

**Klassische Skalierung**

Name:

**Forschungsauftrag 1**

Skaliere die drei Itemgruppen *Zuhause*, *Familie* und *Einschätzungen* auf Basis klassischer Testtheorie! Kann man aus allen Itemgruppen Skalen bilden?

Die drei Itemgruppen *Zuhause*, *Familie* und *Einschätzungen* sind von den Entwicklern des Fragebogens so zusammengestellt worden. Doch bilden sie auch Skalen? Hier wird die Skalierung zunächst einmal detailliert am Beispiel der Itemgruppe *Familie* beschrieben, danach könnt ihr die anderen beiden Itemgruppen selbst skalieren. Dabei werden alle Schritte am Beispiel des Freeware-Programms PSPP beschrieben, sie sind in SPSS jedoch nahezu identisch.

**Vorgehen zur Skalierung der Itemgruppe *Familie***

1. Öffnet die Datei „HarmoS\_Fragebogen\_2008.sav“.
2. Wählt unter *Analysieren Reliabilitätsanalyse* aus. Dort müsst ihr jetzt die Item Familie\_1 bis Familie\_7 markieren und zur Analyse hinzufügen. Das gewählte Modell soll Alpha sein (für Cronbachs Alpha), die Markierung bei *Deskriptive Statistiken für Skalen anzeigen, wenn das Item gelöscht wurde* muss gesetzt sein. Ein kleiner Tipp: PSPP zeigt etwas unkomfortabel nicht die Itemkürzel, sondern das volle Item an. Familie\_1 fängt mit „Hast du...“ an, Familie\_7 lautet „Ich lese Tageszeitungen“.
3. Nun müsst ihr im *Output viewer* die Ausgabe betrachten. Es zeigt sich, dass die Items ein Cronbachs alpha von  $\alpha=0,62$  erreichen – das ist unter der per Konvention angenommenen Grenze von  $\alpha \geq 0,7$ . Aber das Item „Ich sehe im Fernsehen Nachrichtensendungen“ weist eine vergleichsweise geringe Korrelation mit der Skala auf (vorletzte Spalte) und würde Cronbachs alpha auf  $\alpha=0,63$  erhöhen, wenn es gestrichen würde. Das heißt, dass jemand, der dem Item zugestimmt hat, vergleichsweise seltener bei den anderen Items zugestimmt hat.
4. Streicht das Item (Familie\_6) und berechnet Cronbachs alpha erneut (Schritte 2 und 3). Jetzt würde Item Familie\_7 den Wert erhöhen. Berechnet Cronbachs alpha erneut. Es ergibt sich schließlich ein Cronbachs alpha von  $\alpha=0,64$  – das ist nicht zu viel, aber durchaus verwertbar.
5. Interessant ist es jetzt, die gestrichenen Items mit den verbliebenen zu vergleichen. Was unterscheidet diese von den verbliebenen?

Nun könnt ihr die anderen beiden Itemgruppen selbst skalieren! Dazu noch ein Tipp:

⇒ Das Item *Einschätzung\_5* ist umgekehrt gepolt – eine volle Zustimmung bei diesem Item bedeutet das Gegenteil einer vollen Zustimmung bei Item *Einschätzung\_7*. Es muss umgepolt werden! Dies geht im Menü *Transformieren* und *Umkodieren in andere Variable*. Man wählt dann *Einschätzung\_5* aus, fügt diese hinzu und benennt die Ausgabevariable nach Wahl – z.B. *Einschätzung\_5\_trans*. Jetzt kann man unter *old and new values* dafür sorgen, dass aus 1 eine 4 wird, aus 2 eine 3, aus 3 eine 2 und aus 4 eine 1.

**Explorative Faktorenanalyse**

Name:

**Forschungsauftrag 2**

Untersuche, inwieweit die Items der Itemgruppen *Einschätzungen* und *Themen* auf mehrere unterscheidbare Faktoren laden! Sollte dies der Fall sein: Wie könnten diese Faktoren interpretiert werden? Ließen sich die Faktoren mithilfe der Items gut skalieren?

Das Freeware-Programm PSPP unterstützt nicht alle Anwendungen der Faktorenanalysen, die das lizenzierte Programm SPSS anbietet. Dazu gehören u.a.

- der *MSA-Eignungstest für einzelne Items* anhand der Anti-Image-Korrelationen,
- die *Maximum-Likelihood-Extraktionsmethode*,
- im Allgemeinen *oblique Rotationen*,
- das *Ausblenden kleiner Faktorladungen* bei der Ausgabe von Faktormatrizen.

Einige Befehle können in PSPP nicht über Dialogfenster angewählt werden und müssen daher programmiert werden. Darunter fällt u.a. der *KMO-Eignungstest* für Itemgruppen. Um das Programm PSPP kennenzulernen, soll die Faktoranalyse an der Itemgruppe *Einschätzungen* angeleitet durchgeführt werden.

**Vorgehen zur Faktoranalyse der Itemgruppe *Einschätzungen***

1. Öffnet die Datei „HarmoS\_Fragebogen\_2008.sav“.
2. **Überprüfung der Normalverteilung:** Öffnet unter *Analysieren* das Auswahlfenster *Nicht-parametrische Tests* und wählt *K-S bei einer Stichprobe* aus. Markiert die Items *Einsch\_1* bis *Einsch\_7* und fügt sie per Anklicken zu den Testvariablen hinzu. Der Output im *Output viewer* zeigt in der letzten Zeile für jedes Item die Signifikanz (Irrtumswahrscheinlichkeit) für die Annahme an, dass die Daten nicht normalverteilt sind. Für alle Items liegt mit mehr als 99.99% Wahrscheinlichkeit keine Normalverteilung vor.
3. **Faktoranalyse:** Öffnet unter *Analysieren* das Fenster *Faktoranalyse*. Fügt die Itemgruppe *Einschätzungen* den Variablen zu. Stellt im Auswahlfenster *Extraktion ...* folgende Wahl ein:
  - Methode: Hauptachsenanalyse
  - Analysieren: Korrelationsmatrix
  - Extrahieren: Eigenwerte über dem 1,00fachen des mittleren Eigenwerts

Stellt im zweiten Auswahlfenster *Rotation...* folgende Wahl ein:

- Methode: Varimax

Die gewählten Einstellungen werden mit dem Anklicken des Felds *Einfügen* als Befehle in den *Syntax editor* eingefügt. Damit alle Berechnungen angezeigt werden, muss die Zeile 9 durch den Befehl „\PRINT ALL“ ersetzt werden. Mit dem Befehl *Alle* im Menü *Ausführen* wird schließlich die Berechnung der Faktoranalyse gestartet, wobei das Ergebnis im Fenster *Output viewer* erscheint.

**Explorative Faktorenanalyse****Name:**

Nun könnt ihr die beiden Itemgruppen *Einschätzungen* und *Themen* selbst analysieren! Dazu folgender Tipp:

- ⇒ Macht anhand der Korrelationsmatrix eine Beurteilung von der Eignung der Itemgruppe für eine Faktorenanalyse. Schätzt anhand einer Inhaltsanalyse und der Korrelationsmatrix die Anzahl interpretierbarer Faktoren ab. Gibt es Items, die man weglassen sollte?
- ⇒ Überprüft anhand des KMO-Tests die Eignung der Itemgruppe.
- ⇒ Überprüft die Normalverteilung der Daten und anhand des Barlett-Tests die signifikante Abweichung der Korrelationsmatrix von einer Nullmatrix.
- ⇒ Überprüft anhand des Scree-Plots mit dem Kaiser-Kriterium (Eigenwerte  $\geq 1$ ) und mit dem Scree-Kriterium (Items mit höherem Eigenwert als das elbow-Item) die von euch geschätzte Anzahl Faktoren. Ergeben die beiden Kriterien dieselbe Anzahl?
- ⇒ Trefft eine begründete Wahl der Extraktionsmethode: Hauptkomponenten- oder Hauptachsenanalyse.
- ⇒ Obwohl PSPP diese Option nicht bietet: Trefft eine begründete Wahl der Rotationsmethode: orthogonale Rotation oder oblique Rotation.
- ⇒ Führt die Faktorenanalyse durch. Entscheidet anhand der rotierten Faktorenmatrix, ob eine Faktorenstruktur vorliegt und wie sie interpretiert werden kann. Gibt es Items, die man aus der Gruppe entfernen könnte?

**Raschskalierung**

Name:

**Forschungsauftrag 3**

Skaliere die drei Itemgruppen *Zuhause*, *Familie* und *Einschätzungen* mithilfe des Rasch-Modells! Welche Items genügen dem Modell, welche nicht?

Auf dem Fragebogen sind die drei Itemgruppen *Zuhause*, *Familie* und *Einschätzungen* zu finden. Bislang sind sie nicht mehr als drei Gruppen von Items, die von den Entwicklern als inhaltlich ähnlich angenommen wurden. Nun soll überprüft werden, ob die Items jeder Gruppe tatsächlich dasselbe latente Konstrukt messen – also eine Skala bilden. Dazu soll das Rasch-Modell verwendet werden. Am Beispiel der Itemgruppe *Zuhause* soll hier das Vorgehen zuerst nachvollzogen, dann an den beiden anderen Itemgruppen selbst durchgeführt werden.

**Vorgehen zur Skalierung der Itemgruppe *Zuhause***

1. Bereitet die Datei so auf, dass sie von ConceptMap importiert werden kann (Endung .xls muss im Dateinamen stehen!). Öffnet ConceptMap (Benutzername: admin, Passwort: bear). Importiert die Daten (*File – Import Wizard*). Markiert dabei „Use first row of the file as a data header“ (das legt die erste Zeile der Excel-Datei als Itembeschriftungen fest) und markiert ID\_2 als *Respondent ID*.
2. In der Datei sind Rohdaten vorhanden. Unter dem Reiter *Answer key* kann man nachvollziehen, welche Daten auftauchen. Die Itemgruppe *Zuhause* besteht aus dichotomen Daten, also nur 1 oder 0 als möglicher Antwort. Es dürfen im *Answer key* deshalb auch nur 1 und 0 auftauchen. Sollten andere Antworten (Spalte *Response*) gegeben worden sein, müssen diese als „N/A“ markiert sein, als nicht gültige Antwort (dies geschieht durch Löschen der Antwortmöglichkeit). Sie werden dann als *Missing data* behandelt. Eingabefehler können immer passieren! Hier kann man sie als *Missing data* kennzeichnen. Das verkleinert die Stichprobe, verfälscht aber nicht die Daten. In der Praxis sollte man in den Rohdaten natürlich solche Fehler nachverfolgen.
3. Geht in den Reiter *Items* und wählt ausschließlich die Items der Gruppe „Zuhause“ als *active* aus. Der *Max score* muss bei allen Zuhause-Items auf 1 eingestellt sein, damit alles andere als *Missing data* behandelt wird. Nur die als *active* ausgewählten Items werden zur Bestimmung der Parameter des Rasch-Modells berücksichtigt.
4. Ruft im Menü *Estimation Tasks Calibration Options* auf und stellt als *Measurement Model Dichotomous Model* (für dichotome Daten! Bei den anderen Itemgruppen muss es *Partial Credit* sein) ein. *Integration Method* soll *Monte Carlo* sein. Dann: *Compute Item Parameter*.
5. Lasst euch über *Reports – Item Reports – Default* die Fitstatistiken für die Items anzeigen! Ist ein Item problematisch? Dazu betrachtet das Diagramm zum *Infit Mean Square*. Hier wird die Passung der Items zu den Annahmen des Raschmodells dargestellt. Ein Infit von 1 gilt als optimale Passung. Per Konvention sollte das Item zwischen  $0,8 < x < 1,2$  liegen. Als Ergänzung kann man sich den T-Wert in der Tabelle darüber ansehen, er sollte zwischen  $-2,0 < T < 2,0$  liegen, sonst weicht das Item zu stark vom Modell ab.
6. Nicht passende Items streicht man idealerweise schrittweise. Das Item *Zuhause\_7* weicht am stärksten von dem angestrebten Intervall ab. Streicht das Item, indem ihr es im Reiter *Items* nicht mehr als *active* markiert und löscht per Hand die Daten für die *Difficulty* und den *Standard Error*. Jetzt müsst ihr Schritt 4 und 5 wieder ausführen.

**Raschskalierung**

Name:

7. Das Item Zuhause\_8 reißt offenbar nach oben aus, es kann davon ausgegangen werden, dass es den Mittelwert stark beeinträchtigt. Es sollte als nächstes gestrichen werden. Nach Neuberechnung der Itemparameter fällt noch Item Zuhause\_1 aus der Berechnung, danach stimmen Infit und T-Werte. Die Skala genügt dem Rasch-Modell in ausreichendem Maße.
8. Berechnet die Personenparameter. Stellt dazu unter *Estimation Tasks – Proficiency Estimation Options* als *Estimation Type EAP* ein. Klickt auf *Generate Full Ability Report*. Wie groß ist Cronbachs Alpha und die EAP/PV-Reliability? Kann der Test verwendet werden? Wenn es kein befriedigendes Ergebnis gibt, ist das ein Hinweis darauf, dass die Struktur mehrdimensional sein könnte – also mehrere latente Konstrukte die Grundlage zur Bearbeitung der items gebildet haben.

**Skalierung der Itemgruppen Familie und Einschätzungen**

Nun ist es an euch, die anderen beiden Itemgruppen zu skalieren! Hier noch ein paar Hinweise:

- Bei *Familie* achtet drauf, dass im *Answer Key* die Antwortmöglichkeit 9 und 0 als *N/A* eingetragen sind und die Werte 1,2,3,4 als 0,1,2,3 bezeichnet wurden. Das liegt daran, dass der Algorithmus nur von 0 anfangende Stufen akzeptiert. Im Reiter *Item* muss der Max Score als 4 angegeben sein. Die Infit- und T-Werte könnt ihr in der Ausgabe des Programms ganz unten unter „Item statistics (EAP)“ sehen.
- Bei *Einschätzungen* müsst ihr analog zu *Familie* verfahren. Es gibt einen wichtigen Unterschied. Das Item *Einschätzung\_5* ist umgekehrt gepolt – wie muss der *Answer key* aussehen?
- Man kann immer testweise kontrollieren, ob die aussortierten Items zusammen eine Skala ergeben. Auch das könnte ein Hinweis auf Mehrdimensionalität sein. Interessant ist das vor allem bei der Gruppe der Einschätzungitems.

---

**Forschungsauftrag 4**

1. Vergleicht die Items, die ihr aussortiert habt, mit den Ergebnissen der klassischen Skalierung.
  2. Für die fachdidaktische Analyse der Ergebnisse ist es wichtig, zu interpretieren, was die latenten Konstrukte sind, die gemessen wurden. Dazu muss man die Items betrachten, die eine Skala bilden und Theorien bilden, was diese gemeinsam haben. Was muss man können oder wissen, um diese Items beantworten zu können? Was unterscheidet diese Items von den aussortierten? Analysiert die Skalen aus fachdidaktischer Sicht!
-