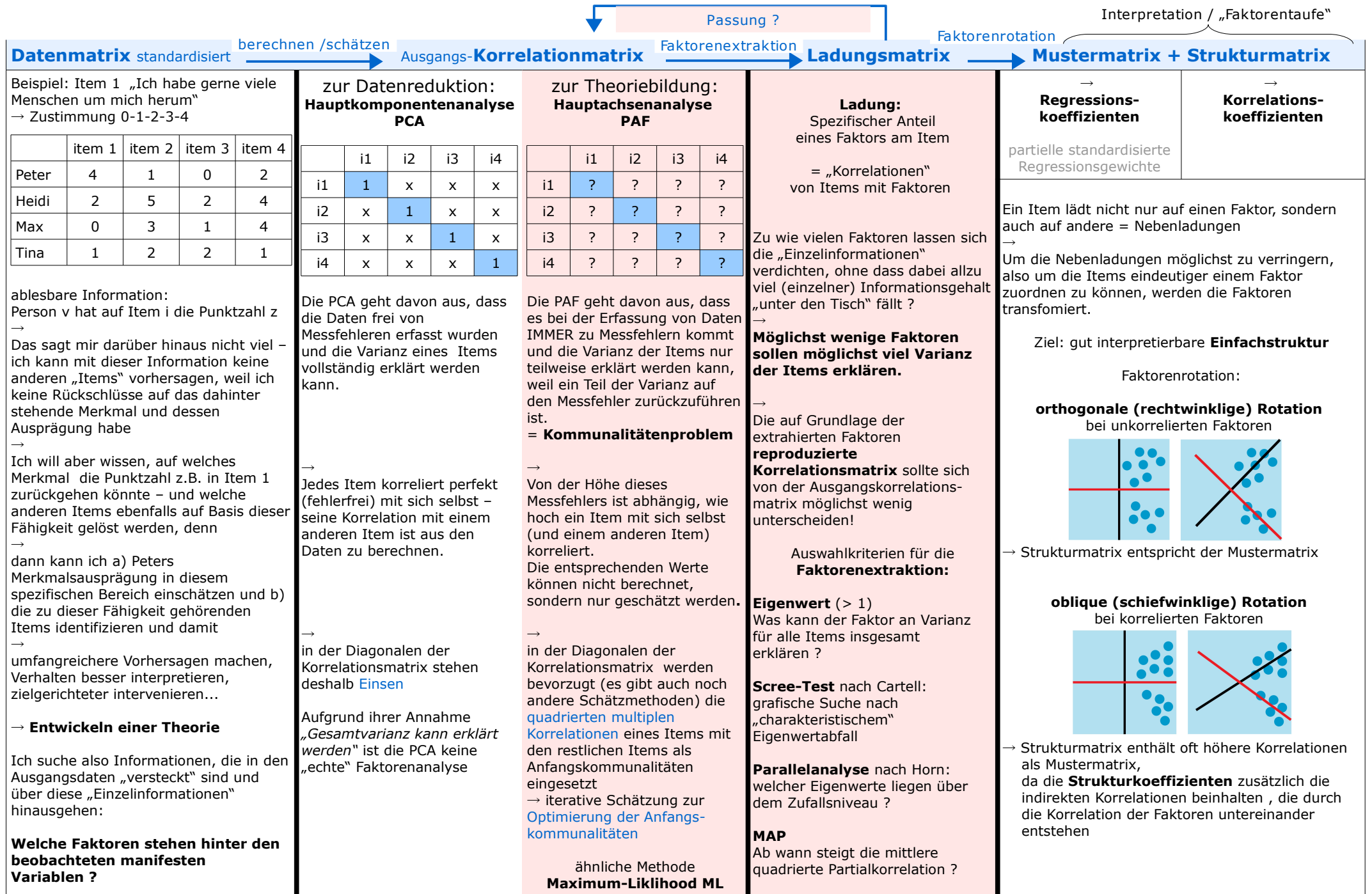


29. Testkonstruktion explorative Faktorenanalyse Spickzettel



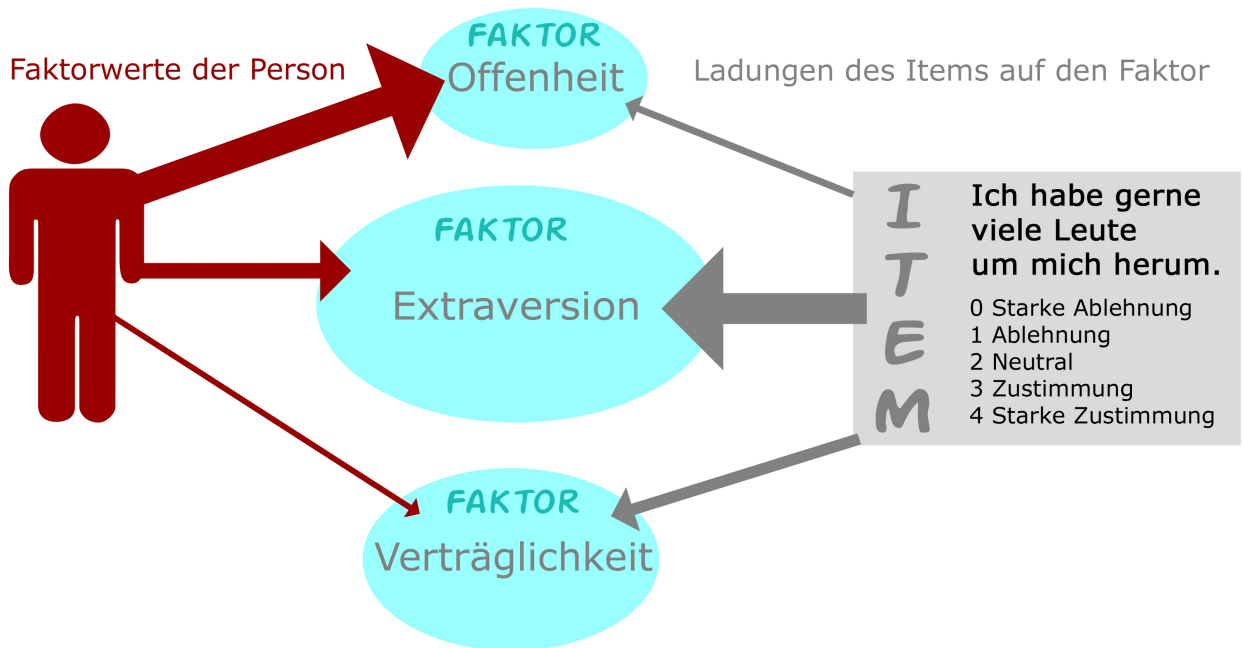
Kern der EFA

| | | | |
|--|--|--|---|
| Was wird untersucht ? | Vergleich | | |
| | Faktorenanalyse | Varianzanalyse Regressionsanalyse | |
| | Korrelationsmuster von mehrere manifesten Variablen | Zusammenhänge zwischen | |
| | | UV + AV | Prädiktor + Kriterium |
| | Faktor = latente Variable , die die Korrelationen der manifesten Variablen erklärt | Faktor = | |
| | | UV, die die Ausprägung der AV bestimmt | Prädiktor, der die Ausprägung des Kriteriums vorhersagt |
| | Items als Effektindikatoren → korrelieren untereinander und zeigen in der Höhe der jeweiligen Korrelationen den Effekt des gemeinsamen Faktors an | Items als Kausalindikatoren → müssen nicht untereinander, sondern nur mit dem Faktor korrelieren und zeigen quasi die Ausprägung des Faktors an (vgl. rationale Testkonstruktion) | |
| 3 Zwecke der Faktorenanalyse Thompson 2004 | Validitätsprüfung Gültigkeit von Modellannahmen bezüglich der empirischen Daten ? | Theorie-Entwicklung + Datenreduktion | |
| | → konfirmatorische Faktorenanalyse CFA | interne Struktur psychologischer Konstrukte ? | Zusammenfassung von Variablen zu Faktoren → Hauptkomponentenanalyse PCA |
| | | → exploratorische Faktorenanalyse EFA | |
| Voraussetzungen → abhängig von der angewandten Methode Inhaltliche Erwägungen vor Faustregeln! | Korrelation der Items ohne die gibt's offensichtlich keine gemeinsamen Faktoren | Stichproben- größe | Verteilungseigenschaften der Items |
| | <p>Tests:</p> <p>Bartlett-Test auf Sphärizität → testet die H0 „totale Unabhängigkeit“ → sollte signifikant werden</p> <p>Kaiser-Meyer-Olkin-Koeffizient → testet Korrelation → sollte möglichst hoch sein nicht < .50</p> <p>Kommunalitäten (= Mindestreliabilität) einzelner Items nicht < .20</p> <p>MSA-Koeffizient (ähnlich wie KMO, für einzelne Items)</p> | <ul style="list-style-type: none"> • Zahl der Items • Verhältnis Items/Faktoren • Reliabilität der Items → sollte um so größer sein, je kleiner die Stichprobe ist | <ul style="list-style-type: none"> • mindestens Intervallskalenniveau für die oben genannten Verfahren – sonst andere Koeffizienten → z.B. <u>tetrachorische, polychorische Korrelation</u> bei ordinal skalierten Variablen • univariate Normalverteilung → Varianz = 1 Für ML: sogar multivariat <p>Meist robust gegen geringe Abweichungen (Schiefe, Exzess) bei großer Abweichung:</p> <ul style="list-style-type: none"> • Transformation • „Parceling“ (z.B. Zusammenfassen von 2 Items mit gegenteiliger Schiefe) • u.a.m. |

Modell mehrerer gemeinsamer Faktoren

Die Ausprägung einer Itemantwort einer Person wird beeinflusst

1. von den hinter diesem Item stehenden Faktoren und
2. von der Ausprägung der Person auf den jeweiligen Faktoren



Definitionsgleichung

$$Z_{im} = f_{i1} \cdot a_{m1} + f_{i2} \cdot a_{m2} + f_{i3} \cdot a_{m3} \dots + f_{ij} \cdot a_{mj} + \dots + f_{iq} \cdot a_{mq} + e_i$$

| | |
|---|--|
| Z_{im} = z-Wert einer Person i auf einem Item m | Grad der Zustimmung 0 - 1 - 2 - 3 - 4 |
| f_{i1} = Faktorwert der Person i auf Faktor 1 | |
| a_{m1} = Ladung des Items m auf Faktor 1 | |
| f_{ij} = Faktorwert der Person i auf Faktor j | |
| a_{mj} = Ladung des Items m auf Faktor j | |
| q = Anzahl der Faktoren | |
| e_i = Fehlerkomponente, die durch die extrahierten Faktoren nicht erklärt werden kann | |

Problem:

in dieser Gleichung ist zunächst nur der **z-Werte der Person** auf dem Item bekannt:
→ Wie setzt sich dieser Wert aus **Ladungen, Faktorwerten** und **Fehlern** zusammen ?

Aufgabe:

Finde Werte für Ladung, Faktorwert (und Fehler), die zu ALLEN z-Werten ALLER Personen über ALLE Items passen !

Versuche, dabei die Anzahl der erklärenden Faktoren möglichst gering zu halten, ohne auf zu viel Information zu verzichten !

Die *aus deinen Werten* berechneten z-Werte sollten möglichst nahe an die ursprünglichen Werte rankommen !

Grundannahme

erklärt wird




erstellt aus der Ausgangsdatenmatrix

(z-Werte der Stichprobe)

| | i1 | i2 | i3 | i4 |
|---|----|----|----|----|
| P | 4 | 1 | 0 | 2 |
| H | 2 | 5 | 2 | 4 |
| M | 0 | 3 | 1 | 4 |
| T | 1 | 2 | 2 | 1 |

| | Hauptkomponentenanalyse P rinciple C omponents A nalasys | Hauptachsenanalyse P rinciple A xes F actoring | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|--|---|--|----|----|----|---|---|---|---|----|---|---|---|---|----|---|---|---|---|----|---|---|---|---|---|--|----|----|----|----|----|---|---|---|---|----|---|---|---|---|----|---|---|---|---|----|---|---|---|---|
| Grundannahme | <p>Varianz der Items ist vollständig erklärbar.</p> <p style="text-align: center;">↓</p> <p>im Rahmen psychologischer Messungen keine „realistische“ Annahme</p> <p>→ keine Faktoren-Analyse, nur „Faktoren-Beschreibung“</p> | <p>Varianz der Items ist nicht vollständig erklärbar.</p> <div style="border: 1px solid black; padding: 5px; margin: 10px 0;"> <p style="text-align: center; color: blue;">Kommunalitätenproblem:</p> <p style="text-align: center;">Es gibt Varianzanteile, die ein Item mit keinem anderen teilt = Einzigartigkeit/Uniqueness</p> </div> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; text-align: center;"> aufgrund von Messfehlern → Fehlervarianz </td> <td style="width: 50%; text-align: center;"> aufgrund im Datensatz nicht enthaltener Faktoren → spezifische Varianz/Spezifität </td> </tr> </table> | aufgrund von Messfehlern → Fehlervarianz | aufgrund im Datensatz nicht enthaltener Faktoren → spezifische Varianz/Spezifität | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| aufgrund von Messfehlern → Fehlervarianz | aufgrund im Datensatz nicht enthaltener Faktoren → spezifische Varianz/Spezifität | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| erklärt wird | Gesamtvarianz (=1) | gemeinsame Varianz (= Kommunalität) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <p>Ausgangs-Korrelationsmatrix: berechnet</p> <table border="1" style="border: 2px solid green; margin: 10px 0;"> <thead> <tr> <th></th> <th>i1</th> <th>i2</th> <th>i3</th> <th>i4</th> </tr> </thead> <tbody> <tr> <th>i1</th> <td style="background-color: #d3d3d3;">1</td> <td>x</td> <td>x</td> <td>x</td> </tr> <tr> <th>i2</th> <td>x</td> <td style="background-color: #d3d3d3;">1</td> <td>x</td> <td>x</td> </tr> <tr> <th>i3</th> <td>x</td> <td>x</td> <td style="background-color: #d3d3d3;">1</td> <td>x</td> </tr> <tr> <th>i4</th> <td>x</td> <td>x</td> <td>x</td> <td style="background-color: #d3d3d3;">1</td> </tr> </tbody> </table> <p>Die paarweisen Korrelationen der Items lassen sich aus den Verteilungen berechnen - jedes Item korreliert mit sich selbst in der Höhe von 1.</p> <p>→ in der Hauptdiagonalen: Einsen</p> <p>Ein-Schritt-Verfahren</p> <p style="background-color: yellow; text-align: center; margin-top: 10px;">→ führt immer zu einer Lösung</p> | | i1 | i2 | i3 | i4 | i1 | 1 | x | x | x | i2 | x | 1 | x | x | i3 | x | x | 1 | x | i4 | x | x | x | 1 | <p>Ausgangs-Korrelationsmatrix: geschätzt</p> <table border="1" style="border: 2px solid green; margin: 10px 0;"> <thead> <tr> <th></th> <th>i1</th> <th>i2</th> <th>i3</th> <th>i4</th> </tr> </thead> <tbody> <tr> <th>i1</th> <td style="background-color: #d3d3d3;">?</td> <td>?</td> <td>?</td> <td>?</td> </tr> <tr> <th>i2</th> <td>?</td> <td style="background-color: #d3d3d3;">?</td> <td>?</td> <td>?</td> </tr> <tr> <th>i3</th> <td>?</td> <td>?</td> <td style="background-color: #d3d3d3;">?</td> <td>?</td> </tr> <tr> <th>i4</th> <td>?</td> <td>?</td> <td>?</td> <td style="background-color: #d3d3d3;">?</td> </tr> </tbody> </table> <p>als Anfangskommunalitäten in der Hauptdiagonalen:</p> <ul style="list-style-type: none"> Einsen (dann entspricht der 1. „Schätzschritt“ der PCA) höchste Korrelation des Items mit <u>einem</u> anderen quadierte multiple Korrelationen des Items mit <u>allen</u> anderen Items (bevorzugt) <p>→ iteratives Schätzverfahren bis zur Konvergenz von reproduzierter Korrelationsmatrix und Ausgangs-Korrelationsmatrix</p> <p>Die Anfangskommunalitäten werden im Laufe der folgenden Faktoren-Extraktion durch die jeweils neuen geschätzten Kommunalitäten immer wieder ersetzt, bis die Residualmatrix (= Differenz zwischen Ausgangskorrelationsmatrix und reproduzierter Korrelationsmatrix) möglichst klein wird</p> <p style="background-color: yellow; text-align: center;">→ im ungünstigsten Fall konvergieren die Schätzungen nicht.</p> | | i1 | i2 | i3 | i4 | i1 | ? | ? | ? | ? | i2 | ? | ? | ? | ? | i3 | ? | ? | ? | ? | i4 | ? | ? | ? | ? |
| | i1 | i2 | i3 | i4 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| i1 | 1 | x | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| i2 | x | 1 | x | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| i3 | x | x | 1 | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| i4 | x | x | x | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | i1 | i2 | i3 | i4 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| i1 | ? | ? | ? | ? | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| i2 | ? | ? | ? | ? | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| i3 | ? | ? | ? | ? | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| i4 | ? | ? | ? | ? | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

33. Testkonstruktion explorative Faktorenanalyse Faktoren-Extraktion 1

| <p>Ladungsmatrix mit extrahierten Faktoren</p> | <p>→ geordnet nach dem sukzessiven Beitrag des Faktors zur Varianzaufklärung</p> <p>→ Jeder neue Faktor erklärt dabei nur Varianz, die von den zuvor extrahierten Faktoren nicht erklärt wurde.</p> | <p>Der erste Faktor erklärt insgesamt die meiste Varianz, der zweite erklärt von der restlichen Varianz das meiste, der dritte Faktor...</p> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|--|--|-------------|---|-------------|---|---|--|--|--|--|--|--|--|--|--|--|---|--|--|----------------------------|--|----------|----------|----------|--|--|--|---------------------------------|-------|-----|-----|-----|---------------------------|-------------|--|--------------------------|-------|-----|-----|-----|---------------------------|-------------|--|----------------------------|-------|-----|-----|-----|---------------------------|-------------|--|-------------------------|-------|-----|-----|-----|---------------------------|-------------|--|----------------------------------|-----------------------|------------|-------------|-------------|---|--|-------------|---|--|--|--|--|--|--|---|----------------------------|
| <p>Anzahl der Faktoren</p> <p>→ zur Berechnung der reproduzierten Korrelations-Matrix</p> <p>→ Berücksichtigung mehrerer Kriterien + inhaltlicher Plausibilität</p> | <p>festzulegen nach dem Kriterium der Varianzmaximierung</p> <div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; padding: 5px;"> <p>PCA:</p> <p>möglichst hohe Varianzaufklärung mit möglichst wenig Faktoren</p> </div> <div style="border: 1px solid black; padding: 5px;"> <p>PAF:</p> <p>möglichst hohe Varianzaufklärung durch möglichst wenige Faktoren bei möglichst wenig Informationsverlust</p> <p>→ Verglichen mit der Ausgangs-Korrelationmatrix</p> </div> </div> <p>Auswahlkriterien</p> <div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; padding: 5px;"> <p>Kaiser-Guttman-Kriterium: Eigenwert > 1</p> <p>(führt oft zur Extraktion vieler Faktoren, wird daher in der PAF eher nicht empfohlen)</p> <p><i>evtl. werden alle Faktoren extrahiert</i></p> </div> <div style="border: 1px solid black; padding: 5px;"> <p>Scree-Test Parallelanalyse MAP</p> <p>evtl. auch nach theoretischen Vorstellungen</p> </div> </div> <div style="border: 1px dashed black; padding: 2px; margin-left: auto; margin-right: auto;"> <p>vlg. Prokrustes-Rotation</p> </div> <p>übrigens: Je höher die gemessenen Variablen untereinander korreliert sind, desto weniger Faktoren sind zur Varianzaufklärung erforderlich.</p> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| <p>Eigenwerte + Kommunalitäten + Varianzaufklärung</p> | <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 10%;"></th> <th style="width: 10%;"></th> <th style="width: 10%;"></th> <th style="width: 10%;"></th> <th style="width: 10%;"></th> <th style="width: 10%;"></th> <th style="width: 10%;"></th> <th style="width: 10%;"></th> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td colspan="2" style="text-align: center;">Kommunalität = gemeinsame mit anderen Items geteilte Varianz des Items</td> <td></td> <td style="text-align: center;">aufgeklärter Varianzanteil</td> </tr> <tr> <td></td> <td style="text-align: center;">Faktor 1</td> <td style="text-align: center;">Faktor 2</td> <td style="text-align: center;">Faktor 3</td> <td colspan="2" style="text-align: center;">→ Zeilen-Summen der quadrierten Faktorladungen</td> <td></td> <td style="text-align: center;">kumulierte % der Kommunalitäten</td> </tr> </thead> <tbody> <tr> <td style="text-align: center;">item1</td> <td style="text-align: center;">0,8</td> <td style="text-align: center;">0,1</td> <td style="text-align: center;">0,1</td> <td style="text-align: center;">$0,8^2 + 0,1^2 + 0,1^2 =$</td> <td style="text-align: center;">0,66</td> <td></td> <td style="text-align: center;">$(0,66/4)*100$ 16,5 %</td> </tr> <tr> <td style="text-align: center;">item2</td> <td style="text-align: center;">0,2</td> <td style="text-align: center;">0,7</td> <td style="text-align: center;">0,0</td> <td style="text-align: center;">$0,2^2 + 0,7^2 + 0,0^2 =$</td> <td style="text-align: center;">0,53</td> <td></td> <td style="text-align: center;">$+(0,53/4)*100$ 29,75 %</td> </tr> <tr> <td style="text-align: center;">item3</td> <td style="text-align: center;">0,6</td> <td style="text-align: center;">0,2</td> <td style="text-align: center;">0,1</td> <td style="text-align: center;">$0,6^2 + 0,2^2 + 0,1^2 =$</td> <td style="text-align: center;">0,41</td> <td></td> <td style="text-align: center;">$+(0,41/4)*100$ 40 %</td> </tr> <tr> <td style="text-align: center;">item4</td> <td style="text-align: center;">0,4</td> <td style="text-align: center;">0,3</td> <td style="text-align: center;">0,1</td> <td style="text-align: center;">$0,4^2 + 0,3^2 + 0,1^2 =$</td> <td style="text-align: center;">0,26</td> <td></td> <td style="text-align: center;">$+(0,26/4)*100$ 46,5 %</td> </tr> <tr> <td style="text-align: center;">Eigenwert des Faktors</td> <td style="text-align: center;">1,2</td> <td style="text-align: center;">0,63</td> <td style="text-align: center;">0,03</td> <td colspan="2" style="text-align: center;">→ → → → → → → → → → bei unkorrelierten Faktoren (nach orthogonaler Rotation) entspricht die Zeilen-Summe der Eigenwerte ebenfalls der aufgeklärten Varianz</td> <td style="text-align: center;">1,86</td> <td style="text-align: center;">46,5 % → bei Extraktion aller Faktoren in der PCA = 100 %</td> </tr> <tr> <td colspan="6" style="text-align: center;">Bei meinen erfundenen Beispiel-Daten wär die aufgeklärte Gesamtvarianz jetzt nicht grad so groß...</td> <td style="text-align: center;"></td> <td style="text-align: center;">gesamt aufgeklärte Varianz</td> </tr> </tbody> </table> | | | | | | | | | | | | | | | | | Kommunalität = gemeinsame mit anderen Items geteilte Varianz des Items | | | aufgeklärter Varianzanteil | | Faktor 1 | Faktor 2 | Faktor 3 | → Zeilen-Summen der quadrierten Faktorladungen | | | kumulierte % der Kommunalitäten | item1 | 0,8 | 0,1 | 0,1 | $0,8^2 + 0,1^2 + 0,1^2 =$ | 0,66 | | $(0,66/4)*100$ 16,5 % | item2 | 0,2 | 0,7 | 0,0 | $0,2^2 + 0,7^2 + 0,0^2 =$ | 0,53 | | $+(0,53/4)*100$ 29,75 % | item3 | 0,6 | 0,2 | 0,1 | $0,6^2 + 0,2^2 + 0,1^2 =$ | 0,41 | | $+(0,41/4)*100$ 40 % | item4 | 0,4 | 0,3 | 0,1 | $0,4^2 + 0,3^2 + 0,1^2 =$ | 0,26 | | $+(0,26/4)*100$ 46,5 % | Eigenwert des Faktors | 1,2 | 0,63 | 0,03 | → → → → → → → → → → bei unkorrelierten Faktoren (nach orthogonaler Rotation) entspricht die Zeilen-Summe der Eigenwerte ebenfalls der aufgeklärten Varianz | | 1,86 | 46,5 % → bei Extraktion aller Faktoren in der PCA = 100 % | Bei meinen erfundenen Beispiel-Daten wär die aufgeklärte Gesamtvarianz jetzt nicht grad so groß... | | | | | |  | gesamt aufgeklärte Varianz |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | Kommunalität = gemeinsame mit anderen Items geteilte Varianz des Items | | | aufgeklärter Varianzanteil | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Faktor 1 | Faktor 2 | Faktor 3 | → Zeilen-Summen der quadrierten Faktorladungen | | | kumulierte % der Kommunalitäten | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| item1 | 0,8 | 0,1 | 0,1 | $0,8^2 + 0,1^2 + 0,1^2 =$ | 0,66 | | $(0,66/4)*100$ 16,5 % | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| item2 | 0,2 | 0,7 | 0,0 | $0,2^2 + 0,7^2 + 0,0^2 =$ | 0,53 | | $+(0,53/4)*100$ 29,75 % | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| item3 | 0,6 | 0,2 | 0,1 | $0,6^2 + 0,2^2 + 0,1^2 =$ | 0,41 | | $+(0,41/4)*100$ 40 % | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| item4 | 0,4 | 0,3 | 0,1 | $0,4^2 + 0,3^2 + 0,1^2 =$ | 0,26 | | $+(0,26/4)*100$ 46,5 % | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Eigenwert des Faktors | 1,2 | 0,63 | 0,03 | → → → → → → → → → → bei unkorrelierten Faktoren (nach orthogonaler Rotation) entspricht die Zeilen-Summe der Eigenwerte ebenfalls der aufgeklärten Varianz | | 1,86 | 46,5 % → bei Extraktion aller Faktoren in der PCA = 100 % | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Bei meinen erfundenen Beispiel-Daten wär die aufgeklärte Gesamtvarianz jetzt nicht grad so groß... | | | | | |  | gesamt aufgeklärte Varianz | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Kaiser-Guttman-Kriterium

| | |
|----|----------|
| 1 | 6,977 |
| 2 | 4,617 |
| 3 | 3,365 |
| 4 | 2,907 |
| 5 | 2,454 |
| 6 | 1,874 |
| 7 | 1,332 |
| 8 | 1,123 |
| 9 | 1,018 |
| 10 | ,981 < 1 |

Eigenwerte > 1

→ Faktor (= zusammengefasste Items → „neue Variable“) erklärt mehr Varianz ein einzelnes Item („alte Variable“)



- differenzierte Aufteilung



- evtl. schwer zu interpretieren
- überschätzt in der Regel tatsächliche Anzahl an Faktoren → methodische Artefakte statt inhaltlicher Konstrukte

Parallelanalyse nach Horn

Vergleich der empirischen Eigenwerte mit simulierten Eigenwerten

sind empirische Eigenwerte größer als (95 % der) zufällige Eigenwerte mit dem gleichen Rangplatz ?

gleiche Stichprobengröße und Variablenzahl !

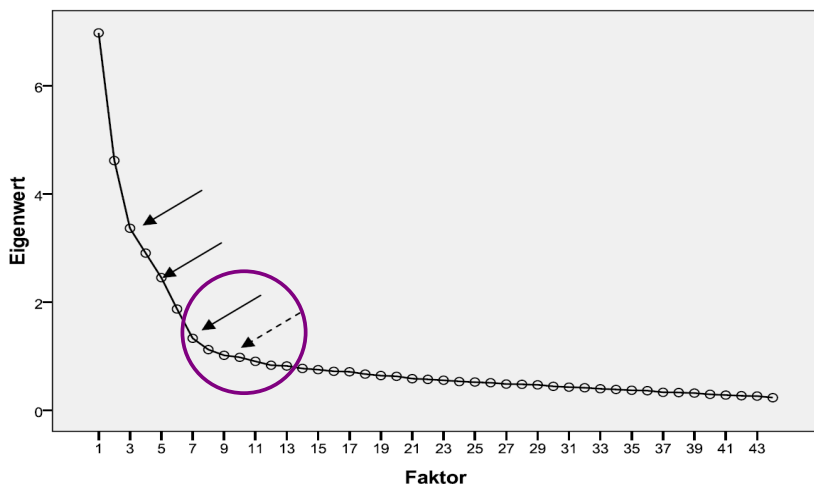
| | empirischer EW | mittlerer zufälliger EW | 95. Percentil |
|----|----------------|-------------------------|---------------|
| 1 | 6,977 | 1,474 | 1,527 |
| 2 | 4,617 | 1,425 | 1,456 |
| 3 | 3,365 | 1,392 | 1,420 |
| 4 | 2,907 | 1,359 | 1,386 |
| 5 | 2,454 | 1,331 | 1,354 |
| 6 | 1,874 | 1,305 | 1,325 |
| 7 | 1,332 | 1,279 | 1,304 |
| 8 | 1,123 | 1,255 | 1,277 |
| 9 | 1,018 | 1,232 | 1,249 |
| 10 | ,981 | 1,211 | 1,232 |

8 ist kleiner als 1,255 und erst recht kleiner als 1,277

Scree-Test nach Cartell

- grafische Analyse des Eigenwerteverlaufs im Scree-Plot
- zu extrahierende Faktoren: vor dem Knick, wenn Eigenwert > 0
- Problem: Wann genau ist ein Knick „auffällig“ ?

Screeplot



- bei sehr großen Stichproben: Überschätzung der Faktorenzahl.
- bei starker erster Hauptkomponente: Unterschätzung der Faktorenzahl

→ grafisch: Eintragen des zufälligen Eigenwerteverlaufs in Screeplot des empirischen → Faktoren links vom Schnittpunkt der beiden Kurven

Minimum-Average-Partial-Test (MAP-Test)

| | squared | power4 |
|--------|---------|--------|
| ,0000 | ,0705 | ,0095 |
| 1,0000 | ,0177 | ,0009 |
| 2,0000 | ,0122 | ,0004 |
| 3,0000 | ,0131 | ,0005 |
| 4,0000 | ,0156 | ,0008 |
| 5,0000 | ,0185 | ,0012 |
| 6,0000 | ,0216 | ,0019 |
| 7,0000 | ,0246 | ,0026 |

Das hat irgendwas mit quadrierten „Partialkorrelationen“ [= Korrelation zweier Variablen, die vom Effekt anderer (spezifizierter) Variablen bereinigt wurden] zu tun...



jedenfalls werden nur die Faktoren mit den „fallenden Werten“ extrahiert – sobald die Werte steigen, fällt der entsprechende Faktor raus.

35. Testkonstruktion explorative Faktorenanalyse Faktoren-Rotation

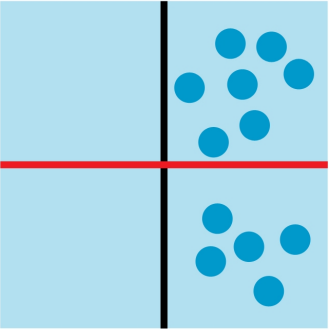
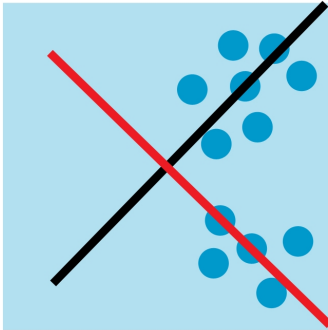
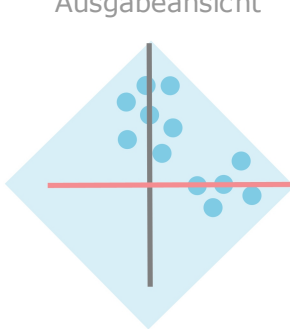
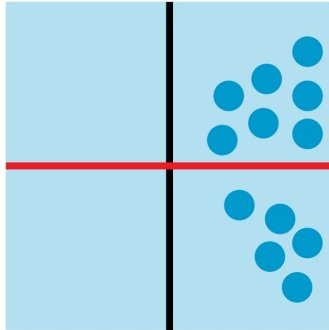
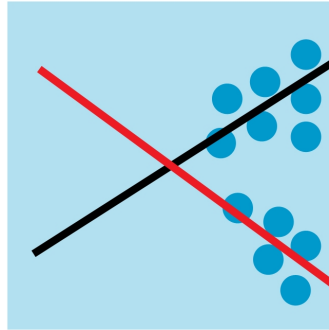
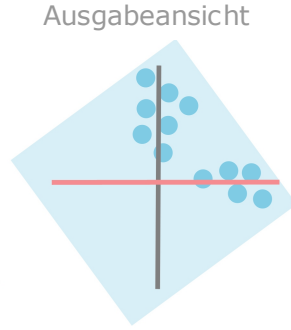
Problem:

Ladungsmuster, die aus der Faktorenextraktion resultieren, sind oftmals inhaltlich schwer zu erklären.

→ Transformation / **Rotation der Faktoren** mit dem Ziel: **Einfachstruktur**

- auf jeden Faktor laden einige Variablen hoch und die übrigen Variablen möglichst niedrig
- jede Variable lädt nur auf einem Faktor hoch und auf den übrigen niedrig.

Häh ? Rotation ? Was ist das und wie sieht das aus, wenn ich mehr als 2 Faktoren extrahiert habe ?  → <http://www.youtube.com/watch?v=I2eDpl6bba8>

| orthogonale Rotation: unkorrelierte Faktoren | | | oblique Rotation: korrelierte Faktoren | | |
|--|---|--|---|--|---|
|  |  |  |  |  |  |
| meistens: Varimax-Rotation → maximaler Unterschied: hohe Ladungen höher, geringere geringer | alternativ: Quartimax Rotation → geringe noch geringer | Equamax-Rotation → Kompromiss aus den anderen beiden | Promax-Rotation → hohe Ladungen verkleinern sich, geringe Ladungen verschwinden fast | Direkte Oblimin-Rotation → Grad der Korreliertheit mit Kennwerten veränderbar | |

→ Ausgabe nach Rotation: 2 **Ladungsmatrizen**

| Mustermatrix Faktormusterkoeffizienten = partielle standardisierte Regressionsgewichte des Items mit den rotierten Faktoren | Strukturmatrix Strukturkoeffizienten Korrelationen zwischen Items und den rotierten Faktoren |
|---|---|
| <ul style="list-style-type: none"> • Im Extremfall: Werte <-1 und >1 → Multikollinearität → Modellinterpretation nicht mehr eindeutig | <ul style="list-style-type: none"> • Werte zwischen -1 und 1 • bei obliquen Rotation (korrelierte Faktoren) → höhere Korrelationen als die Mustermatrix • bei orthogonaler Rotation → mit Mustermatrix identisch |

Einschub: Interkorrelationen → woher kommen bei der obliquen Rotation die Unterschiede in den Matrizen ?

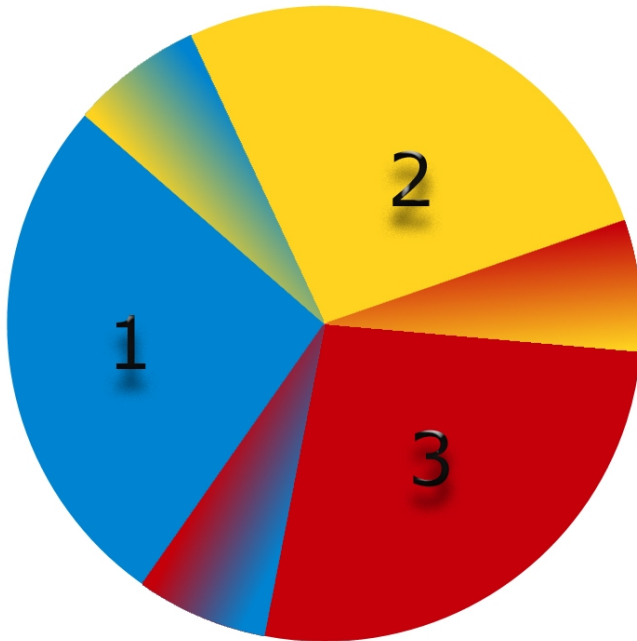
Regressionskoeffizienten







Welcher Varianzanteil wird **nur** durch diesen einen Faktor erklärt ?



Korrelationskoeffizienten

Welcher Varianzanteil wird **auch** durch diesen einen Faktor erklärt ?



-  Korrelation mit Faktor 1
-  gemeinsame Korrelation mit Faktor 1 + 2
-  Korrelation mit Faktor 2
-  gemeinsame Korrelation mit Faktor 2 + 3
-  Korrelation mit Faktor 3
-  gemeinsame Korrelation mit Faktor 3 + 1

36. Testkonstruktion explorative Faktorenanalyse Interpretation + Faktorwert

| | |
|--|--|
| <p>Interpretation „Faktorentaufe“</p> | <p>= aus den Variablenbündeln eine Bezeichnung für den Faktor finden → subjektive Interpretation!</p> <ul style="list-style-type: none"> • Nebenladungen beachten • Muster- und Strukturmatrix vergleichen • evtl. orthogonal + oblique rotieren und vergleichen • ... |
| <p>Faktorwerte</p> | <p>= gewichtete Summation von Einzelitems → individuelle Ausprägung einer Person auf den Faktoren</p> <ul style="list-style-type: none"> • stichprobenabhängig → zwischen Studien kaum vergleichbar • evtl. Grundlage für Faktorenanalyse höherer Ordnung (lassen sich 2 Faktoren vielleicht auf einen „höheren“ Faktor zusammenfassen ?) |