem [b] zwei der oben genannten Merkmale teilt (Artikulationsort

und -dauer). Nur hinsichtlich der Stimmhaftigkeit gibt es einen markanten Unterschied. Wenn die Phoneme der deutschen Sprache als Variablen z. B. für eine Beurteilung ihrer Ähnlichkeiten/Verschiedenheiten verwendet werden, dann sollte eine Faktorenanalyse der Phonem-Interkorrelationen, welche Ähnlichkeitsmaße darstellen, mit anschließender Variminrotation Faktoren erkennen lassen, die den bekannten drei distinktiven Merkmalkategorien entsprechen. Die Phoneme [b] und [p] könnten z. B. bei den Faktoren F_1 und F_2 gleiche Ladungen aufweisen, bei F_3 aber müssten die Ladungen markant auseinander treten. Die Aufgabe des Faktorenanalytikers bestünde lediglich darin, den ‚minimalen‘ Unterschied zwischen [b] und [p] für F_3 zu deuten: F_3 wäre dann als Effekt der Stimmhaftigkeit/Stimmlosigkeit anzusehen. Die Deutung würde sich sichern lassen, wenn weitere minimale Paare aus dem Datensatz gebildet werden, z. B. die Paare [d] vs. [t] sowie [g] vs. [k], bei denen der in gleicher Weise zu deutende minimale Unterschied für F_3 auftritt wie bei [b] und [p], während für F_1 und F_2 bei den Paaren [d] vs. [t] sowie [g] vs. [k] wieder unterschiedslose Ausprägungen vorliegen würden, obgleich für diese übrigen Phonempaare beliebige andere F_1 und F_2 -Ausprägungen vorliegen dürfen (nicht müssen) als für [b] und [p].⁵

Empirischer Beitrag 1:

Beurteilung von Phonem-Ähnlichkeiten

Diese Überlegungen wurden für ein Experiment verwendet, an dem zwei deutschsprachige Studierende teilnahmen, eine Psychologiestudentin mit offensichtlicher Sprachbegabung und eine Linguistikstudentin höheren Semesters, zu deren Ausbildung Phonologie gehörte. Die Probandinnen hatten 10 Phoneme der deutschen Sprache ([b], [d], [f], [g], [k], [m], [n], [p], [t], [v] auf bipolaren Likert-Skalen mit sieben Stufen (Mittelstufe = 0) nach Ähnlichkeit zu beurteilen ([v] wird ausgesprochen wie „w“ in z. B. ‚wenn‘). Als Skalenpole wurden die zu beurteilenden 10 Phoneme paarweise in allen Kombinationen (45 Kombinationen) verwendet, die zufällig aufeinander folgten, z. B. waren die Phoneme auf den Skalen

[d] 3---2---1---0---1---2---3 [m]

[b] 3---2---1---0---1---2---3 [d]

und weiteren 43 Skalen nach Ähnlichkeit einzustufen. Wenn das zu beurteilende Phonem gleichzeitig einen Skalenpol definiert, läuft das auf ein Identitätsurteil hinaus, z. B. ist [d] auf der Skala

[d] 3—2—1—0—1—2—3 [m] dem linken Skalenpol maximal ‚ähnlich‘ (mit ihm identisch). In solchen Fällen war laut Instruktion immer die Stufe 3 (maximale Ähnlichkeit) anzukreuzen.

Die zehn Urteilsprofile mit je 45 Urteilen wurden sodann für jede Probandin individuell interkorreliert

Tabelle 1: Korrelationen der Faktorladungen mit objektiven Rangwerten sowie die Tucker-Kongruenzen der Faktoren von zwei Probandinnen

Faktoren		Korrelationen der Faktorladungen mit objektiven Rangwerten			Kongruenz der Faktoren der Vpn
		Vp Psych	Vp Lingui	beide Vpn	
F1	kurz-lang	.972	.990	.996	.943
F2	stimmhaft-stimmlos	.917	.957	.966	.936
F3	Artikulationsort	.671	.976	.922	.574

und einer Principal Component-Analyse (PCA) unterzogen. Die ersten drei extrahierten Faktoren (= Komponenten) wurden für die Varimin-Rotation berücksichtigt, da drei interpretierbare Faktoren erwartet wurden (Eigenwerte der ersten fünf Faktoren: 2.23, 1.58, 1.40, 1.12, 1.05 für Vp 1 und 4.86, 1.68, 1.09, 1.03, 0.59 für Vp2).

Die Güte der Varimin-Rotation wird danach bemessen, wie gut die Faktorladungen mit den zu erwartenden Klassifizierungen übereinstimmen. Nach phonologischer Klassifizierung haben die folgenden Phoneme eine kurze Artikulationsdauer: [b], [d], [g], [k], [p], [t] (man nennt sie auch ‚plosiv‘), die Phoneme [f], [m], [n], [v] werden mit anhaltendem Luftstrom, also länger artikuliert). Die Phoneme [b], [d], [g], [m], [n], [v] sind stimmhaft, stimmlos sind [f], [k], [p], [t]. Während sich nach Artikulationsdauer und Stimmbeteiligung eine bipolare Unterteilung ohne zusätzliche Differenzierung ergibt, liegen beim Artikulationsort vier Stufen vor, von der Artikulation ‚weit vorn‘ bis ‚weit hinten‘ sind aufzuführen (1) die bilabialen Phoneme [b], [m], [p], (2) die labiodentalen Phoneme [f], [v], (3) die alveolaren [t], [d], [n], und (4) die velaren [k], [g]. Die Ladungen der aus den Urteilsdaten der Probandinnen ermittelten Varimin-rotierten Faktoren werden sodann punktbiserial korreliert mit den linguistischen Variablen Artikulationsdauer (kurz = 1 und lang = 2) und Stimmhaftigkeit (stimmlos = 1 und stimmhaft = 2). Mit den vier ordinalen Stufen des Artikulationsortes (ganz vorn, 1, bis ganz hinten, 4) werden die Faktorladungen Produktmoment-korreliert.

Die Korrelationen zeigt Tabelle 1. Man sieht, dass mit Ausnahme der Korrelation für den Artikulationsort bei der Psychologie-Probandin ($r = .671$) die Korrelationen sehr hoch sind, sie liegen über .90. F_1 repräsentiert bei beiden Probandinnen die Artikulationsdauer, F_2 repräsentiert die Stimmhaftigkeit. Offenbar hat die Linguistik-Studentin aufgrund ihrer Ausbildung ein

feineres Empfinden erworben, was die Wahrnehmung des Artikulationsortes im Mund betrifft (die Tucker-Kongruenz für F_3 bei den beiden Versuchsteilnehmerinnen beträgt nur .574. Da aber die Ladungen bei F_3 für die Linguistik-Studentin wesentlich höher ausfallen als bei der Psychologie-Studentin, ergibt sich nach einer Mittelung der beiden Datensätze mithilfe von Fishers Z-Transformation auch für F_3 eine Korrelation mit dem Artikulationsort von $r = .922$, so dass weiterhin die Ergebnisse der vereinten beiden Datensätze für alle drei Faktoren dargestellt werden dürfen.

Abbildung 4 zeigt die Varimin-Ergebnisse. Positive Faktorladungen werden durch dunkle Kreise, negative durch hellere Kreise dargestellt, die variablen Kreisflächen entsprechen den variablen Ladungshöhen. Eine Null-Ladung würde als ausdehnungsloser Punkt wiedergegeben. Wäre die Deutung der drei Faktoren nicht schon durch die Korrelationen mit Expertenurteilen erfolgt, hätte man für F_1 die minimalen Paare [t] vs. [f] und [d] vs. [n] bilden können, wodurch die Artikulationsdauer als distinktives Merkmal in Erscheinung tritt. Bei F_2 wären die minimalen Paare ‚Stimmbeteiligung ja‘ vs. ‚Stimmbeteiligung nein‘ aufgefallen: [b] vs. [p], [d] vs. [t], [g] vs. [k] und [v] vs. [f]. Bei F_3 hätten die minimalen Paare [p] vs. [k] und [b] vs. [g] den vorn-hinten Kontrast des Artikulationsortes manifestiert.

Das Ergebnis der Varimax-Rotation der Faktoren wird ohne Abbildung mitgeteilt, das der Linguistik-Probandin auch nur, weil eine Deutung weniger schwierig ist als das der Psychologie-Probandin. Durch Varimax- F_1 werden bei der Linguistik-Probandin die *langen stimmhaften* Phoneme [m], [n] und [v], die hier ein positives Vorzeichen haben, geclustert, mit negativem Vorzeichen gruppieren sich bei F_1 die *kurzen stimmlosen* Phoneme [t] und [k]. F_2 gruppiert mit positivem Vorzeichen die *kurzen stimmhaften* Phoneme [g] und [d].

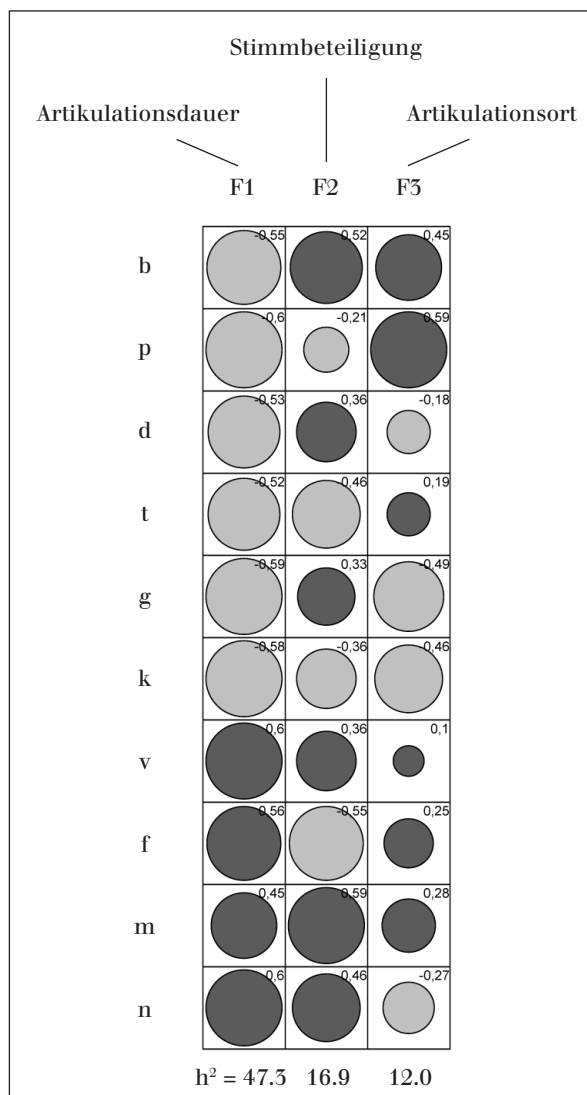


Abbildung 4: Varimin-transformierte Faktorladungen von zehn deutschen Phonemen (aufgrund von Ähnlichkeitsbeurteilungen)

Eine negative F_2 -Ladung hat das *lange stimmlose* [f]. Der unipolare Faktor F_3 gruppiert die bilabialen plosiven Phoneme [p] und [b]. Es ist offenkundig, dass die Varimax-Rotation zwar einander ähnliche Phoneme clustert (die mit Plus-Ladung) und auch die jeweils gegensätzlichen Cluster zusammen bringt (die mit Minus-Ladung). Doch an solchen Clustern, zu denen immer mehrere phonetische Merkmale beitragen, kann man kaum interessiert sein. Denn gesucht werden die latenten Quellen der Kovarianz. Bei der Herstellung von Einfachstrukturen, also Clustern, werden die Varianzquellen *nicht* aufgefunden, sie werden verschleiert.

Eine Deutung von Varimin-transformierten Faktoren ist demnach unproblematisch, sofern minimale Paare vorliegen. Die Deutung von Faktoren der Einfachstruktur-orientierten Faktorenanalysen ist in der Regel ungleich schwieriger, weil bei der Deutung dieser Faktoren *globale Ähnlichkeiten* unter den faktoriell geclusterten Variablen beschrieben werden müssen, wobei immer mehrere zugrunde liegende Merkmale zur Ähnlichkeit bzw. Verschiedenheit beitragen. Bei der Deutung von Komplexstruktur-orientierten Faktoren sind nur *Unterschiede* zwischen jeweils zwei gepaarten Variablen bezüglich *eines einzigen* Merkmals zu ermitteln. ‚Einfachheit‘ im methodischen Vorgehen kommt also – scheinbar paradoxerweise – eher der Complex Structure-Modellierung zu als der Simple Structure-Modellierung.

Frage II

Wird die Komplexität der Kovarianzquellen nicht schon durch die Initiallösung hinreichend repräsentiert?

Man kann bezweifeln, dass es nötig sein soll, die in der Initiallösung schon vorliegende Komplexität noch zu steigern. Man könnte einwenden, das dafür vorgebrachte Argument reiche nicht aus (das Argument war: realitätsverzerrende Bevorteilung der zuerst extrahierten Faktoren gegenüber später extrahierten). Die bislang vorgebrachte Kritik könne zwar das Konzept der Einfachstruktur schwächen, die Forscher würden auf den bislang vernachlässigten Informationswert der Initialstruktur aufmerksam werden, nicht jedoch würde eine Rotation der Initialfaktoren erforderlich sein mit dem Ziel, die Komplexität der initialen Lösung noch zu erhöhen.

Diesen Einwand würde man nur mit empirischen Prüfergebnissen entkräften können. Die Ergebnisse Varimin-rotierter Lösungen, die sich von den Ergebnissen initialer Lösungen unterscheiden, müssten sich als durchweg leichter interpretierbar erweisen, d.h. als befriedigender integrierbar in den jeweils vorliegenden Wissenskontext, im Vergleich mit den initialen Lösungen, also nicht nur im Vergleich mit Varimax-rotierten Lösungen. Die relative ‚Modelliergüte‘ der Lösungen wäre für die Entscheidung über die anstehende Frage letztendlich maßgebend.

Unter zahlreichen empirischen Belegen, bei denen die Modelliergüte einer Varimin-Rotation die der Initiallösung übertrifft, werden die Ergebnisse einer Analyse der Merkmale beurteilter Münzen ausgewählt.

Farbe	Pence	Größe	Form
nicht silbern	1	kleiner	rund
nicht silbern	2	größer	rund
silbern	5	kleiner	rund
silbern	10	größer	rund
silbern	20	kleiner	nicht rund
silbern	50	größer	nicht rund
Pound			
nicht silbern	1	kleiner	rund
nicht silbern	2	größer	rund

Abbildung 5: Das System der britischen Münzen mit den Eigenschaften Farbe, Größe und Form

Empirischer Beitrag 2:
Ähnlichkeitsbeurteilung britischer Münzen
Britische Münzen als Urteilsobjekte eignen sich als manifeste Variablen einer faktoriellen Analyse insbesondere deshalb, weil sie eindeutige Merkmale (Varianzquellen ihres phänomenalen Eindrucks) besitzen, nämlich Größe, Form und Farbe. Diese lassen sich mit physikalischen Reizeigenschaften in Verbindung bringen, ihre Feststellung ist nicht auf subjektive Urteile angewiesen (s. Abbildung 5).

Die nach Münzenwert benachbarten Münzen sind größenverschieden (z.B. 1 Pence klein, 2 Pence groß, 5 Pence klein, 10 Pence groß usw.). Dadurch ergeben

sich natürliche Paare, die sich wiederum von anderen Paaren entweder durch die Metallfarbe (5 und 10 Pence z.B. sind silbern, 1 und 2 Pence sind nicht silbern) oder durch ihre Form voneinander unterscheiden (z.B. sind 5 und 10 Pence rund, 20 und 50 Pence sind siebeneckig). Diese Merkmale liegen der Kovarianz der Ähnlichkeitsurteile über die untersuchten Objekte zugrunde. Durchmesser und das Gewicht der Münzen sind aus Tabelle 2 ersichtlich.

Experiment: Eine in Cambridge studierende deutsche Studentin wurde gebeten, die acht Münzen der gegenwärtigen britischen Währung nach Ähnlichkeit einzustufen (1, 2, 5, 10, 20, 50 Pence und 1 und 2 Pound). Die Münzen wurden auf Karton aufgeklebt und in Paaren dargeboten: Jede Münze kam mit jeder anderen Münze gepaart vor, angefangen mit den Paaren 1 Penny⁴ – 2 Pence, 1 Pence – 5 Pence usw. bis zum Paar 1 Pound – 2 Pound, insgesamt 28 Paare.

Die Aufgabe der Studentin bestand darin, jede der acht Münzen in die Hand zu nehmen und mit den aufgeklebten Münzen der Karton-Vorlage zu vergleichen. Auf einer bipolaren Siebenpunkt-Skala hatte sie anzugeben, ob die Münze in ihrer Hand eher der links oder der rechts aufgeklebten Münze ähnlich war. Zum Beispiel könnte ihr die 50 Pence-Münze eher der 5-Pence Münze als der 1 Pound-Münze ähnlich erscheinen, worauf sie dann auf der Skala 5 Pence – 1 Pound einen Skalenpunkt in der Nähe des 5 Pence-Pols ankreuzen würde. Ein ganzheitliches Urteil wurde erbeten, alle Ähnlichkeits- und Verschiedenheitsgründe sollten gleichzeitig beachtet werden. Vom finanziellen Münzenwert sollte abgesehen werden.

Auf diese Weise kam für jede Münze ein Urteilsprofil mit 28 Einzelurteilen zustande. Die Urteilsprofile der acht Münzen wurden interkorreliert, die Korrelationsmatrix wurde einer PCA unterzogen, die extrahierten Faktoren anschließend der Varimin- und der Varimax-Rotation zugeführt. Die Erwartung war, dass sich das Merkmalsystem der britischen Münzen nach einer Varimin-Rotation, nicht aber nach einer Varimax-Rotation, faktoriell durchsetzen würde, und dass das Varimin-rotierte Ergebnis Münzen-Merk-

Tabelle 2: Durchmesser und Gewicht der Münzen

	1 Penny	2 Pence	5 Pence	10 Pence	20 Pence	50 Pence	1 Pound	2 Pound
Gewicht (g)	3.28	7.10	3.52	6.50	5.00	8.20	9.48	11.92
Durchmesser (mm)	20	25	18	24	20	28	22	28

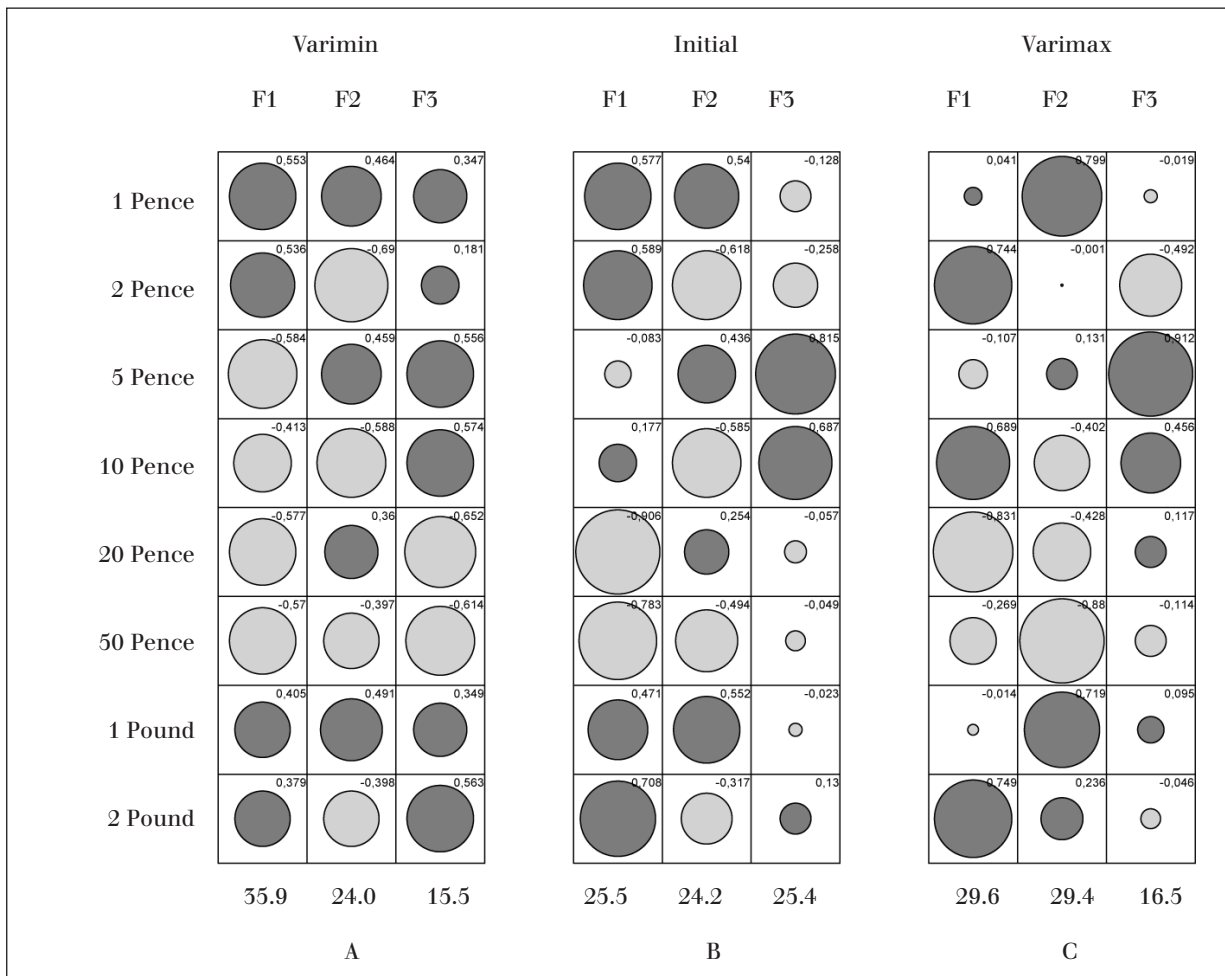


Abbildung 6: Varimin- (A), initiale (B) und Varimax-transformierte Lösung (C) mit Ähnlichkeitsdaten britischer Münzen

male besser hervortreten lassen würde als die initiale Lösung. Das Ergebnis zeigt Abbildung 6.

Das Varimin-Ergebnis fällt erwartungsentsprechend aus: Die Münzen von silberner Farbe heben sich auf dem ersten bipolaren Faktor von den Münzen mit nicht-silberner Farbe eindeutig ab. Dies drücken die Ladungsvorzeichen aus, die Ladungshöhen unterscheiden sich nur wenig. Dasselbe gilt für den zweiten Faktor, der offensichtlich die Größen-Unterschiede der Münzen wiedergibt. Im dritten Faktor schlägt sich offenbar der Varianzanteil der Münzenform nieder.⁵ Eine Deutung mithilfe minutiöser Minimalpaar-Vergleiche kann man sich bei der hier vorliegenden Transparenz der ‚latenten‘ Bedingungslage ersparen.

Zur Initiallösung: Unter den Faktoren der Initiallösung lässt sich F_2 als Manifestation der Münzengröße deuten. In F_1 könnte man den Farbfaktor vermuten, doch lassen sich die Ladungen von 5 und 10 Pence nicht mit der Münzenfarbe in Zusammenhang

bringen. Eine Deutung von F_3 als Formmerkmal ist ausgeschlossen, F_3 bleibt rätselhaft. Das hier deutlich gewordene defizitäre Ergebnis der Initiallösung führt mit vielen anderen nicht berichteten Ergebnissen aus anderen Untersuchungen zu der praktischen Schlussfolgerung, dass PCA-Faktoren grundsätzlich nach Varimin rotiert werden sollten, auch dann, wenn die Varimin-rotierte Lösung von der initialen nicht sehr abweichen sollte – was bei Faktor-Lösungen mit mehr als zwei Faktoren ohnehin selten vorkommt.

Zur Varimax-Lösung: Das Ideal einer einfaktoriellen Ladung wird nur etwa bei der Hälfte der Variablen erreicht. Die 10 Pence und 20 Pence-Münzen weichen vom solitären Ladungsideal erheblich ab. Multiple Faktorladungen sind in der Simple Structure-Praxis seit jeher bekannt, sie wurden als ein permanent unleidiges Problem hingenommen. Schwerer wiegt die Tatsache, dass Varimax sein Ziel, deskriptive Einfachheit auf der inhaltlichen Ebene zu erreichen, nicht erreicht. Weder Farbe, noch Form, noch Größe

werden durch die Varimax-Faktoren repräsentiert, nicht einmal in Andeutungen.

Um die Generalisierbarkeit des Ergebnisses zu prüfen, wurde das Experiment wiederholt mit acht Studierenden, die mit britischen Münzen keinen Umgang hatten. Die über die Versuchsteilnehmer gemittelten Urteilstwerte wurden genau so analysiert wie die individuellen Werte der Cambridge-Studentin (T. S.). Es zeigte sich, dass bei einem Eigenwertverlauf 2.07, 1.90, 1.30, 1.20 vier Faktoren extrahierbar waren, bei T. S. waren es nur drei Faktoren (Eigenwerte: 2.87, 1.92, 1.24, 0.97). Die faktoriellen Kongruenzen der zu vergleichenden Lösungen, von T. S. und Studenten, waren für F_1 (Farbfaktor) und F_2 (Größenfaktor) beachtlich, sie betrugen .981 bzw. .973. Doch fand sich für F_3 von T. S. weder bei F_3 noch bei F_4 der Studenten eine Spur von Kongruenz (.312 bzw. .224). Es stellte sich heraus, dass bei den Studenten die acht F_3 -Gewichte mit den acht Rangplätzen *des Geldwertes* der Münzen hoch korrelierten (.88), während für T. S. die höchste Korrelation mit dem Geldwert nur .14 betrug (bei F_1). Offensichtlich haben sich die Studenten bei ihren Ähnlichkeitsurteilen außer von Farbe und Größe noch vom Geldwert der Münzen beeinflussen lassen, während die Münzenform, die sich bei T. S. als F_3 manifestiert hatte, von ihnen unbeachtet blieb.⁶

Frage III

Lassen sich faktorielle Simple Structure-Lösungen nicht doch zumeist befriedigend interpretieren, wie könnte man sie sonst ständig verwenden?

Die Variablen der bisher berichteten Untersuchungen waren Objekte, deren wahrgenommene Merkmale den Ähnlichkeitsurteilen der Probanden zugrunde lagen. Für multivariate Untersuchungen dieser Art werden Faktorenanalysen selten verwendet. Man weiß, dass Faktorenanalysen mit Daten beurteilter Objekte unbefriedigende Ergebnisse liefern. Man weicht auf alternative Verfahren wie MDS (mehrdimensionale Skalierung) aus, mit denen beurteilte Objekte befriedigender dimensional geordnet werden können.

Ein Haupt-Einsatzgebiet für Faktorenanalysen in der Psychologie sind Urteilsdaten über Menschen (Selbst- und Fremdbeurteilungen), Urteile über Persönlichkeitseigenschaften, Verhaltensdispositionen, Einstellungen usw.. Die Varianz der dabei verwendeten Variablen, d. h. der für solche Erhebungen verwendeten Items, ergibt sich durch Messwiederholungen über urteilende Personen. Faktorenanalysen mit solchen Daten führen, wenn sie Simple Structure-orientiert sind, zur faktoriellen Gruppierung verbal-semanticischer Urteileinheiten. Bei diesen kann man sich zur Einschätzung der Modelliergüte nicht wie bei Objektbeurteilungen auf anschauliche Kriterien stützen. Semantische Urteileinheiten aus der Erleb-

nis- und Verhaltensdomäne sind weniger streng voneinander geschieden, sie lassen sich leicht kombinieren. Irgendein Sinn wird in den faktoriell gruppierten Variablen fast immer gefunden, das Ungenügen der Simple Structure-Modellierung solcher Variablen wird nicht erkannt. Zwar wird mitunter von signifikanten Korrelationen zwischen Faktorscores aus Selbstbeurteilungen und objektiven Verhaltensdaten berichtet, doch wird dem Faktorenanalytiker die konzeptuelle Arbeit, die er beim Ermitteln latenter Bedingungen der nach Simple Structure gruppierten Urteilsvariablen leisten muss, auch bei anschließender Nutzung von Außenkriterien nicht erspart.

In der folgenden Untersuchung werden Korrelationen von Variablen faktorisiert, bei denen Messwiederholungen über urteilende Personen eine Rolle spielen. Doch handelt es sich hier um inhaltsarme adverbale Variablen (*Formwörter*). Sie stammen aus einer Untersuchung von Carl (1968) über das Antwortstil-Verhalten. Mit den an diesen Daten gewonnenen Ergebnissen lässt sich aufzeigen, wie Simple Structure-Faktoren und ihre Deutungen entstehen. Zu diesem Zweck werden die mit denselben Daten ermittelten Variminfaktoren herangezogen.

Empirischer Beitrag 3:

Differenzierung von Antwortstilen bei der Beantwortung von Fragebögen (Daten von Carl)

Das Ziel Carls (1968) war die Ermittlung von Antwort-Tendenzen (response sets) bei der Verwendung von Stufenskalen. Er hatte von 100 Personen Fünf-Punkt-Likert-Einstufungen für 580 Items aus dem *Minnesota Multiphasic Personality Inventory* (MMPI) erhoben und für jede Person über alle Items, ohne Rücksicht auf deren Inhalt zu nehmen, die Anzahl der fünf Einstufungen summiert, die von ‚stark zustimmend‘ bis ‚stark ablehnend‘ reichten (die Zwischenstufen der Skalen waren nicht verbalisiert worden). Carl hatte ausgeschlossen, dass aufgrund der gegenseitigen Abhängigkeit der Stufen-Urteile Artefakt-Korrelationen vorkamen: Er ermittelte die Verwendungshäufigkeiten für die fünf Antwortstufen mit jeweils aus dem Gesamtpool gewonnenen Teilstichproben von Items, die inhaltlich parallelisiert worden waren. Aus der Interkorrelationsmatrix Carls, die auch Variablen aus Skalen mit mehr Stufen enthielt (er hatte z.B. auch Sieben-Punkt-Skalen verwendet), wurde für den vorliegenden Zweck die Matrix für das Format mit fünf Urteilsstufen separiert. Für diese wurde mithilfe einer PCA die initiale Faktorenstruktur ermittelt, die dann mit Varimin und zum Vergleich auch mit Varimax rotiert wurde.

Bei dieser Untersuchung sind mit fachlichem Vorwissen bestimmte Ergebnisse erwartbar. Denn bekannt ist die *Ja-Sage-Tendenz* (Akquieszenz), die ei-

nen Faktor hervorbringen müsste, bei dem die beiden zustimmenden Antwortstufen positiv und die beiden ablehnenden Antwortstufen negativ geladen sind. Bekannt ist auch die *Extremantwort-Tendenz*, die einen Faktor hervorbringen müsste, bei dem die beiden Extrem-Stufen der Zustimmung und Ablehnung positiv und die dazwischen liegenden Stufen negativ geladen sein müssten (evtl. nach Vorzeichenumkehr, die Polungsrichtung ist irrelevant). Mindestens diese beiden Faktoren also müsste eine Nachbehandlung der initialen Faktoren durch Varimin zutage fördern.

Das *Varimin-Ergebnis* zeigt Abbildung 7B. Man erkennt in F_3 den erwarteten Akquieszenz-Faktor und in F_1 den Extremantwort-Faktor. Allerdings ist ein weiterer substantieller Faktor F_2 vorhanden, der nicht erwartet wurde und der zur Deutung herausfordert. Man kann hier bei Herrmann (1965) anknüpfen, der die von ihm so genannte *Urteilsnuanciertheit* als Variante eines Antwortstils bei Fragebogen-Untersuchungen identifizierte. Demnach unterscheiden sich Personen beim Ausfüllen von Fragebögen auch darin, inwieweit sie über eine zustimmende vs. ablehnende Entscheidung hinausgehend zwischen ‚Zustimmung‘ und ‚starker Zustimmung‘ sowie zwischen ‚Ablehnung‘ und ‚starker Ablehnung‘ und auch zwischen ‚Zustimmung/Ablehnung‘ einerseits und ‚Unentschieden‘ bzw. ‚Weiß nicht‘ andererseits differenzieren. Diese Antwortstil-Variante wird in der Literatur seltener diskutiert (gelegentlich unter der Bezeichnung ‚Urteilsvariabilität‘, ‚response variance‘), vermutlich weil sie die Ergebnis-

se weniger verzerrt als die anderen Antwort-Tendenzen (Hinz et al., 2003). Der zur Klärung anstehende Faktor F_2 der Varimin-Analyse lässt sich m. E. risikolos als Ausdruck der ‚Urteilsnuanciertheit‘ deuten.

Sind die drei Response Sets schon in der *initialen* Faktorenstruktur zu finden (s. Grafik 7A)? Die initiale Faktorstruktur ist der Varimin-Faktorstruktur sehr ähnlich. Faktor F_1 ist in den beiden Lösungen nahezu identisch, also ist auch der *initiale* F_1 -Faktor als Ausdruck der Extremantwort-Tendenz zu deuten. Doch bei F_3 , Akquieszenz, ist die initiale Lösung weniger gut, da hier die mittlere Kategorie der Urteils-Enthaltung eine beachtliche negative Ladung aufweist, während doch einer Urteils-Enthaltung andere Bedingungen zugrunde liegen sollten als einem Urteil, das eine Ablehnung beinhaltet. Auch für F_2 ist die Varimin-Lösung prägnanter. Bei der initialen Lösung wird bei F_2 zwischen ‚Ablehnung‘ und ‚starke Ablehnung‘ nicht differenziert. Auch ist die numerische Differenz zwischen ‚Zustimmung‘ und ‚starke Zustimmung‘ in der initialen Lösung deutlich schwächer als bei der Varimin-Lösung. Von Urteilsnuanciertheit kann man beim initialen F_2 -Faktor kaum sprechen.

Unser Hauptinteresse aber gilt hier dem *Varimax-Ergebnis* (Grafik 7C). Wie lässt es sich interpretieren? Darüber gibt ein Vergleich der Varimax- mit der Varimin-Lösung (vgl. Abbildung 7C mit 7B) Aufschluss. Die Varimax-Rotation hat zur bipolaren Clusterung der Variablen ‚starke Zustimmung‘ (positive Ladung

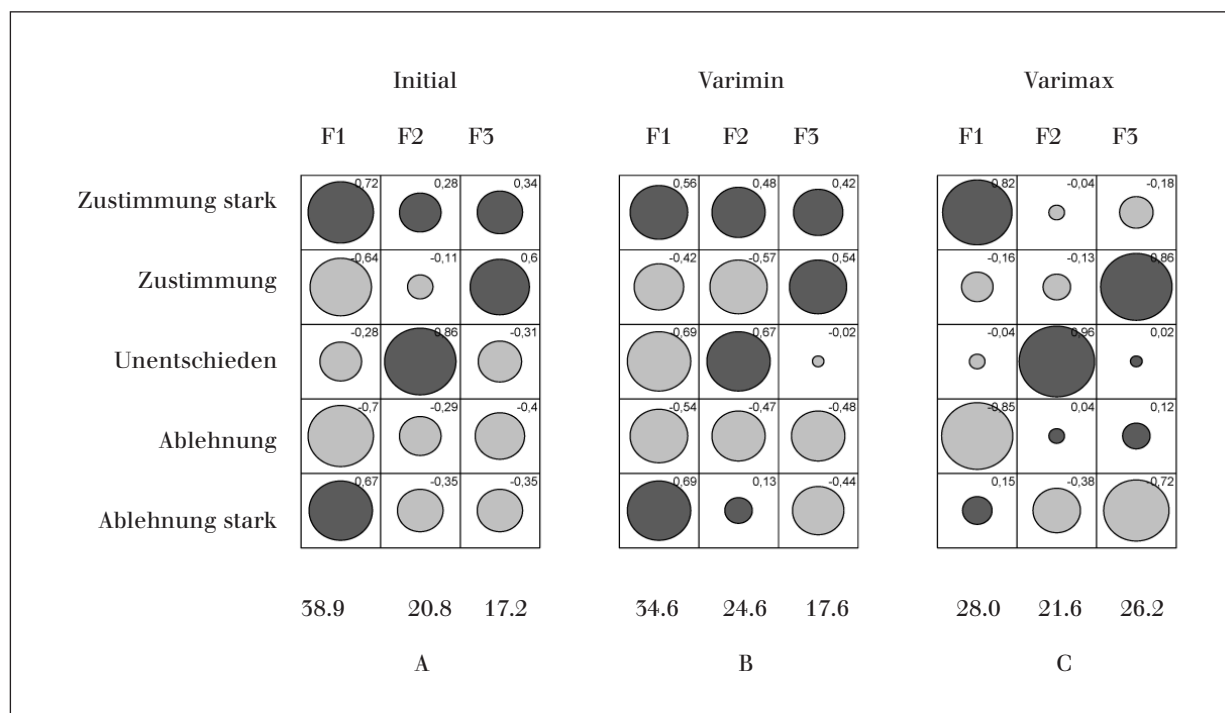


Abbildung 7: Initiale (A), Varimin- (B) und Varimaxstruktur (C) der Response Set-Daten von Carl (1968)

bei F_1) und ‚Ablehnung‘ geführt (negative Ladung bei F_1). Warum? Es fällt auf, dass die Vorzeichen der Ladungen von ‚starke Zustimmung‘ und ‚Ablehnung‘ bei den drei Merkmalen, die durch Varimin differenziert wurden, kontrastieren. Auch beim Varimax-Faktor F_2 , der ebenfalls bipolar ist, finden sich bei den Varimin-Merkmalprofilen Kontraste, diesmal bei den Variablen ‚Zustimmung‘ und ‚starke Ablehnung‘. Für die Urteilstufen im Varimin-Ergebnis keinen polaren Gegensatz, so bleibt ‚unentschieden‘ auf Varimax- F_2 ziemlich isoliert.

Diese Ergebnisse entsprechen dem, was der Varimax-Varimin Faktorenvorgleich bei den Münzen erbracht hat. Kurz: Varimax-Transformationen clustern Variablen mit gleichartigen Merkmalprofilen.⁷ Im Falle bipolarer Faktoren clustern sie zusätzlich, mit dem jeweils anderen Vorzeichen, die Variablen mit gegensätzlich ausgeprägten Profilen. Die Merkmalsgrundlage selbst wird nicht analysiert, eine Differenzierung auf der Merkmalebene wird durch eine Varimax-Clustering systematisch verunmöglicht.⁸

Was würde ein konventioneller Faktorenanalytiker nach Durchführung einer Varimax-Analyse der Response-Set-Daten möglicherweise publizieren? Er könnte behaupten, Akquieszenz sei nicht, wie man bislang dachte, ein einfaktorielles Konstrukt, man habe vielmehr zwischen einer Akquieszenz I mit Extremheitstendenz (F_1) und einer Akquieszenz II ohne Extremheitstendenz (F_3) zu unterscheiden. Dann

könnte er noch F_2 als ‚Unentschiedenheit‘ interpretieren und sich darüber hinweg setzen, dass F_2 auch bei ‚starke Ablehnung‘ eine beachtliche negative Ladung aufweist. Er könnte der statistischen Datenverarbeitung voll vertrauen und glauben, drei neue psychologische Konstrukte entdeckt zu haben (Akquieszenz I und II und Unentschiedenheit). Da diese nicht sinnlos sind, würde man nicht bemerken, dass sie aber ziemlich nutzlose Variablengruppierungen darstellen, wie sie durch Anwendung des Modells der Einfachstruktur fortwährend hervorgebracht werden.

Mit der folgenden Reanalyse *verbaler* Daten aus einer MDS-Untersuchung wird die unverbindliche Deutungspraxis, die bei Verwendung semantischer Einheiten nach ihrer Faktorisierung durch Varimax möglich ist, noch deutlicher.

Empirischer Beitrag 4:
Semantische Merkmale bei Verwandtschaftsbezeichnungen (Daten von Marx & Hejj)

Für 16 Verwandtschaftsbezeichnungen, die den Probanden auf Wortkarten dargeboten und von ihnen hierarchisch sortiert worden waren, ermittelten Marx & Hejj (1989) eine Matrix der Ähnlichkeiten (s. Tabelle 3, bei den Autoren Tabelle 2.5, S. 112). Die Autoren selbst verwendeten ein NMDS-Verfahren, um aus den Häufigkeiten der zusammen gelegten Wortkarten die semantischen Merkmale der Verwandtschaft zu gewinnen, was ihnen auch einigermaßen gelang.⁹ Für den vorliegenden Zweck wurde die Originalmatrix der

Tabelle 3: Daten des hierarchischen Sortierens von Marx & Hejj (1989)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	ANH 2.5 Gruppiematrix der Distanz-Koeffizienten
1 Bruder																	
2 Cousin	759																
3 Cousine	858	451															
4 Enkel	777	770	866														
5 Enkelin	842	865	792	454													
6 Großmutter	860	882	850	702	570												
7 Großvater	783	822	877	564	695	444											
8 Mutter	747	942	895	856	803	679	854										
9 Nefte	787	558	674	802	882	925	865	970									
10 Nichte	900	685	542	875	783	869	926	927	465								
11 Onkel	819	594	711	804	875	856	772	913	592	719							
12 Schwester	432	857	764	841	781	774	853	673	884	797	892						
13 Sohn	632	892	953	721	869	864	764	710	900	969	912	743					
14 Tante	905	717	602	896	821	766	845	840	725	576	446	810	960				
15 Tochter	728	872	890	863	735	775	869	561	971	902	972	625	476	906			
16 Vater	614	896	947	787	894	823	693	477	915	981	831	744	571	924	718		

Sortierhäufigkeiten, welche als Indikatoren der Distanz oder Unähnlichkeit (UÄ) unter den 16 Begriffen zu betrachten sind, diagonal gespiegelt und zu einer Vierecksmatrix aufgefüllt. Sodann wurden die UÄ-Maße durch $\ddot{A} = 1 - U\ddot{A}/1000$ in Ähnlichkeitsmaße transformiert.¹⁰ In die Diagonale wurde der aus der jeweils zugehörigen Spalte/Zeile zu entnehmende höchste Ähnlichkeitswert eingesetzt. Die Spalten der Ähnlichkeitsmatrix wurden sodann untereinander korreliert und die Interkorrelationsmatrix einer PCA unterworfen. Fünf substantielle Faktoren wurden extrahiert¹¹ und nach Varimin und Varimax transformiert.

Abbildung 8A zeigt das Ergebnis der Varimin-Lösung. Der Generalfaktor F_1 hat keine inhaltlichdifferenzierende Funktion (minimale Ladungsvarianz), er ist im wesentlichen verfahrensbedingt und kann ignoriert werden.¹² Die Faktoren F_2 bis F_5 sind bipolar, sie lassen mit kontrastierenden Ladungsvorzeichen die erwar-

teten semantischen Merkmale erkennen: Linealität (F_2), Kernfamilie (F_3), Geschlecht (F_4) und Alter oder Generation (F_5).¹⁵ Das minimale Wortpaar *Bruder* und *Schwester* z. B. zeigt einen Unterschied der Ladungsvorzeichen lediglich bei Faktor F_4 , der als Geschlechtssfaktor (männlich vs. weiblich) zu deuten ist. Andere minimale Wortpaare lassen sich leicht finden, so z. B. *Vater* und *Sohn* oder *Mutter* und *Tochter*, die beide auf F_5 , dem Faktor 'Generation', mit ihrer Ladungsrichtung kontrastieren. In der Faktorenstruktur kommen zwar subtilere Unterschiede nicht zum Ausdruck, so etwa nicht die zwischen der jüngsten, mittleren und der ältesten Generation (etwa Sohn, Vater, Großvater). Auch werden Unterschiede, die durch die Ego-Perspektive bedingt sind, faktoriell ignoriert, z. B. die zwischen (*mein*) *Bruder* und (*meines Vaters*) *Sohn* – die gleiche Person ist gemeint. Doch sind die Hauptmerkmale der Verwandtschaft im Resultat aufzufinden.

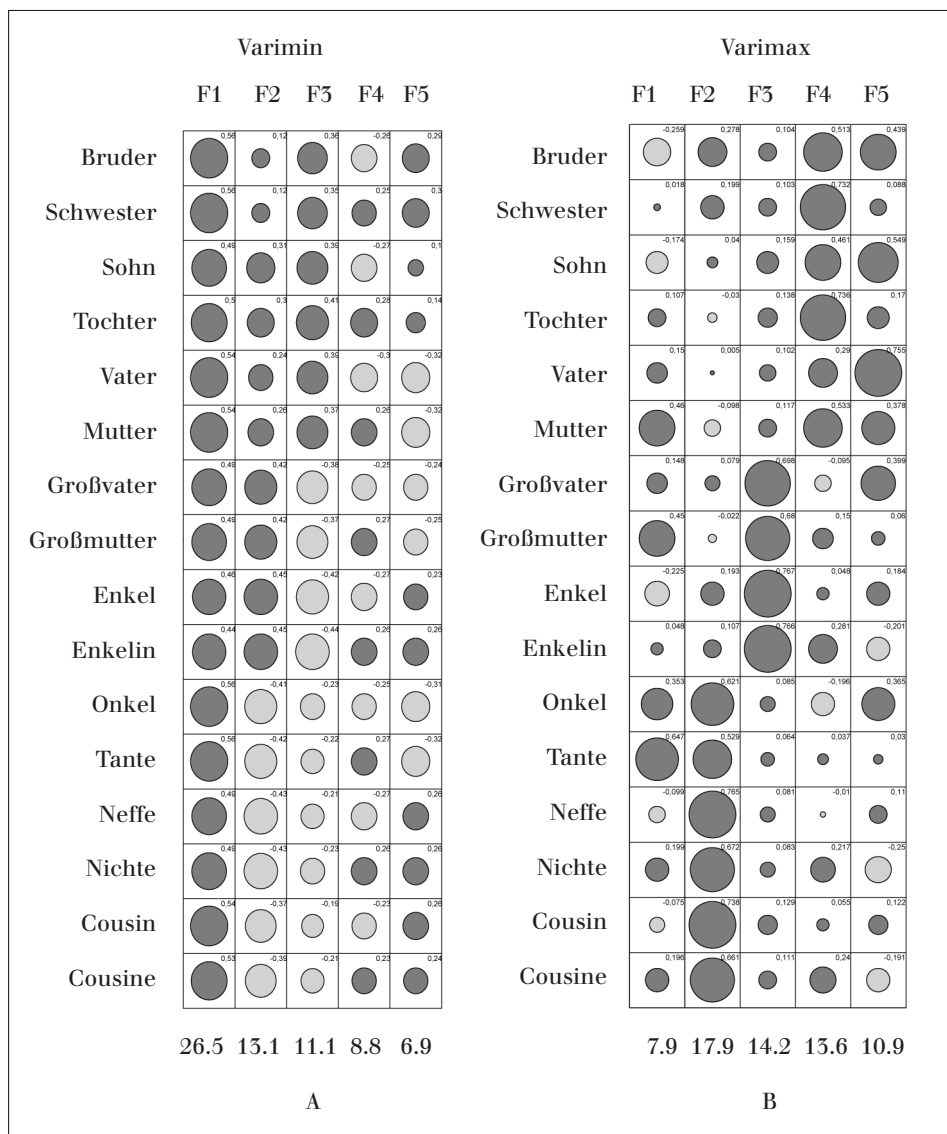


Abbildung 8:
Varimin- (A) und
Varimax-rotierte
Faktoren (B) der
Analyse von Verwandtschaftsähnlichkeiten
(Daten von Marx & Heij, 1989)

Abbildung 8B zeigt das Ergebnis einer *Varimax*-Rotation der Faktoren, die für Abbildung 8A Variminrotiert wurden. Von den fünf rotierten Faktoren lassen sich F_1 und F_5 gar nicht und F_3 nur mit einiger Willkür interpretieren. F_3 vereint Großeltern und Enkel, worin man einen Sinn hineinlegen kann, denn damit werden die generationsmäßig extremen Vertreter innerhalb derselben Linie gruppiert. Mit ‚sehr jung oder sehr alt in der dominanten Abstammungslinie‘ würde man diese Einheit bezeichnen können.

Bei den Faktoren F_2 und F_4 hat das Varimax-Ergebnis eine gewisse Ähnlichkeit mit dem Varimin-Ergebnis. Mit Varimax- F_2 sind die Verwandtschaftswörter der *nicht-linealen* Abstammung geladen (‚entferntere Verwandte‘). Im Unterschied zur Varimin-Lösung fehlt aber den Wörtern der linealen Abstammung (‚nähere Verwandte‘) als faktorielle Kennzeichnung der Vorzeichen-Gegensatz ([+] ‚Merkmal vorhanden‘ vs. [-] ‚Merkmal nicht vorhanden‘), alle Varimax-Faktoren sind unipolar. F_4 gruppiert die Vertreter der Kernfamilie, doch die Nicht-Mitglieder bleiben bei F_4 wieder ohne Kennzeichen (kein negatives Vorzeichen). Außerdem haben *Enkelin* und *Nichte* und *Cousine* noch beträchtliche F_4 -Ladung, womit man wenig anfangen kann.

Es ist also nicht so, dass sich Varimax-Faktoren einer Interpretation völlig entziehen. Doch wegen semantischer Überlappungen lässt ihre Interpretation meist zu wünschen übrig, was mithilfe der transparenten Semantik der Verwandtschaftswörter leicht zu erkennen ist. Vor allem bleiben im Varimax-Ergebnis auch bei den deutbaren Faktoren die *Quellen* der semantischen Ähnlichkeit und Verschiedenheit (Generation, Geschlecht usw.), die eine Faktorenanalyse aufdecken sollte, verborgen. Großeltern und Enkel werden aufgrund der Merkmale Linealität und Generation zusammen geclustert, doch werden diese Merkmale faktoriell nicht manifest. Eklatante Beispiele für die mit Varimax oft einhergehende schwere Interpretierbarkeit von Item- und Skalen-Clustern, die ein weitgehend willkürliches Sinngeben und Benennen ermöglichen, bietet die Big-Five-Persönlichkeitsforschung: Bei einem der fünf Faktoren schwankte man anfangs sogar zwischen drei semantisch höchst heterogenen Bezeichnungen: ‚intellect‘, ‚culture‘ und ‚openness‘ (mehr dazu in Ertel a). Auch ist ‚agreeableness‘ eine ziemlich ungeeignete Bezeichnung, da sie doch eher die Wirkung einer Persönlichkeit A auf eine Persönlichkeit B beschreibt als eine Erlebens- oder Verhaltenstendenz der Persönlichkeit A selbst.

Frage IV

Enthält die regelmäßig anzutreffende Bipolarität von Varimin-Faktoren substantielle Information?

Die Bipolarität in Varimin-Faktorlösungen, die schon im Vorläufer-Artikel diskutiert wurde (Ertel, 2009, p. 42f) und in den beiden letzten empirischen Beiträgen festzustellen war, verdient besondere Beachtung. Bipolarität tritt in Varimin-Lösungen regelmäßig auf, weit häufiger als in Varimax- und anderen Simple Structure-Lösungen. Das Vorkommen von negativen Ladungen – hier sind nur die mit der Komplexstruktur gewonnenen gemeint und die initialen – kann bedeuten, dass die negativ geladenen Variablen gegenüber den positiv geladenen für die betreffende Quelle der Kovarianz (für den betreffenden Faktor) eine funktional abträgliche Wirkung haben. Bei den Response Set-Daten von Claus z. B. musste man bei den Variablen mit Minus-Faktorgewichten eine hemmende Wirkung annehmen: Wenn jemand viele extreme Ja-Nein-Antworten abgibt, dann gibt er plausiblerweise weniger gemäßigte Antworten ab, und umgekehrt. Das lässt sich seiner Motivation zuschreiben, die gemäßigten Urteile entweder zu meiden oder zu bevorzugen.¹⁴ Eine motivational-funktionale Deutung der Bipolarität bei diesen Response-Set-Faktoren ist erlaubt, weil Carl Vorkehrungen zur Ausschaltung von Artefakt-Korrelationen getroffen hatte. Für die bipolaren Akquieszenz- und Nuanciertheitsfaktoren gilt die motivational-funktionale Deutung des Vorzeichen-Kontrasts bei Faktoraladungen entsprechend.

Anders ist die Bipolarität bei den Verwandtschaftsdaten zu interpretieren. Beim Faktor Geschlecht z. B. kann Bipolarität nicht auch funktional gedeutet werden. Das Vorliegen des Merkmals *männlich* setzt keine Hemmung durch *weiblich* voraus, sondern eine organismische Gegebenheit, die das Vorkommen des Merkmals *weiblich* ausschließt (in der Regel), wenn *männlich* vorliegt. Das Merkmal Linealität setzt für sein Vorkommen oder Nichtvorkommen nur die richtige Antwort auf die Frage nach der entsprechenden Platzierung des Verwandten im Abstammungsbaum voraus, die ein Ja oder Nein ergibt. In solchen Fällen muss man sich darauf beschränken, mit den Vorzeichen das Vorkommen bzw. Nichtvorkommen eines Merkmals ausgedrückt zu finden. Dies wird übrigens auf anderen Gebieten, z. B. in linguistischen Diskursen, viel häufiger so gehandhabt als in der Psychologie.

Der Fall, dass eine negative Ladung eines Faktors das Nichtvorkommen eines Einflusses oder einer Funktion ausdrückt, kann darauf schließen lassen, dass der Faktor für die negativ geladenen Variablen irrelevant ist. Die Entstehung des Faktors ist in solchen Fällen denjenigen Variablen des untersuchten Samples zu verdanken, bei denen er positive Ladungen aufweist.¹⁵

Man sollte sich bei Faktorenanalysen mit Komplexorientierung darauf einstellen, negative Vorzeichen – wenn sie bei einem Faktor mit positiven zusammen vorkommen – zunächst immer als *Merkmal nicht vorhanden* zu deuten und erst in einem zweiten Schritt, aufgrund eines von anderswoher bezogenen Kontextwissens, eine funktionale Deutung (Hemmung, polar gerichtete Ausprägung usw.) zusätzlich einbringen, wenn man eine solche mithilfe seines Kontextwissen begründen kann.^{16, 17}

Zum besseren Verständnis der Bipolarität von Faktoren ist es nützlich, ihr Entstehen anhand eines besonders transparenten Beispielfalls zu verfolgen:

Empirischer Beitrag 5:

Intelligenzentwicklung in der Kindheit

(Daten von Humphreys & Davey)

In einer Längsschnittstudie testeten Humphreys & Davey (1988) Kinder vom dritten Lebensmonat bis zum 9. Lebensjahr mit insgesamt vier jeweils altersangepassten vergleichbaren Intelligenztests. Sie wollten die Beständigkeit der intellektuellen Entwicklung in diesem Zeitraum ermitteln. Die von ihnen mitgeteilte Interkorrelationsmatrix für 14 Testwiederholungen im Längsschnitt wurde von uns PCA-faktoriert und Varimin- und Varimax-rotiert. Abbildung 9 gibt die Ergebnisse wieder (die nicht-wiedergegebene Initiallösung ist mit der Variminlösung nahezu identisch). Nur zwei Faktoren sind substantiell (Eigenwerte: 5.705, 1.473, 0.827, 0.788 ...).

Der erste Variminfaktor repräsentiert mit einer Kommunalität von 40.7% den stabil bleibenden Anteil der allgemeinen Intelligenz, die Faktorladungen bleiben innerhalb der Untersuchungszeit auf ungefähr gleichem Level. Der zweite Faktor ist bipolar, er repräsentiert mit 10.6 % Kommunalität die zeitabhängige Varianz im Laufe der Intelligenzentwicklung. Diese ist plausiblerweise zurückzuführen auf interindividuell variierende, günstige oder ungünstige Lebensumstände: auf Erziehungs- und Ausbildungseinflüsse, psychische oder körperliche Erkrankungen usw..

Die Bipolarität von F_2 und die monotone Abfolge der Faktorgewichte werden wie folgt interpretiert: Der mit F_1 nicht ausgeschöpfte Kovarianzanteil verteilt sich ziemlich gleichmäßig über die Altersspanne, d. h. die Veränderungen von einem Messzeitpunkt zum nächsten sind geringfügiger als von einem Messzeitpunkt zum übernächsten. Bis zum *über-übernächsten* sind die Veränderungen noch größer, und vom ersten bis zum letzten Messzeitpunkt sind sie am größten. Dementsprechend korrelieren die Testergebnisse der einander benachbarten Messzeitpunkte höher als die der einander entfernteren Messzeitpunkte. Die im mitt-

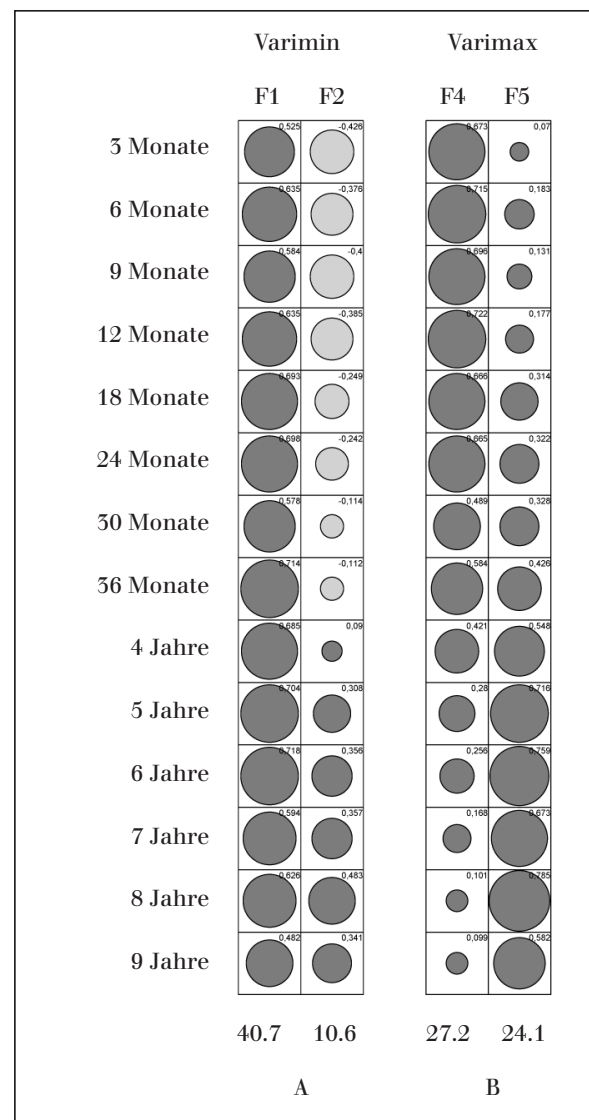


Abbildung 9: Varimin- (A) und Varimaxlösung (B) der Längsschnittdaten zur Intelligenzentwicklung, gewonnen von Humphreys & Davy (1988)

leren Bereich des Längsschnitts liegenden Testwerte haben zu den Testanfangs- und -endwerten ungefähr gleiche Differenzen, die entsprechenden Korrelationen des Testergebnisses von der Mitte des Längsschnitts mit dem ersten und letzten Testergebnis sind ungefähr gleich hoch. Den Testzeitpunkt mit mittlerer Differenz gegenüber dem Anfangs- und Endwert macht das Varimin-Modell zum Nullpunkt des F_2 -Ladungsvektors. Faktor F_2 repräsentiert somit im vorliegenden Fall das Ausmaß an Veränderung der Intelligenz gegenüber dem Mittelwert den Veränderungen.

Das Beispiel ist insofern lehrreich, als es zeigt, dass die Faktorgewichte von F_2 auf die von F_1 bezogen sind, wobei die Art des Zusammenhangs durch einschlägiges Kontextwissen spezifiziert werden muss. Negative

F_2 -Ladungen sind in diesem Fall weder Anzeichen von fehlendem oder gegensätzlichem Einfluss noch von logischem Ausschluss, sondern Anzeichen von Differenzen der Intelligenz-Veränderungswerte zum mittleren Veränderungswert. Da die F_2 -Faktorgewichte positiv und negativ gepolt sind, entsteht eine Skala, die anzeigt, dass die Veränderung der Testergebnisse im Laufe der Längsschnitt-Untersuchung monoton zunimmt. Denn wenn z. B. ein genügend hoher konstanter Wert zu den Faktorladungen addiert wird, können die Bipolarität und der Nullpunkt verloren gehen, der Anstieg der Faktorladungen mit zunehmendem Alter bliebe erhalten, dieser würde die seit Beginn der Testperiode allmählich größer werdende Veränderung der Intelligenzleistungen repräsentieren.

Die Varimaxlösung beseitigt den Generalfaktor und ignoriert somit die Tatsache, dass die Intelligenz im getesteten Längsschnitt relativ konstant bleibt. Stattdessen wartet sie mit zwei Faktoren auf, die auf den ersten Blick wie zwei unabhängige Intelligenzen erscheinen, die eine, die sich in der frühesten Kindheit auswirkt, während eine andere Intelligenz die primäre allmählich ablöst – eine absurde Vorstellung, zu der auch Einfachstruktur-Faktoristen kaum neigen werden. Sie würden ihr Verfahren im Falle von Längsschnitt-Wiederholungsdaten vermutlich nur für ungeeignet ansehen.

Ein andersartiges Beispiel, bei dem die negative Ausprägung eines Faktors nicht lediglich Abwesenheit eines Merkmals, sondern ein skalometrisch eigenständiges Merkmal erkennen lässt, bietet die folgende Untersuchung.

Empirischer Beitrag 6:

Körpervolumen und Körperform von Rindern

(Daten von Rasch)

Von 107 weiblichen Rindern (Färsen) erhob Rasch (1962) 12 Maße ihrer Körperausdehnung. Es handelte sich um Höhen-, Breiten- und Längenmaße. Unsere übliche faktorielle Verarbeitung der von E. Weber mitgeteilten Interkorrelationsmatrix führte zu der in Abbildung 10 wiedergegebenen zweifaktoriellen Lösung (Eigenwerte: 7.69, 1.20, 0.74...). Die nicht wiedergegebene Initialstruktur ist mit der Variminstruktur nahezu identisch.

Der erste Faktor (Kommunalität 64.0%) repräsentiert das Körpervolumen (die Körperausdehnung) ohne Spezifizierung einer Raumdimension. Er besagt, dass ein Rind, das über- oder unterdurchschnittlich hoch gewachsen ist, meist auch über- bzw. unterdurchschnittlich in die Länge und Breite gewachsen ist. F_1 ist dem Generalfaktor der Intelligenz analog zu deuten.

Der zweite Faktor (Kommunalität 10.1%) repräsentiert mit seiner Bipolarität die Varianz, die die Körperform mit sich bringt. Es gibt unter sehr voluminösen wie unter wenig voluminösen Rindern *relativ* schmal gewachsene (schlanke) und *relativ* breit gewachsene (dicke). Das Auftreten von Polarität zeigt an, dass mitmäßigem Spielraum das Wachstum in die Länge „auf Kosten der Breite“ und das in die Breite „auf Kosten der Länge“ vonstatten geht.

In der Varimaxlösung geht der Generalfaktor des Körpervolumens verloren. Die Variablen der Schlankheit und Dicke sind zwar auch erkennbar, allerdings als nicht-polare orthogonale Dimensionen. Die empirische Gegenläufigkeit der Längen- und Breitenausdehnung, die sich durch Varimin aufgrund der Dominanz des Volumenfaktors zu einer bipolaren Einflussgröße der ‚Körperform‘ herausbildet, kommt im Modell der Einfachstruktur nicht zum Tragen.¹⁸

Die Deutung negativer Ladungen eines Faktors als Ausdruck eines bipolar aufzufassenden Charakteristikums kann auch bei psychologischen Daten nahe liegen, demonstrierbar mit folgendem Beitrag.

Empirischer Beitrag 7:

Intelligenz- und Leistungstests

(Daten von Holzinger & Swineford)

Die Holzinger & Swineford-Daten werden gelegentlich in Lehrtexten zur Demonstration von Modellrechnungen verwendet, so auch von Jöreskog & Sörbom (2003), von deren Website die Interkorrelationstabelle entnommen wurde (zu ergoogeln mit ‚LISREL 8.52 Jöreskog‘). Es handelt sich um je drei Tests für visuelle, für verbale und für Tempo-Leistungen (speed tests). Die Ergebnisse der vorliegenden Standard-Auswertung für Varimin und Varimax zeigt Abbildung 11.

Varimin- F_1 (Abbildung 11 A) lässt sich als Generalfaktor ‚g‘ der allgemeinen Intelligenz ansehen, der für Intelligenztestbatterien generell zu erwarten ist. Der Faktor F_2 kontrastiert durch gegensätzliche Vorzeichen die drei Speed-Test-Leistungen mit den übrigen Testleistungen, bei denen es offensichtlich weniger auf Tempo als auf Konzentration oder ‚power‘ ankommt. Hier liegt ein Fall vor für Bipolarität im Intelligenzbereich, bei dem – mit psychologischen Kontextkenntnissen – negative Ladungen mit funktionaler Bedeutung in Zusammenhang gebracht werden dürfen. Eine erhöhte Fähigkeit und Neigung zu Tempoleistungen wirkt sich vor allem bei geeigneten Speed-Tests förderlich aus. Bei den Tests, die Konzentration und Vertiefung in Probleme erfordern, werden sich Speed-Neigungen und -Fähigkeiten vermutlich eher nachteilig auswirken. Eine entsprechende Gegenwirkung lässt sich auch für die Fähigkeit und Neigung zur Konzentration denken, die sich bei Tests, die Tempofähigkeit und -neigung

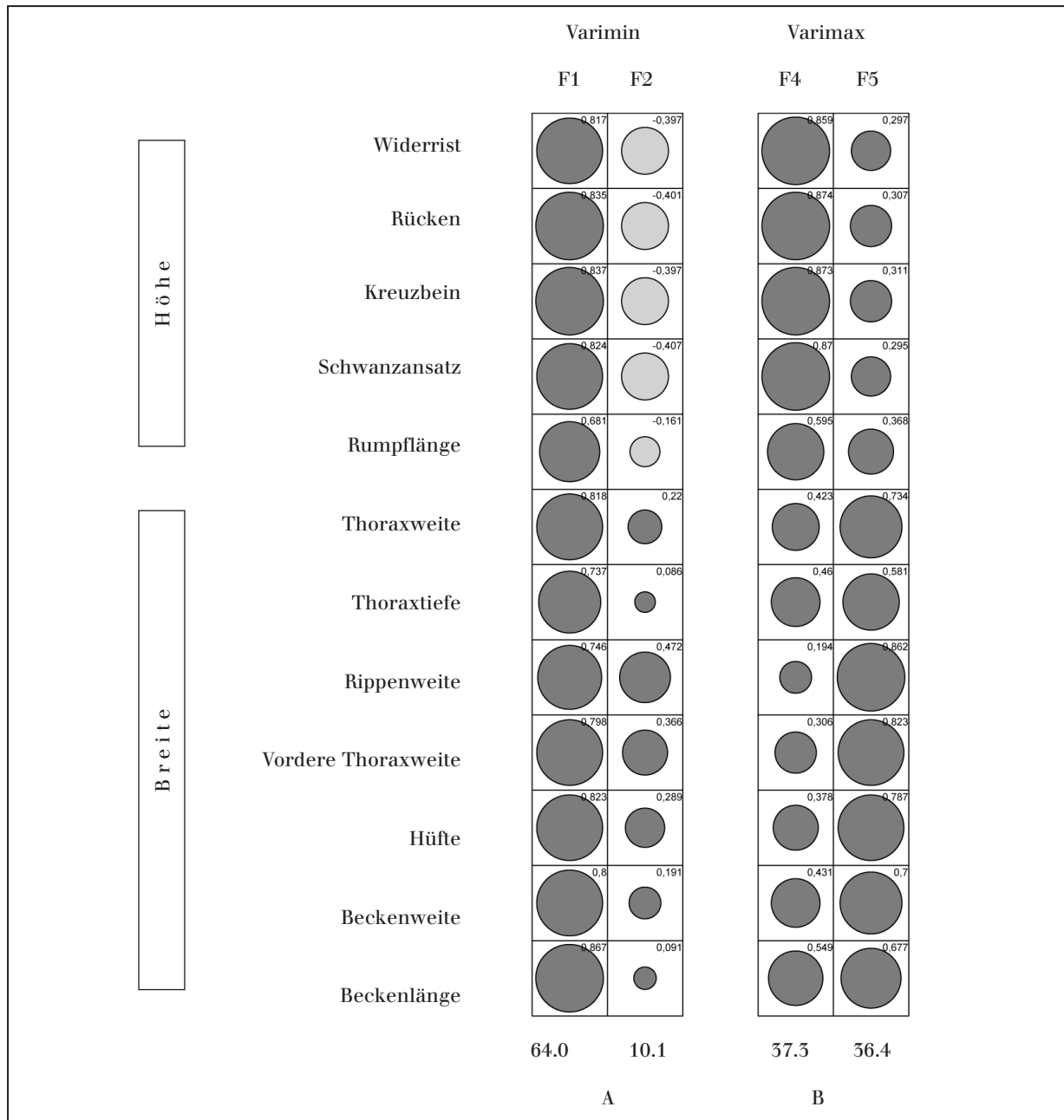


Abbildung 10: Varimin- (A) und Varimaxfaktor (B) von 12 Körpermaßen weiblicher Rinder (Färsen)

erfordern, vermutlich nachteilig auswirken. Diese Gegenläufigkeit kann sich mit der Bipolarität von F_2 manifestieren.

Die Bipolarität von Varimin- F_5 hebt eine andere Kovarianzquelle ans Licht: Die Besonderheit verbaler Testleistungen. Bemerkenswert ist, dass sich die Bipolarität von F_5 nicht auf verbale vs. visuelle Tests beschränkt, die beide zur Power-Kategorie gehören, wegen der unterschiedlichen Aufgabeninhalte (Wörter vs. Bildmaterial) bilden verbale und visuelle Tests gute minimale Kontraste. Doch auch die rechnerischen Leistungen aus der Speed-Testserie kontrastieren,

wenn auch etwas schwächer, mit den verbalen Testleistungen. Offenbar liegen den verbalen Leistungen übergreifende Besonderheiten zugrunde. Auch hier lässt sich eine antagonistische Beziehung denken: Die verbalen Fähigkeiten und Neigungen, wenn sie im individuellen Intelligenzprofil dominieren, könnten die rechnerischen als auch die visuellen Aktivitäten etwas benachteiligen und/oder umgekehrt. Doch kann die funktionale Relevanz bei negativen Varianzquell-Ladungen hier wie oft in anderen ähnlichen Fällen nur Hypothese sein.¹⁹

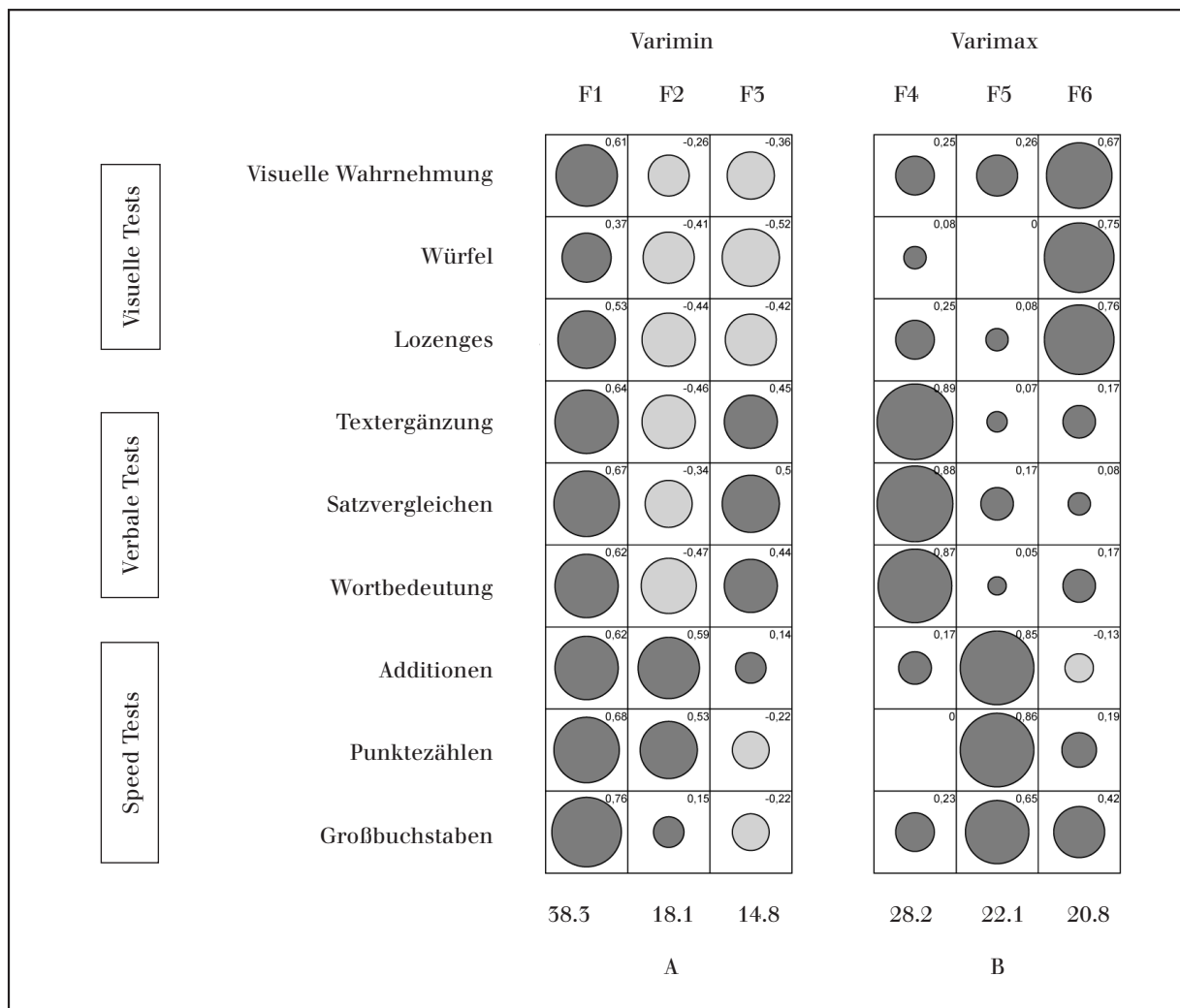


Abbildung 11: Varimin- (A) und Varimaxlösung (B) der Holzinger & Swineford-Daten (1939)

Die Varimax-Lösung (Abbildung 11 B) bringt, zum Verschwinden und separiert die drei Testsorten voneinander, so dass sich mit dieser Lösung kein Antagonismus zwischen verschiedenen Fähigkeiten und Neigungen manifestieren kann. Die LISREL-Lösung von Jöreskog & Sörbom kommt mit größerem Aufwand und mit mehr Eleganz zum gleichen Ergebnis wie die Varimax-Lösung von 11 B. Die Autoren hatten das Simple Structure-Modell mit den drei latenten Varianzquellen, deren Unabhängigkeit vorausgesetzt wurde, konfirmatorisch überprüfen wollen und waren dabei auch erfolgreich – allerdings auf Kosten der naheliegenderen und theoretisch plausibleren Lösung, zu der Varimin geführt hat.

Dass in bipolaren Lösungen nicht immer Variablen nur entweder mit positiven oder negativen Ladungen vorkommen, sondern dass sie auch Nahe-Null-Ladungen haben können, und dass dies informativ sein kann, zeigt das folgende Beispiel.

Empirischer Beitrag 8:

Psychophysiologische Aktivationsindikatoren (Daten von Köhler & Troester)

Den Autoren Köhler & Troester (1991) ging es um die Validierung des Maßes der Handflächen-Schweißproduktion (PSI, Palmar Sweat Index) als Indikator für psychophysiologische Aktivierung. Sie testeten 50 Personen unter drei Ruhebedingungen und einer Anstrengungsbedingung (von der Zahl 2007 hatten die Probanden sukzessive die Zahl 7 zu subtrahieren) und erhoben psychophysiologische Werte pro Person 16 mal im Laufe der vier Testphasen, nämlich PSI-Werte am Mittel- und Zeigefinger (PSI-M, PSI-F), spontane Fluktuationen der Schweißproduktion (SF), den Hautwiderstand (SCL, skin conductance level) und die Herzschlagfrequenz (HR). Die fünf erhobenen Werte wurden pro Person über die 16 Messwiederholungen intraindividuell korreliert, die Korrelationen wurden über die 50 Probanden gemittelt. Die Ergebnisse un-

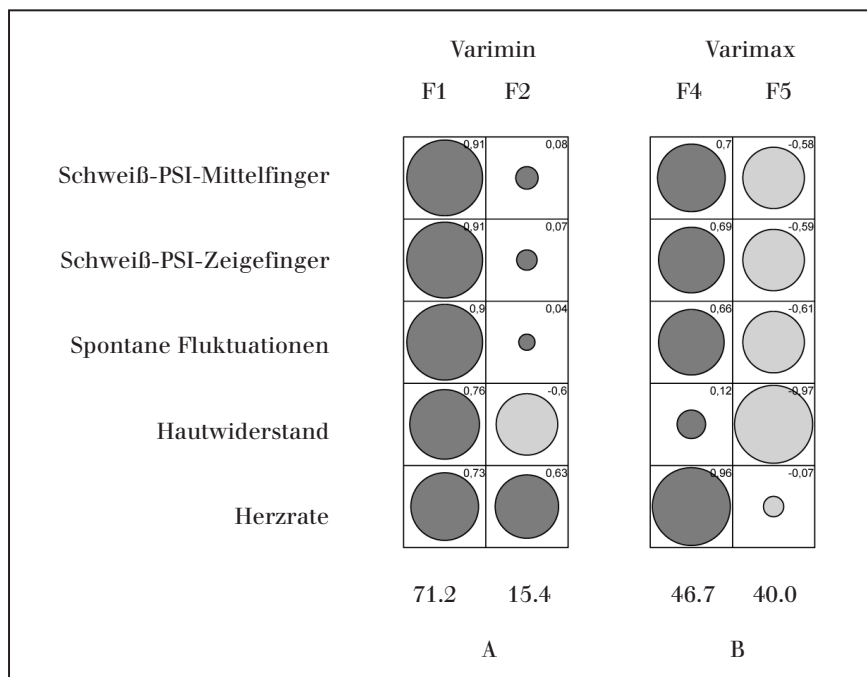


Abbildung 12:
Varimin- (A) und Varimax-
lösung (B) der Köhler &
Troester-Daten (1991)

serer Standardauswertung der faktorisierten Interkorrelationen liefert Abbildung 12. Zwei Faktoren wurden nach dem Eigenwertverlauf 3.50, 0.76, 0.33, 0.20 ... und nach dem Ladungsmuster für interpretierbar gehalten, auf die Wiedergabe der Initiallösung, die der Variminlösung sehr ähnlich ist, wird verzichtet.

Ergebnisse: Varimin- F_1 ist ein Generalfaktor, der besagt, dass den individuellen Differenzen der fünf Maße offensichtlich der gleiche Indikatorwert – Aktivierung – zukommt. Varimin- F_2 repräsentiert Zusatzvarianz, die sich allerdings auf den Hautwiderstand und die Herzrate beschränkt, die Schweißvariablen haben mit dieser Varianzquelle offenbar nichts zu tun, dies zeigen ihre Um-Null-Ladungen an. Die F_2 -Ladungen mit gegensätzlichen Vorzeichen bei SCL und HR besagen, dass die beiden Variablen offensichtlich ein leicht antagonistisches funktionales Verhältnis zueinander haben: Es gibt offenbar Probanden, bei denen der Hautwiderstand auf Aktivierung stärker reagiert als die Herzfrequenz und andere, bei denen umgekehrt die Herzfrequenz stärker reagiert als der Hautwiderstand. Obgleich beide Variablen mit F_1 im wesentlichen die gleiche Aktivationsfunktion indizieren, bleibt ein geringer Rest an Varianz übrig, der durch eine leichte Präferenz entweder des Hautwiderstands oder der Herzrate aufgeklärt wird. Hier scheint ein Anteil ‚gegabelter Wirkung‘ vorzuliegen: Je mehr Wirkung in Richtung X vorliegt, umso weniger Wirkung ergibt sich für Richtung Y, und umgekehrt. Die Schweißproduktion ist von diesen koordinierten Präferenzwirkungen nicht betroffen.

Die Varimaxlösung stellt den Auswerter vor zwei unlösbare Rätsel. Die Aufspaltung des Aktivationseffekts auf zwei unabhängige Äste ist das eine Rätsel –

aus physiologischen Gründen sind zwei unabhängige Aktivationsquellen kaum vorstellbar. Das zweite Rätsel ist die Zugehörigkeit von SRL zum einen Aktivationsast und von HR zum anderen Aktivationsast. Varimax vermischt hier die Varianzquellen bis zur Unkenntlichkeit.

Die bisherigen Beispiele zeigten: Die faktorenanalytische Auswertung von Variablen, die der hypothetischen Komplexität ihrer latenten Varianzquellen gerecht werden will, führt zu Ergebnissen, die leichter interpretiert werden können als Ergebnisse, denen man eine Einfachstruktur aufzuzwingen sucht. Die Ebene der latenten Bedingungen wird im Folgenden differenzierter in den Blick genommen.

Frage V

Sind mit Hilfe der Komplexstruktur-orientierten Faktorenanalyse auch methodenbedingte Einflüsse erfassbar?

Campbell & Fiske legten 1959 einen neuen methodologischen Ansatz vor, der in den darauf folgenden Jahrzehnten die Entwicklung zahlreicher methodischer Verfahren unter dem Label MTMM (multi trait multi method) nach sich zog. Die Autoren nahmen systematisch in Angriff, worauf man schon vorher, aber weniger nachdrücklich, aufmerksam gemacht hatte: Persönlichkeitsforscher sehen sich in ihrer Forschung nicht nur einer Vielheit von Varianz-erzeugenden latenten Traits gegenüber. Mit zusätzlicher Varianz haben sie zu rechnen, wenn sie zu deren Erfassung unterschiedliche Verfahren verwenden. Bald darauf bezog man auch die Varianzanteile bei Fragebögen mit ein, die durch unterschiedliche Informanten (Selbst- vs. Fremdbeurteilung) entstehen. Andere mögliche Vari-

anzeinflüsse, die an den Kovarianzen der gemessenen Variablen einen Anteil haben können, kamen hinzu (Situationseinflüsse bei Testwiederholungen usw.). Da sich das hier favorisierte komplexitätsoffene Verfahren beim Aufspüren multipler Varianzquellen bewährt hat, liegt es nahe, eine entsprechende Leistung auch für MTMM-Daten zu erwarten.

Empirischer Beitrag 9:

Wissensprüfung mit Variation der Prüfmethode

(Daten von Campbell & Fiske)

Zur Differenzierung methodischer Varianzquellen entwickelten Campbell & Fiske (1959) ein Verfahren, das auf eine systematische Inspektion der Interkorrelationsstabellen hinausläuft. Sie demonstrierten ihr Vorgehen anhand zahlreicher Korrelationsdaten, die sie in der Fachliteratur fanden. Ein Datensatz, aus Campbell und Fiske, der schon von Cronbach und Vernon für andere Zwecke reanalysiert worden war – ihre Quelle geben sie nicht an – stammt von einem Wissenstest, bei dem im Physik-Unterricht vermittelte Inhalte zur Mechanik und Elektrizität durch rein verbale Fragen sowie durch Fragen mit Hilfe bildlicher Darstellungen geprüft wurden.

Die Ergebnisse unserer Auswertung für Varimin und Varimax zeigt Abbildung 13. Nach dem Eigenwertverlauf (2.67, 0.62, 0.45, 0.27) wäre nach üblichen Kriterien nur F_1 , allenfalls noch F_2 heranzuziehen. Doch Campbell & Fiske erkannten mithilfe ihrer Korrelationsinspektion recht deutlich die Auswirkung auch der beiden Varianzquellen, die in unserer Faktorenanalyse als F_2 und F_3 in Erscheinung treten.²⁰

Ergebnis: Varimin- F_1 repräsentiert mit hohem Kommunalitätsanteil den Generalfaktor, zu dem wohl nicht nur die unterschiedliche Intelligenz der Probanden, sondern auch ihr unterschiedlicher Fleiß für das

Unterrichtsfach beigetragen haben. Hinzu kommt mit weitaus geringerem Gewicht F_2 als Methodenfaktor (Sprache vs. Bilder) und F_3 als Faktor der Lehrinhalte (Elektrizität vs. Mechanik). Unter den Probanden gab es offenbar solche, die mehr mit verbal ausformulierten Fragen als mit Abbildungen zurecht kamen und solche, denen Probleme mit Abbildungen leichter fielen (F_2). Auch hatten einige wohl mehr für die Mechanik als für die Elektrizität gelernt und andere umgekehrt (F_3). Wieder erweist sich hier Bipolarität als ein Indikator für konkurrierende Gegebenheiten, nicht etwa nur für zwei das Vorliegen oder Nicht-Vorliegen einer einzigen Gegebenheit.

Die Gegenläufigkeit der Wirkungen bei verbalem vs. bildlichem Prüfmaterial geht in der Varimaxlösung verloren. Ferner fehlt in der Varimaxlösung eine faktorielle Repräsentanz der Elektrizitätsinhalte, nur die Mechanikinhalte werden durch Varimax faktoriell repräsentiert. In einem letzten Beitrag soll die Frage geklärt werden, ob sich durch Varimin auch die Varianzquelle ausdrücken wird, die durch den Wechsel befragter Personenstichproben gegeben ist.

Empirischer Beitrag 10:

Selbst- und Fremdbeurteilung von Kindern

(Daten von Matson & Nieminen)

In einer Fragebogen-Untersuchung zu Verhaltensstörungen, Depression und Angst bei Kindern von Matson & Nieminen (1987) verwendeten die Autoren sechs Skalen, die den Kindern selbst sowie ihren Lehrern/Lehrerinnen vorgelegt wurden, welche die Kinder zu beurteilen hatten. Das Ergebnis unserer Auswertung zeigt Abbildung 14. Die Eigenwerte betragen 3.98, 2.14, 1.41, 0.99, 0.86..., die Prozentanteile an aufgeklärter Varianz bei den Initialfaktoren 33.1, 17.8, 11.8, 8.2..., wonach drei Faktoren als substantiell zu bezeichnen sind.

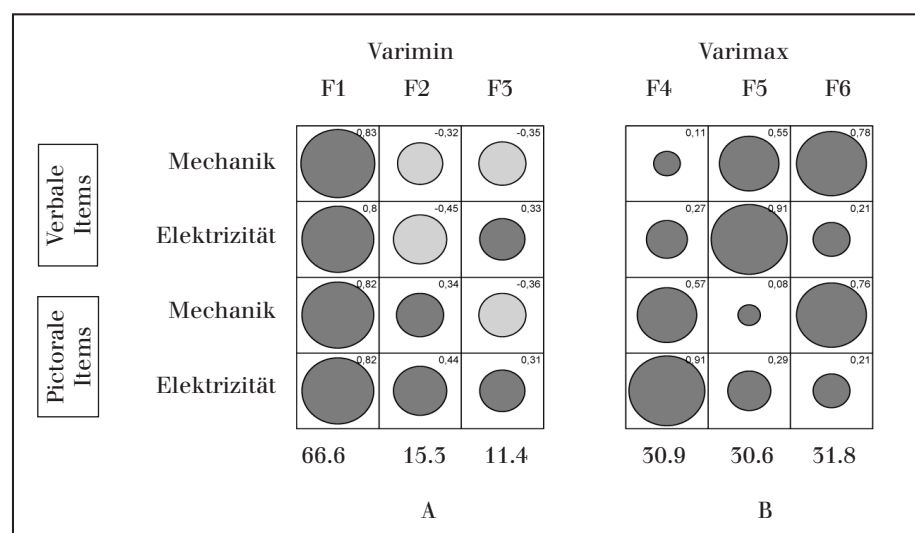


Abbildung 13:
Varimin- (A) und
Varimaxlösung (B) der
Campbell & Fiske-Daten
(1959)

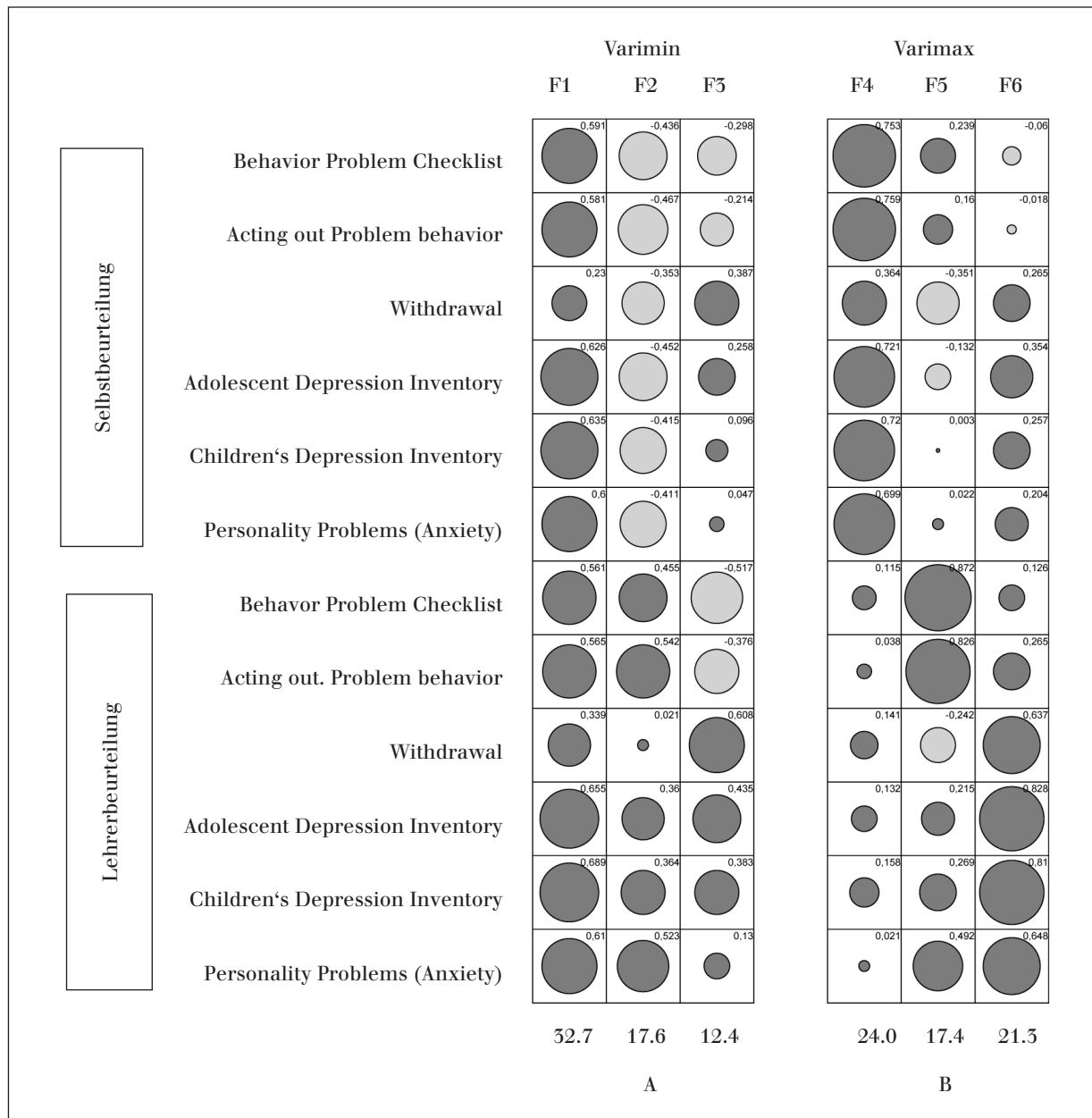


Abbildung 14: Varimin-(A) und Varimaxlösung (B) der Matson & Nieminen-Daten (1987)

Zum Ergebnis der Variminlösung, Abbildung 14A: Die im Generalfaktor F_1 sich ausdrückende Gemeinsamkeit der Kovarianzquelle bei Kindern und Lehrern lässt mit etwas Zurückhaltung darauf schließen, dass in den Urteilen tatsächlich vorliegende dysfunktionale Symptome bei den Kindern mit ihrer interindividuellen Varianz zum Ausdruck gebracht werden. Zurückhaltung ist geboten, weil sich bei den meisten Fragebogenuntersuchungen, so auch wohl in dieser, im ersten Faktor zusätzlich die Akquieszenz als Antwortungsstil geltend machen kann. Dass in dieser Untersuchung die Variablen mit F_1 bei Kindern und Lehrern unterschiedslos geladen sind, ist kein stichhaltiger

Einwand, da Akquieszenz bei Kindern und Lehrern gleichermaßen ausgeprägt sein kann. Zusätzlich kann zu F_1 bei Kindern und Lehrern ein Einfluss der *social desirability* (SD) beigetragen haben der nicht erkennbar ist, da alle Skalen negativ bewertetes Erleben und Verhalten zum Inhalt haben.

Mit gegensätzlichen Vorzeichen beim bipolaren F_2 -Faktor macht sich offenbar der Informanten-Einfluss geltend. Dieser ist als solcher zwar nicht zweifelhaft, seine Herkunft aber bleibt mehrdeutig. Es mag sein, dass die Kinder sich selbst tendenziell als weniger erlebens- und verhaltensgestört beurteilen als die Lehrer, oder auch als gestörter, oder dass sie zu mehr

oder zu weniger Akquieszenz oder SD neigen als die Lehrer. Auch können sich diese Einflüsse ähnlich wie die Teil-Einflüsse bei F_1 addieren.

Mit dem Vorzeichen-Gegensatz des bipolaren Faktors F_3 werden die dysfunktionalen Symptome inhaltlich differenziert, bei Kindern und Lehrern in gleicher Weise. Auf der einen Seite gruppieren sich die Störungen mit enthemmter Auswirkung (acting out, conduct disorder), auf der anderen Seite die Störungen mit gehemmter Wirkung (withdrawal, depression, anxiety). Auch hier ist eine antagonistische Beziehung anzunehmen, dysfunktionale Dispositionen können sich je nach Veranlagung entweder als Hemmungen ausdrücken oder aber zu einem enthemmenden ‚acting out‘ führen (Aggressionen z. B.), wobei Angst ‚abgeführt‘ wird, was psychoanalytisch so gedeutet werden könnte.

Die Varimaxlösung (Abbildung 14 B) führt zu drei Clustern der Schüler- und Lehrerskalen. In einem Cluster sind die sechs von den Schülern beurteilten Skalen vereint, in einem zweiten die drei von den Lehrern beurteilten Skalen der Enthemmungsstörung, in einem dritten Cluster die zwei von den Lehrern beurteilten Skalen der Hemmungsstörung, wobei eine Reihe von Ausreißern und mehrfaktorielle Ladungen das Bild beeinträchtigen. Mit diesen Clustern lässt sich konzeptuell kaum etwas anfangen.

Die Reanalyse der Daten von Matson & Nieminen hat nochmals deutlich gemacht, dass der Ausdruck ‚latent‘ (‚unsichtbar‘, ‚verborgen‘) zur Kennzeichnung der Faktoren, die den Variablen zugrunde liegen, nicht immer zutrifft. Die Verwendung wechselnder Beurteiler liegt als Bedingungsfaktor offen zutage. Ihr Einfluss, wenn ein solcher vorliegt, sollte sich durch eine Analyse, welche Varianzquellen aufzudecken hat, ebenso zeigen wie der Einfluss zunächst unbekannter Varianzquellen, oder auch nicht zeigen, wenn kein solcher Einfluss vorliegt. Im vorliegenden Fall konnte die Interpretation von Varimin- F_2 (Informanten) auf einen transparenten Einfluss zurückgreifen, die Interpretation von F_3 (Enthemmungs- vs. Hemmungsstörung) musste sich mit einer weniger transparenten inhaltlich-psychologischen Unterscheidung und Abstraktion begnügen.

3 Diskussion

Fünf Fragen zur Methodologie wurden behandelt, die sich mit der Einführung des Rotationsverfahrens Varimin ergeben. Zehn empirische Anwendungsbeispiele sollten helfen, sie zu klären.

Gegen die Auswahl der zehn Beiträge könnte man vorbringen, dass sie zu wenig Datensätze aus psychologischen Untersuchungen enthielt. Doch sollte die

Güte einer neuen Methode zunächst mit Daten geprüft werden, bei denen zweifelsfrei bestimmte Ergebnisse zu erzielen sind, sofern die Methode funktioniert. Wenn ein zu erwartendes Ergebnis nicht erzielt wird, soll man schlussfolgern dürfen, dass mit der Methode etwas nicht stimmt. Von dieser Prüfstrategie machen Methodiker unter den Psychologen m. E. generell zu wenig Gebrauch. Psychologische Daten, zumal einschlägige sprachliche Daten, deren Ergebnisse oft mehrdeutig bleiben, eignen sich zur Methodenprüfung kaum. Die Sprache hat zwar zur Differenzierung von Erlebens- und Verhaltensweisen eine Fülle an Wörtern und Wendungen hervorgebracht. Doch auf eine transparente sprachliche Semantik wie die der Verwandtschaftsnamen kann sich die Kommunikation über psychologisch relevante Sachverhalte nicht stützen.

Natürlich soll die Faktorenanalyse, wenn sie ihre methodischen Prüfungen bestanden hat, bei der Lösung inhaltlicher Probleme unserer Disziplin eingesetzt werden. Aus den inzwischen vorliegenden Ergebnissen der Methodenforschung lässt sich schon jetzt eine bedeutsame Schlussfolgerung voraussagen: Auch dem manifesten verbalen Urteilmaterial, das Psychologen zur Analyse einbringen, liegen auf latenter Ebene generell Komponenten- oder Merkmalstrukturen zugrunde. Die Merkmale oder Komponenten psychisch funktionaler Einheiten konkreteren Bedingungen sind nur viel schwerer zu identifizieren und voneinander zu separieren als die Konzeptseinheiten, die die Erfassung der Gegebenheiten unserer anschaulich gegliederten Lebenswelt unterstützen.²¹ Deshalb wird man sich bei der Interpretation von Varimin-Faktoren einer psychologischen Domäne nicht, wie gewöhnlich bei der Interpretation von Simple Structure-Faktoren, mit Faktorbenennungen begnügen dürfen, die unter den verwendeten *manifesten Variablen* (Items, Statements) ausgewählt werden oder mit diesen synonym sind. Man wird sich auf die Ebene der latenten Merkmale begeben und sich trennen müssen von Begriffen umgangssprachlich festgeschriebener psychologischer Konzepte, z.B. von Trait-Konzepten wie *conscientious*, *agreeable*, *open-minded* usw., weil nicht sie, sondern deren Varianzquellen gefragt sind. Man wird sich ähnlich verhalten müssen wie bei der Analyse der Verwandtschaftsdomäne, bei der man zur Entschlüsselung der Faktoren auf eine ‚latente‘ Ebene hinabsteigen und sich von den manifesten Begriffen *Mutter*, *Schwester*, *Tante* usw. trennen musste, weil nicht diese Begriffe selbst oder deren Cluster, sondern deren *distinctive features* gefragt waren.

Die empirischen Beispielfälle, die relativ transparente Domänen zur Analyse wählten, haben gezeigt, dass man Varimin-Faktorenlösungen durchaus deuten kann, obgleich die den manifesten Variablen zugrunde liegenden Varianzquellen komplexe Strukturen

darstellen. Die Deutungen waren für Variminlösungen weit befriedigender als für Varimaxlösungen. Mit Anwendung des minimalen Paarvergleichs, der bei schwierigen Variminfaktor-Deutungen vorgenommen werden kann, beschränkt man sich auf die Feststellung von Unterschieden zwischen gepaarten Variablen hinsichtlich nur eines Merkmals, was durchweg leichter fällt als das Auffinden von Gemeinsamkeiten bei unanalysierten Merkmalkomplexen, die durch Varimax gebildet werden.

Drei Einschränkungen sind indes nachzutragen. Erstens, die Anwendung des Minimalpaar-Vergleichs setzt voraus, dass im Datensatz geeignete Variablenpaare vorhanden sind. Oft gewinnt man sichere Deutungen eines Faktors erst mit mehreren Minimalpaaren, die verschiedene Merkmalprofile aufweisen. Da man dieses Desiderat bei der Variablenselektion im voraus meist nicht berücksichtigen kann, wird mitunter ein Faktor nicht sicher gedeutet werden können, weil minimale Paare fehlen. Die noch fehlenden Paare können sich erst mit einer repräsentativeren Variablen-Stichprobe einstellen.

Das Vorkommen minimaler Paare ist umso eingeschränkter, je weniger Variablen verwendet wurden und je mehr substantielle Faktoren extrahiert wurden, so dass der Bedarf an Kombinationen gegensätzlicher Faktorgewichte zur Bildung von Minimalpaaren nicht gedeckt ist. Deshalb empfiehlt es sich, bei der Zusammenstellung von Variablenstichproben für eine faktorielle Untersuchung einer Domäne die Zahl der vermuteten Varianzquellen eher etwas zu beschränken, aber innerhalb des gewählten Vermutungsbereichs Bedingungsvariationen großzügig zuzulassen.

Die zweite Einschränkung resultiert aus dem Umstand, dass die Varianzquellen, die mit Varimin gefunden werden, immer einen höheren Grad an Abstraktheit besitzen als die Begriffe, die man gewöhnlich zur Beschreibung von Simple Structure-Faktoren verwendet. Mit den Verwandtschaftswörtern lässt sich das andeuten: Das Merkmal ‚*Generation*‘, mit Varimin gewonnen, ist abstrakter als ein Begriff wie ‚*die Jüngsten und Ältesten in der Abstammungslinie*‘, der bei der Varimax-Analyse notwendig war, um Großeltern und Enkel unter einen faktoriellen Hut zu bringen. Man wird in der psychologischen Domäne mit der Identifizierung abstrakter Merkmale größere konzeptuelle Schwierigkeiten haben als mit den konkret typisierenden Begriffen, wie sie in der Simple Structure-Praxis verwendet werden. Auch unter *Möbeln* z. B. kann man sich nicht so schnell etwas vorstellen als unter *Kleiderschrank* und *Polstersessel*.²²

Schließlich ist die Minimalpaarbildung in der Regel bei Generalfaktoren nicht anwendbar. Die Generalfaktor-Ladungen der Variablen einer Domäne sind durchweg positiv, und sie weisen meist wenig Varianz

auf. Andererseits repräsentiert ein Generalfaktor in der Regel die für den Datensatz jeweils dominante Varianzquelle, an der ein Forscher in erster Linie interessiert ist und die meist transparenter ist als hinzukommende schwächere Varianzquellen. Man wird einen Generalfaktor bei der Analyse von z. B. psychophysiologischen Variablen selbstverständlich auf z. B. Aktivationsvarianz zurückführen und nicht etwa auf Intelligenzvarianz.

Vor allem bei Generalfaktoren ist indes immer auch daran zu denken, dass sich mit ihnen mehr als nur eine Varianzquelle ausdrücken kann. Zum Beispiel kann sich in einem Generalfaktor bei Intelligenztests *Intelligenz plus Leistungsehrgeiz* ausdrücken, bei Fragebögen kann sich im Generalfaktor die mit den Items erfragte Traitdisposition zusammen mit einer mehr oder weniger ausgeprägten Akquieszenzneigung ausdrücken. Über die relativen Anteile der an einem Faktor beteiligten ‚Teil-Varianzquellen‘ ist aus den Daten selbst oder aus ihrer formalen Repräsentation keine Information zu gewinnen. Die Komplexität latenter Bedingtheiten kann sich mit Varimin nur in dem Maße differenziert ausdrücken, wie sich die relevanten Varianzquellen mit *unterschiedlichen* Beiträgen an der Varianz der verwendeten Variablen beteiligen.

Dass sich die Interpretierbarkeit initialer komplexer Strukturen, wenn diese mehr als zwei Faktoren aufweisen, durch Varimin-Rotationen verbessert, haben einige Beispiele verdeutlicht. Mir ist bei zahlreichen faktoriellen Reanalysen kein Fall begegnet, bei dem das Ergebnis einer Varimin-Transformation *weniger* gut interpretierbar gewesen wäre als das Ergebnis der vorausgehenden Initiallösung.²³

Was die Interpretierbarkeit von *Simple Structure*-Strukturen betrifft, so ließ diese bei den ausgewählten Beispielen viel zu wünschen übrig. In der Praxis würde man bei manchen der hier verwendeten Datensätze auf Faktorenanalysen nicht zurückgreifen, man würde auf andere Verfahren, vornehmlich auf MDS (bzw. NMDS) oder MTMM oder Circumplex-Verfahren ausweichen. Der Vorzug des Komplexität-modellierenden Verfahrens ist unter anderem darin zu sehen, dass es sich zur Analyse der verschiedensten Datensätze eignet, während die Anwendbarkeit des Simple Structure-Verfahrens aufgrund seiner Modellermängel von vorne herein begrenzt ist.

Analysen mit Simple Structure-Rotationen können indes *auch* nützlich sein. Ihre Bedeutung besteht nur nicht in dem, was man bisher glaubte, dass mit ihnen ‚Dimensionen‘ der untersuchten Domänen aufgedeckt würden, sondern darin, dass sie die Variablen einer Domäne clustern, während gleichzeitig die clusterbildenden Basismerkmale verborgen bleiben. Nur wenn eine Fragestellung zu bearbeiten sein sollte, die das Clustern von kovariierenden Variablen erfordert,

ohne dass die Quellen ihrer Kovarianz interessieren, würden sich Varimax und die obliquen Rotationsverfahren – diese insbesondere – als hilfreich erweisen.

Varimax könnte darüber hinaus nützlich bleiben für den Fall, dass Variablen zur Analyse anstehen, von denen man vorweg weiß, dass ihnen nur *eine* Varianzquelle zugrunde liegt. Das wäre z. B. der Fall, wenn die mit Varimin aufgedeckten latenten Merkmale, z. B. durch Itembildung für Fragebögen, zu manifesten Variablen erhoben werden, wenn diese dann für eine Korrelationsstudie eingesetzt werden, und wenn die so gewonnene Matrix einer Faktorenanalyse übergeben wird.

Komplikationen können für Varimin-Untersuchungen auch dann entstehen, wenn a priori unbekannt ist, ob den zur Analyse ausgewählten Variablen tatsächlich eine größere Zahl von Varianzquellen zugrunde liegt. Bislang wurde davon ausgegangen, dass Variablen in aller Regel multifaktoriell bedingt sind. Dies aber schließt Ausnahmen nicht aus, und man weiß nicht, wie viele Ausnahmen man zulassen darf und wie sie als solche zu erkennen sind.

Von Ungewissheit bei der Deutung Varimin-rotierter Faktoren war schon die Rede bei der Behandlung der Frage, ob und wie *negative* Faktorladungen zu deuten sind. Ob bei den mit Plus- und Minusvorzeichen bipolar gekennzeichneten Strukturen funktional antagonistische Verhältnisse vorliegen oder ob sich in den Vorzeichen lediglich das Vorhandensein bzw. Nichtvorhandensein einer latenten Bedingung ausdrückt, darüber sagt die Analyse selbst nichts aus, darüber entscheidet man mit einschlägigem Kontextwissen und den daraus sich ergebenden Mutmaßungen.

Bei der Verwendung des neuen Analyseverfahrens ist zu beachten – wie zuvor angedeutet –, dass die Variablen-Stichprobe für die Domäne, deren Kovarianzquellen differenziert werden sollen, repräsentativ ist. Die Variablen sollten von vorne herein als Ausdruck einer Mehrzahl latenter Bedingungen verdächtig sein, die unter den Domänvariablen Kovariationen hervorrufen. Zumindest sollte man Variablen nicht aufnehmen, die plausiblerweise nur wenig Aufschluss über Kovariationsquellen liefern. Würde man z. B. altersabhängige Variablen unter die Variablen eines Persönlichkeitsfragebogens mischen („ich werde immer vergesslicher“, „mir geht es gesundheitlich nicht so gut wie früher“), dann würden diese möglicherweise einen eigenen Faktor („Altersanzeichen“) bilden. Diese würden zwar mit Plusladungen einen Beitrag zur Varianz bekunden – wenn ein solcher vorliegt –, sie würden aber bei den Variablen, die mit Alter nichts zu tun haben, vermutlich Minusladungen hervorbringen („Einfluss nicht vorhanden“). Mit aufschlussarmen Minus-Ladungen bei diesen Variablen würde Kommunalität ‚verschwendet‘ werden.²⁴ Re-Analysen

mit Varimin-Rotation führen mitunter durchaus auch zu unbefriedigenden Ergebnissen. Diese sind jedoch meist auf unzureichende Variablen-Stichproben zurückzuführen. Mithilfe eines Measures of Sampling Adequacy (Kaiser, 1970) ließe sich ggf. die Adäquatheit von Datensätzen für exploratorische Faktorenanalysen routinemäßig ermitteln.

Zur weiteren Agenda gehört die Klärung des Invarianzproblems. Die Frage bleibt zu klären, ob das Varimin-Verfahren zu Lösungen führt, die gegenüber Hinzufügungen zur Variablenstichprobe und gegenüber Streichungen aus dieser oder auch gegenüber einem Wechsel der untersuchten Personenstichproben stabil genug sind. Eine nicht-publizierte Untersuchung (Ertel, b) ergab deutlich mehr Stabilität für Varimin als für Varimax beim Hinzufügen einer artifiziellen Variablen zu einer zuvor analysierten Stichprobe naturalistischer Variablen.²⁵

Zu den noch zu bearbeitenden Fragen gehört auch die nach der optimalen Faktoren-Extraktionsmethode. Im bisherigen Projekt wurde ausschließlich die Hauptkomponentenmethode (PCA) verwendet, weil sie sich in der faktorenanalytischen Forschung zum Standard entwickelt hat. Es mag sein, dass das Einsetzen von Einsen in die Diagonalfelder der Korrelationsmatrizen, das zur PCA gehört, auf die Endergebnisse u. U. eine leicht verzerrende Wirkung hat, vor allem dann, wenn überwiegend niedrige Interkorrelationen vorliegen. Zu den weiteren Aufgaben der Methodenforschung gehört, die Hauptachsen- und das Maximum-Likelihood-Verfahren vergleichsweise auf ihren Beitrag zur Modelliergüte zu prüfen und ggf. weitere Verbesserungen des Vorgehens vorzunehmen, falls durch eine andere Wahl der Extraktionsmethode am Ergebnis noch etwas zu verbessern sein sollte.²⁶

Die vorliegende Arbeit beschränkte sich auf die Entwicklung von Varimin als einem Verfahren, das für die exploratorische Forschung eine neue Richtung zeigt. Nachdem sich das Verfahren in zahlreichen Prüfungen mit verschiedenen Datensätzen bewährt hat, wird sein Einsatz bei der systematischen Bearbeitung psychologischer Sachfragen zum Desiderat. Zwei abgeschlossene Projekte zur Frage der Varianzquellen der Persönlichkeit (Ertel, c) und der Intelligenz (Ertel, b) liegen vor, sie würden grünes Licht vonseiten innovationsbereiter Zeitschriften-Herausgeber und ihrer Manuskript-Begutachter begrüßen.²⁷

Abschließend sei ein Einwand bedacht, der sich richten könnte gegen die dem vorliegenden Ansatz zugrunde liegende Strategie, über formale Fragen der Datenverarbeitung (hier: Faktorenrotation) anhand von Ergebnissen ihrer Anwendung auf inhaltliche Problemfelder zu urteilen. Mit ‚weichen‘ Argumenten des alltäglichen Lebens, wie man sie abwertend nennen

könnte, würde man einem präzisen mathematischen Instrumentarium nichts anhaben können.

Diesem Einwand würde man entgegen halten dürfen, dass es in dieser Arbeit nur darum ging, einen einzigen Fehler der Methodiker zu korrigieren: die Dogmatisierung von *Simple Structure* als Modell für latente Bedingungen bei multivariaten Analysen. Dieses Modell ist ein formales, doch es wurde aus *nicht-formalen* Gründen eingeführt, denn die Behauptung seiner Notwendigkeit (simple structure – Transformation mache Faktorstrukturen erst interpretierbar) ergibt sich, wenn man die Güte der Methode auf die Inhalte der mit ihr gewonnenen Ergebnisse (auf deren ‚Interpretierbarkeit‘) bezieht.²⁸ Dieser Bezug erfordert psychologisches *Verstehen*, die Mitverwendung von Erkenntnis, in welcher sich neben bisherigen Ergebnissen wissenschaftlicher Forschung, die oft kontrovers bleiben, eine ungleich größere Menge an ‚weiche- ren‘ Welt- und Menschen-Erfahrungen aufgebaut haben, von denen die meisten im Leben erfolgreich genutzt werden. Thurstonianer geraten mit ihrer Begründung des faktoriellen Rotationsprinzips, mit der sie dem nicht-formalisierten ‚Verstehen‘ Konzessionen einräumen, in die Nachbarschaft anderer Forscher, die sich auch – vielleicht mehr noch als sie – dem ‚Verstehen‘ verpflichtet fühlen. Auf dem Forum, auf dem sich die Positionen der Modellformalisten mit denen der weniger streng und eng denkenden Forscher verträglich überschneiden, sollten Letztere ein Mitspracherecht geltend machen dürfen. Sie sollten auf mögliche Fehler der Formalisten achten dürfen, soweit sie deren Sprache verstehen, sollten die Fehler prüfen, die sie vermuten, sie durchdenken, wenn sie nicht wegzudenken sind, und, soweit es in ihren Kompetenzbereich fällt, auch helfen dürfen, sie zu korrigieren.

Anmerkungen

- 1 Die Faktorenextraktion durch PCA führt dazu, dass die Interkorrelationen der Initialfaktoren bzw. die Kongruenzindizes, mit denen man ihre Ähnlichkeit untereinander ermittelt, alle exakt Null ergeben. Wenn aber bei suboptimalen Variablen-Stichproben die latenten Komponenten, die die Korrelationen unter den manifesten Variablen herbeiführen, miteinander korreliert auftreten, führt der Zwang zur Null-Korrelation der Faktoren, die die Komponenten manifestieren sollen, zu einer gewissen Verzerrung dieser Manifestationsleistung. Diese soll durch die Faktorenrotation Varimin korrigiert werden, d.h. nach Anwendung von Varimin sind ≈ 0 -Inter-Kongruenzen der Faktoren feststellbar.
- 2 Die zusätzlich übliche Unterscheidung hinsichtlich Artikulationsmodus (plosiv, frikativ, nasal usw.) ist spezieller und kann mit der beabsichtigten Phonem-Stichprobe nicht ermittelt werden. Das [m] und das [f] sind beides Phoneme mit längerer Artikulationsdauer (im Unterschied z. B. zu den kurz artikulierten plosiven Lauten [b], [p], [t], [k]), doch [m] gehört nach seinem Artikulationsmodus zu den nasalen Lauten, [f] zu den frikativen Lauten. Da alle plosiven Phoneme kurz sind, hätte man auch eine bipolare Kategorie ‚plosive‘ vs. ‚nicht-plosive‘ Phoneme bilden und diese als Faktor im Beurteilungsversuch erwarten können.
- 3 Die hier bevorzugte Forschungsstrategie bei der faktorenanalytischen Arbeit lässt sich unmittelbar verknüpfen mit theoretischen und empirischen Ansätzen, in denen Begriffe und andere kognitiv repräsentierte Objekte als Bündel oder Strukturen mehr oder weniger latenter Komponenten aufgefasst werden (‚Komponentenmodell‘ von Einstellungsobjekten, Feger, 1971, ‚Feature Pattern Analysis‘, Feger & Brehm, 2001). Auch läuft sie in die gleiche Richtung wie die Bemühungen um das Phänomen des Ähnlichkeitserlebens, zu dessen Erklärung bei Shepard, 1980, ‚hidden structures‘ und bei Tversky, 1977, ‚collections of features‘ analysiert wurden.
- 4 Umgangssprachlich ersetzt oft *one pence* das eigentlich richtige *one penny*.
- 5 Die leicht variierenden Ladungshöhen bei F_2 (am meisten fällt das 2 Pence-Stück aus der Reihe) könnten mit münzspezifischen subjektiven Gewichtungen der Merkmale zusammenhängen. Tatsächlich wirkt das 2-Pence Stück angesichts seines geringen Wertes übermäßig groß, so dass bei ihm im Vergleich mit anderen Münzen wahrscheinlich der Größeneindruck stärker als der Formeindruck zur Geltung kommt.
- 6 Die Interpretation von F_2 als Faktor der anschaulichen Münzengröße wird gestützt durch Korrelationen der F_2 -Ladungen mit der gemessenen Münzengröße: $r = .81$ (T. S.), $r = .85$ (Studenten). Das Gewicht der Münzen in Gramm, das mit ihrer Größe in cm bei $r = .80$ korreliert, korreliert mit den Faktorladungen nur geringfügig.
- 7 Zimmermann (1953) bemerkte, dass Variablen, die in der unrotierten Lösung mit ähnlichem Profil mehrfaktoriell geladen sind, sich durch Rotation zur Einfachstruktur in einem Faktor vereinigen: „... a test which actually contains variance on two or more factors may appear with all of that variance confined on a single factor“. Diesen nennt er ‚composite factor‘. Der Autor hält also an

- der inhaltlichen Bedeutung der initialen Faktoren (Centroidfaktoren) fest, die er im rotierten Faktor als Faktorenbündel („composed“) vereint sieht. „*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, 389).
- 8 Dass die Rotation zur Einfachstruktur nicht zur Analyse zugrunde liegender Dimensionen, sondern nur zur Clusterung ähnlicher Testvariablen führt, wurde explizit von Overall (1964) nicht nur behauptet, sondern anhand eines Analysebeispiels demonstriert: „*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). „... *there is no need to assume that simple structure factors will correspond to any particular set of fundamental dimensions of the objects ...*“ (p. 276).
 - 9 Die Probanden hatten 16 Verwandtschaftswörter nach selbst gewählten Kriterien zunächst in zwei beliebig große Teilgruppen zu teilen, dann jede der beiden Teilgruppen wieder in zwei Teile usw., bis jedes Wort an einem unteren Ende in der Hierarchie vereinzelt übrig blieb. Die Ähnlichkeit zweier Worte für einen Probanden wurde definiert als die Anzahl der Schritte, die nötig war, um aus der isolierten untersten Position der beiden zu vergleichenden Wörter in der Gruppierungshierarchie aufsteigend einen ersten gemeinsamen Knoten zu erreichen. Tabelle 2 enthält die über die Probanden summierten ‚Schritte bis zum ersten gemeinsamen Knoten‘.
 - 10 Diese lineare Transformation wurde lediglich zum leichteren Verständnis der aus der Tabelle abgelesenen Einzelwerte vorgenommen.
 - 11 Eigenwerte: 4.31, 2.34, 1.54, 1.09, 1.05, 0.90, 0.79, 0.75...
 - 12 Man könnte F_1 als Ausdruck der Verwandtschaftskategorie deuten. Seine differenzierende Funktion würde faktoriell vermutlich erst in Erscheinung treten, wenn man in dieser Untersuchung zusätzlich Wörter einer anderen Kategorie verwendet hätte.
 - 13 Die Interpretationen der Varimin-Faktoren F_2 , F_4 und F_5 sind auch in einer vergleichbaren MDS-Untersuchung von Romney & D’Andrade (1964) anzutreffen (gender, generation, consanguinity). Die vorliegende Untersuchung differenziert noch innerhalb F_2 (Linealität) zwischen ‚Kernfamilie‘ und ‚Nicht-Kernfamilie‘ (durch F_3). Die Untersuchung von Romney & D’Andrade bietet ansonsten einen vorzüglichen Einstieg in die begrifflichen und methodischen Grundlagen der ‚componential analysis‘ von Begriffen und ein Verständnis für die Schlussfolgerungen, die man aus der Existenz der ‚discriminative stimuli‘ (der ‚Sememe‘) für den Begriffserwerb ziehen kann.
 - 14 Es wäre denkbar, dass Probanden, die viele extreme Ablehnungen abgeben, auch viele Unentschieden-Urteile oder viele gemäßigte Zustimmungen abgeben, was motivational aber nicht nachvollziehbar wäre.
 - 15 Aus diesem Grund sind multivariate Kausalmodelle, wie sie mit Methoden der konfirmatorischen Faktorenanalyse und mit Strukturgleichungsverfahren Eingang fanden, als fragwürdig zu betrachten. Die Modellierung durch multivariate Analysen zielt in erster Linie auf multiple Quellen der Kovarianz manifester Variablen ab. Das Zusammenwirken der latenten Quellen lässt sich nicht ohne weiteres als ein kausales verstehen.
 - 16 Da die ‚Zuweisung‘ von positiven und negativen Vorzeichen durch das Rotationsprogramm vom zugeordneten Inhalt der Messungen unabhängig erfolgt, ist auch darüber zu entscheiden, ob die initial gegebenen Vorzeichen beibehalten oder der jeweiligen Deutung entsprechend ggf. umgekehrt werden sollten. Die Variablen, bei denen der Faktor funktional hemmenden Einfluss ausdrückt bzw. bei denen das faktoriell repräsentierte Merkmal als nicht vorhanden zu interpretieren ist, können zugunsten eines leichteren Verstehens ein negatives Vorzeichen erhalten. Es hat sich gezeigt (bislang eine Daumenregel), dass das Nichtvorliegen eines Merkmals oder der negative Einfluss durch eine Varianzquelle innerhalb eines faktoriellen Vektors tendenziell eher bei denjenigen Variablen vorkommt, die niedrigere Ladungen haben und deren Vorzeichen (Plus oder Minus) überwiegen.
 - 17 Ein aus der Frühzeit der Centroid-Faktorenanalyse stammender Hinweis, der den unrotierten bipolaren Faktoren einen Informationswert einräumt, der damals schon vernachlässigt wurde, findet sich in einer Notiz von 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).
- 18 Die Simple Structure-Modellierung psychologischer Daten hat vielfach zur Auflösung bipolarer Konstrukte geführt: Urteile über negativ getönte Gefühle scheinen nach faktorieller Einfachstruktur-Modellierung statistisch unabhängig zu sein von Urteilen über positiv getönte Gefühle und sollen dann entgegen aller Lebenserfahrung generell so erlebt werden (Diener & Emmons, 1985, Watson & Clark, 1988, späte Korrektur nach Modellvergleichen durch Crawford & Henry, 2004). Ähnlich zusammenhangslos erscheinen in Faktorenanalysen mit Einfachstruktur-Rotation Optimismus und Pessimismus, die normalerweise bipolar sich auswirkenden Einstellungen zur Zukunft (Marshall et al., 1992). Auch geht die Polarität der Geschlechter-Typologie durch Faktorenanalysen verloren: Femininität und Maskulinität soll aufgrund faktorieller Orthogonalität nicht polar aufeinander bezogen sein (Bem 1981).
 - 19 Man könnte die Tests mit zwei alternativen Zusatzinstruktionen durchführen lassen, eine Speed-Instruktion würde die Probanden generell zu mehr Tempo ermuntern, eine Power-Instruktion zu mehr bedachtsamer Konzentration. Vermutlich wird man je nach Testart differenzielle Steigerungen bzw. Verminderungen ihrer Leistung beobachten.
 - 20 Die Eigenwertkriterien der Mindesthöhe (>1) bzw. der Eigenwertknick (zu ermitteln durch den Scree Plot) sind schon von anderer Seite als unzuverlässig erkannt worden (Fabrigar, 1999, p. 287). Der u.a. auch von Fabrigar empfohlene ‘Paralleltest’ ist mit den vorliegenden Korrelationsdaten nicht durchführbar. Nach eigenen Erfahrungen mit faktoriellen Reanalysen von Datensätzen haben sich hier als informativ erwiesen die Prozentanteile, die die Initialfaktoren an der Summe der Kommunalitäten haben. Ein Faktor, dessen Anteil an der Summe h^2 weniger als 10% beträgt, ist in der Regel nicht mehr interpretierbar. Im vorliegenden Datensatz von Campbell & Fiske betragen die Kommunalitätsprozente der vier Initialfaktoren 66.7, 15.4, 11.5, 6.7. Wenn ein Mindestanteil von 10% Kommunalität als Kriterium angesetzt wird, sind F_2 und F_3 noch zu berücksichtigen.
 - 21 „*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*“ (1984, Just & Carpenter, p. 63 f).
 - 22 In einem Projekt mit Reanalysen von NEO PI-R-Daten, das mithilfe der Varimin-Transformation von Faktoren zu einer neuen Konzeptualisierung von Persönlichkeitskomponenten führen sollte (Ertel b), wurden die Begriffe ‘Aktivationslevel’, ‘Aktivationsgradient’, ‘Regulationsquelle’, ‘Repräsentationsmodus’, ‘Funktionsbilanz’ zur Deutung der Varimin-Faktoren vorgeschlagen. Abstrakte Begriffe dieser Art liegen nicht immer fertig vor, sie sind schwer zu bilden, doch sind sie leichter miteinander verknüpfbar.
 - 23 Die Gesamtzahl der Datensätze, an denen das Varimin-Programm erprobt wurde, liegt inzwischen bei über 500.
 - 24 Die Faktorladung einer Varianzquelle, die bei einer Variablen nur ein ‘nicht vorhanden’ aussagt, wird dennoch bei der betreffenden Variablen als Beitrag zu ihrer Kommunalität mit aufgenommen eine Art Informationsverschwendung. Dies ist, zumal wenn es sich um eine Varianzquelle handelt, die nur für eine Minorität von Variablen als ‘vorhanden’ relevant ist.
 - 25 Einen Mangel an Invarianz von Simple Structure-Lösungen beklagte schon 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)
 - 26 Dies ist indessen nach bislang vorliegenden Vergleichen der Extraktionsverfahren kaum zu erwarten: „*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).
 - 27 Der weiteren faktorenanalytischen Forschung mit Varimin durch Interessenten stelle ich auf Anfrage das von Dr. Uwe Engeland in meinem Auftrag erstellte Varimin-Varimax-Programm faktor.exe zur Verfügung, wobei eine kurze Beschreibung des Ziels der jeweiligen Untersuchung, in der das Programm eingesetzt werden soll, möglicherweise hilfreichen Empfehlungen von mir nach sich ziehen. Was seine Nutzerfreundlichkeit angeht, ist faktor.exe überarbeitungsbedürftig. Derzeit sind für seine erstmalige Nutzung noch beträchtliche Instruktionen erforderlich.
 - 28 „... *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, 6). „The arguments in favor 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).

Literatur

- Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review*, 88, 354-364.
- Burroughs, G. E. R. & Miller, H. W. L. (1961). The rotation of principal components. *The British Journal of Statistical Psychology*, 14, 35-49.
- Butler, J. M. (1969). Simple structure reconsidered: Distinguishability and invariance in factor analysis. *Multivariate Behavioral Research*, 4, 5.
- 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.
- Comrey, A. L. (1978). Common methodological problems in factor analytic studies. *Journal of Consulting and Clinical Psychology*, 46 (4), 648-659.
- 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.
- Diener, E. & Emmons, R. A. (1983). The independence of positive and negative affect. *Journal of Personality and Social Psychology*, 47, 5, 1105-1117.
- Ertel, S. (2009). Kritik am Dogma „Simple Structure“. *Journal Psychologie des Alltagshandelns*, 1, Vol. 2, 36-63.
- Ertel, S. (a) Intelligenz (g) und Lernkapital. Komplexstrukturelle Faktorenanalyse des I-S-T (unveröffentlicht).
- Ertel, S. (b) Faktorielle Komplexstruktur durch Varimin-Rotation. Manifestation simulierter Zusatzvarianz in I-S-T-Daten. (unveröffentlicht).
- Ertel, S. (c) Varianzquellen der Persönlichkeit. Exploratorische Faktorenanalyse mit komplextheoretischer Ausrichtung (unveröffentlicht).
- 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. (1971). 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.
- Gorsuch, R. L. (1974) *Factor analysis*. Philadelphia: Saunders.
- Guilford, J. P. (1952). When not to factor analyze. *Psychological Bulletin*, 49, 26-37.
- Guilford, J. P. & Hoepfner, R. (1969). Comparison of varimax rotations with rotations to theoretical targets. *Educational and Psychological Measurement*, 29, 3-23.
- Harman, H. H. (1967). *Modern factor analysis*. Chicago: The University of Chicago Press.
- Herrmann, T. (1965). *Psychologie der kognitiven Ordnung*. Phänomenologisch-Psychologische Forschungen. In C. F. Graumann & J. Linschoten (Eds.), (Vol. 6). Berlin: Walter de Gruyter & Co.
- Hinz, A., Brähler, E., Geyer, M., & Körner, A. (2003). Urteileffekte beim NEO-FFI. *Diagnostica*, 49, 157-163.
- 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.
- Humphreys, L. G. & Davey, T. C. (1988). Continuity in intellectual growth from 12 months to 9 years. *Intelligence*, 12, 183-197.
- Jacobson, R. & Halle, M. (1956). *Fundamentals of Language*. Den Haag: Mouton.
- Jöreskog, K. G. & Sörbom, D. (2003) LISREL 8.52. *Website*: www.ssicentral.com.
- Just, M. A. & Carpenter, P. A. (1984). *The psychology of reading and language comprehension*. Boston: Allyn and Bacon.
- Kaiser, H. F. (1970). A second generation Little Jiffy. *Psychometrika*, 35, 401-416.
- Köhler, T. & Troester, U. (1991). Changes in the palmar sweat index during mental arithmetic. *Biological Psychology*, 32, 143-154.
- Marx, A. & Heij, A. (1989). *Subjektive Strukturen. Ergebnisse aus der Sprach- und Einstellungsfor-schung*. Göttingen: Hogrefe.

- 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 (6), 1067-1074.
- Matson, J. L. & Nieminen, G. S. (1987). Validity of measures of conduct disorder, depression, and anxiety. *Journal of Clinical Child Psychology*, 16 (2), 151-157.
- Overall, J. E. (1964). Notes on the scientific status of factors. *Psychological Bulletin*, 61, 270-276.
- 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 (3), 287-297.
- Romney, A. K. & D'Andrade R. G. (1964). *Transcultural Study of Cognition. Special publication of the American Anthropologist*. Vol. 66, Part 2.
- Shepard, R. N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science*, 210, 390-398.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 6, 327-348.
- Überla, K. (1971). *Faktorenanalyse. Eine systematische Einführung für Psychologen, Mediziner, Wirtschafts- und Sozialwissenschaftler*. 2. Auflage, Berlin: Springer.
- Velicer, W. F. & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, 25 (1), 1-28.
- 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 (6), 1063-1070.
- Wellek, A. (1959). *Die Polarität im Aufbau des Charakters*. Bern: Francke.
- Zimmermann, W. S. (1953). A note on the recognition and interpretation of composite factors. *Psychological Bulletin*, 50, 387-389.
- Korrespondenz-Adresse:
 Prof. Dr. em. Suitbert Ertel
 Georg-Elias-Müller-Institut für Psychologie
 Georg-August-Universität Göttingen
 Gossler Straße 14
 D-37073 Göttingen
 sertel@uni-goettingen.de