II. Test auf Trennschärfe

Auch bei der sogenannten Itemtrennschärfe geht es darum, wie gut ein einzelnes Item in der Lage ist, zwischen den Befragungspersonen zu unterscheiden, also zwischen Personen mit einer positiven und solchen mit einer negativen Einstellung zu einem gegebenen Objekt zu diskriminieren.[664] Die Trennschärfe gibt an, wie stark die Differenzierung des jeweiligen Items mit der Differenzierung der zusammengefassten übrigen Items übereinstimmt.[665] Ein gängiges Verfahren zur Ermittlung der Trennschärfe ist die Ermittlung der Item-zu-Rest-Korrelation, also der Korrelation zwischen jedem einzelnen Item-Wert und den jeweiligen gemittelten restlichen Item-Werten.[666] Es existieren keine einheitlichen Vorgaben dazu, wie hoch der Item-zu-Rest-Korrelationskoeffzient mindestens sein sollte, teils wird eine Untergrenze von 0,2 angenommen,[667] teils auch von 0,4.[668] Allgemein kann ein Trennschärfekoeffizient von 0,5 oder höher als hoch eingestuft werden, was dafür spricht, das entsprechende Item zu verwenden.[669] Die in dieser Arbeit verwendeten Indizes bestehen bis auf einen Fall alle nur aus Items, die Trennschärfekoeffizienten von 0,5 oder darüber aufweisen; bei einem Index beträgt die niedrigste Item-zu-Rest-Korrelation 0,449.[670]

III. Test auf Homogenität

Wenn die einzelnen Indikatoren Operationalisierungen desselben Konstrukts darstellen, so muss sich dies darin zeigen, dass sie untereinander korrelieren, die Höhe dieser sogenannten Item-Interkorrelation ergibt die Homogenität eines Index.[671] Bei eindimensionalen Messinstrumenten, also bei Indizes, deren Einzelindikatoren letztlich dasselbe theoretische Konstrukt messen sollen, sind hohe Homogenitäten erstrebenswert.[672] Weist ein Item eine auffallend geringere Item-Interkorrelation auf als die anderen Items, dann deutet dies darauf hin, dass dieses Item nicht nur die Zieldimension

664 Vgl. *Rohwer / Pötter*, Datenkonstruktion, S. 104.

665 Vgl. *Moosbrugger / Kelava*, Testtheorie und Fragebogenkonstruktion, S. 84.

666 Vgl. *Bortz / Döring*, Forschungsmethoden und Evaluation, S. 219; *Asendorpf / Neyer*, Psychologie der Persönlichkeit, S. 53f., 92.

667 Vgl. *Asendorpf / Neyer*, Psychologie der Persönlichkeit, S. 54.

668 Vgl. *Diekmann*, Empirische Sozialforschung, S. 246, die Ausführungen beziehen sich jedoch auf die Item-zu-Summenscore-Korrelation, die höher ausfällt als die Item-zu-Rest-Korrelation, da das getestete Item selbst Teil des Summenscores ist.

669 Vgl. *Bortz / Döring*, Forschungsmethoden und Evaluation, S. 220.

670 Vgl. Index „Beurteilung der eigenen Subjektstellung durch die Klägerin / den Kläger“, unten, Kapitel 11, B.

671 Vgl. *Bortz / Döring*, Forschungsmethoden und Evaluation, S. 220.

672 Vgl. *Bortz / Döring*, Forschungsmethoden und Evaluation, S. 220.

https://doi.org/10.5771/9783845245584-155-1 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

misst, sondern auch oder sogar weitgehend eine Fremddimension.[673] Es sollte dann nicht in den Index aufgenommen werden.[674] Wie beim Trennschärfekoeffizienten besteht auch hier kein allgemeiner Standard im Sinne einer festen Mindestgröße. Bei allen nachfolgend dargestellten Indizes beträgt die mittlere Item-Interkorrelation mindestens 0,4, was als ausreichender Wert anzusehen ist.[675]

IV. Reliabilität des Gesamtindex

Nachdem im Wege der Itemanalyse die weniger geeigneten Items aus dem vorläufigen Itempool ausgeschieden wurden, ist der so entstandene Index seinerseits auf Reliabilität zu testen. Die Reliabilität eines Messinstruments ist ein Maß für die Reproduzierbarkeit der damit erzielten Messergebnisse.[676] Zur Bestimmung dieser Zuverlässigkeit werden verschiedene Reliabilitätskoeffizienten verwendet. Ein sehr gebräuchlicher Reliabilitätskoeffizient ist Cronbachs Alpha, ein Koeffizient, der aus der Korrelation aller Items untereinander und der Anzahl der Items berechnet wird.[677] Daneben gibt es noch mehrere Koeffizienten, die auf der Methode der Testhalbierung beruhen. Hier wird ein Messinstrument mit multiplen Indikatoren in zwei Hälften aufgeteilt und aus der Korrelation zwischen den beiden Testhälften der sogenannte Split-Half-Reliabilitätskoeffizient errechnet.[678] Aus dieser Gruppe von Reliabilitätskoeffizienten wurden in dieser Arbeit der sogenannte Spearman-Brown-Koeffizient sowie Guttmans Split-Half-Koeffizient verwendet.[679] Schließlich wurde noch die Ermittlung von Guttmans Lambda als viertes Verfahren angewendet.[680] Auch bei den Reliabilitätskoeffizienten existiert keine allgemeingültige Vorgabe eines Wertes, ab dem ein Index als hinreichend zuverlässig angesehen werden kann, jedoch ist es üblich, Mindestwerte von 0,7 oder 0,8 zu verlangen.[681]

---

673 *Diekmann*, Empirische Sozialforschung, S. 244; *Bortz / Döring*, Forschungsmethoden und Evaluation, S. 221.

674 Vgl. *Janssen / Laatz*, Statistische Datenanalyse mit SPSS, S. 589.

675 Vgl. *Janssen / Laatz*, Statistische Datenanalyse mit SPSS, S. 589.

676 *Diekmann*, Empirische Sozialforschung, S. 250; *Schnell / Hill / Esser*, Methoden der empirischen Sozialforschung, S. 143.

677 Vgl. *Janssen / Laatz*, Statistische Datenanalyse mit SPSS, S. 590.

678 Vgl. *Diekmann*, Empirische Sozialforschung, S. 250.

679 Vgl. dazu *Janssen / Laatz*, Statistische Datenanalyse mit SPSS, S. 590f.

680 Vgl. dazu *Janssen / Laatz*, Statistische Datenanalyse mit SPSS, S. 591.

681 Vgl. *Janssen / Laatz*, Statistische Datenanalyse mit SPSS, S. 589; einige Autoren differenzieren auch nach der Art des Koeffizienten und verlangen Werte von 0,8, 0,85 oder 0,9, vgl. *Stier*, Empirische Forschungsmethoden, S. 56; *Schelten*, Testbeurteilung und Testerstellung, S. 116.

https://doi.org/10.5771/9783845245584-155-1 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY