

# Forschungspraktikum

## Gruppenbezogene Menschenfeindlichkeit

### Hauptkomponenten- und Faktorenanalyse

- Hauptkomponentenanalyse
- Von der Hauptkomponentenanalyse zur exploratischen Faktorenanalyse
- Bestimmung der Zahl der Faktoren
- Rotation der Ausgangslösung

# Hauptkomponentenanalyse

Bei der Reliabilitätsanalyse von Summenscores wird die Reliabilität der Summe von Variablen (Items) betrachtet, die als mehrfache Messung der gleichen Eigenschaft aufgefasst werden. Es stellt sich die Frage, ob es nicht anderen Kombinationen der Items gibt, die die interessierende Eigenschaft besser repräsentieren als ein Summenscore.

Eine Antwort darauf bietet die Hauptkomponentenanalyse. In der Hauptkomponentenanalyse werden mehrere Variablen in dem Sinne zu optimalen linearen Indizes zusammengefasst, dass die Korrelationsmatrix der Variablen möglichst gut reproduzierbar ist

Wenn etwa eine Hauptkomponentenanalyse von 4 (standardisierten) Variablen  $Z_1$  bis  $Z_4$  durchgeführt werden soll, dann wird zunächst ein ebenfalls standardisierter linearer Index

$$H_1 = a_{11} \cdot Z_1 + a_{12} \cdot Z_2 + a_{13} \cdot Z_3 + a_{14} \cdot Z_4$$

so gebildet, dass die Regression der Variablen  $Z_1$  bis  $Z_4$  auf den Index  $H_1$  die Korrelationsmatrix bestmöglich reproduziert:

$$r_{ij} = r(Z_i, Z_j) \approx b_{i1} \cdot b_{j1} \cdot s^2(H_1) = b_{i1} \cdot b_{j1} = \hat{r}_{ij} \text{ mit } i, j = 1, 2, 3, 4$$

Die Gewichte  $a_{11}$  bis  $a_{14}$  und  $b_{11}$  bis  $b_{41}$  werden so festgelegt, dass diese Bedingung erfüllt ist. Der Index  $H_1$  ist die **erste Hauptkomponente** der Hauptkomponentenanalyse der Variablen  $Z_1$  bis  $Z_4$ .

# Hauptkomponentenanalyse

Zur Verdeutlichung wird die erste Hauptkomponente der Variablen ff03d, ff04d, ff05d, und ff08d berechnet.

Korrelationsmatrix

		ff03d Ausländer sind Belastung fürs soziale Netz	ff04d Es leben zu viele Ausländer in D.	ff05d Ausländerkind er verhindern gute Bildung der dt. Kinder	ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp
Korrelation	ff03d Ausländer sind Belastung fürs soziale Netz	1.000	.675	.388	.531
	ff04d Es leben zu viele Ausländer in D.	.675	1.000	.394	.589
	ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.388	.394	1.000	.381
	ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.531	.589	.381	1.000

Komponentenmatrix<sup>a</sup>

	Komponente
	1
ff03d Ausländer sind Belastung fürs soziale Netz	.838
ff04d Es leben zu viele Ausländer in D.	.861
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.646
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.799

Extraktionsmethode: Hauptkomponentenanalyse.

a. 1 Komponenten extrahiert

Ausgangspunkt ist die links wiedergegebene Korrelationsmatrix.

Die Komponentenmatrix enthält die Regressionsgewichte  $b_{i1}$  der Regressionen der standardisierten Ausgangsvariablen auf die erste Hauptkomponenten. Die Vorhersagegleichungen sind somit im Beispiel:

$$\hat{Z}_{f003d} = 0.838 \cdot H_1 ; \hat{Z}_{f004d} = 0.861 \cdot H_1 ; \hat{Z}_{f005d} = 0.646 \cdot H_1 ; \hat{Z}_{f008d} = 0.799 \cdot H_1$$

# Hauptkomponentenanalyse

Korrelationsmatrix

	ff03d Ausländer sind Belastung fürs soziale Netz	ff04d Es leben zu viele Ausländer in D.	ff05d Ausländerkind er verhindern gute Bildung der dt. Kinder	ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp
Korrelation	ff03d Ausländer sind Belastung fürs soziale Netz	ff04d Es leben zu viele Ausländer in D.	ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp
	1.000	.675	.388	.531
	.675	1.000	.394	.589
	.388	.394	1.000	.381
	.531	.589	.381	1.000

Komponentenmatrix<sup>a</sup>

	Komponente
	1
ff03d Ausländer sind Belastung fürs soziale Netz	.838
ff04d Es leben zu viele Ausländer in D.	.861
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.646
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.799

Extraktionsmethode: Hauptkomponentenanalyse.

a. 1 Komponenten extrahiert

Die erste Hauptkomponenten ergibt die folgenden *reproduzierten Korrelationen*:

$$\hat{r}_{21} = b_{21} \cdot b_{11} = 0.861 \cdot 0.838 = .721 ; \hat{r}_{31} = b_{13} \cdot b_{11} = 0.646 \cdot 0.838 = .541$$

$$\hat{r}_{32} = b_{31} \cdot b_{21} = 0.646 \cdot 0.861 = .556 ; \hat{r}_{41} = b_{41} \cdot b_{11} = 0.799 \cdot 0.838 = .669$$

$$\hat{r}_{42} = b_{41} \cdot b_{21} = 0.799 \cdot 0.861 = .688 ; \hat{r}_{43} = b_{41} \cdot b_{31} = 0.799 \cdot 0.646 = .516$$

Die erklärten Varienzen, die hier als *Kommunalitäten* bezeichnet werden, berechnen sich nach:

$$s^2(\hat{Z}_{ff03d}) = b_{11}^2 \cdot s^2(H_1) = b_{11}^2 = .838^2 = .702$$

$$s^2(\hat{Z}_{ff04d}) = b_{21}^2 = .861^2 = .741 ; s^2(\hat{Z}_{ff05d}) = b_{31}^2 = .646^2 = .418 ; s^2(\hat{Z}_{ff08d}) = b_{41}^2 = .799^2 = .639$$

# Hauptkomponentenanalyse

Korrelationsmatrix

		ff03d Ausländer sind Belastung fürs soziale Netz	ff04d Es leben zu viele Ausländer in D.	ff05d Ausländerkind er verhindern gute Bildung der dt. Kinder	ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp
Korrelation	ff03d Ausländer sind Belastung fürs soziale Netz	1.000	.675	.388	.531
	ff04d Es leben zu viele Ausländer in D.	.675	1.000	.394	.589
	ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.388	.394	1.000	.381
	ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.531	.589	.381	1.000

Komponentenmatrix<sup>a</sup>

	Komponente
	1
ff03d Ausländer sind Belastung fürs soziale Netz	.838
ff04d Es leben zu viele Ausländer in D.	.861
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.646
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.799

Extraktionsmethode: Hauptkomponentenanalyse.

a. 1 Komponenten extrahiert

Die Differenzen zwischen den reproduzierten und den beobachteten Korrelationen sind:

$$\begin{pmatrix} 1.00 & & & \\ .676 & 1.00 & & \\ .388 & .394 & 1.00 & \\ .531 & .589 & .381 & 1.00 \end{pmatrix} - \begin{bmatrix} .702 & & & \\ .721 & .741 & & \\ .541 & .556 & .418 & \\ .669 & .861 & .516 & .639 \end{bmatrix} = \begin{bmatrix} .298 & & & \\ -.046 & .259 & & \\ -.154 & -.162 & .582 & \\ -.139 & -.099 & -.135 & .361 \end{bmatrix}$$

Die reproduzierten Korrelationen überschätzen die empirischen Werte in diesem Beispiel.

## Hauptkomponentenanalyse

Die Berechnung einer Hauptkomponentenanalyse mit SPSS erfolgt über folgenden Befehl:

```
FACTOR VARIABLES=varlist  
    /PRINT DEF COR REPR FSCORE  
    /CRITERIA FACTORS(nh)  
    /EXTRACTION PC /ROTATE NOROTATE  
    /SAVE (n rootname).
```

Die erste Option `"/VARIABLES=..."` benennt die Variablen, für die eine Hauptkomponentenanalyse berechnet werden soll.

Mit `"/PRINT=..."` wird die angeforderte Ausgabe gesteuert. Neben den Voreinstellungen ist es oft sinnvoll, die Korrelationsmatrix (COR), die Matrix der reproduzierten Korrelationen (REPR) und die Koeffizienten  $a_{ji}$ , mit deren Hilfe die Hauptkomponenten aus den standardisierten Variablen der Variablenliste berechnet werden (FSCORE), auszudrucken.

Die Anzahl der zu berechnenden Hauptkomponenten wird durch die Zahl "nh" in Klammern hinter dem Schlüsselwort "FACTORS" in der Option `"/CRITERIA"` angegeben.

das Schlüsselwort "PC" oder alternativ "PA1" in der Option `"/EXTRACTION"` fordert die Hauptkomponentenanalyse an, das Schlüsselwort "NOROTATE" in der Option `"/ROTATE"` verhindert die sog. Rotation der Lösung.

Über die Option `"/SAVE"` können die Hauptkomponenten als zusätzliche Variablen in den Datensatz aufgenommen werden. In Klammern wird die Zahl der zu speichernden Komponenten und der Beginn des Variablennamens aufgeführt. SPSS ergänzt den Namen um die Nummer der Hauptkomponente (und ggf. der Analyse).

# Hauptkomponentenanalyse

Für das Beispiel ergibt sich folgender Befehl:

```
FACTOR
/VAR ff03d ff04d ff05d ff08d
/PRINT DEF CORR REPR FSC
/CRITERIA FACTORS(1) /EXTR PC /ROT NOR.
```

**Kommunalitäten**

	Anfänglich	Extraktion
ff03d Ausländer sind Belastung fürs soziale Netz	1.000	.702
ff04d Es leben zu viele Ausländer in D.	1.000	.741
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	1.000	.418
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	1.000	.639

Extraktionsmethode: Hauptkomponentenanalyse.

Nach der Korrelationsmatrix werden die Kommunalitäten der Extraktion, also die erklärten Varianzen, ausgegeben. In der zweiten Spalte "Anfänglich" sind Ausgangsschätzungen angegeben, die bei der Hauptkomponentenanalyse gleich den Diagonalelementen der Korrelationsmatrix sind.

## Hauptkomponentenanalyse

Es folgt eine Tabelle, die im linken Teil die sog. Eigenwertzerlegung der Korrelationsmatrix wiedergibt und im rechten die Summe der durch die Hauptkomponenten erklärten Varianzen:

Erklärte Gesamtvarianz

Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	2.499	62.479	62.479	2.499	62.479	62.479
2	.700	17.496	79.976			
3	.484	12.094	92.070			
4	.317	7.930	100.000			

Extraktionsmethode: Hauptkomponentenanalyse.

Der Wert 2.499 ist gleich der Summe der Kommunalitäten ( $.702 + .741 + .418 + .639$ ) und entspricht 62.479 der Gesamtvarianz von 4, die als Folge der Standardisierung der Ausgangsvariablen gleich der Zahl der Ausgangsvariablen ist.

Da formal die Hauptkomponentenanalyse eine sog. Eigenwertzerlegung der Korrelationsmatrix ist, entspricht der Wert dem ersten Eigenwert.

Nach der Komponentenmatrix mit den standardisierten Regressionskoeffizienten der Regression der Ausgangsvariablen auf die Hauptkomponente wird die Tabelle der reproduzierten Korrelationen und der Residuen ausgegeben.

# Hauptkomponentenanalyse

## Reproduzierte Korrelationen

		ff03d Ausländer sind Belastung fürs soziale Netz	ff04d Es leben zu viele Ausländer in D.	ff05d Ausländerkind er verhindern gute Bildung der dt. Kinder	ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp
Reproduzierte Korrelation	ff03d Ausländer sind Belastung fürs soziale Netz	.702 <sup>b</sup>	.721	.541	.669
	ff04d Es leben zu viele Ausländer in D.	.721	.741 <sup>b</sup>	.556	.688
	ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.541	.556	.418 <sup>b</sup>	.516
	ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.669	.688	.516	.639 <sup>b</sup>
Residuum <sup>a</sup>	ff03d Ausländer sind Belastung fürs soziale Netz		-.046	-.154	-.139
	ff04d Es leben zu viele Ausländer in D.	-.046		-.162	-.099
	ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	-.154	-.162		-.135
	ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	-.139	-.099	-.135	

Extraktionsmethode: Hauptkomponentenanalyse.

a. Residuen werden zwischen beobachteten und reproduzierten Korrelationen berechnet. Es liegen 5 (83.0%) nicht redundante Residuen mit absoluten Werten größer 0,05 vor.

b. Reproduzierte Kommunalitäten

## Hauptkomponentenanalyse

Schließlich folgt die Koeffizientenmatrix, die die Koeffizienten angibt, mit deren Hilfe die Hauptkomponenten berechnet werden. Die Koeffizienten sind gleichzeitig die standardisierten Regressionsgewichte  $a_{1j}$  der Regression der Hauptkomponente auf die Ausgangsvariablen.

Koeffizientenmatrix der Komponentenwerte

	Komponente
	1
ff03d Ausländer sind Belastung fürs soziale Netz	.335
ff04d Es leben zu viele Ausländer in D.	.345
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.259
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.320

Extraktionsmethode: Hauptkomponentenanalyse.

Die erste Hauptkomponenten berechnet sich also nach:

$$\begin{aligned}
 H_1 &= a_{11} \cdot Z_{ff03d} + a_{12} \cdot Z_{ff04d} + a_{13} \cdot Z_{ff05d} + a_{14} \cdot Z_{ff08d} \\
 &= .335 \cdot Z_{ff03d} + .345 \cdot Z_{ff04d} + .259 \cdot Z_{ff05d} + .320 \cdot Z_{ff08d} .
 \end{aligned}$$

Die erste Hauptkomponente ist in dem Sinne eine optimale Zusammenfassung der Ausgangsvariablen, als sie die in der Korrelationsmatrix enthaltenen Informationen über die Beziehung zwischen den Variablen berücksichtigt.

Die erste Hauptkomponenten kann die Korrelationsmatrix allerdings nicht vollständig reproduzieren. Es ist jedoch möglich, weitere Hauptkomponenten zu berechnen.

Die zweite Hauptkomponente  $H_2$  wird analog der ersten Hauptkomponente für die Residuen der ersten Komponente berechnet. Dieses Verfahren kann solange fortgesetzt werden, bis die Korrelationsmatrix aus der Summe der Hauptkomponenten vollständig reproduziert werden kann.

Normalerweise ist das der Fall, wenn genau so viele Hauptkomponenten berechnet werden, wie es Variablen gibt.

# Hauptkomponentenanalyse

Die vollständige Hauptkomponentenanalyse wird durch folgenden SPSS-Befehl angefordert:

```
FACTOR  
  /VAR ff03d ff04d ff05d ff08d  
  /PRINT DEF CORR REPR FSC  
  /CRITERIA FACTORS(4) /EXTR PC /ROT NOR  
  /SAVE (4 H).
```

**Kommunalitäten**

	Anfänglich	Extraktion
ff03d Ausländer sind Belastung fürs soziale Netz	1.000	1.000
ff04d Es leben zu viele Ausländer in D.	1.000	1.000
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	1.000	1.000
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	1.000	1.000

Extraktionsmethode: Hauptkomponentenanalyse.

Die Tabelle mit den Kommunalitäten zeigt, dass tatsächlich die gesamte Varianz (von 1) der standardisierten Ausgangsvariablen "erklärt" wird.

# Hauptkomponentenanalyse

Dies zeigt sich auch in der Tabelle der erklärten Varianzen:

**Erklärte Gesamtvarianz**

Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	2.499	62.479	62.479	2.499	62.479	62.479
2	.700	17.496	79.976	.700	17.496	79.976
3	.484	12.094	92.070	.484	12.094	92.070
4	.317	7.930	100.000	.317	7.930	100.000

Extraktionsmethode: Hauptkomponentenanalyse.

Die erste Hauptkomponenten "erklärt" 62,5% der Gesamtvariation, die zweite 17.5%, die dritte 12.1% und die letzte den Rest.

**Komponentenmatrix<sup>a</sup>**

	Komponente			
	1	2	3	4
ff03d Ausländer sind Belastung fürs soziale Netz	.838	-.213	-.363	.348
ff04d Es leben zu viele Ausländer in D.	.861	-.224	-.153	-.430
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.646	.762	-.033	-.008
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.799	-.152	.572	.105

Extraktionsmethode: Hauptkomponentenanalyse.

a. 4 Komponenten extrahiert

Die Komponentenmatrix enthält die Koeffizienten  $b_{ij}$  für alle Hauptkomponenten.

Da die Korrelationsmatrix bis auf Rundungsfehler exakt reproduziert wird, macht es hier eigentlich keinen Sinn, die reproduzierten Korrelationen und die Residuen auszu-drucken.

# Hauptkomponentenanalyse

## Reproduzierte Korrelationen

		ff03d Ausländer sind Belastung fürs soziale Netz	ff04d Es leben zu viele Ausländer in D.	ff05d Ausländerkind er verhindern gute Bildung der dt. Kinder	ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp
Reproduzierte Korrelation	ff03d Ausländer sind Belastung fürs soziale Netz	1.000 <sup>b</sup>	.675	.388	.531
	ff04d Es leben zu viele Ausländer in D.	.675	1.000 <sup>b</sup>	.394	.589
	ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.388	.394	1.000 <sup>b</sup>	.381
	ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.531	.589	.381	1.000 <sup>b</sup>
Residuum <sup>a</sup>	ff03d Ausländer sind Belastung fürs soziale Netz		-3.33E-016	-4.44E-016	-1.11E-016
	ff04d Es leben zu viele Ausländer in D.	-3.33E-016		-8.33E-016	-1.11E-016
	ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	-4.44E-016	-8.33E-016		-5.00E-016
	ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	-1.11E-016	-1.11E-016	-5.00E-016	

Extraktionsmethode: Hauptkomponentenanalyse.

a. Residuen werden zwischen beobachteten und reproduzierten Korrelationen berechnet. Es liegen 0 (.0%) nicht redundante Residuen mit absoluten Werten größer 0,05 vor.

b. Reproduzierte Kommunalitäten

# Hauptkomponentenanalyse

Koeffizientenmatrix der Komponentenwerte

	Komponente			
	1	2	3	4
ff03d Ausländer sind Belastung fürs soziale Netz	.335	-.305	-.750	1.097
ff04d Es leben zu viele Ausländer in D.	.345	-.320	-.317	-1.356
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.259	1.089	-.068	-.025
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.320	-.217	1.183	.331

Extraktionsmethode: Hauptkomponentenanalyse.  
Komponentenwerte.

Kovarianzmatrix des Komponentenwerts

Komponente	1	2	3	4
1	1.000	.000	.000	.000
2	.000	1.000	.000	.000
3	.000	.000	1.000	.000
4	.000	.000	.000	1.000

Extraktionsmethode: Hauptkomponentenanalyse.  
Komponentenwerte.

Die Komponentenwerte geben wieder an, wie die Hauptkomponenten berechnet werden.

In die erste Hauptkomponente gehen alle Ausgangsvariablen mit annähernd gleichen Gewichten ein; nur das Gewicht von ff05d ist merklich kleiner. Dafür ist es bei der zweiten Hauptkomponente besonders groß, während die dritte Hauptkomponente vor allem ff08d hoch gewichtet. Die letzte Hauptkomponente erfasst vor allem die Differenz zwischen ff03d und dd04d

Die Kovarianzmatrix der Hauptkomponenten zeigt, dass die Komponenten standardisiert und untereinander unkorreliert sind.

## Hauptkomponentenanalyse

Die in den Datensatz aufgenommene erste Komponente kann als optimaler Index aus den vier Ausgangsvariablen verwendet werden.

Über Korrelations- und Regressionsanalysen kann aber auch verdeutlicht werden, wie die Beziehungen zwischen den Hauptkomponenten, einem einfachen Summenscore und den Ausgangsvariablen sind:

```
comp ffsum=ff03d+ff04d+ff05d+ff08d.  
corr var=ff03d ff04d to ff08d ffsum with ffsum H1 to H4.  
reg stat def cha /dep=ff03d /ent H1 /ent H2 /ent H3 /ent H4.  
reg dep=ff04d /ent H1 to H4.  
reg dep=ff05d /ent H1 to H4.  
reg dep=ff08d /ent H1 to H4.  
reg dep=H1 /ent ff03d to ff08d.  
reg dep=H2 /ent ff03d to ff08d.  
reg dep=H3 /ent ff03d to ff08d.  
reg dep=H4 /ent ff03d to ff08d.
```

# Hauptkomponentenanalyse

## Korrelationen

		ffsum	H1 REGR factor score 1 for analysis	H2 REGR factor score 2 for analysis	H3 REGR factor score 3 for analysis	H4 REGR factor score 4 for analysis	
ff03d	Ausländer sind Belastung fürs soziale Netz	Korrelation nach Pearson Signifikanz (2-seitig) N	.824 .000 8987	.838 .000 8987	-.213 .000 8987	-.363 .000 8987	.348 .000 8987
ff04d	Es leben zu viele Ausländer in D.	Korrelation nach Pearson Signifikanz (2-seitig) N	.850 .000 8987	.861 .000 8987	-.224 .000 8987	-.153 .000 8987	-.430 .000 8987
ff05d	Ausländerkinder verhindern gute Bildung der dt. Kinder	Korrelation nach Pearson Signifikanz (2-seitig) N	.685 .000 8987	.646 .000 8987	.762 .000 8987	-.033 .002 8987	-.008 .448 8987
ff08d	Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	Korrelation nach Pearson Signifikanz (2-seitig) N	.790 .000 8987	.799 .000 8987	-.152 .000 8987	.572 .000 8987	.105 .000 8987
ffsum		Korrelation nach Pearson Signifikanz (2-seitig) N	1.000 .000 8987	.999 .000 8987	.052 .000 8987	.000 .970 8987	-.003 .760 8987

Die Korrelationen der Hauptkomponenten mit den Ausgangsvariablen sind mit den Werten in der Komponentenmatrix identisch.

Sichtbar wird auch, dass in dieser Analyse der Summenscore beinahe perfekt mit der ersten Hauptkomponente korreliert. Im Beispiel ist der Summenscore (bzw. der Mittelwert über die Variablen) ein Index, der ebenfalls bereits fast die maximale Information der Korrelationsmatrix berücksichtigt.

# Hauptkomponentenanalyse

## Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers	Änderungsstatistiken				
					Änderung in R-Quadrat	Änderung in F	df1	df2	Änderung in Signifikanz von F
1	.838 <sup>a</sup>	.702	.702	.528	.702	21122.870	1	8985	.000
2	.864 <sup>b</sup>	.747	.747	.486	.046	1617.477	1	8984	.000
3	.937 <sup>c</sup>	.879	.879	.337	.132	9774.793	1	8983	.000
4	1.000 <sup>d</sup>	1.000	1.000	.000	.121	.	1	8982	.

a. Einflußvariablen : (Konstante), H1 REGR factor score 1 for analysis

b. Einflußvariablen : (Konstante), H1 REGR factor score 1 for analysis, H2 REGR factor score 2 for analysis

c. Einflußvariablen : (Konstante), H1 REGR factor score 1 for analysis, H2 REGR factor score 2 for analysis, H3 REGR factor score 3 for analysis

d. Einflußvariablen : (Konstante), H1 REGR factor score 1 for analysis, H2 REGR factor score 2 for analysis, H3 REGR factor score 3 for analysis, H4 REGR factor score 4 for analysis

Die schrittweise Regression von ff03d auf die Hauptkomponenten zeigt den Anstieg der Kommunalitäten bis schließlich die gesamte Varianz erklärt wird.

Summiert man die Änderungen in R-Quadrat über alle vier Variablen ff03d bis ff08d auf, ergibt sich die Tabelle der erklärten Varianzen.

Da die Hauptkomponenten unkorreliert sind, ändern sich die Regressionskoeffizienten in der schrittweisen Regression nicht.

Die standardisierten Koeffizienten sind dann Korrelationen und gleichzeitig die Koeffizienten in der Regression der Ausgangsvariablen auf die Komponenten.

# Hauptkomponentenanalyse

Koeffizienten<sup>a</sup>

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
		B	Standardfehler	Beta		
1	(Konstante)	2.543	.006		456.343	.000
	H1 REGR factor score 1 for analysis	.810	.006	.838	145.337	.000
2	(Konstante)	2.543	.005		495.696	.000
	H1 REGR factor score 1 for analysis	.810	.005	.838	157.870	.000
	H2 REGR factor score 2 for analysis	-.206	.005	-.213	-40.218	.000
3	(Konstante)	2.543	.004		716.261	.000
	H1 REGR factor score 1 for analysis	.810	.004	.838	228.116	.000
	H2 REGR factor score 2 for analysis	-.206	.004	-.213	-58.113	.000
	H3 REGR factor score 3 for analysis	-.351	.004	-.363	-98.868	.000
4	(Konstante)	2.543	.000		.	.
	H1 REGR factor score 1 for analysis	.810	.000	.838	.	.
	H2 REGR factor score 2 for analysis	-.206	.000	-.213	.	.
	H3 REGR factor score 3 for analysis	-.351	.000	-.363	.	.
	H4 REGR factor score 4 for analysis	.337	.000	.348	.	.

a. Abhängige Variable: ff03d Ausländer sind Belastung fürs soziale Netz

## Hauptkomponentenanalyse

Die Regression der Hauptkomponenten auf die Ausgangsvariablen ergibt schließlich die Werte in der Koeffizientenmatrix:

Koeffizienten<sup>a</sup>

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
		B	Standardfehler	Beta		
1	(Konstante)	-3.387	.000		.	.
	ff03d Ausländer sind Belastung fürs soziale Netz	.347	.000	.335	.	.
	ff04d Es leben zu viele Ausländer in D.	.339	.000	.345	.	.
	ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.268	.000	.259	.	.
	ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.340	.000	.320	.	.

a. Abhängige Variable: H1 REGR factor score 1 for analysis

# Faktorenanalyse

## Von der Hauptkomponentenanalyse zur Faktorenanalyse

**Hauptkomponenten** sind **Indizes**, d.h. sie werden für jeden Fall aus den Werten der Ausgangsvariablen berechnet.

Demgegenüber sind **Faktoren** latente Variablen, d.h. theoretische Konstrukte, deren Werte prinzipiell nicht direkt beobachtet werden können.

Stattdessen wird angenommen, dass diese latenten Variablen über Regressionsgleichungen beobachtbare Variablen, die **Indikatoren** kausal beeinflussen.

Geschätzt werden zunächst nicht die Werte dieser latenten Variablen, sondern die Regressionsgleichungen, nach denen die Faktoren die Indikatoren beeinflussen, sowie gegebenenfalls die Korrelationen unter den Faktoren.

In der **explorativen Faktorenanalyse** wird üblicherweise davon ausgegangen, dass Faktoren standardisiert sind und auch die Indikatoren standardisiert vorliegen. Analysiert wird dann wie bei der Hauptkomponentenanalyse die Korrelationsmatrix.

Im Unterschied zur Hauptkomponentenanalyse wird in der Faktorenanalyse grundsätzlich unterstellt, dass sich jeder Indikator  $Z_i$  als gewichtete Summe der durch die  $K$  **Faktoren**  $F_k$  erklärten Komponente, eines **itemspezifischen Faktors**  $U_i$ , der sog. **Uniqueness**, und einer **Residualvariable**  $\zeta_i$  zusammensetzt:

$$Z_i = \lambda_{i1} \cdot F_1 + \lambda_{i2} \cdot F_2 + \dots + \lambda_{iK} \cdot F_K + U_i + \zeta_i$$

## Faktorenanalyse

Da es unmöglich ist, die Uniqueness von den Residuen zu trennen, werden diese beiden Komponenten einfach zu nur einer Restgröße  $U_i$  zusammengefasst.

Es wird außerdem unterstellt, dass diese Restgröße weder mit den Faktoren noch mit der Uniqueness der übrigen Indikatoren korreliert ist.

Die Gesamtvarianz (von 1) eines Indikators ist dann die Summe der durch die Faktoren erklärten Varianz, die wie in der Hauptkomponentenanalyse als **Kommunalität** bezeichnet wird, und der Varianz der Uniqueness. Die Uniqueness ist entsprechend eins minus der Kommunalität.

Ein Problem der Faktorenanalyse besteht darin, dass es für jede empirisch mögliche Korrelationsmatrix beliebig viele Schätzungen der Regressionskoeffizienten  $\lambda_{ik}$  für die oben beschriebene Modellgleichung gibt, wenn es mehr als einen Faktor gibt und alle Faktoren alle Indikatoren beeinflussen.

Um zu einer Lösung dieses **Extraktionsproblems** zu kommen, werden bei der Schätzung zunächst zusätzliche Bedingungen eingeführt: Wie bei der Hauptkomponentenanalyse soll der 1. Faktor möglichst viel erklären, der zweite Faktor möglichst viel von dem, was der erste Faktor nicht erklärt usw.. Daraus folgt dann auch, dass die Faktoren unkorreliert sind.

Es gibt verschiedene Schätzmethoden, um die Koeffizienten der explorativen Faktorenanalyse zu bestimmen. Am häufigsten wird vermutlich die sog. **Hauptachsenfaktorenanalyse** berechnet, die auf der Hauptkomponentenanalyse aufbaut.

## Faktorenanalyse

Um zu berücksichtigen, dass die Faktoren nicht die gesamte Varianz der Indikatoren erklären, sondern Uniqueness und Residuen zum Modell gehören, wird die Hauptkomponentenanalyse auf eine reduzierte Korrelationsmatrix angewendet. Die Diagonalelemente enthalten nicht die jeweilige Gesamtvarianz, also 1.0, sondern eine Schätzung der Kommunalität.

In der Regel ist die erste Schätzung der Kommunalität der Anteil der Varianz eines Indikators bei der multiplen Regression auf die übrigen Indikatoren. Im Beispiel der vier Items zur Fremdenfeindlichkeit werden also zunächst vier multiple Regressionen berechnet:

```
reg stat r /dep=ff03d /ent ff04d ff05d ff08d.  
reg stat r /dep=ff04d /ent ff03d ff05d ff08d.  
reg stat r /dep=ff05d /ent ff03d ff04d ff08d.  
reg stat r /dep=ff08d /ent ff03d ff04d ff05d.
```

**Modellzusammenfassung**

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	.702 <sup>a</sup>	.493	.493	.689

a. Einflußvariablen : (Konstante), ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp, ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder, ff04d Es leben zu viele Ausländer in D.

# Faktorenanalyse

## Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	.733 <sup>a</sup>	.537	.537	.692

- a. Einflußvariablen : (Konstante), ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp, ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder, ff03d Ausländer sind Belastung fürs soziale Netz

## Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	.453 <sup>a</sup>	.206	.205	.860

- a. Einflußvariablen : (Konstante), ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp, ff03d Ausländer sind Belastung fürs soziale Netz, ff04d Es leben zu viele Ausländer in D.

## Modellzusammenfassung

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	.630 <sup>a</sup>	.397	.397	.729

- a. Einflußvariablen : (Konstante), ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder, ff03d Ausländer sind Belastung fürs soziale Netz, ff04d Es leben zu viele Ausländer in D.

## Faktorenanalyse

Die Hauptkomponentenanalyse wird dann auf folgende "reduzierte Korrelationsmatrix" angewendet:

$$\begin{pmatrix} .493 & & & & \\ .676 & .537 & & & \\ .388 & .394 & .206 & & \\ .531 & .589 & .381 & .397 & \end{pmatrix}$$

```
factor /var ff03d to ff08d
      /crit fac(1) iter(1)
      /extr paf /rot nor.
```

Anstelle einer Hauptkomponentenanalyse wird eine Hauptachsenfaktorenanalyse über ***/extr paf*** angefordert, wobei allerdings mittels "iter(1)" nur eine Iteration berechnet wird.

### Kommunalitäten

	Anfänglich	Extraktion
ff03d Ausländer sind Belastung fürs soziale Netz	.493	.577
ff04d Es leben zu viele Ausländer in D.	.537	.634
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.206	.253
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.397	.476

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Die "anfänglichen" Kommunalitäten sind die erklärten Varianzen, die Kommunalitäten der Extraktion Ergebnis der Hauptkomponentenanalyse der reduzierten Korrelationsmatrix. Mit Hilfe dieser Kommunalitäten kann eine weitere Hauptkomponentenanalyse berechnet werden.

## Faktorenanalyse

Die Hauptkomponentenanalyse wird in der zweiten Iteration auf folgende "reduzierte Korrelationsmatrix" angewendet:

$$\begin{pmatrix} .577 & & & \\ .676 & .634 & & \\ .388 & .394 & .253 & \\ .531 & .589 & .381 & .476 \end{pmatrix}$$

```
factor /var ff03d to ff08d
      /diag= .577,.634,.253,.476
      /crit fac(1) iter(1)
      /extr paf /rot nor.
```

Die Option "/diag" ermöglicht die Eingabe der Ausgangsschätzung der Kommunalitäten.

### Kommunalitäten

	Anfänglich	Extraktion
ff03d Ausländer sind Belastung fürs soziale Netz	.577	.602
ff04d Es leben zu viele Ausländer in D.	.634	.671
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.253	.254
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.476	.494

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Die resultierenden Schätzungen der Kommunalitäten können wiederum als Ausgangspunkt einer weiteren Iteration verwendet werden.

Diese Vorgehensweise wird fortgesetzt, bis sich in zwei aufeinanderfolgenden Iterationen die neu berechneten Kommunalitäten nicht mehr von den anfänglichen Kommunalitäten unterscheiden.

## Faktorenanalyse

$$\begin{pmatrix} .613 & & & \\ .676 & .706 & & \\ .388 & .394 & .248 & \\ .531 & .589 & .381 & .493 \end{pmatrix}$$

```
factor /var ff03d to ff08d
      /diag= .613,.706,.248,.493
      /crit fac(1) /extr paf /rot nor.
```

Kommunalitäten

	Anfänglich	Extraktion
ff03d Ausländer sind Belastung fürs soziale Netz	.613	.613
ff04d Es leben zu viele Ausländer in D.	.706	.707
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.248	.248
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.493	.493

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Faktorenmatrix<sup>a</sup>

	Faktor
	1
ff03d Ausländer sind Belastung fürs soziale Netz	.783
ff04d Es leben zu viele Ausländer in D.	.841
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.498
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.702

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

a. 1 Faktoren extrahiert. Es werden 1 Iterationen benötigt.

Diese endgültige Lösung ergibt sich für diese Daten nach 8 Iterationen:

```
factor /var ff03d to ff08d /print def rep fscor
      /crit fac(1) /extr paf /rot nor /save (1 F).
```

# Faktorenanalyse

**Kommunalitäten**

	Anfänglich	Extraktion
ff03d Ausländer sind Belastung fürs soziale Netz	.493	.613
ff04d Es leben zu viele Ausländer in D.	.537	.706
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.206	.248
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.397	.493

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

**Faktorenmatrix<sup>a</sup>**

	Faktor
	1
ff03d Ausländer sind Belastung fürs soziale Netz	.783
ff04d Es leben zu viele Ausländer in D.	.840
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.498
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.702

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

a. 1 Faktoren extrahiert. Es werden 8 Iterationen benötigt.

**Erklärte Gesamtvarianz**

Faktor	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	2.499	62.479	62.479	2.060	51.500	51.500
2	.700	17.496	79.976			
3	.484	12.094	92.070			
4	.317	7.930	100.000			

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

# Faktorenanalyse

## Erklärte Gesamtvarianz

Faktor	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	2.499	62.479	62.479	2.060	51.500	51.500
2	.700	17.496	79.976			
3	.484	12.094	92.070			
4	.317	7.930	100.000			

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Da in der Faktorenanalyse Uniqueness berücksichtigt wird, ist die insgesamt erklärte Varianz mit 2.060 geringer als die erklärte Varianz der korrespondierenden Hauptkomponentenanalyse.

### Faktorenmatrix<sup>a</sup>

	Faktor
	1
ff03d Ausländer sind Belastung fürs soziale Netz	.783
ff04d Es leben zu viele Ausländer in D.	.840
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.498
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.702

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

a. 1 Faktoren extrahiert. Es werden 8 Iterationen benötigt.

Der Vergleich der Ladungen der Faktorenanalyse mit den korrespondierenden Komponentenwerten der Hauptkomponentenanalyse zeigt, dass die Unterschiede zwischen den Variablen bei der Faktorenanalyse deutlicher zutage treten.

### Komponentenmatrix<sup>a</sup>

	Komponente
	1
ff03d Ausländer sind Belastung fürs soziale Netz	.838
ff04d Es leben zu viele Ausländer in D.	.861
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.646
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.799

Extraktionsmethode: Hauptkomponentenanalyse

a. 1 Komponenten extrahiert

# Faktorenanalyse

Der Vergleich der Korrelationen mit den reproduzierten Korrelationen ergibt, dass die Abweichungen bei der Hauptfaktorenanalyse geringer sind als bei der Hauptkomponentenanalyse:

<i>Empirische Korrelationen</i>	<i>Reproduzierte Korrelationen</i>	
	<i>Faktorenanalyse</i>	<i>Hauptkomponentenanalyse</i>
$\begin{pmatrix} 1.00 & & & \\ .676 & 1.00 & & \\ .388 & .394 & 1.00 & \\ .531 & .589 & .381 & 1.00 \end{pmatrix}$	$\begin{bmatrix} .613 & & & \\ .658 & .706 & & \\ .390 & .418 & .248 & \\ .550 & .590 & .350 & .493 \end{bmatrix}$	$\begin{bmatrix} .702 & & & \\ .721 & .741 & & \\ .541 & .556 & .418 & \\ .669 & .861 & .516 & .639 \end{bmatrix}$

**Koeffizientenmatrix der Faktorwerte**

	Faktor
	1
ff03d Ausländer sind Belastung fürs soziale Netz	.316
ff04d Es leben zu viele Ausländer in D.	.449
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.112
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.228

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Methode für Faktorwerte: Regression.

**Kovarianzmatrix für Faktorwerte**

Faktor	1
1	.840

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Methode für Faktorwerte: Regression.

Die Koeffizientenmatrix gibt die unstandardisierten Regressionsgewichte der Regression des Faktors auf die standardisierten Indikatoren wieder.

Die Varianz des Faktors ist dann kleiner 1.

# Faktorenanalyse

```
descri ff03d(z03d) ff04d(z04d) ff05d(z05d) ff08d(z08d) .
reg dep F1 / ent z03d to z08d.
```

**Koeffizientenmatrix der Faktorwerte**

	Faktor
	1
ff03d Ausländer sind Belastung fürs soziale Netz	.316
ff04d Es leben zu viele Ausländer in D.	.449
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.112
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.228

Extraktionsmethode: Hauptachsen-Faktor

Methode für Faktorwerte: Regression.

**Koeffizienten<sup>a</sup>**

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
		B	Standardfehler	Beta		
1	(Konstante)	-.018	.000		-4E+008	.000
	Z03d Z-Wert: Ausländer sind Belastung fürs soziale Netz	.316	.000	.345	5E+009	.000
	Z04d Z-Wert: Es leben zu viele Ausländer in D.	.449	.000	.489	6E+009	.000
	Z05d Z-Wert: Ausländerkinder verhindern gute Bildung der dt. Kinder	.112	.000	.122	2E+009	.000
	Z08d Z-Wert: Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.230	.000	.249	4E+009	.000

a. Abhängige Variable: F1 REGR factor score 1 for analysis

# Faktorenanalyse

Korrelationen

		ffsum	ff03d Ausländer sind Belastung fürs soziale Netz	ff04d Es leben zu viele Ausländer in D.	ff05d Ausländerkind er verhindern gute Bildung der dt. Kinder	ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp
ffsum	Korrelation nach Pearson	1.000	.824	.850	.685	.790
	Signifikanz (2-seitig)	.000	.000	.000	.000	.000
	N	8987	8987	8987	8987	8987
H1 REGR factor score 1 for analysis	Korrelation nach Pearson	.999	.838	.861	.646	.799
	Signifikanz (2-seitig)	.000	.000	.000	.000	.000
	N	8987	8987	8987	8987	8987
F1 REGR factor score 1 for analysis	Korrelation nach Pearson	.980	.854	.917	.544	.766
	Signifikanz (2-seitig)	.000	.000	.000	.000	.000
	N	8987	8987	8987	8987	8987

Faktorenmatrix<sup>a</sup>

	Faktor
	1
ff03d Ausländer sind Belastung fürs soziale Netz	.783
ff04d Es leben zu viele Ausländer in D.	.840
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.498
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.702

Die Korrelation des geschätzten Faktors mit den Indikatoren sind geringer als die Korrelationen mit den Faktorscores, die in einer multiplen Regression auf der Basis der folgenden Korrelationsmatrix geschätzt wird:

ff03d	1.000				
ff04d	0.676	1.000			
ff05d	0.388	0.394	1.000		
ff08d	0.531	0.589	0.381	1.000	
Faktor	0.783	0.840	0.498	0.702	1.000

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

a. 1 Faktoren extrahiert. Es werden 8 Iterationen benötigt.

# Faktorenanalyse

## Die Festlegung der Faktorenanzahl

Das Modell der Faktorenanalyse unterstellt, dass die Korrelationen zwischen den Indikatoren letztlich Scheinkausalitäten sind, die durch Faktoren als gemeinsame Ursache(n) hervorgerufen werden.

Eine zentrale Frage der explorativen Faktorenanalyse ist dann, wie viele Faktoren sind notwendig, um die Beziehungen zwischen den Indikatoren praktisch vollkommen zu erklären.

Eine Antwort darauf gibt das sogenannte *Eigenwertkriterium*, nachdem so viele Faktoren extrahiert werden, wie es bei der Korrelationsmatrix der Indikatoren Eigenwerte größer "1" gibt. Die Logik dahinter ist, dass bei standardisierten Variablen jede Variable eine Varianz von "1" hat und ein Faktor mindestens so viel Erklärungskraft aufweisen soll, wie ein Indikator varianz hat.

Erklärte Gesamtvarianz

Faktor	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	2.499	62.479	62.479	2.060	51.500	51.500
2	.700	17.496	79.976			
3	.484	12.094	92.070			
4	.317	7.930	100.000			

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

## Faktorenanalyse

Das Eigenwertkriterium, das auch das Default-Kriterium bei SPSS zur Festlegung der Faktorenzahl ist, legt im Beispiel nahe, dass es nur einen Faktor gibt.

Eine Faktorenanalyse der beiden ersten Items der Fremdenfeindlichkeit mit den beiden Heterophobie-Items, die sich auf Behinderte beziehen, legt dagegen zwei Faktoren nahe:

```
factor var=ff03d ff04d he01b he02b /print def cor rep  
/plot eigen /extr paf /rot nor.
```

### Erklärte Gesamtvarianz

Faktor	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	1.717	42.932	42.932	1.365	34.116	34.116
2	1.421	35.518	78.450	.909	22.731	56.847
3	.531	13.266	91.716			
4	.331	8.284	100.000			

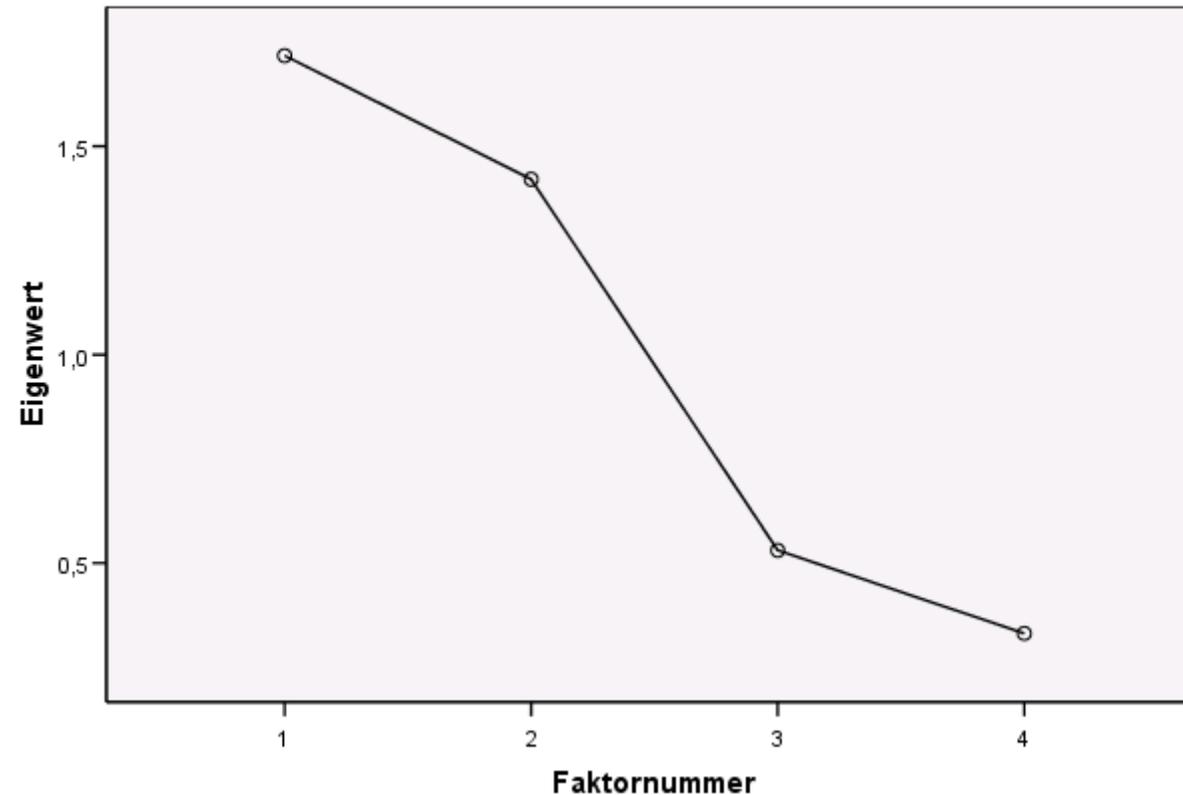
Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Auf zwei Faktoren weist auch das sog. "Screeplot" hin, das erzeugt wird, wenn Schlüsselwort "eigen" in der Option "/plot" angegeben ist.

Ein Screeplot plottet die Eigenwerte gegen ihre Rangnummer, was stets zu einer von links nach rechts abfallenden Kurve führt:

# Faktorenanalyse

Screepplot



Die Faktorenzahl wird hierbei grafisch so bestimmt, dass nur so viele Faktoren extrahiert werden, wie es einen starken Abfall der Kurve gibt.

Im Beispiel gibt es einen starken Abfall der Kurve zwischen den 2. und 3. Eigenwert (hier als "Faktor" bezeichnet) und anschließend einen flachen Verlauf. Die Kurve legt daher ebenfalls zwei Faktoren nahe.

# Faktorenanalyse

**Korrelationsmatrix**

	ff03d Ausländer sind Belastung fürs soziale Netz	ff04d Es leben zu viele Ausländer in D.	he01b Unwohles Gefühl in Gegenwart von Behinderten	he02b Unsicheres Verhalten gegenüber Behinderten
Korrelation	ff03d Ausländer sind Belastung fürs soziale Netz	ff04d Es leben zu viele Ausländer in D.	he01b Unwohles Gefühl in Gegenwart von Behinderten	he02b Unsicheres Verhalten gegenüber Behinderten
	1.000	.668	.073	.033
	.668	1.000	.069	.043
	.073	.069	1.000	.469
	.033	.043	.469	1.000

Zwei Faktoren legt auch bereits die Korrelationsmatrix nahe, da die beiden Fremdenfeindlichkeits-Items ff03d und ff04d mit 0.668 und die beiden Heterophobie-Items he01b und he02b mit 0.469 hoch miteinander korrelieren, zwischen den beiden Items aber nur geringe Korrelationen bestehen.

## Faktorenanalyse

Bei zwei Faktoren führt die Hauptachsen-Faktorenanalyse zu folgenden Ladungen:

### Erklärte Gesamtvarianz

Faktor	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	1.717	42.932	42.932	1.365	34.116	34.116
2	1.421	35.518	78.450	.909	22.731	56.847
3	.531	13.266	91.716			
4	.331	8.284	100.000			

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

### Faktorenmatrix<sup>a</sup>

	Faktor	
	1	2
ff03d Ausländer sind Belastung fürs soziale Netz	.800	-.169
ff04d Es leben zu viele Ausländer in D.	.800	-.164
he01b Unwohles Gefühl in Gegenwart von Behinderten	.224	.649
he02b Unsicheres Verhalten gegenüber Behinderten	.185	.658

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

a. 2 Faktoren extrahiert. Es werden 9 Iterationen benötigt.

Auf den ersten Faktor, der 34.1% der Varianz erklärt, laden vor allem die beiden Indikatoren ff03d und ff04d hoch, auf den zweiten Faktor, der 22.7% der Varianz erklärt, die Indikatoren he01 und he02b. Der erste Faktor lässt sich daher als "Xenophobie", der zweite als "Heteriphobie" identifizieren.

# Faktorenanalyse

## Rotation

Die Ausgangslösung nach der Extraktion ist immer so, dass analog zur Hauptkomponentenanalyse der erste Faktor am meisten, der zweite am zweitmeisten usw. erklärt. Dies ist aber ein technisches Artefakt der Schätzung und kein empirischer Befund.

Hinzu kommt, dass sehr oft die Lösung kaum zu interpretieren ist, weil es kein klares Muster der Ladungen gibt.

Faktorenmatrix<sup>a</sup>

	Faktor	
	1	2
ff03d Ausländer sind Belastung fürs soziale Netz	.771	-.039
ff04d Es leben zu viele Ausländer in D.	.865	-.188
ff05d Ausländerkinder verhindern gute Bildung der dt. Kinder	.514	.258
ff08d Ausländer nach Hause schicken, wenn Arbeitsplätze knapp	.697	.086

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

- a. Es wurde versucht, 2 Faktoren zu extrahieren. Es werden mehr als 25 Iterationen benötigt. (Konvergenz=.002). Die Extraktion wurde abgebrochen.

So lässt sich der zweite Faktor bei einer "erzwungenen" Zwei-Faktoren-Lösung der vier Fremdenfeindlichkeits-Items kaum sinnvoll interpretieren.

```
factor var=ff03d to ff08d /crit fac(2)  
      /plot eigen /extr paf /rot nor.
```

## Faktorenanalyse

Eine Lösung bietet die sog. "**Rotation**" der Faktoren. Statistisch gesehen sind Rotationen lineare Transformationen der Faktoren, die so beschaffen sind, dass die resultierenden **rotierten Faktoren** leichter zu interpretieren sind, was bedeutet, dass sie ein klareres Muster der Ladungen aufweisen.

Die Bezeichnung Rotation kommt daher, dass geometrisch eine Rotation auch so interpretiert werden kann, dass bei einer grafischen Darstellung der Ladungen als Punkte in einem Koordinatensystem, die Rotation als Drehung der Koordinatenachsen verstanden werden kann.

Es gibt verschiedene Rotationsalgorithmen. Bei **orthogonalen Rotationen** bleiben die Faktoren auch nach der Rotation unkorreliert. In SPSS ist die Default-Rotation die sog. **Varimax-Rotation**:

```
factor var=ff03d ff04d he01b he02b /print def rep /extr paf /rot.
```

**Faktor-Transformationsmatrix**

Faktor	1	2
1	.968	.252
2	-.252	.968

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Rotationsmethode: Varimax mit Kaiser-Normalisierung.

Die Faktortransformationsmatrix gibt die Transformationsgleichung wieder. Der rotierte erste Faktor ist

$$0.968 \cdot F_1 + 0.253 \cdot F_2$$

der Ausgangslösung.

# Faktorenanalyse

## Erklärte Gesamtvarianz

Faktor	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion			Rotierte Summe der quadrierten Ladungen		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	1.717	42.932	42.932	1.365	34.116	34.116	1.336	33.396	33.396
2	1.421	35.518	78.450	.909	22.731	56.847	.938	23.451	56.847
3	.531	13.266	91.716						
4	.331	8.284	100.000						

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Die Rotation verändert die Aufteilung der erklärten Varianzen (Kommunalitäten) auf die Faktoren.

Im Beispiel sinkt die durch den ersten Faktor erklärte Varianz, während die durch den zweiten Faktor erklärte Varianz leicht ansteigt.

Die insgesamt erklärten Varianzen bleiben gleich.

Durch die Rotation ändern sich vor allem die Werte der Ladungen.

Der Vergleich der unrotierten mit den rotierten Faktorladungen zeigt, dass die Interpretation dadurch erleichtert wird, dass die ersten beiden Indikatoren noch stärker auf den ersten und die letzten beiden Indikatoren noch stärker auf den zweiten Faktor laden. Gleichzeitig sind die Ladungen der ersten beiden Indikatoren auf den zweiten und der letzten beiden Indikatoren auf den ersten Faktor geringer und nahe Null.

# Faktorenanalyse

Faktorenmatrix<sup>a</sup>

	Faktor	
	1	2
ff03d Ausländer sind Belastung fürs soziale Netz	.800	-.169
ff04d Es leben zu viele Ausländer in D.	.800	-.164
he01b Unwohles Gefühl in Gegenwart von Behinderten	.224	.649
he02b Unsicheres Verhalten gegenüber Behinderten	.185	.658

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

a. 2 Faktoren extrahiert. Es werden 9 Iterationen benötigt.

Rotierte Faktorenmatrix<sup>a</sup>

	Faktor	
	1	2
ff03d Ausländer sind Belastung fürs soziale Netz	.817	.037
ff04d Es leben zu viele Ausländer in D.	.815	.042
he01b Unwohles Gefühl in Gegenwart von Behinderten	.054	.684
he02b Unsicheres Verhalten gegenüber Behinderten	.013	.683

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Rotationsmethode: Varimax mit Kaiser-Normalisierung.

a. Die Rotation ist in 3 Iterationen konvergiert.

Noch deutlicher als in der Ausgangslösung ist also hier der erste Faktor als Fremdenfeindlichkeit und der zweite Faktor als Heterophobie gegenüber Behinderten zu interpretieren.

Bei einer *nicht-orthogonalen* oder *obliquen Rotation* können die Faktoren nach der Rotation korreliert sein. Dies ermöglicht bisweilen ein noch klareres Muster der Faktorladungen.

Bei nicht-orthogonalen Rotationen wird in letzter Zeit oft der Algorithmus der *Promax-Methode* benutzt.

```
factor var=ff03d ff04d he01b he02b /print def rep /extr paf
/rot promax.
```

# Faktorenanalyse

**Faktorenmatrix<sup>a</sup>**

	Faktor	
	1	2
ff03d Ausländer sind Belastung fürs soziale Netz	.800	-.169
ff04d Es leben zu viele Ausländer in D.	.800	-.164
he01b Unwohles Gefühl in Gegenwart von Behinderten	.224	.649
he02b Unsicheres Verhalten gegenüber Behinderten	.185	.658

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

a. 2 Faktoren extrahiert. Es werden 9 Iterationen benötigt.

**Rotierte Faktorenmatrix<sup>a</sup>**

	Faktor	
	1	2
ff03d Ausländer sind Belastung fürs soziale Netz	.817	.037
ff04d Es leben zu viele Ausländer in D.	.815	.042
he01b Unwohles Gefühl in Gegenwart von Behinderten	.054	.684
he02b Unsicheres Verhalten gegenüber Behinderten	.013	.683

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Rotationsmethode: Varimax mit Kaiser-Normalisierung.

a. Die Rotation ist in 3 Iterationen konvergiert.

**Mustermatrix<sup>a</sup>**

	Faktor	
	1	2
ff03d Ausländer sind Belastung fürs soziale Netz	.818	-.003
ff04d Es leben zu viele Ausländer in D.	.816	.002
he01b Unwohles Gefühl in Gegenwart von Behinderten	.020	.684
he02b Unsicheres Verhalten gegenüber Behinderten	-.020	.685

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Rotationsmethode: Promax mit Kaiser-Normalisierung.

a. Die Rotation ist in 3 Iterationen konvergiert.

Im Beispiel wird noch stärker als bei der Varimax-Rotation die Lösung akzentuiert, obwohl die Korrelation zwischen den beiden Faktoren recht gering ist:

**Korrelationsmatrix für Faktor**

Faktor	1	2
1	1.000	.098
2	.098	1.000

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Rotationsmethode: Promax mit Kaiser-Normalisierung.

## Faktorenanalyse

Die Matrix der Ladungen, also der standardisierten Regressionskoeffizienten der Regression dre Indikatoren auf die Faktoren wird bei einer nicht-orthogonalen Rotation als **Mustermatrix** bezeichnet.

Zusätzlich zur Mustermatrix wird auch die sog. **Strukturmatrix** ausgedruckt, die die Korrelationen der Faktoren mit den Indikatoren wiedergibt. Je höher die Korrelation zwischen den Faktoren, desto stärker unterscheiden sich Mustermatrix und Strukturmatrix.

Bei unkorrelierten Faktoren sind die Werte von Ladungen und Korrelationen stets identisch.

**Mustermatrix<sup>a</sup>**

	Faktor	
	1	2
ff03d Ausländer sind Belastung fürs soziale Netz	.818	-.003
ff04d Es leben zu viele Ausländer in D.	.816	.002
he01b Unwohles Gefühl in Gegenwart von Behinderten	.020	.684
he02b Unsicheres Verhalten gegenüber Behinderten	-.020	.685

Extraktionsmethode: Hauptachsen-Faktorenanalyse.  
Rotationsmethode: Promax mit Kaiser-Normalisierung.

a. Die Rotation ist in 3 Iterationen konvergiert.

**Strukturmatrix**

	Faktor	
	1	2
ff03d Ausländer sind Belastung fürs soziale Netz	.818	.077
ff04d Es leben zu viele Ausländer in D.	.816	.082
he01b Unwohles Gefühl in Gegenwart von Behinderten	.087	.686
he02b Unsicheres Verhalten gegenüber Behinderten	.046	.683

Extraktionsmethode: Hauptachsen-Faktorenanalyse.  
Rotationsmethode: Promax mit Kaiser-Normalisierung.

## Faktorenanalyse

Alle Rotationen sind empirisch äquivalent, was bedeutet, dass die reproduzierte Korrelationsmatrix immer (bis auf Rundungsfehler) identisch ist.

Ganz generell berechnen sich die reproduzierten Korrelationen nach der Formel:

$$\hat{r}_{ij} = \sum_{i=1}^K \sum_{j=1}^K \lambda_{ik} \cdot \lambda_{jk} \cdot r(F_i, F_j)$$

Da die Korrelation einer Variable mit sich selbst "1" ist, ist der Wert  $r(F_k, F_k)$  bei jedem Faktor "1". Orthogonalen Faktoren sind unkorreliert, daher ist die Berechnung sehr viel einfacher als bei korrelierten Faktoren.

Im Beispiel berechnen sich die modellimplizierten Korrelationen bei der Ausgangslösung und den rotierten Lösungen nach:

Faktorenmatrix<sup>a</sup>

	Faktor	
	1	2
ff03d Ausländer sind Belastung fürs soziale Netz	.800	-.169
ff04d Es leben zu viele Ausländer in D.	.800	-.164
he01b Unwohles Gefühl in Gegenwart von Behinderten	.224	.649
he02b Unsicheres Verhalten gegenüber Behinderten	.185	.658

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

a. 2 Faktoren extrahiert. Es werden 9 Iterationen benötigt.

$$\hat{r}_{11} = .800^2 + (-.169)^2 = 0.669$$

$$\hat{r}_{21} = .800 \cdot .800 + (-.169) \cdot (-.164) = 0.668$$

$$\hat{r}_{22} = .800^2 + (-.164)^2 = 0.667$$

$$\hat{r}_{31} = .224 \cdot .800 + .649 \cdot (-.169) = 0.070$$

$$\hat{r}_{32} = .224 \cdot .800 + .649 \cdot (-.164) = 0.073$$

$$\hat{r}_{33} = .224^2 + .649^2 = 0.471$$

$$\hat{r}_{41} = .185 \cdot .800 + .658 \cdot (-.169) = 0.037$$

$$\hat{r}_{42} = .185 \cdot .800 + .658 \cdot (-.164) = 0.040$$

$$\hat{r}_{43} = .185 \cdot .224 + .658 \cdot .649 = 0.468$$

$$\hat{r}_{44} = .185^2 + .658^2 = 0.467$$

# Faktorenanalyse

Rotierte Faktorenmatrix<sup>a</sup>

	Faktor	
	1	2
ff03d Ausländer sind Belastung fürs soziale Netz	.817	.037
ff04d Es leben zu viele Ausländer in D.	.815	.042
he01b Unwohles Gefühl in Gegenwart von Behinderten	.054	.684
he02b Unsicheres Verhalten gegenüber Behinderten	.013	.683

Extraktionsmethode: Hauptachsen-Faktorenanalyse.  
 Rotationsmethode: Varimax mit Kaiser-Normalisierung.  
 a. Die Rotation ist in 3 Iterationen konvergiert.

$$\hat{r}_{11} = .817^2 + .037^2 = 0.669$$

$$\hat{r}_{21} = .815 \cdot .817 + .042 \cdot .037 = 0.667$$

$$\hat{r}_{22} = .815^2 + .042^2 = 0.666$$

$$\hat{r}_{31} = .054 \cdot .817 + .684 \cdot .037 = 0.069$$

$$\hat{r}_{32} = .054 \cdot .815 + .684 \cdot .042 = 0.073$$

$$\hat{r}_{33} = .054^2 + .684^2 = 0.471$$

$$\hat{r}_{41} = .013 \cdot .817 + .683 \cdot .037 = 0.036$$

$$\hat{r}_{42} = .013 \cdot .815 + .683 \cdot .042 = 0.039$$

$$\hat{r}_{43} = .013 \cdot .054 + .683 \cdot .684 = 0.468$$

$$\hat{r}_{44} = .013^2 + .683^2 = 0.467$$

Mustermatrix<sup>a</sup>

	Faktor	
	1	2
ff03d Ausländer sind Belastung fürs soziale Netz	.818	-.003
ff04d Es leben zu viele Ausländer in D.	.816	.002
he01b Unwohles Gefühl in Gegenwart von Behinderten	.020	.684
he02b Unsicheres Verhalten gegenüber Behinderten	-.020	.685

Extraktionsmethode: Hauptachsen-Faktorenanalyse.  
 Rotationsmethode: Promax mit Kaiser-Normalisierung.  
 a. Die Rotation ist in 3 Iterationen konvergiert.

$$\hat{r}_{11} = .818^2 + (-.003)^2 = 0.669$$

$$\hat{r}_{21} = .816 \cdot .818 + .002 \cdot (-.003) + (.818 \cdot .002 + .816 \cdot (-.003)) \cdot .098 = 0.667$$

$$\hat{r}_{22} = .816^2 + .002^2 = 0.666$$

$$\hat{r}_{31} = .020 \cdot .818 + .684 \cdot (-.003) + (.818 \cdot .684 + .020 \cdot (-.003)) \cdot .098 = 0.085$$

$$\hat{r}_{32} = .020 \cdot .816 + .684 \cdot .002 + (.816 \cdot .684 + .020 \cdot .002) \cdot .098 = 0.072$$

$$\hat{r}_{33} = .020^2 + .684^2 = 0.468$$

$$\hat{r}_{41} = -.020 \cdot .818 + .685 \cdot (-.003) + (.818 \cdot .685 + (-.020) \cdot (-.003)) \cdot .098 = 0.037$$

$$\hat{r}_{42} = -.020 \cdot .816 + .685 \cdot .002 + (-.020 \cdot .002 + .816 \cdot .685) \cdot .098 = 0.040$$

$$\hat{r}_{43} = -.020 \cdot .020 + .685 \cdot .684 + (-.020 \cdot .684 + .020 \cdot .685) \cdot .098 = 0.468$$

$$\hat{r}_{44} = (.020)^2 + .685^2 = 0.470$$