

Markus Bühner

**Einführung in die Test- und
Fragebogenkonstruktion**

3., aktualisierte und erweiterte Auflage

PEARSON

Studium

ein Imprint von Pearson Education
München • Boston • San Francisco • Harlow, England
Don Mills, Ontario • Sydney • Mexico City
Madrid • Amsterdam

Vorwort zur 3. Auflage	11
Kapitel 1 Einführung	13
1.1 Ziel des Buches	14
1.2 Testanwendungsbereiche	18
1.3 Arten von Tests	20
1.3.1 Raven Progressive Matrices	23
1.3.2 NEO-FFI und NEO-PI-R	24
1.3.3 Thematischer Apperzeptionstest (TAT)	26
1.4 Diagnosemöglichkeiten mit Tests	27
Kapitel 2 Testtheoretische Grundlagen	29
2.1 Gegenstand einer Testtheorie	30
2.2 Klassische Testtheorie	39
2.2.1 Grundannahmen der Klassischen Testtheorie	41
2.2.2 Kritische Anmerkungen zur Klassischen Testtheorie	53
2.3 Kurzer Ausblick auf die Probabilistische Testtheorie	57
2.4 Haupt- und Nebengütekriterien	58
2.4.1 Hauptgütekriterien	58
2.4.2 Nebengütekriterien	71
2.4.3 Kurzchecklisten zur Testbeurteilung	77
2.4.4 Kurzcheckliste zur Testbewertung	78
Kapitel 3 Erstellung eines Testentwurfs	83
3.1 Festlegung der Art der Indikatoren	85
3.2 Festlegen der Zielgruppe	87
3.3 Testziel und Entscheidung für eine Konstruktionsstrategie	92
3.3.1 Rationale Testkonstruktion	93
3.3.2 Externale Testkonstruktion	93
3.3.3 Induktive Testkonstruktion	94
3.3.4 Prototypenansatz	95
3.3.5 Vergleich der Methoden	95
3.4 Generieren von Indikatoren und Eingrenzen des Konstrukts	97
3.4.1 Erfahrungsgeleitet-intuitiver Ansatz	99
3.4.2 Sammlung und Analyse von Definitionen/Literaturrecherche	100
3.4.3 Analytisch-empirischer Ansatz	101
3.4.4 Personenbezogen-empirische Methode	103
3.5 Erstellen einer Definition des Messgegenstandes	105

3.6	Wahl des Itemformats	108
3.6.1	Gebundene Aufgabenbeantwortung	110
3.6.2	Allgemeine Probleme gebundener Itemformate	125
3.6.3	Die freie Aufgabenbeantwortung	130
3.6.4	Atypische Aufgabenbeantwortung	132
3.7	Richtlinien zur Itemformulierung	133
Kapitel 4 Reliabilität		i«
4.1	Wie ist die Reliabilität eines Tests definiert?	142
4.2	Voraussetzungen für die Reliabilitätsschätzung	147
4.2.1	Modell paralleler und im Wesentlichen paralleler Messungen	149
4.2.2	Modell Tau-äquivalenter und im Wesentlichen Tau-äquivalenter Messungen	150
4.2.3	Modell Tau-kongenerischer Messungen	151
4.3	Methoden der Reliabilitätsschätzung	153
4.3.1	Konsistenzmethode	157
4.3.2	Paralleltestmethode	158
4.3.3	Retestmethode	159
4.4	Formeln zur Schätzung der Reliabilität	161
4.4.1	Reliabilitätsschätzung durch Testhalbierungskoeffizienten	161
4.4.2	Reliabilitätsschätzung durch Konsistenzkoeffizienten	166
4.5	Trennschärfen	171
4.5.1	Berechnung von Eigentrennschärfen	172
4.5.2	Fremdtrennschärfen	177
4.6	Faktoren, die die Reliabilität beeinflussen	178
4.6.1	Homogenität	178
4.6.2	Verteilungsmerkmale der Testkennwerte	178
4.6.3	Verschiedene Arten von Messfehlern	179
4.7	Möglichkeiten der Reliabilitätsverbesserung	182
4.8	Reliabilitätsschätzungen als Ausgangspunkt der psychometrischen Einzelfalldiagnostik	184
4.8.1	Vertrauensintervalle um den beobachteten Wert einer individuellen Testleistung	193
4.8.2	Bedeutsamkeit von Untertestdifferenzen	199
4.8.3	Eine messfehler- und valenzkritische Analyse von Testwertdifferenzen	203
4.8.4	Unterscheiden sich die Leistungen einer Person ^x bei einer wiederholten Messung?	206
4.8.5	Unterscheiden sich zwei Probanden in ihrer Leistung?	209
4.8.6	Richtlinien zur Interpretation von diskrepanten Testbefunden	210
4.9	Minderungskorrekturen	212

Kapitel 5	Empirische Überprüfung des Testentwurfs und Normierung	215
5.1	Itemcodierung und Schwierigkeitsanalyse mit SPSS	217
5.1.1	Durchführung einer Schwierigkeitsanalyse mit SPSS	227
5.1.2	Ergebnisse einer Schwierigkeitsanalyse mit SPSS	228
5.2	Reliabilitätsanalyse mit SPSS	235
5.2.1	Beispiel einer Item- und Reliabilitätsanalyse mit SPSS	240
5.2.2	Itemanalyse mit allen Items der Skala Extraversion	240
5.2.3	Itemanalyse der Skala Extraversion unter Ausschluss von Items mit geringer inhaltlicher Passung	249
5.2.4	Bewertung der inhaltlichen Passung der Items	252
5.3	Zusammenhang zwischen Schwierigkeit, Standardabweichung, Trennschärfe und Reliabilität	257
5.4	Norm- und kriteriumsorientierte Testauswertung mit SPSS	261
5.4.1	Normorientierte Testauswertung	261
5.4.2	Kriteriumsorientierte Testauswertung	281
Kapitel 6	Exploratorische Faktorenanalyse	295
6.1	Grundgedanke der Faktorenanalyse	299
6.1.1	Definitionsgleichung	299
6.1.2	Strukturgleichungen	305
6.2	Methoden der Faktorenanalyse	309
6.2.1	Kommunalitätenproblem	310
6.2.2	Methode der Hauptachsenanalyse (PAF)	313
6.2.3	Maximum-Likelihood-Faktorenanalyse (ML)	316
6.2.4	Vergleich der faktorenanalytischen Methoden mit der Hauptkomponentenmethode	318
6.3	Extraktionskriterien für Faktoren	320
6.3.1	Explizite Theorie zur Faktorenextraktion	321
6.3.2	Eigenwertkriterium größer eins	321
6.3.3	Scree-Test nach Cattell	322
6.3.4	Parallelanalyse nach Hom	323
6.3.5	Minimum-Average-Partial-Test (MAP-Test)	325
6.3.6	Modelltest der ML-Faktorenanalyse	326
6.3.7	Reduktion der Itemspezifität	327
6.4	Rotation	329
6.4.1	Geometrische Darstellung der Faktorenanalyse	330
6.4.2	Orthogonale Rotation	336
6.4.3	Oblique Rotation	338
6.5	Faktorwerte	340
6.6	Voraussetzungen für die Berechnung einer Faktorenanalyse	342
6.7	Kriterien zur Bewertung der Durchführbarkeit einer Faktorenanalyse	346
6.8	Faktorenanalyse mit SPSS	350
6.9	Beispiel einer Faktorenanalyse mit SPSS	354

Kapitel 7	Konfirmatorische Faktorenanalyse	379
7.1	Modell der konfirmatorischen Faktorenanalyse	381
7.2	Definitionsgleichungen . . . j.	384
7.2.1	Messmodell	385
7.2.2	Strukturmodell	388
7.3	Strukturgleichungen	391
7.4	Identifikationsgleichungen	394
7.4.1	Fixierung der Ladungen einer Referenzvariablen bzw. der Fehlervariablen	396
7.4.2	Fixierung der Varianz der latenten Variablen	398
7.4.3	Parameterfixierung durch Modellarmahmen	399
7.4.4	Prüfung der Identifizierbarkeit	401
7.5	Schätzmethoden	407
7.6	Modell-Fit	418
7.6.1	Wie erkenne ich, ob ein Modell passt?	418
7.6.2	Signifikanztests	419
7.6.3	Approximativer Modell-Fit: Fit-Indizes	423
7.6.4	Beurteilung von Modellen	427
7.6.5	Modifikation von Modellen	428
7.7	Voraussetzungen von konfirmatorischen Faktorenanalysen	431
7.8	Durchführung einer konfirmatorischen Faktorenanalyse mit AMOS	434
7.9	Beispiel einer konfirmatorischen Faktorenanalyse mit AMOS	445
7.9.1	Prüfung eines einfaktoriellen Modells	445
7.9.2	Prüfung eines zweifaktoriellen Modells	456
7.10	Multi-Trait-Multi-Method-Ansatz	462
7.11	Das Latent-State-Trait-Modell (LSTM)	464
7.11.1	Wodurch verändern sich Messwerte über die Zeit?	464
7.11.2	Annahmen	471
7.11.3	LSTM mit AMOS	471
Kapitel 8	Probabilistische Testtheorie	477
8.1	Messen	479
8.1.1	Unabhängige Messungen	484
8.1.2	Messinstrumente, die einem Messmodell genügen	487
8.2	Probabilistische Testmodelle	494
8.2.1	Probabilistische Testmodelle für dichotome Itemantworten	495
8.2.2	Probabilistische Testmodelle für ordinale Itemantworten	515
8.3	Modelltests	528
8.3.1	Likelihood-Quotienten-Tests	529
8.3.2	Andersen-Test	531
8.3.3	Nonparametrische Teststatistik TU	532
8.3.4	Likelihood-Quotienten-Test gegen ein saturiertes Modell	533
8.3.5	Pearson- χ^2 -Test	534
8.3.6	Martin-Löf-Test	538
8.3.7	Grafischer Modelltest	539
8.3.8	Informationstheoretische Maße	541

8.3.9	Axiomatische Modellprüfungen	543
8.3.10	Q-Index	543
8.4	Adaptives Testen	549
8.4.1	Branched-Testing	549
8.4.2	Tailored-Testing	550
8.5	Einführung in WINMIRA	557
8.6	Anwendungsbeispiele des Rasch-Modells	561
8.6.1	Beispiel eines ordinalen Rasch-Modells mit WINMIRA	562
8.6.2	Erstes Beispiel eines ordinalen Mixed-Rasch-Modells mit WINMIRA. *	588
8.6.3	Zweites Beispiel eines ordinalen Mixed-Rasch-Modells mit WINMIRA. !!	591
Kapitel 9	Probleme der Testkonstruktion	603
	Literaturverzeichnis	609
	Stichwortverzeichnis	637