

## Impulse zum Gestalten, Durchführen und Auswerten von Fragebögen mit SPSS

Stand: 25.06.2024

Diese Liste gibt einen Überblick über einige zentrale Schritte der Fragebogengestaltung und -auswertung, orientiert an Prä-Post-Testungen zu zwei Messzeitpunkten. Wenn Ihr Fragebogen nur einen Messzeitpunkt umfasst, ignorieren Sie die Anmerkungen zum Umgang mit Post-Daten.

**Diese Übersicht umfasst Impulse für ein mögliches, methodisches Vorgehen und formuliert keinen Anspruch auf Vollständigkeit, ebenso wenig wie auf Aktualität. Weiterführende Hinweise zum empirischen Arbeiten finden Sie insbesondere in:** Döring, N. & Bortz, J. (2016). Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften (5. Aufl.). Heidelberg: Springer. <https://doi.org/10.1007/978-3-642-41089-5>

Sie benötigen einen zentralen, strukturellen Überblick darüber, wie man einen Fragebogen schrittweise konzipiert und an welchen Stellen man relevante Entscheidungen zur methodischen Ausgestaltung trifft? Dann schauen Sie gerne in meine [grafische Übersicht](#).

### Grundeinstellungen und Vorgehensweisen in SPSS:

- Wenn **keine Variablenamen** angezeigt werden: Bearbeiten > Optionen > Allgemein > Variablenlisten > Namen anzeigen (oder Rechtsklick auf den Namensbereich)
- **Syntax öffnen und auf Einfügen** (statt ok und direkter Ausführung) klicken, um Syntax nach einem Befehl zu erhalten – so Syntax fortlaufend anlegen, dann kann man rekonstruieren, was passiert ist, und die Syntax auch für spätere Auswertungen (bspw. nur nach einer Teilnehmer-Subgruppe o.ä.) nutzen
- Tipp: in der Syntax auch die Hypothesen etc. notieren, das gibt Orientierung innerhalb von Syntax und Ausgabe / Auswertung

### Zentrale Schritte der quantitativen Befragung mittels Fragebogen

#### 1. Fragebogen konzipieren

- **Items literaturbasiert zusammenstellen**, ggf. vorhandene Fragebögen oder Skalen hinzunehmen inkl. derer Originaldaten zum Vergleich, [Skalenniveaus](#) definieren [Notiz: einzelne Likert Items sind ordinalskaliert, Likert Skalen sind intervallskaliert]
- **Fragebögen gestalten**
  - in Orientierung an **Muster-Aufbau**, Struktur durchdenken, zudem schriftlich begründen – Instruktion / Erhebung soziodemografischer Daten / Items sortieren in Blöcken und Fragebogenseiten / Dank zum Schluss usw. (bspw. bei Reinders, 2022)
  - logische, inhaltliche **Abfolge** der Items und Frageblöcke durchdenken / begründen;
  - **Standardisierung** von einzelnen Items: Wahl und Gestaltung geschlossener (bspw. Menold & Bogner, 2015) bzw. offener Items durchdenken / begründen (bspw. Züll, 2015)

- Nutzung von **Rating-Skalen** inkl. Items durchdenken / begründen (bspw. Menold & Bogner, 2015), Skalen-Niveaus wählen und begründen (bspw. Döring & Bortz, 2016; Schell, Hiller & Esser, 2011), Umgang mit Tendenz zur Mitte begründen usw.)

Menold, Natalja & Bogner, Kathrin (2015). Gestaltung von Ratingskalen in Fragebögen. Mannheim, GESIS – Leibniz-Institut für Sozialwissenschaften (SDM Survey Guidelines). DOI: 10.15465/sdm-sg\_015

Züll, Cornelia (2015). Offene Fragen. Mannheim, GESIS – Leibniz-Institut für Sozialwissenschaften (SDM Survey Guidelines). DOI: 10.15465/sdm-sg\_002

Schell, Rainer; Hill, Paul B.; Esser, Elke (2011): Methoden der empirischen Sozialforschung. 9. Auflage. München: Oldenbourg Wissenschaftsverlag.

Reinders, H. (2022). Fragebogen. In: Reinders, H., Bergs-Winkels, D., Prochnow, A., Post, I. (eds) *Empirische Bildungsforschung*. Springer VS, Wiesbaden. [https://doi.org/10.1007/978-3-658-27277-7\\_10](https://doi.org/10.1007/978-3-658-27277-7_10)

- bei Prä-Post-Test am Anfang einen **individuellen Code** durch die Teilnehmer\*innen generieren lassen, um am Ende Prä- und Post-Antworten zusammenführen zu können (sonst keine Vergleichbarkeit zwischen den Messzeitpunkten (MZP) gegeben!)
- Grundsätzliche, inhaltliche und formale Prüfung des Fragebogaufbaus (nach Kallus, 2010):
  - „Ist das **Sprachniveau** angemessen? – Überprüfen Sie, ob Ihre Formulierungen für Ihre Zielgruppe verständlich sind. Fachbegriffe sollten Sie vermeiden, z. B. *Ich bin intrinsisch motiviert – bei einer Umfrage unter Schüler\*innen*.
  - Sind die **Bezüge klar und eindeutig**? – Ein problematisches Beispiel wäre: *In meiner Schule gehöre ich zu den Besten im Lesen*. Hier ist nicht eindeutig, auf wen sich der Vergleich bezieht. Es könnten sowohl ältere Schüler\*innen als auch die Lehrpersonen in den Vergleich einbezogen werden. Eine klarere Formulierung wäre: *Im Vergleich mit den Mitschüler\*innen in meiner Klasse lese ich gut*.
  - Entspricht **jedes Item einem Aspekt**? – Items, die mehrere Aspekte abfragen, können nicht eindeutig beantwortet werden, z. B.: *Ich lese gerne Zeitung und Romane*. Solche Items sollten Sie trennen und dafür zwei Items formulieren. Auch Bedingungen innerhalb eines Items oder zwischen Items sollten Sie vermeiden, z. B. *Wenn es Winter ist und ich ein gutes Buch zur Hand habe, lese ich manchmal den ganzen Tag*.
  - Sind die **Antwortkategorien passend**? – Bspw. passt die Antwortskala von 1 (*schlecht*) bis 5 (*sehr gut*) nicht zu dem Item *Meistens macht mir lesen Spaß*.“

zitiert aus: Bohndick, 2015: Fragebögen. [Blog der Uni Paderborn](#). In Anlehnung an: Kallus. (2010). Erstellung von Fragebogen (1. Auflage). Facultas. wuv.

- jetzt oder später: ggf. **Fragebogen Pre-testen**, dabei alle Items auf **Gütekriterien** testen (Objektivität, Reliabilität, Validität), zudem dabei **Skalen** bilden (ausführliche Anleitung s. unten)

Gütekriterien:

- **Objektivität:** Beobachterunabhängigkeit (Rey 2020)
  - **Durchführungsobjektivität:** Ergebnisse unabhängig vom Testleiter
  - **Auswertungsobjektivität:** Ergebnisse unabhängig vom Testauswerter
  - **Interpretationsobjektivität:** Ergebnisinterpretation unabhängig von der Person, die diese vornimmt
  
- **Reliabilität:** Zuverlässigkeit bzw. Genauigkeit einer Messung (Rey 2020)
  - **Interne Konsistenz:** Vergleich der einzelnen Aufgaben bzw. Items eines Tests (Prüfung auf Homogenität)
  - **Paralleltestreliabilität:** Korrelation zwischen den Ergebnissen zweier ähnlicher Testformen, die zeitnah an derselben Stichprobe erhoben wurden
  - **Retestreliaibilität:** Korrelation zwischen zwei Ergebnissen des gleichen Tests, die zu zwei unterschiedlichen Zeitpunkten an derselben Stichprobe erhoben wurden
  - **Testhalbierungsreliabilität:** Korrelation zwischen zwei Hälften des gleichen Tests
  - für Skalen: **Crombachs Alpha**
  
- **Trennschärfe:** bei Items berechnet, Korrelation zwischen einem Item und dem Gesamtwert des Tests (Rey 2020)
  - Trennschärfe gibt an, wie stark die Differenzierung des jeweiligen Items mit der Differenzierung des Gesamtwertes übereinstimmt
  - Itemschwierigkeit: kann in SPSS nur berechnet werden, wenn die geringste Antwort-Ausprägung mit 0 codiert ist; daher ggf. über Daten in Excel berechnen: Die Schwierigkeit wird bei Ratingskalen in der Regel wie der Mittelwert (deskriptive Statistik) berechnet: Man summiert die Werte dieses einen Items über alle Teilnehmenden auf und teilt diese Summe durch die Anzahl der Teilnehmenden.  
$$=(\text{MITTELWERT-ZELLE} - 1)/\text{AnzMaxPunkte} * 100$$
Bei klassischer Testkonstruktion haben Items mit mittlerer Itemschwierigkeit in der Regel die beste Trennschärfe.
  - Gängige Konvention: Bei  $0.4 \leq r_{it} \leq .7$  gilt die Trennschärfe als gut.
  
- **Validität** (Rey 2020):
  - Validität: Gültigkeit der Messung – Misst der Test das, was er messen soll? Nur bei einem validen Test sind die Messergebnisse interpretierbar!
    - **inhaltliche (oder logische) Validität:** Aufgaben des Tests sind inhaltlich identisch mit den Merkmalen, die durch den Test erfasst werden sollen – diese Validitätsform wird argumentativ und nicht empirisch-numerisch begründet
    - **Konstruktvalidität:** Test erfasst alle Facetten des theoretischen Konstrukts, die erfasst werden sollen

- Konvergente Validität: (möglichst hohe) Korrelation zwischen verschiedenen Tests, die dasselbe Konstrukt messen
- Diskriminante (bzw. divergente) Validität: (möglichst niedrige) Korrelation zwischen verschiedenen Tests, die verschiedene Konstrukte messen
- **kriterienbezogene Validität:** Testergebnis stimmt mit anderen, praktisch relevanten Kriterien (Außenkriterien) überein, die das Merkmal ebenfalls erfassen
  - Konkurrente Validität: Übereinstimmungsvalidität
  - Prognostische (prädiktive) Validität: Vorhersagevalidität

Rey, G. D. (2020). Methoden der Entwicklungspsychologie. Datenerhebung und Datenauswertung. Norderstedt BoD.; zudem: Rey, G. D. (o.J.). [Testgütekriterien](#).

- **Nebengütekriterien** (in Auszügen, Rey 2020):
  - Normierung: geeignete, aktuelle Referenzstichprobe
  - Testfairness: keine systematische Diskriminierung
  - Ökonomie: Dauer und Kosten der Erhebung gering
  - Nützlichkeit: praktische Relevanz des Merkmals; Beantwortung der Fragestellung möglich
  - Zumutbarkeit: Nutzen überwiegt zeitliche, psychische und körperliche Belastung der Testpersonen
  - Unverfälschbarkeit: bspw. durch soziale Erwünschtheit gefährdet
  - Transparenz: Verständlichkeit der Instruktion; Übungsitens im Vorfeld; angemessenes Feedback im Anschluss des Tests
  - Akzeptanz: Erhebung bzw. Test von Laien akzeptiert
  - äußere Gestaltung: sprachlich und optisch ansprechend; Anpassung an die Zielgruppe

Rey, G. D. (2020). Methoden der Entwicklungspsychologie. Datenerhebung und Datenauswertung. Norderstedt BoD.; zudem: Rey, G. D. (o.J.). [Testgütekriterien](#).

## 2. Fragebogen digital realisieren, Befragung durchführen

- bspw. per **Limesurvey**, die Uni Münster / EW hat einen eigenen Server
- dabei beachten, dass der Fragebogen ggf. zum Prä- und Post-Zeitpunkt verwendet werden muss (1. Messzeitpunkt durchführen, dann Antwort-Daten zum 1. MZP exportieren, Umfrage deaktivieren und online Datentabelle archivieren, dann ggf. Änderungen in Fragebogen für Post-Befragung einbauen, Umfrage wieder aktivieren, 2. Messzeitpunkt durchführen, dann Antwort-Daten zum 2. MZP exportieren) / oder als zwei getrennte, digitale Umfragen anlegen, sofern sich die Befragungen Prä zu Post inhaltlich unterscheiden
- ggf. **Teilnehmer-Daten** hochladen (bspw. E-Mail-Adressen für individuelle Zusendung)

## 3. Daten bearbeiten und in SPSS zusammenführen:

- **Antwortdateien** aus SPSS herunterladen
  - Die heruntergeladene Datei „syntax file.sps“ wird geöffnet, im oberen Bereich wird der Pfad zur Datei „.dat“. eingefügt ([Tutorial](#)). / Strg+A, dann Klick auf Ausführen

(grüner Pfeil) / In einem neuen Fenster werden die Variablen und die Ergebnisse sichtbar. Diese Datei wird als Ergebnisdatei (Gruppenname, .sav) gespeichert.

- SPSS-Prä-Daten öffnen, allen Variablen-Namen ein \_Prä anfügen
- SPSS-Post-Daten öffnen, allen Variablen-Namen ein \_Post anfügen
- **Textfelder** können Probleme verursachen. Zu Beginn in SPSS für Textfelder-Variablen die Wertebreit auf 2000 setzen.
- **Manche Antworten können mit A001 statt mit 1 codiert sein, ggf. korrigieren ([Tutorial](#)).**
- **SPSS-Daten** in einer SPSS-Datei zusammenführen / ggf. unter Hinzunahme von Excel-Zwischenkopien → **ganz, ganz wichtig:**
  - Die **Prä- und Post-Antworten der einzelnen Personen müssen „hintereinander“** zu einem einzigen Fall pro Person gesetzt werden (nicht in die gleiche Variable, sondern es muss pro Item pro Fall eine Variable \_Prä und eine Variable \_Post geben); beim Zusammenführen auf Passung der **individuellen Codes** achten, damit jeder Fall seine zugehörige Post-Antwort bekommt.
  - Wenn mehreren Gruppen (Experimental-/Kontrollgruppe u.a.) vorhanden sind, dann müssen **alle Befragten in eine gemeinsame Datei zusammengeführt** werden. Hierzu muss vorher eine Variable implementiert werden, die die zusammengeführten Fälle unterscheidbar macht (1 = Experimentalgruppe, 2 = Kontrollgruppe usw.).
  - Teilnehmern, die nur am Prätest teilgenommen haben, können natürlich keine Post-Daten zugeordnet werden, diese bleiben frei (und werden später als Missings definiert). Zusätzlich eine Spalte / Variable in SPSS anlegen, in der Teilnehmer per Wert = 1 markiert werden, wenn sie nur am Prä-Test teilgenommen haben, so kann man sie später einfacher herausfiltern.
  - Teilnehmern, die nur am Posttest teilgenommen haben, können natürlich keine Prä-Daten zugeordnet werden, diese bleiben frei (und werden später als Missings definiert). Zusätzlich eine Spalte / Variable in SPSS anlegen, in der Teilnehmer per Wert = 1 markiert werden, wenn sie nur am Post-Test teilgenommen haben, so kann man sie später einfacher herausfiltern.

#### 4. Daten aufbereiten

- prüfen, ob Antworten **umcodiert** werden müssen (sowohl hinsichtlich der Antworten an sich in ihrem Format, hinsichtlich **inverser Antworten** / negativ gepolte Items als auch hinsichtlich der hinterlegten Werte)

per Transformieren-Befehl, alternativ per SYNTAX, Beispiel vierstufig inkl. Missing (-99):

```
RECODE Item1, Item2 (1=4) (2=3) (3=2) (4=1) (-99=-99).  
EXECUTE.
```

Achtung, Missings oder 0 mit eingeben, alle theoretisch vorhandenen Werte müssen angegeben werden!

- **Missings** definieren – zuerst die Missings über Syntax überall in den Variablen anlegen lassen  
-> minus 99 wird überall ergänzt und als Missing definiert, auch im Datensatz

SYNTAX:

RECODE

Item1

Item2

Item3

(-99=SYSMIS) (SYSMIS=-99).

EXECUTE.

MISSING VALUES

Item1

Item2

Item3

(-99).

EXECUTE.

ACHTUNG: Immer nur Variablen zusammen mit Missings versehen, die dem **gleichen Typ** entstammen (z. B. alle numerischen Variablen zusammen).

ACHTUNG: NIEMALS Missings definieren für kodierte offene Items (sonst funktioniert das Addieren / Zählen von Antwort-Ausprägungen später nicht) sowie NIEMALS für die Variable „hat nur am Prätest teilgenommen“ (s. oben; sonst können Fälle nicht über diese Variable gefiltert werden)

- **Plausibilitätstest:** Berechnen und Sichten von Streuungsmaßen, Berechnen und Sichten von Häufigkeitsverteilungen, Berechnen und Vergleichen von Häufigkeitsverteilungen
- **Fehlerdiagnose und Fehlerkorrektur**
- spätestens jetzt: **alle Items auf Gütekriterien testen** (u.a. Objektivität, Reliabilität und Validität, Details s. oben)
- **offene Items codieren** (ggf. dafür offene Felder in Items oder MAXQDA übernehmen, dort codieren), Codierungen als Code-Variablen in SPSS einfügen
  - ggf. für Mehrfachantworten Variable erzeugen; alternativ, wenn die Antworten nur gezählt werden sollen, kann dies später auch über deskriptive Statistik „von Hand für jede codierte Antwortausprägung“ erfolgen (Mehrfachantwort-Variable funktioniert manchmal nicht solide im Abruf)
- **bzgl. Angaben zum Geschlecht:** Entscheidung zur Codierung treffen, bspw.: Geschlecht diverse – als Missing definiert, damit nur 2 Ausprägungen im Geschlecht, theoriebasiert begründen (bspw. Döring 2013)
 

Döring, Nicola (2013): Zur Operationalisierung von Geschlecht im Fragebogen: Probleme und Lösungsansätze aus Sicht von Mess-, Umfrage-, Gender- und Queer-Theorie. In: GENDER, 2/2013, S. 94-113. Leverkusen: Verlag Barbara Budrich. [Link](#).
- **Skalen bilden** (bzw. vorhandene, inhaltlich angenommene Skalen / Item-Gruppen bestätigen) über **Faktorenanalyse** → **ausführliche Anleitung dazu siehe unten, wichtig:** die Faktorenanalyse getrennt durchführen für Prä-Antworten der Items und für Post-Antworten

der Items – gemessen wird zwar das gleiche, angenommene Konstrukt, aber an zwei Zeitpunkten zu unterschiedlichen Einflüssen; für Skalen ebenfalls Missings definieren (zugehörige Syntax s. oben)

- ggf. [Globalskalen](#) anlegen, Missings definieren (zugehörige Syntax s. oben)
- für alle **Prä- zu Post-Skalen Differenz-Diff-Werte-Variablen** bilden, um die Veränderung messbar zu machen (grundsätzlich für abhängigen t-Test nötig), Missings definieren
- ggf. **Skalen (Prä, Post, Diff) als z-standardisiert anlegen**, insb. dann, wenn später Skalen bspw. korreliert werden sollen, die unterschiedliche Antwortstufungen haben (Item mit 4stufiger Skala mit Item mit 6stufiger Skala bspw., dann für beide z-standardisierte Items anlegen; Missings für z-standardisierte Skalen definieren)

## 5. statistische Methoden festlegen in Abhängigkeit von ggf. gebildeten Hypothesen; Voraussetzungen prüfen, zentrale Entscheidungen treffen

- Voraussetzungen prüfen:
  - bspw. **Verteilung eines Merkmals** (Geschlecht, studierte Schulform u.a.): Wenn eine Teilstichprobe größer 50 Personen ist, sind starke Unterschiede ( $m = 51$ ,  $w = 290$ ) unkritisch. Dann sind auch die t-Tests etc. sehr robust, d. h. die Voraussetzung einer Normalverteilung kann dann ignoriert werden. Bei kleineren Stichproben ist im Einzelfall zu begründen (bspw. Normalverteilung betrachten), wobei t-Tests auch dort als recht robust eingestuft werden. Allgemein: verschobene Verhältnisse berichten bzw. resultierende Einschränkungen begründen.
  - ggf. für **Prä-Post-Vergleich** festlegen: nur Fälle nutzen, die Prä und Post teilgenommen haben wg. Vergleichbarkeit der einzelnen Fälle und Gruppen
  - ggf. **Umgang mit fehlenden Werten** – Listenweiser Fallausschluss, ...
- **statistische Methoden festlegen:**
  - [Entscheidungsbaum](#) mit SPSS-Klick-Anleitung und Berichtssätzen
  - [Entscheidungsassistent](#)

## 6. Berechnungen durchführen – auch hierfür wieder fortlaufend Syntax anlegen und abspeichern!

- Entscheidungsbaum und [Anleitungen](#)

## 7. Ergebnisse berichten

- Mustertexte am Ende der [Anleitungen](#)
- Ergebnisdarstellung gemäß [APA-Konventionen](#)
- Tipp: Wenn man Tabellendaten aus SPSS kopieren möchte (mehrere zusammenhängende Felder) und diese in einer vorstrukturierten Word-Tabelle einfügen möchte, dann muss man in Word erst die passende Anzahl an Zellen mit der Maus markieren und dann einfügen.
- wichtig beim Ablesen der Daten im SPSS-Output:
  - t-Tests:
    - Wenn es bei einem Hypothesentest darum geht, **ob sich** die Wahrscheinlichkeit eines Ereignisses geändert hat, handelt es sich um einen einseitigen Signifikanztest. / insb. ja / nein

- Wenn es bei einem Hypothesentest darum geht, ob die Wahrscheinlichkeit eines Ereignisses **anders ist als der bislang angenommene Wert**, spricht man von einem zweiseitigen Signifikanztest. / insb. Gruppenunterschiede, Prä-Post-Veränderungen u.a.
- Cohens  $d$ -Effekte werden immer positiv berichtet und auch dann, wenn ein Ergebnis nicht signifikant ist
- Wenn  $p$  größer 0.05 ist (Sig.), dann werden die Werte der ersten Zeile (Zeile „Varianzen sind gleich“) berichtet, ansonsten Werte aus zweiter Zeile („Varianzen sind nicht gleich“)

#### einseitiger t-Test, $d$ berichten:

Die **Beurteilung der Stärke des Effekts** wird in Anlehnung an Eid, Gollwitzer & Schmitt (2017, S. 301) mittels  $d$  ( $d = 0.14$ : kleiner Effekt,  $d = 0.35$ : mittlerer Effekt,  $d = 0.57$ : großer Effekt).

Gollwitzer, Eid, M., & Schmitt, M. (2017). Statistik und Forschungsmethoden: Lehrbuch. Beltz Verlagsgruppe.

#### abhängiger oder unabhängiger t-Test, $d$ berichten:

Als **Maß für die Effektstärke** wird Cohen's  $d$  (Cohen, 1988) herangezogen ( $d = 0.2$ : kleiner Effekt,  $d = 0.5$ : mittlerer Effekt,  $d = 0.8$ : starker Effekt).

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed). Hillsdale, N.J.: L. Erlbaum Associates.

#### Korrelation, $r$ berichten:

Muster: Es zeigt sich eine signifikant positive Korrelation zwischen X und Y,  $r(258) = .29$ , 95% KI [.18, .39],  $p < .001$  (einseitig), wobei die Stärke des Effekts als mittel zu beurteilen ist.

Muster: Es zeigt sich eine signifikant xxx Korrelation zwischen X und Y,  $r(\text{Freiheitsgrade, das ist das } n \text{ minus } 2) = \text{Pearson-Korrelation, Level } 95\% \text{ [Lower CI bis Upper CI], Signifikanz } p < \dots$

Nach Cohen (1988) lassen sich **Korrelationskoeffizienten** grob wie folgt klassifizieren:

$|r| \approx 0,10$ : schwacher Zusammenhang,  $|r| \approx 0,30$ : mittlerer Zusammenhang,  $|r| \approx 0,50$  starker Zusammenhang

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed). Hillsdale, N.J.: L. Erlbaum Associates.



## 8. Diagramme erstellen, hier in SPSS:

WICHTIG! Für ein direktes Kopieren werden die Grafiken mit zu geringer Auflösung in SPSS generiert. VOR Bearbeitung der Diagramme im Viewer muss dort die Auflösung der Grafiken erhöht werden, damit diese beim Exportieren in ein Dateiformat oder in Word nicht unscharf werden!

> Doppelklick auf die Grafik, dann Bearbeiten > Eigenschaften. Diagrammbreite: Seitenbreite angeben, bspw. 42cm > anschließend Schriftgrößen entsprechend anpassend (mind. 12er). Die damit überarbeitete Grafik ist bei Übernahme bspw. in Word scharf.

**Alternativ / Notfalls:** Wenn die Diagramme bereits erstellt und überarbeitet sind, ohne dass die Diagrammbreite zuvor erhöht wurde, kann ein Word-Export helfen: Datei > Exportieren > als Word > Seitengröße festlegen A4 quer mit schmalem Rand > Optionen: Inhalte an Seitenbreite anpassen > ok > Damit wird ein Word-Dokument erstellt, was die passende Vermaßung hat und vergrößerte Grafiken enthält. Wichtig: Werden die Grafiken nun in ein anderes Dokument eingefügt, können beim PDF-Exports dieses anderen Dokuments Fehler entstehen / Fragezeichen-Sperrsymbole in die Grafiken integriert werden. Um dies zu vermeiden, empfiehlt es sich, das Word-Dokument mit den Grafiken über PDF24 in Bilddateien (eine Bilddatei pro Diagramm bzw. Grafik) zu zerlegen und diese Bilddateien weiterzuverarbeiten.

### Bearbeitung der Diagramme in SPSS:

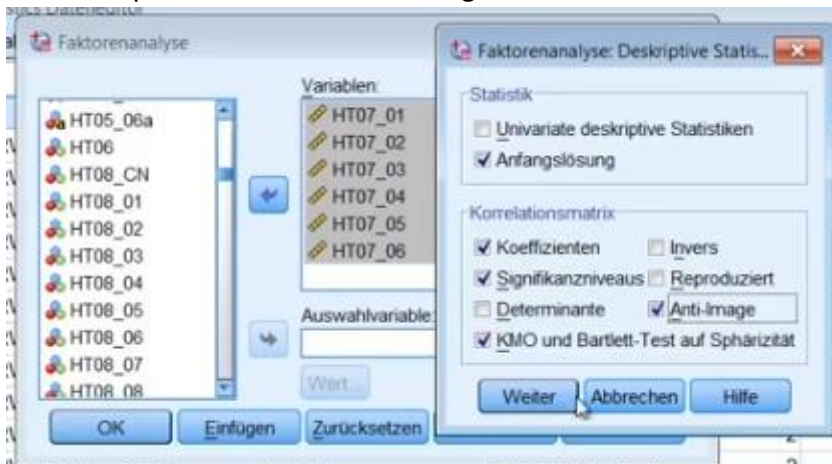
- **technisch nicht möglich** sind
  - das Eingeben von 0,00 mit anführender Null in Skalenbeschriftungen gemäß APA
  - die Anpassung von , auf . in Skalenbeschriftungen gemäß APA
  - die Anpassung der Legenden (M statt Mittelwert usw.)
- Beschriftungstexte der Achsen können durch Einfach-, Doppel- oder Trippel-Klick geändert werden. (Wenn das nicht möglich ist, Gerät wechseln, dann unterdrückt eine Windows-Funktion das Hineinklicken und Markieren von Text!)

## 9. Daten interpretieren in Textform

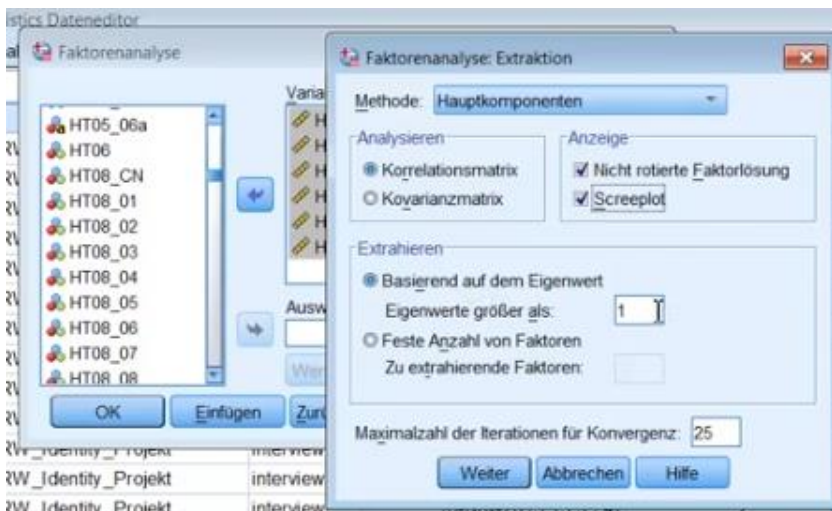
**Skalenbildung via Dimensionsreduktion und Faktorenanalyse – pro (angenommener,) aus mindestens drei Items bestehender Skala:**

- Durchführung der Faktorenanalyse (hier am Beispiel der konfirmatorischen FA, wenn N recht klein); Dimensionalität der angenommenen Skala wird geprüft („Wie viele Faktoren sind tatsächlich in der Skala enthalten, wirklich nur der eine, inhaltlich vermutete, oder misst die Skala noch etwas anderes?“)

- Analyse – Dimensionsreduktion – Faktorenanalyse > Variablen auswählen
- weitere Optionen auswählen wie folgt:

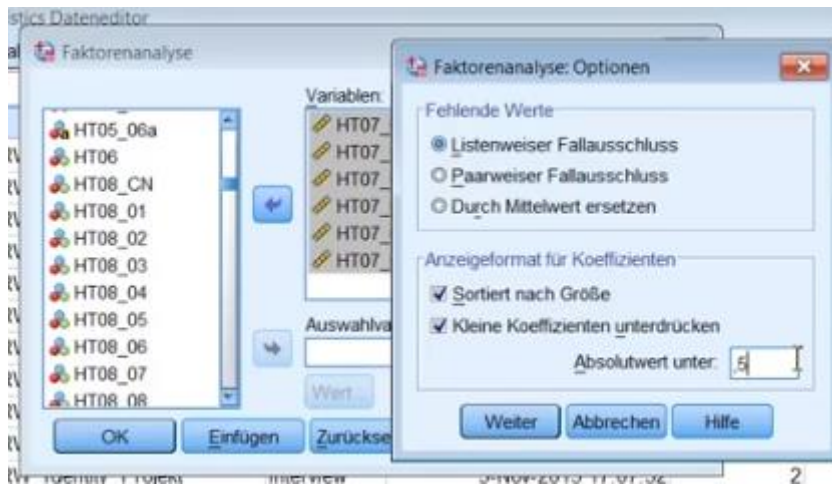


Deskr.: KMO aktivieren



Extr.: Screplot aktivieren

Rotation: Varimax → erzeugt größtmögliche Varianz zwischen den Faktoren



Auswahl s. Grafik inkl. Ziffernwert

➔ zeigt genau einen Faktor pro (vermuteter) Skala

○ Interpretation der SPSS-Ausgabe:

- **Korrelationsmatrix:** unteres Dreieck der Daten sollte über 0,3 sein, dann besteht gute Korrelation unter den Variablen
- **Sig 1seitig:** ,000 = signifikante Korrelationen zwischen Items, siehe ggf. auch Hervorhebungen durch SPSS
- **KMO:** mind. größer 0,5, möglichst nahe 1,0
- **Barlett:** mögl. klein, Nullhypothese ablehnen: möglichst klein (bis zu ,000) = hohe Signifikanz
  
- **Anti-Image-Korrelation (unten):** Diagonale überprüfen, Werte sollten mögl. hoch sein, über 0,7
- **Kommunalitäten:** Extraktion zeigt Streuung pro Frage, sollte über 0,3 oder größer sein, besser über 0,5, dann kann man pro Item die Streuung erklären, ansonsten ein Item weglassen
  
- **Erklärte Gesamtvarianz:** zeigt im rechten Bereich relevante / ermittelte Faktoren, die aus den Items hervorgehen (zeigt sich auch noch einmal im Screeplot, „in Screeplot nur alles nutzen, was über dem größtem Ellbogen-Knick liegt“); rechteste Zahl in Tabelle zeigt % der Varianz, die wir mit den Items erklären können (kumulativ)
  
- **Komponentenmatrix:** zeigt bei mehreren Faktoren, welches Item welchen Faktor bedient bzw. welches Item welchen Faktor hat; rotierte Komponentenmatrix, gibt mehr Informationen hierzu
- **analyisierte Faktoren:** Ziel: am Ende sollte jedes Item genau einem Faktor zuzuordnen sein. **Pro Skala soll es nur eine inhaltliche Dimension / einen Faktor geben, den diese Skala misst.**

- wenn alle Tests ok, den Faktor als Variable abspeichern – Skala anlegen für Mittelwert (oder Summe): per Mittelwertbildung: Transformieren, Variable berechnen > neue Zielvariable benennen [NAME\_MEAN], numerischer Ausdruck: MEAN(Items per Doppelklick rein, mit Komma trennen) > neue Variable wird damit angelegt

Alternativ SYNTAX:

```
COMPUTE Skalename_Mean=MEAN(Item1,Item2,Item3).
```

Execute.

### • Prüfung auf hohe Reliabilität / Cronbachs Alpha

- Analysieren – Metrisch – Reliabilitätsanalyse > Variablen einfügen, Modell: alpha wählen; Statistiken: ITEM, METRISCH, SKALA WENN ITEM GELÖSCHT aktivieren, Mittelwert, Varianzen aktivieren
- Interpretation der Ausgabe: aus Items gebildete Skala wird dann analysiert,
  - auf Cronbachs  $\alpha$

#### Cronbachs $\alpha$ interpretieren und (später) berichten:

>.9 exzellent

>.8 gut

>.7akzeptabel

>.6 fragwürdig

>.5 schlecht

<.5 inakzeptabel (vgl. Cronbach, 1951).

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–333.

Blanz, Mathias. *Forschungsmethoden und Statistik für die Soziale Arbeit: Grundlagen und Anwendungen*. Stuttgart: W. Kohlhammer Verlag, 2015. Print.

- Item-Skala-Statistik: „Korrigierte Item-Skala-Korrelation“ gibt **Trennschärfen pro Item** aus – größer 0, nicht nahe 0, nicht negativ

Die Trennschärfe sollte einen Wert von über .5 aufweisen (Bortz & Döring, 2016, S. 478). Alle Items mit Trennschärfen-Werten unterhalb .3 benötigen daher einer Überarbeitung der Skala, z.B. durch Streichen der betroffenen Items.

Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Aufl.). Heidelberg: Springer. <https://doi.org/10.1007/978-3-642-41089-5>

- Item-Skala-Statistik: „Cronbachs Alpha, wenn Item weggelassen“ gibt an, ob sich Cronbachs  $\alpha$  verbessern würde, wenn man das jeweilige Item entfernt / nicht der Skala hinzufügt