

LÖSUNG 13

a.

- Die Variablen sollten hoch miteinander korrelieren. Deshalb sollten die einfachen Korrelationskoeffizienten hoch ausfallen. Sie sollten sich außerdem wechselseitig weitgehend überdecken. Deshalb sollten die partiellen Korrelationskoeffizienten gering ausfallen.

Die Höhe der Korrelationen kann man aus der Korrelationsmatrix entnehmen. Das KMO = Kaiser-Meyer-Olkins -Maß ist ein Maß dafür, inwieweit die beiden oben angegebenen Bedingungen erfüllt sind. Es kann maximal den Wert 1 annehmen. Ein Wert von 0,9 und höher gilt als großartig, von 0,8 bis < 0,9 als lobenswert, von 0,7 bis < 0,8 als mittel, von 0,6 bis < 0,7 als zweitklassig, von 0,65 bis > 0,6 als sehr schlecht und von weniger als 0,6 als inakzeptabel. Das KMO wird für die gesamte Korrelationsmatrix ausgegeben, aber auch für die einzelnen Items, letzteres in Form der Diagonalen in der Anti-Image-Korrelationsmatrix.

Der „Bartlett-Test auf Sphärizität“ wird häufig ebenfalls zur Prüfung herangezogen. Er prüft die Hypothese, dass die Daten aus einer Population stammen könnten, bei der alle Korrelationskoeffizienten Null sind. Diese Hypothese sollte zurückgewiesen werden können.

- Die Befehle „Analysieren“, „Dimensionsreduktion“, „Faktorenanalyse...“ öffnen die Dialogbox „Faktorenanalyse“. Es werden WOHLSTAND bis POLEINST als Analysevariablen ausgewählt.



- Klicken auf die Schaltfläche „Deskriptive Statistik...“ öffnet die Unterdialogbox „Faktorenanalyse: Deskriptive Statistiken“. Markieren Sie dort die Optionen „Koeffizienten“, „Reproduziert“, „Anti-Image“ und „KMO und Bartlett-Test auf Sphärizität“. Beenden Sie mit „Weiter“ und „OK“.



Die ausgegebene Korrelationsmatrix sei hier nicht dargestellt. Sie ist wegen ihrer Größe etwas schwierig zu beurteilen. Insgesamt finden sich wenige sehr große Korrelationskoeffizienten, der größte ist 0,56. Dies spricht nicht unbedingt für eine sinnvolle Anwendung der Faktorenanalyse. Aber entscheidend ist, ob es einige Gruppen von Variablen mit hinreichend großen Korrelationskoeffizienten gibt.

Leichter anzuwenden ist das KMO-Maß. Es zeigt mit 0,82 einen recht guten Wert (ab 0,8 erhält das Maß die Einstufung lobenswert).

KMO- und Bartlett-Test

Maß der Stichprobeneignung nach Kaiser-Meyer-Olkin.		,820
Bartlett-Test auf Sphärizität	Ungefähres Chi-Quadrat	4086,321
	df	78
	Signifikanz nach Bartlett	,000

Sehen wir uns nun in der Anti-Image-Matrix die unter Hälfte und dort die Werte in der Diagonalen an, so haben wir die KMO-Werte für die einzelnen Items. Auch sie fallen gut aus. 0,573 ist der geringste Wert, die meisten Werte liegen über 0,8.

Anti-Image-Matrizen

		Wohlstand	Bildelt	Bildselbst	Ehrgeiz	Fähigkeit	Fleiss	Beziehungen	Polverb	Rasse	Religion	Region	Gender	PolEinst
Anti-Image-Kovarianz	Wohlstand	,686	-,207	,013	-,019	-,014	,052	-,115	-,069	-,078	,023	-,016	-,053	-,012
	Bildelt	-,207	,737	-,212	,027	-,050	,018	-,038	-,029	-,003	-,001	-,015	,006	-,039
	Bildselbst	,013	-,212	,803	-,170	,020	-,143	-,045	-,007	-,009	,048	,003	,006	-,001
	Ehrgeiz	-,019	,027	-,170	,839	-,068	-,210	-,024	,001	-,027	,048	-,002	,012	,031
	Fähigkeit	-,014	-,050	,020	-,068	,861	-,142	-,089	-,056	-,051	,022	-,033	,019	-,019
	Fleiss	,052	,018	-,143	-,210	-,142	,824	-,017	,039	,088	-,064	,011	-,005	-,022
	Beziehungen	-,115	-,038	-,045	-,024	-,089	-,017	,618	-,235	-,036	-,040	,004	-,021	,022
	Polverb	-,069	-,029	-,007	,001	-,056	,039	-,235	,564	-,054	-,007	-,009	-,018	-,133
	Rasse	-,078	-,003	-,009	-,027	-,051	,088	-,036	-,054	,630	-,114	-,112	-,151	,002
	Religion	,023	-,001	,048	,048	,022	-,064	-,040	-,007	-,114	,630	-,168	,039	-,179
	Region	-,016	-,015	,003	-,002	-,033	,011	,004	-,009	-,112	-,168	,610	-,120	-,096
	Gender	-,053	,006	,006	,012	,019	-,005	-,021	-,018	-,151	,039	-,120	,667	-,150
	PolEinst	-,012	-,039	-,001	,031	-,019	-,022	,022	-,133	,002	-,179	-,096	-,150	,563
Anti-Image-Korrelation	Wohlstand	,850 ^a	-,290	,018	-,026	-,018	,069	-,177	-,111	-,119	,035	-,025	-,078	-,019
	Bildelt	-,290	,780 ^a	-,276	,034	-,063	,023	-,056	-,045	-,004	-,002	-,022	,009	-,061
	Bildselbst	,018	-,276	,644 ^a	-,207	,024	-,176	-,065	-,010	-,012	,067	,004	,008	-,001
	Ehrgeiz	-,026	,034	-,207	,637 ^a	-,080	-,253	-,033	,002	-,037	,066	-,003	,016	,046
	Fähigkeit	-,018	-,063	,024	-,080	,835 ^a	-,168	-,122	-,080	-,069	,030	-,046	,025	-,027
	Fleiss	,069	,023	-,176	-,253	-,168	,573 ^a	-,024	,057	,122	-,088	,016	-,007	-,032
	Beziehungen	-,177	-,056	-,065	-,033	-,122	-,024	,821 ^a	-,398	-,057	-,064	,007	-,033	,037
	Polverb	-,111	-,045	-,010	,002	-,080	,057	-,398	,835 ^a	-,091	-,011	-,016	-,030	-,236
	Rasse	-,119	-,004	-,012	-,037	-,069	,122	-,057	-,091	,874 ^a	-,181	-,180	-,232	,004
	Religion	,035	-,002	,067	,066	,030	-,088	-,064	-,011	-,181	,810 ^a	-,272	,060	-,301
	Region	-,025	-,022	,004	-,003	-,046	,016	,007	-,016	-,180	-,272	,869 ^a	-,188	-,164
	Gender	-,078	,009	,008	,016	,025	-,007	-,033	-,030	-,232	,060	-,188	,859 ^a	-,245
	PolEinst	-,019	-,061	-,001	,046	-,027	-,032	,037	-,236	,004	-,301	-,164	-,245	,840 ^a

a. Maß der Stichprobeneignung

Das Ergebnis des „Bartlett-Tests auf Sphärizität“ findet sich ebenfalls im Output des KMO. Die Hypothese, die Daten könnten aus einer Population mit lauter Nullkorrelationen stammen, kann zurückgewiesen werden: Die Abweichung von der Nullhypothese ist signifikant. Allerdings sollte man auf diesen Test nicht viel geben. Er leidet unter dem Mangel aller Tests, bei denen die Nullhypothese die eigentliche Zielhypothese ist. Bei großen Stichproben kommt praktisch immer, auch bei minimalen Abweichungen, eine signifikante Abweichung von der Nullhypothese zustande. Da wir mit 1335 Personen eine recht große Stichprobe haben, kann man sich diesen Test praktisch sparen. Er wird auch bei sehr kleinen Korrelationskoeffizienten auf jeden Fall eine signifikante Abweichung von der Nullhypothese indizieren.

Insgesamt spricht aber insbesondere der KMO-Test für die Brauchbarkeit der Daten für eine Faktorenanalyse.

Zu klären wäre u.U. auch noch, ob die Stichprobengröße für eine Faktorenanalyse ausreicht. Da die benötigte Stichprobengröße von vielen Komponenten abhängt, kann dies hier nur überschlägig geprüft werden.

Gängige Kriterien sind: 10 Fälle pro Variable oder mindestens 300 Fälle. Bei 13 Variablen wären dies nach dem ersten Kriterium mindestens 130 Fälle. Dies ist in unserem Falle weit überschritten, ebenso das zweite Kriterium.

b.

- Das betrachtete Modell lässt sich wie folgt beschreiben: Jede Variable kann aus Common Factors und Unique Faktor erklärt werden:

$$V_x = a_1F_1 + a_2F_2 + a_3F_3 \dots + U_x$$

a = Koeffizienten oder Ladungen.

Umgekehrt ist ein Faktor erklärt aus den Variablen:

$$F_i = w_1V_1 + w_2V_2 + \dots + w_xV_x$$

w = Faktor Score Koeffizient.

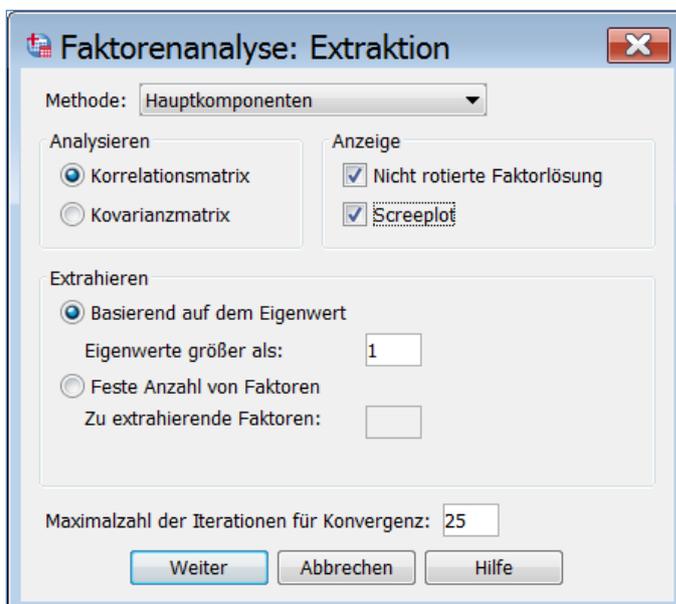
- Wir wollen mit dem voreingestellten Extraktionsmodell „Hauptkomponenten“ arbeiten und als Rotationsverfahren „Varimax“ verwenden.

Zur Bestimmung der Anzahl der Faktoren benutzt man in der Regel die Eigenwerte. Da bei der Anfangslösung (unabhängig von der verwendeten Extraktionsmethode) immer die „Hauptkomponentenmethode“ verwendet wird, die so viele Faktoren wie Variablen ausgibt, benutzt man in der Regel für die weitere Analyse so viele Faktoren, wie bei der Anfangslösung Eigenwerte größer gleich 1 besitzen. Alle anderen Faktoren erklären weniger von der Varianz als es die Variablen alleine erklären würde. (Allerdings kann man von dieser Zahl abweichen, wenn man z.B. aus inhaltlichen Gründen eine Lösung mit weniger oder mehr Faktoren vorzieht. Dann wäre die Zahl der gewünschten Faktoren einzustellen oder aber man stellt einen Eigenwert $\neq 1$ ein). Zur Bestimmung der Zahl der Faktoren wird auch häufig ein Screeplot herangezogen. Dazu wählen Sie:

- „Analysieren“, „Dimensionsreduktion“, „Faktorenanalyse“

Wählen Sie die zu analysierenden Variablen WOHLSTAND bis POLEINST aus.

Klicken Sie auf die Schaltfläche „Extraktion...“. Es öffnet sich die Unterdialogbox „Faktorenanalyse: Extraktion“. Stellen Sie, falls nicht bereits ausgewählt „Hauptkomponenten“ als Extraktionsmethode ein. Klicken Sie dann auf das Kontrollkästchen „Screeplot“. Bestätigen Sie die Auswahl mit „Weiter“ und „OK“.



Als Ergebnis zeigt sich u.a.:

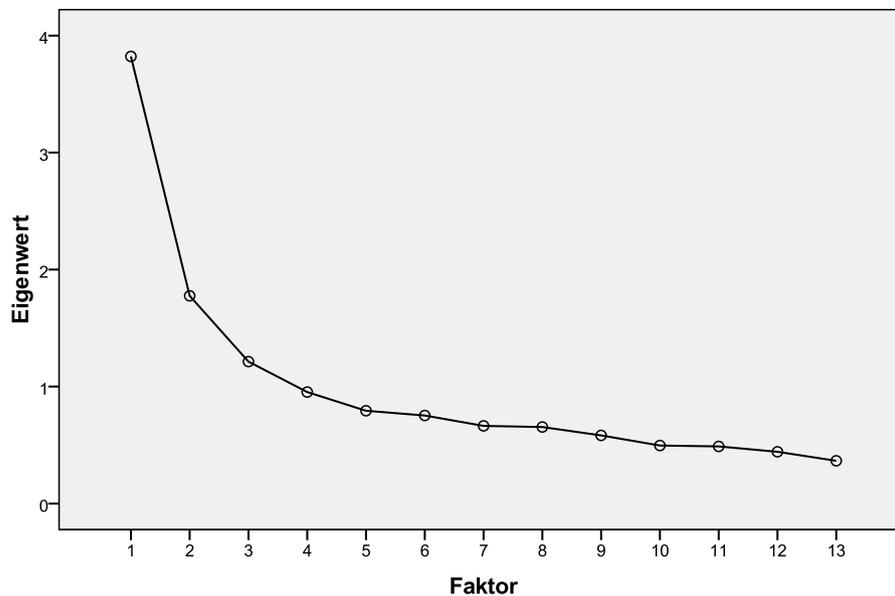
Erklärte Gesamtvarianz

Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	3,822	29,402	29,402	3,822	29,402	29,402
2	1,776	13,660	43,062	1,776	13,660	43,062
3	1,213	9,332	52,394	1,213	9,332	52,394
4	,953	7,329	59,723			
5	,793	6,099	65,822			
6	,752	5,788	71,610			
7	,664	5,104	76,714			
8	,654	5,029	81,743			
9	,582	4,479	86,222			
10	,496	3,814	90,036			
11	,488	3,755	93,790			
12	,442	3,401	97,191			
13	,365	2,809	100,000			

Extraktionsmethode: Hauptkomponentenanalyse.

Aus der Tabelle kann man entnehmen, dass 3 Faktoren einen Eigenwert größer 1 besitzen, die zusammen ca. 52 % der Gesamtvarianz erklären.

Screepplot



Auch der Screepplot geht nach 3 Faktoren deutlich in eine flachere Kurve über, so dass auch dieser für eine Drei-Faktoren-Lösung spricht. (Allerdings könnte man sich u.U. auch mit einer 4-Faktoren-Lösung anfreunden, da der Eigenwert des vierten Faktors noch nahe bei 1 liegt und der Verlauf des Screepplots nach 4 Faktoren noch deutlich flacher wird).

Wir wollen mit der Drei-Faktoren-Lösung weiter arbeiten.

c.

- Dazu wählen Sie:

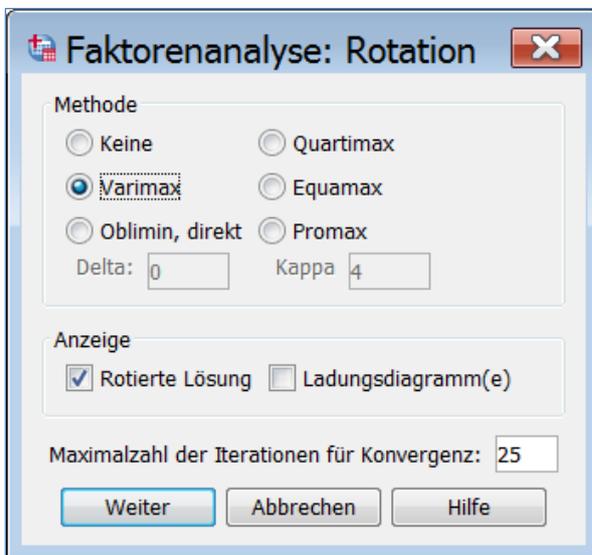
„Analysieren“, „Dimensionsreduktion“, „Faktorenanalyse“.

Wählen Sie die zu analysierenden Variablen WOHLSTAND bis POLEINST aus.

(Schalten Sie in der Dialogbox „Faktorenanalyse: Extraktion“ die Optionen „Screeplot“ und „Nicht rotierte Faktorenlösung“ aus).

- Klicken Sie auf die Schaltfläche „Rotation...“. Es erscheint die Dialogbox „Faktorenanalyse: Rotation“.

Wählen Sie als Methode „Varimax“. Bestätigen Sie die Auswahl mit „Weiter“ und „OK“.



Von der sich ergebenden Ausgabe interessiert uns hier nur die „rotierte“ Komponentenmatrix.

Rotierte Komponentenmatrix^a

	Komponente		
	1	2	3
Wohlstand Man benötigt eine wohlhabende Familie um voranzukommen	,193	,734	-,055
Bildet Man benötigt gut ausgebildete Eltern um voranzukommen	,040	,704	,134
Bildselbst Man benötigt gute Ausbildung um voranzukommen	-,132	,356	,568
Ehrgeiz Man muss ehrgeizig sein um voranzukommen	-,078	,033	,716
Fähigkeit Man muss natürliche Fähigkeiten besitzen um voranzukommen	,206	,262	,404
Fleiss Man muss hart arbeiten um voranzukommen	,040	-,184	,781
Beziehungen Man muss die richtigen Leute kennen um voranzukommen	,280	,658	,122
Polverb Man muss politische Verbindungen haben um voranzukommen	,432	,601	,019
Rasse Man muss zur richtigen Rasse gehören um voranzukommen	,631	,316	-,055
Religion Man muss zur richtigen Religion gehören um voranzukommen	,764	,006	,017
Region Man muss in der richtigen Region wohnen um voranzukommen	,768	,103	,019
Gender Man muss das richtige Geschlecht haben um voranzukommen	,641	,216	-,034
PolEinst Man muss eine bestimmte politische Einstellung haben um voranzukommen	,743	,199	,024

Extraktionsmethode: Hauptkomponentenanalyse.
Rotationsmethode: Varimax mit Kaiser-Normalisierung.

a. Die Rotation ist in 5 Iterationen konvergiert.

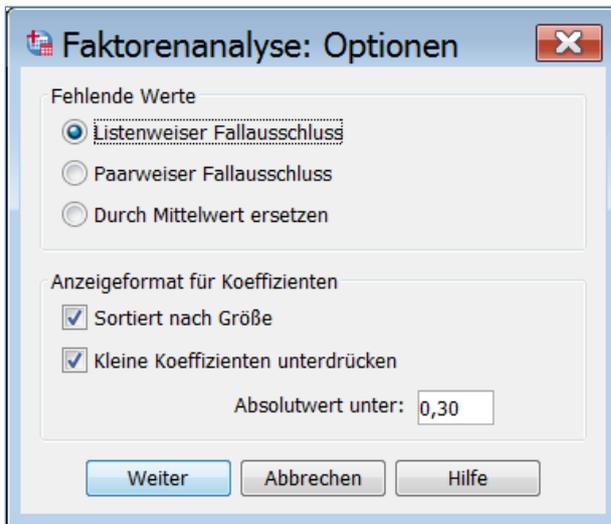
Ihr entnehmen wir, welche Variable wie hoch auf welchem Faktor lädt.

Zum Beispiel lädt die erste Variable „Wohlstand“ besonders hoch auf dem zweiten Faktor, nämlich mit der Ladung 0,734, die Ladung auf den beiden anderen Faktoren ist gering. Generell ist eine Lösung gut, wenn jede Variable nur auf einem Faktor hoch lädt.

Etwas übersichtlicher gestaltet sich das, wenn wir die Variablen nach der Höhe der Faktorladung ordnen lassen und gleichzeitig die Ausgabe von geringen Ladungen unterdrücken (wir unterdrücken Ladungen unter 0,3).

- Übernehmen Sie dazu die vorige Einstellung und ergänzen Sie diese wie folgt:
 - Klicken Sie auf die Schaltfläche „Optionen“. Das Dialogfenster „Faktorenanalyse: Optionen“ öffnet sich.
 - Markieren Sie im Feld „Anzeigeformat für Koeffizienten“ das Kontrollkästchen „Sortiert nach Größe“. Danach das Kontrollkästchen „Kleine Koeffizienten unterdrücken“ und tragen im Eingabefeld „Absolutwert unter:“ in das zugehörige Eingabefeld 0.3 ein.

Bestätigen Sie mit „Weiter“ und „OK“.



Die rotierte Komponentenmatrix sieht jetzt wie folgt aus:

Rotierte Komponentenmatrix^a

	Komponente		
	1	2	3
Region Man muss in der richtigen Region wohnen um voranzukommen	,768		
Religion Man muss zur richtigen Religion gehören um voranzukommen	,764		
PolEinst Man muss eine bestimmte politische Einstellung haben um voranzukommen	,743		
Gender Man muss das richtige Geschlecht haben um voranzukommen	,641		
Rasse Man muss zur richtigen Rasse gehören um voranzukommen	,631	,316	
Wohlstand Man benötigt eine wohlhabende Familie um voranzukommen		,734	
Bildelt Man benötigt gut ausgebildete Eltern um voranzukommen		,704	
Beziehungen Man muss die richtigen Leute kennen um voranzukommen		,658	
Polverb Man muss politische Verbindungen haben um voranzukommen	,432	,601	
Fleiss Man muss hart arbeiten um voranzukommen			,781
Ehrgeiz Man muss ehrgeizig sein um voranzukommen			,716
Bildselbst Man benötigt gute Ausbildung um voranzukommen		,356	,568
Fähigkeit Man muss natürliche Fähigkeiten besitzen um voranzukommen			,404

Extraktionsmethode: Hauptkomponentenanalyse.
 Rotationsmethode: Varimax mit Kaiser-Normalisierung.

a. Die Rotation ist in 5 Iterationen konvergiert.

Wir sehen, dass fast alle Variablen nur auf einem Faktor hoch laden. Nur 3 haben Ladungen höher als 0,3 auf zwei Faktoren.

Außerdem gewinnen wir aus den Bezeichnungen der Variablen, die auf demselben Faktor laden, Hinweise auf die Art des Faktors und können evtl. eine geeignete Bezeichnung für den jeweiligen Faktor daraus gewinnen.

Wir können z.B. die Faktoren wie folgt benennen:

Faktor 1: Soziale Eigenschaften

Faktor 2: Familienfaktor

Faktor 3: Persönliche Fähigkeiten.

Erklärte Gesamtvarianz

Komponente	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	3,822	29,402	29,402	3,822	29,402	29,402	2,905	22,344	22,344
2	1,776	13,660	43,062	1,776	13,660	43,062	2,256	17,356	39,699
3	1,213	9,332	52,394	1,213	9,332	52,394	1,650	12,695	52,394
4	,953	7,329	59,723						
5	,793	6,099	65,822						
6	,752	5,788	71,610						
7	,664	5,104	76,714						
8	,654	5,029	81,743						
9	,582	4,479	86,222						
10	,496	3,814	90,036						
11	,488	3,755	93,790						
12	,442	3,401	97,191						
13	,365	2,809	100,000						

Extraktionsmethode: Hauptkomponentenanalyse.

Gegenüber der Anfangslösung ändert sich im Übrigen der Anteil der durch die Faktoren gemeinsam erklärten Varianz nicht, wohl aber die durch die einzelnen Faktoren erklärte Varianz.

d.

- Jetzt wollen wir die drei Faktoren als neue Variable der Datei speichern. Jeder Fall bekommt also für jeden der Faktoren einen Wert zugeordnet, den wir später z.B. für weitere Analysen verwenden können. Zur Kalkulation der Faktorscores benutzen wir die voreingestellte Methode „Regression“.

Wählen Sie zusätzlich zu der letzten Einstellung Folgendes:

Klicken Sie auf die Schaltfläche „Scores“. Die Dialogbox „Faktorenanalyse: Faktorscores“ öffnet sich. Aktivieren Sie das Kontrollkästchen „Als Variable speichern“ und wählen Sie den Optionsschalter für die Methode „Regression“. Aktivieren Sie auch das Kontrollkästchen „Koeffizientenmatrix für Faktorwerte anzeigen“. Bestätigen Sie mit „Weiter“ und „OK“.



Aus der Koeffizientenmatrix der Faktoren kann man entnehmen, mit welchen Gewichten die Werte der einzelnen Variablen in die Berechnung eines Faktorwerts eingehen.

So wird der Faktor 1 berechnet aus: $-.091 * \text{Wohlstand} + -.143 * \text{Bildelt} + \dots + 0,281 * \text{PolEinst}$

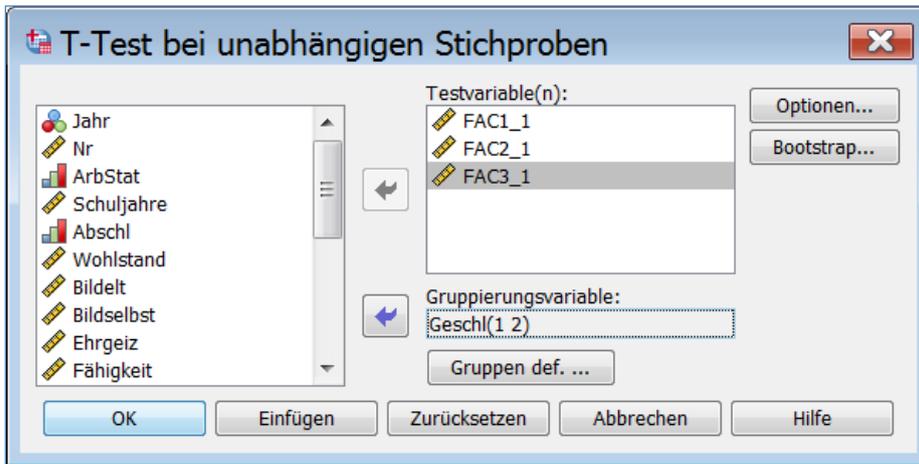
Koeffizientenmatrix der Komponentenwerte

	Komponente		
	1	2	3
Wohlstand Man benötigt eine wohlhabende Familie um voranzukommen	-,091	,388	-,107
Bildelt Man benötigt gut ausgebildete Eltern um voranzukommen	-,143	,386	,008
Bildselbst Man benötigt gute Ausbildung um voranzukommen	-,118	,176	,311
Ehrgeiz Man muss ehrgeizig sein um voranzukommen	-,011	-,041	,442
Fähigkeit Man muss natürliche Fähigkeiten besitzen um voranzukommen	,046	,059	,233
Fleiss Man muss hart arbeiten um voranzukommen	,094	-,202	,512
Beziehungen Man muss die richtigen Leute kennen um voranzukommen	-,027	,304	,016
Polverb Man muss politische Verbindungen haben um voranzukommen	,049	,246	-,035
Rasse Man muss zur richtigen Rasse gehören um voranzukommen	,201	,041	-,042
Religion Man muss zur richtigen Religion gehören um voranzukommen	,335	-,179	,043
Region Man muss in der richtigen Region wohnen um voranzukommen	,314	-,123	,034
Gender Man muss das richtige Geschlecht haben um voranzukommen	,229	-,021	-,018
PolEinst Man muss eine bestimmte politische Einstellung haben um voranzukommen	,281	-,063	,025

Extraktionsmethode: Hauptkomponentenanalyse.
 Rotationsmethode: Varimax mit Kaiser-Normalisierung.
 Komponentenwerte.

Als Letztes benutzen wir die Faktorscores zum Vergleich der Geschlechter. Wir vergleichen die Mittelwerte auf den drei Faktoren und benutzen dazu den t-Test für unabhängige Stichproben. („Analysieren“, „Mittelwerte vergleichen“, „T-Test bei unabhängigen Stichproben“. „Testvariable(n)“ FAC1-1, FAC2-1, FAC3-1, „Gruppenvariable“ GESCHL. Klicken auf „Grup-

pen definieren...“. Optionsschalter „Angegebene Werte verwenden“, in Eingabefeld „Gruppe 1“ 1, für Gruppe 2“ 2 eintragen.) Beenden Sie mit „Weiter“ und „OK“.



Gruppenstatistiken

	Geschl	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
FAC1_1	1 Männlich	580	,0049705	,99692487	,04139505
	2 Weiblich	755	-,0038184	1,00299961	,03650289
FAC2_1	1 Männlich	580	-,1277951	,99817278	,04144687
	2 Weiblich	755	,0981737	,99091692	,03606316
FAC3_1	1 Männlich	580	,0614535	1,02136029	,04240968
	2 Weiblich	755	-,0472093	,98134290	,03571472

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
									Untere	Obere
FAC1_1	Varianzen sind gleich	,411	,521	,159	1333	,874	,00878885	,05523468	-,09956753	,11714522
	Varianzen sind nicht gleich			,159	1249,425	,874	,00878885	,05519068	-,09948779	,11706549
FAC2_1	Varianzen sind gleich	,007	,935	-4,117	1333	,000	-,22596885	,05488736	-,33364386	-,11829383
	Varianzen sind nicht gleich			-4,113	1241,245	,000	-,22596885	,05493992	-,33375421	-,11818349
FAC3_1	Varianzen sind gleich	,050	,823	1,970	1333	,049	,10866272	,05515497	,00046273	,21686272
	Varianzen sind nicht gleich			1,960	1220,192	,050	,10866272	,05544477	-,00011493	,21744037

Die Analyse ergibt für den ersten Faktor keinen signifikanten Unterschied zwischen Männern und Frauen.

Bei den beiden anderen Faktoren zeigen sich signifikante Unterschiede. Den Faktor "Familie" (FAC2_1) betonen Männer tendenziell eher als Frauen, dagegen betonen die Frauen eher den Faktor "persönliche Fähigkeiten" (FAC3_1).