6. Faktorenanalyse (FA) von Tests

Das Prinzip der Faktorenanalyse besteht darin, die in einem Satz von Variablen enthaltene Information auf eine möglichst geringe Zahl von hypothetischen Dimensionen oder Faktoren zu reduzieren. Es ist damit ein Verfahren zur Datenreduktion.

Beispiel:
Ein Untersucher hat ein Intelligenztest entwickelt, der aus 8 Subtests zur Erfassung von Intelligenzkomponenten besteht. Der Test wird einer grossen Personenstichprobe vorgegeben. Die Leistungen in den Subtests werden miteinander korreliert. Es entsteht eine Korrelationsmatrix (**R**), die die Beziehung zwischen den Variablen wiederspiegelt.

Tabelle 7: (Diehl & Kohr, 1989, Tabelle 48)

Tabelle 7: (Diehl & Kohr, 1989, Tabelle 48)

Tabelle 48 : Korrelationsmatrix für acht Subtests eines Testverfahrens zur Erfassung numerischer und verbaler Intelligenz[1]

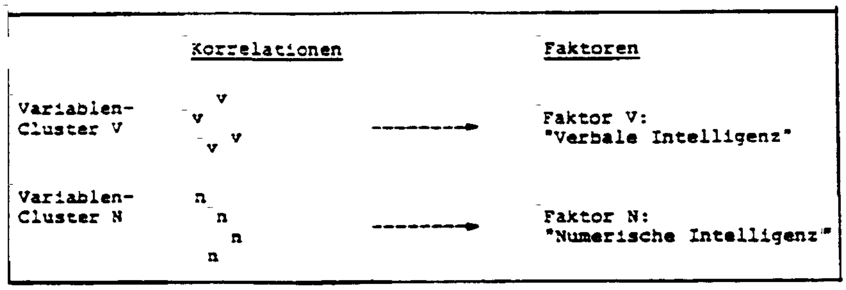
Subtest	"Numerische Subtests"				"Verbale Subtests"			
	GR	SCH	ZN	ER	GW	AL	SP	WG
	A							B
GR	1.00	.44	.34	.44	-.05	.02	.13	.14
SCH	.44	1.00	.44	.44	.04	.32	.27	.08
ZN	.34	.44	1.00	.48	.06	.34	.05	.30
ER	.44	.44	.48	1.00	-.05	.16	.19	.15
	C							D
GW	-.05	.04	.06	-.05	1.00	.45	.48	.27
AL	.02	.32	.34	.16	.45	1.00	.24	.15
SP	.13	.27	.05	.19	.48	.24	1.00	.20
WG	.14	.08	.30	.15	.27	.15	.20	1.00

Anmerkung: Aus *Deskriptive Statistik* (S. 341) von J. M. Diehl & H. U. Kohr, 1989, Eschborn bei

Der Untersucher zeichnet die Variablen als Punkte in eine Ebene. Wenn zwei Variablen hoch miteinander korrelieren, liegen sie nah beieinander. Er entdeckt zwei Variablencluster, wobei in einem Cluster hauptsächlich verbale Fähigkeiten und im anderen Cluster eher numerische Fähigkeiten verlangt werden. Er postuliert zwei relativ unabhängige Faktoren.

Abbildung 13: Variablencluster im Raum (Diehl & Kohr, 1989, Abbildung 54)

Abbildung 13: Variablencluster im Raum (Diehl & Kohr, 1989, Abbildung 54)



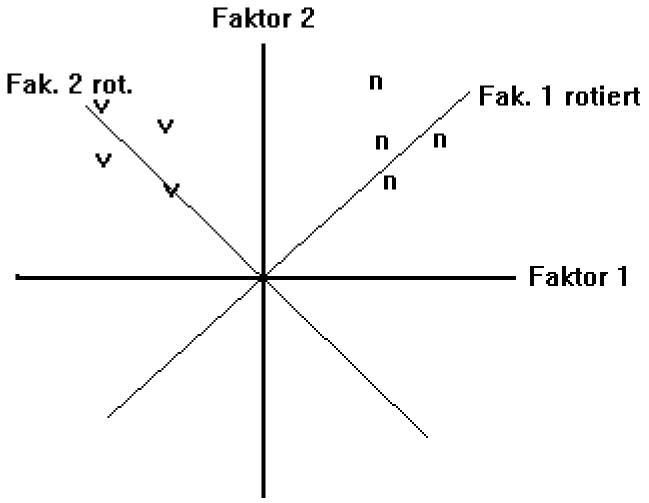
***Anmerkung*: Aus *Deskriptive Statistik* (S. 340) von J. M. Diehl & H. U. Kohr, 1989, Eschborn bei Frankfurt am Main: Klotz. Copyright 1977 bei Verlag Dietmar Klotz GmbH. Wiedergabe mit Genehmigung.**

Der Untersucher entwickelt also die Hypothese: "Es handelt sich um je vier Verfahren (Subtests) zur Erfassung der Faktoren numerische und verbale Intelligenz". Die Prüfung der Hypothese erfolgt durch die Faktorenanalyse (FA), die die zwei Faktoren nachweisen müsste und die Interkorrelation der Variablen müsste durch die Faktoren hinreichend erklärt werden.

6.1. Grundzüge der FA nach der Haupkomponentenmethode (PCA) mit anschliessender VARIMAX-Rotation:

In der Hauptkomponentenmethode (principal component analysis: PCA) sollen die Korrelationen der Variablen durch möglichst wenige Faktoren erklärt werden (Faktorenproblem: Extraktion und Anzahl). Die Varimax-Rotation rotiert die Faktoren orthogonal (rechtwinklig) so, dass ein Cluster möglichst durch einen Faktor erklärt wird, wodurch die Interpretation leichter möglich wird:

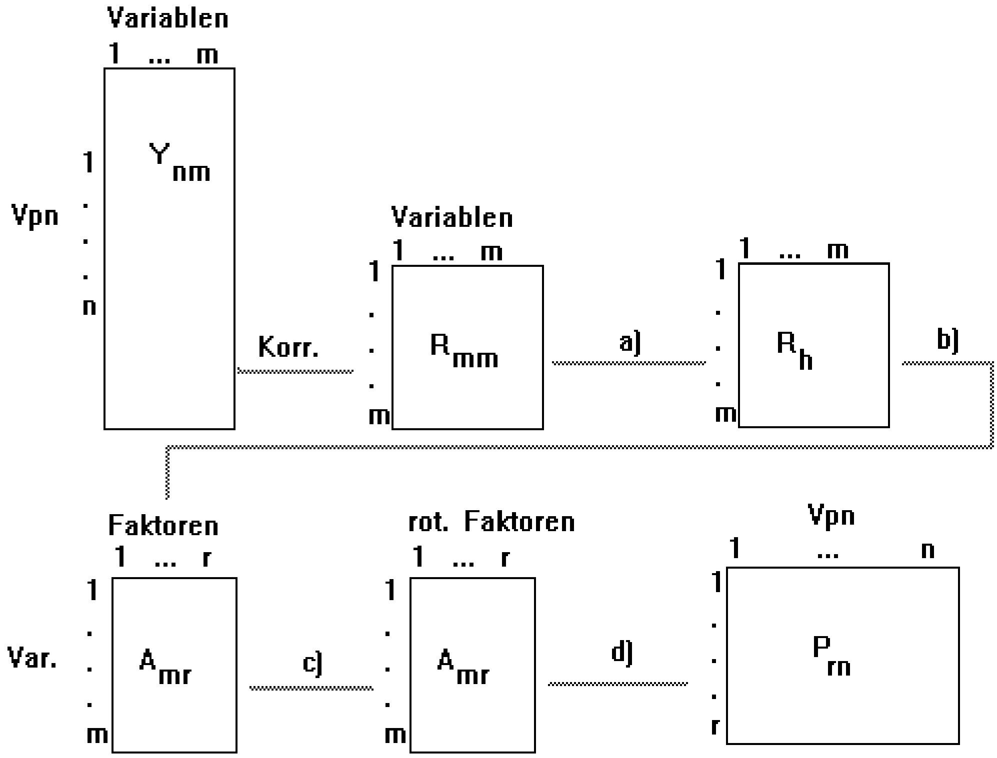
Abbildung 14: Varimax-Rotation



Legende: Faktor 1 und Faktor 2 bestimmen den Faktorenraum, indem die Variablen aufgezeichnet waren. Mit v sind die Variablen oder Aufgaben bezeichnet, die verbale Fähigkeiten verlangen. Mit n sind die Variablen oder Aufgaben bezeichnet, die numerische Fähigkeiten verlangen. Die Faktoren werden durch Rotation in die Variablencluster gelegt.

6.2. Die Matrizen der FA

Abbildung 15: Die Matrizen der Faktorenanalyse



Legende: Die Kasten stellen die Matrizen im Ablauf der Faktorenanalyse dar. Es werden n Personen und m Variablen untersucht. Aus den m Variablen werden r Faktoren extrahiert. a) bis d) bezeichnet die Stellen an denen Probleme auftreten, für die es unterschiedliche Lösungen gibt.

Die FA geht von der bekannten normalen Datenmatrix (Vpn x Variablen) aus. Es wird die Korrelation berechnet, um so die Interkorrelation der Variablen ($\mathbf{R_{mm}}$) zu erhalten.

a) Kommunalitätenproblem: Es wird die gemeinsame Varianz, ohne die spezifische und ohne die Fehlervarianz gesucht (vgl. Abschnitt 6.4).

b) Faktorenproblem: Hier müssen die Variablencluster durch Faktoren ersetzt werden. Probleme sind die Extraktion (Auswahl) und die Anzahl der Faktoren.

c) Rotationsproblem: Versucht wird von den mathematisch erhaltenen Achsen auf inhaltliche zu rotieren.

d) Faktorenwerteproblem: Es soll nun wieder ein Rückschluss auf die Personen erfolgen. So kann man aussagen, welchen Wert eine Person auf einem bestimmten Faktor besitzt.

6.3. Die Faktorenladungsmatrix "A_{mr}"

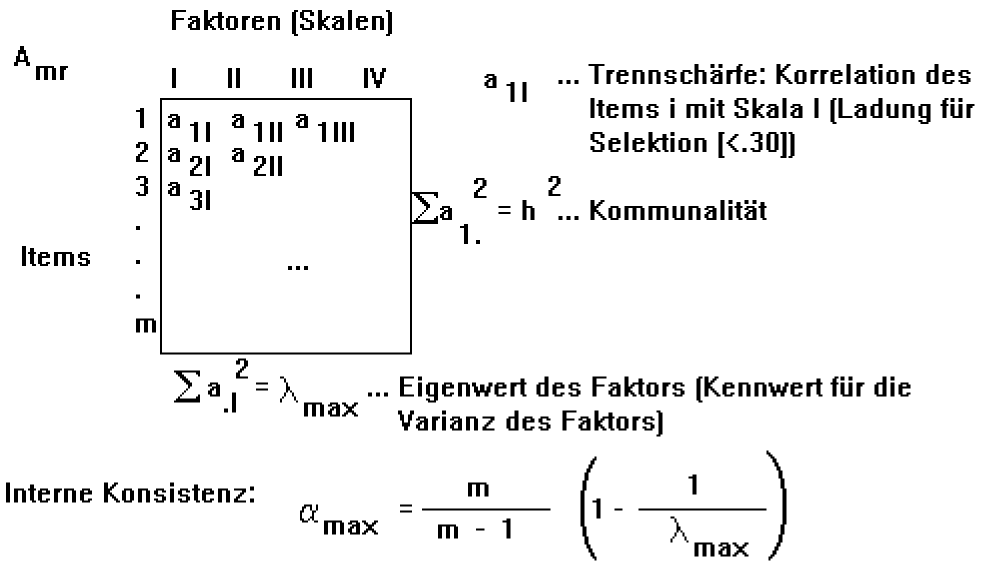
Sie zeigt den Zusammenhang zwischen den Variablen und den Faktoren. Je höher die Ladung einer Variablen auf einem Faktor, desto grösser ist die Bedeutung der Variablen für den Faktor. Diese Matrix wird von den Computerprogrammen ausgegeben, denn sie hat für die Interpretation eine zentrale Bedeutung. Im folgenden ist die rotierte Ladungsmatrix für das Beispiel dargestellt.

Tabelle 8: (Diehl & Kohr, 1989, Tabelle 51)

	Kürzel	Faktor I	Faktor II
Grundrechnen	GR	0,74	-0,04
Gleiche Wortbedeutung	GW	-0,17	0,87
Analogien	AL	0,20	0,68
Schätzen	SCH	0,73	0,22
Sprichwörter	SP	0,11	0,69
Zahlenreihen	ZN	0,72	0,21
Eingekleidetes Rechnen	ER	0,79	0,06
Wortgewandtheit	WG	0,21	0,45

Faktor 1 bildet eher die numerische Komponente der Intelligenz ab, während auf Faktor 2 hauptsächlich verbale Subtests laden.
Eine unrotierte Faktorenladungsmatrix beinhaltet einige Konzepte der klassischen Testtheorie, wie Abbildung 16 zeigt.

Abbildung 16: Unrotierte Faktorenladungsmatrix und die klassische Testtheorie



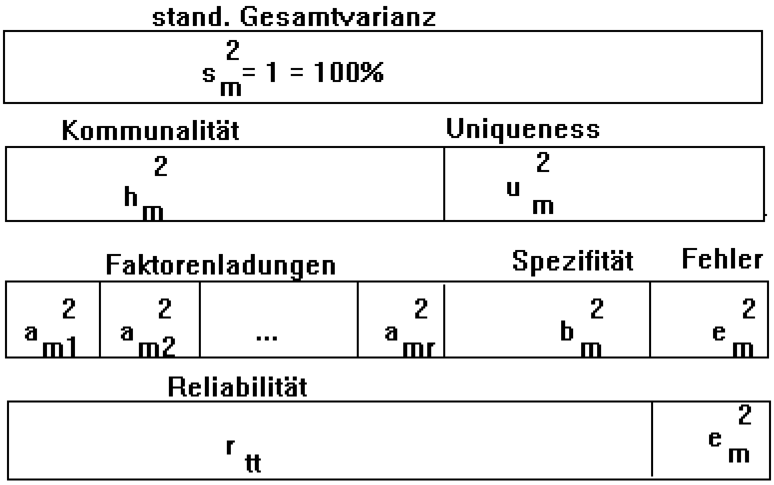
Legende: Aus der unrotierten Faktorenladungsmatrix $\mathbf{A}_{mr}$ können die Trennschärfekoeffizienten a_{mr} direkt abgelesen werden. Die Zeilensumme der quadrierten Ladungen ($\Sigma a_{1.}^2$) ergibt die Kommunalität. Die Spaltensumme der quadrierten Ladungen ($\Sigma a_{.I}^2$) ergibt den Eigenwert des Faktors. Der Eigenwert des Faktors geht in die Berechnung der internen Konsistenz mit ein

$$\left(a_{max} = \frac{m}{m-1} * \left(1 - \frac{1}{l_{max}}\right)\right).$$

6.4. Der Zusammenhang zwischen Kommunalität und Reliabilität

Nach dem Modell der FA teilt sich die Variablenvarianz wie in Abbildung 17 auf:

Abbildung 17: Die Variablenvarianz in der Faktorenanalyse



Legende: Die Kasten stellen die Variablenvarianz dar. Die Gesamtvarianz unterteilt sich in Kommunalität und Uniqueness. Die Kommunalität setzt sich aus der Summe der Ladungsquadrate zusammen. Die Uniqueness unterteilt sich weiter in Spezifität und Fehler. Die Reliabilität ist aus Kommunalität und Spezifität zusammengesetzt.

Die standardisierte Gesamtvarianz setzt sich additiv zusammen aus Kommunalität (Varianz, die auf gemeinsame Faktoren zurückgeht) plus Uniqueness (Varianz, die nicht auf gemeinsame Faktoren zurückzuführen ist). Wie oben gezeigt ist die Kommunalität die Summe der Ladungsquadrate. Die Varianz, die nicht auf die gemeinsamen Faktoren zurückzuführen ist, lässt sich weiter unterteilen in die Spezifität (Varianz, die auf spezifische Faktoren beruht) und den Fehleranteil (Fehler der Messung). Die Reliabilität war definiert als wahre Varianz geteilt durch die Gesamtvarianz. Da die Gesamtvarianz 1 ist, ist die Reliabilität gleich der Kommunalität plus der Spezifität. Die Kommunalität kann somit als untere Schätzung der Reliabilität verwendet werden.

6.5. Voraussetzungen der FA

Die Stichprobe sollte mindestens 100 Vpn enthalten. Die Zahl der Vpn sollte grösser als die der Variablen sein (Verhältnis c.a. 3:1). Diese Voraussetzungen beeinflussen nicht den Rechenablauf und werden deshalb in der Praxis kaum beachtet. Man muss allerdings bedenken, dass die Ergebnisse für wenig Personen sehr instabil sind. Zudem sollte die FA nicht unbedingt mit dichotomen (0/1) Variablen durchgeführt werden, da der Produkt-Moment-Korrelationskoeffizient nicht entsprechend variieren kann, wie bei kontinuierlichen Variablen.

Die FA ist für die KTT ein wichtiges Verfahren, da es eine Möglichkeit zur Konstruktvalidierung bietet. Zudem sei noch einmal auf die Multitrait-Multimethod Matrix von Campbell & Fiske (1959) verwiesen.

6.6. Uebungsaufgaben zur Faktorenanalyse

1. Folgende Ladungsmatrix sei gegeben:

$$A = \begin{pmatrix} .80 & .30 \\ .15 & .65 \\ .90 & .01 \\ .38 & .78 \end{pmatrix}$$

Reliabilität der Items:

1: .80
2: .50
3. .85
4. .80

a) Berechne die Spezifität und Fehlervarianz jedes Items.
b) Berechne die wahre Varianz eines jeden Items.

2. Die klassische Testtheorie und die Faktorenanalyse enthalten folgende Konzepte, die man aufgrund ihres Bedeutungsgehaltes als nahezu identisch zuordnen kann:

- ❑ Rotation und Messfehler
- ❑ Einfachstruktur und differentielle Vorhersagbarkeit
- ❑ Faktorenladung und Trennschärfe
- ❑ Hauptachsen und Testlänge
- ❑ Reliabilität und Kommunalität

3. Welche der folgenden Aussagen trifft für die Methode der Faktorenanalyse <u>nicht</u> zu.

- ❑ Die Aufgabe der Faktorenanalyse besteht in der Ordnung und Beschreibung einer Datenvielfalt.
- ❑ Die Aufgabe besteht darin, die gemeinsamen Quellen der Variabilität einer beobachteten Vielfalt von Merkmalen aufzuklären.
- ❑ Die Aufgabe besteht in der Bestimmung der Korrelation von Variablenpaaren.
- ❑ Die Aufgabe besteht in der Reduktion mehrerer gemessener Variablen auf einige gemeinsame Faktoren.