

Inhaltsverzeichnis

1	Einführung und zusammenfassender Überblick	1
	<i>Helfried Moosbrugger und Augustin Kelava</i>	
1.1	Zielgruppen und Gliederungsüberlegungen	2
1.2	Teil I: Konstruktionsgesichtspunkte	3
1.3	Teil II: Testtheorien	5
1.4	Teil III: Validität und Möglichkeiten ihrer Überprüfung	8
1.5	Ergänzende Materialien	10
1.6	Zusammenfassung	10
I	Konstruktionsgesichtspunkte	
2	Qualitätsanforderungen an Tests und Fragebogen („Gütekriterien“)	13
	<i>Helfried Moosbrugger und Augustin Kelava</i>	
2.1	Vom Laienfragebogen zum wissenschaftlichen Messinstrument	15
2.2	Unterschiedliche Qualitätsanforderungen	16
2.3	Allgemeine Gütekriterien für Tests und Fragebogen	17
2.4	Spezielle testtheoriebasierte Gütekriterien für wissenschaftliche Tests und Fragebogen	27
2.5	Dokumentation der erfüllten Qualitätskriterien	36
2.6	Zusammenfassung	36
2.7	Kontrollfragen	36
	Literatur	37
3	Planungsaspekte und Konstruktionsphasen von Tests und Fragebogen	39
	<i>Holger Brandt und Helfried Moosbrugger</i>	
3.1	Spezifikation des interessierenden Merkmals	41
3.2	Testarten	44
3.3	Geltungsbereich und Zielgruppe	50
3.4	Testlänge und Testzeit	51
3.5	Testadministration	53
3.6	Struktureller Testaufbau	55
3.7	Konstruktionsphasen im Überblick	57
3.8	Zusammenfassung	63
3.9	Kontrollfragen	63
	Literatur	64
4	Itemkonstruktion und Antwortverhalten	67
	<i>Helfried Moosbrugger und Holger Brandt</i>	
4.1	Ziele und Aspekte der Itemkonstruktion	69
4.2	Itemstamm und Zielgruppe	69
4.3	Vorgehensweisen bei der Itemgenerierung	71
4.4	Kategorisierung von Frageformen	73
4.5	Gesichtspunkte der Itemformulierung	75
4.6	Kognitive und motivationale Prozesse bei der Itembearbeitung	78
4.7	Response-Bias als Fehlerquelle beim Antwortverhalten	81
4.8	Computerunterstützte Itemkonstruktion	86
4.9	Zusammenfassung	86
4.10	Kontrollfragen	87
	Literatur	87

5	Antwortformate und Itemtypen	91
	<i>Helfried Moosbrugger und Holger Brandt</i>	
5.1	Antwortformate im Überblick	93
5.2	Aufgaben mit freiem Antwortformat	94
5.3	Aufgaben mit gebundenem Antwortformat	96
5.4	Aufgaben mit atypischem Antwortformat	112
5.5	Entscheidungshilfen für die Wahl des Aufgabentyps	114
5.6	Computerunterstützte Antwortformate	114
5.7	Zusammenfassung	115
5.8	Kontrollfragen	115
	Literatur	115
6	Computerbasiertes Assessment	119
	<i>Frank Goldhammer und Ulf Kröhne</i>	
6.1	Computerbasiertes Assessment: Definition und Übersicht	121
6.2	Itementwicklung: Antwortformat, Stimulus und Antwortbewertung	124
6.3	Testentwicklung: Testzusammenstellung und -sequenzierung	130
6.4	Testadministration	132
6.5	Datenanalyse und Rückmeldung	135
6.6	Zusammenfassung	137
6.7	EDV-Hinweise	137
6.8	Kontrollfragen	138
	Literatur	138
7	Deskriptivstatistische Itemanalyse und Testwertbestimmung	143
	<i>Augustin Kelava und Helfried Moosbrugger</i>	
7.1	Einleitung	145
7.2	Erstellung der Datenmatrix	145
7.3	Schwierigkeitsanalyse	146
7.4	Itemvarianz	151
7.5	Vorläufige Testwertermittlung	153
7.6	Trennschärfe	153
7.7	Itemselektion auf Basis von Itemschwierigkeit, Itemvarianz und Itemtrennschärfe	155
7.8	Testwertbestimmung und Itemhomogenität	156
7.9	Zusammenfassung	157
7.10	EDV-Hinweise	157
7.11	Kontrollfragen	158
	Literatur	158
8	Testwertverteilung	159
	<i>Augustin Kelava und Helfried Moosbrugger</i>	
8.1	Einleitung	160
8.2	Zentrale Tendenz der Testverteilung	160
8.3	Streuung der Testwertverteilung	161
8.4	Beurteilung der Verteilungsform	161
8.5	Ursachen für die Abweichung der Testwertverteilung von der Normalverteilung	163
8.6	Normalisierung der Testwertverteilung	164
8.7	Zusammenfassung und weiteres Vorgehen	168
8.8	EDV-Hinweise	168
8.9	Kontrollfragen	168
	Literatur	168

9	Testwertinterpretation, Testnormen und Testeichung	171
	<i>Frank Goldhammer und Johannes Hartig</i>	
9.1	Testwertbildung und Testwertinterpretation	172
9.2	Normorientierte Testwertinterpretation	173
9.3	Kriteriumsorientierte Testwertinterpretation	179
9.4	Integration von norm- und kriteriumsorientierter Testwertinterpretation ...	187
9.5	Normdifferenzierung	188
9.6	Testeichung	189
9.7	Zusammenfassung mit Anwendungsempfehlungen	193
9.8	EDV-Hinweise	194
9.9	Kontrollfragen	194
	Literatur	195
10	Standards für psychologisches Testen	197
	<i>Volkmar Höfling und Helfried Moosbrugger</i>	
10.1	Ziele von Teststandards	198
10.2	Standards für die Entwicklung und Evaluation psychologischer Tests	198
10.3	Standards für die Übersetzung und Anpassung psychologischer Tests	203
10.4	Standards für die Anwendung psychologischer Tests	204
10.5	Standards für die Qualitätsbeurteilung psychologischer Tests	210
10.6	Zusammenfassung	213
10.7	Kontrollfragen	213
	Literatur	213
11	Standards für pädagogisches Testen	217
	<i>Sebastian Brückner, Olga Zlatkin-Troitschanskaia und Hans Anand Pant</i>	
11.1	Die „Standards for Educational and Psychological Testing“ im Überblick	219
11.2	Domänen, Ziele und Designs pädagogischen Testens	220
11.3	Validitätsstandards und pädagogisches Testen (Standards 1.0–1.25)	229
11.4	Standards zur Reliabilität (Standards 2.10–2.20)	234
11.5	Schwellenwerte und ihre Bedeutung für die Testwertinterpretation	234
11.6	Weitere Implikationen der Standards für pädagogisches Testen	238
11.7	Standards zum Management und zur Archivierung von Daten pädagogischen Testens	242
11.8	Standards für Forschungsethik	244
11.9	Zusammenfassung	245
11.10	Kontrollfragen	245
	Literatur	245
II	Testtheorien	
12	Testtheorien im Überblick	251
	<i>Helfried Moosbrugger, Karin Schermelleh-Engel, Jana C. Gäde und Augustin Kelava</i>	
12.1	Einleitung	252
12.2	Klassische Testtheorie (KTT)	254
12.3	Item-Response-Theorie (IRT)	260
12.4	Klassische Testtheorie (KTT) vs. Item-Response-Theorie (IRT)	268
12.5	Zusammenfassung	271
12.6	Kontrollfragen	271
	Literatur	271

13	Klassische Testtheorie (KTT)	275
	<i>Helfried Moosbrugger, Jana C. Gäde, Karin Schermelleh-Engel und Wolfgang Rauch</i>	
13.1	Einleitung	277
13.2	Grundannahmen der KTT	277
13.3	Zerlegung einer Itemvariablen in True-Score- und Messfehlervariable	279
13.4	Testwertvariable Y und Testwerte Y_p	280
13.5	Das Gütekriterium der Reliabilität	281
13.6	Messmodelle zur Schätzung der Reliabilität	282
13.7	Empirisches Beispiel	291
13.8	Schätzung individueller Merkmalsausprägungen	294
13.9	Erweiterung der KTT	298
13.10	Zusammenfassung	302
13.11	EDV-Hinweise	302
13.12	Kontrollfragen	302
	Literatur	303
14	Klassische Methoden der Reliabilitätsschätzung	305
	<i>Jana C. Gäde, Karin Schermelleh-Engel und Christina S. Werner</i>	
14.1	Was ist Reliabilität?	307
14.2	Grundlagen	309
14.3	Cronbachs Alpha	314
14.4	Test-Test-Korrelation	322
14.5	Vergleichbarkeit der Reliabilitätsmaße	329
14.6	Einflüsse auf die Reliabilität	330
14.7	Anzustrebende Höhe der Reliabilität	330
14.8	Auswahl eines geeigneten Reliabilitätsmaßes	331
14.9	Zusammenfassung	332
14.10	EDV-Hinweise	333
14.11	Kontrollfragen	333
	Literatur	333
15	Modellbasierte Methoden der Reliabilitätsschätzung	335
	<i>Karin Schermelleh-Engel und Jana C. Gäde</i>	
15.1	Klassische vs. modellbasierte Reliabilitätsschätzung	337
15.2	Eindimensionale Modelle	339
15.3	Mehrdimensionale Modelle	350
15.4	Omega-Koeffizienten im Rahmen weiterer Faktormodelle	360
15.5	Bewertung der modellbasierten Reliabilitätsschätzung	361
15.6	Reliabilitätsschätzung ordinalskalierten Variablen	363
15.7	Erste Empfehlungen zur Beurteilung der Omega-Koeffizienten	364
15.8	Zusammenfassung	365
15.9	EDV-Hinweise	366
15.10	Kontrollfragen	366
	Literatur	366
16	Einführung in die Item-Response-Theorie (IRT)	369
	<i>Augustin Kelava und Helfried Moosbrugger</i>	
16.1	Grundüberlegungen zur IRT	371
16.2	Latent-Trait-Modelle	372
16.3	Dichotomes Rasch-Modell (1PL-Modell)	373
16.4	2PL-Modell nach Birnbaum	399
16.5	3PL-Modell nach Birnbaum	401
16.6	Weitere IRT-Modelle	402
16.7	Zusammenfassung	406

16.8	EDV-Hinweise	407
16.9	Kontrollfragen	407
	Literatur	407
17	Interpretation von Testwerten in der Item-Response-Theorie (IRT)	411
	<i>Dominique Rauch und Johannes Hartig</i>	
17.1	Vorbemerkungen	412
17.2	Grundlagen kriteriumsorientierter Testwertinterpretation in IRT-Modellen ..	414
17.3	Definition von Kompetenzniveaus zur kriteriumsorientierten Testwertinterpretation	417
17.4	Verwendung von Post-hoc-Analysen und A-priori-Merkmalen zur Testwertbeschreibung	418
17.5	Zusammenfassung	422
17.6	EDV-Hinweise	423
17.7	Kontrollfragen	423
	Literatur	423
18	Überblick über Modelle der Item-Response-Theorie (IRT)	425
	<i>Augustin Kelava, Stefano Noventa und Alexander Robitzsch</i>	
18.1	Modelle mit eindimensionalen latenten Merkmalen	426
18.2	Modelle mit mehrdimensionalen latenten Merkmalen	438
18.3	Ausblick auf weitere Modelle	443
18.4	Weiterführende Literatur	444
18.5	EDV-Hinweise	444
18.6	Kontrollfragen	445
	Literatur	445
19	Parameterschätzung und Messgenauigkeit in der Item-Response-Theorie (IRT)	447
	<i>Norman Rose</i>	
19.1	Verfahren der Parameterschätzung in der IRT: Überblick	449
19.2	Maximum-Likelihood-Schätzung (ML-Schätzung)	450
19.3	Bayes'sche Schätzverfahren	466
19.4	Weitere Schätzverfahren	483
19.5	Personenparameterschätzung in der IRT	484
19.6	Reliabilitätsbeurteilung in der IRT	490
19.7	Zusammenfassung	496
19.8	EDV-Hinweise	497
19.9	Kontrollfragen	498
	Literatur	499
20	Computerisiertes adaptives Testen	501
	<i>Andreas Frey</i>	
20.1	Was ist computerisiertes adaptives Testen?	502
20.2	Grundgedanke	503
20.3	Elementare Bausteine	506
20.4	Auswirkungen des adaptiven Testens	516
20.5	Multidimensionales adaptives Testen	520
20.6	Zusammenfassung und Anwendungsempfehlungen	521
20.7	EDV-Hinweise	522
20.8	Kontrollfragen	522
	Literatur	523

III Validität und Möglichkeiten ihrer Überprüfung

21	Validität von Testwertinterpretationen	529
	<i>Johannes Hartig, Andreas Frey und Nina Jude</i>	
21.1	Einleitung	530
21.2	Validität im fachgeschichtlichen Wandel	530
21.3	Argumentationsbasierter Ansatz der Validierung	535
21.4	Beispiele für Validierungsprozesse	539
21.5	Zusammenfassung	544
21.6	Kontrollfragen	544
	Literatur	544
22	Latent-Class-Analyse (LCA)	547
	<i>Mario Gollwitzer</i>	
22.1	Einleitung und Überblick	549
22.2	Herleitung der Modellgleichung	552
22.3	Parameterschätzung und Überprüfung der Modellgüte	556
22.4	Exploratorische und konfirmatorische Anwendungen der LCA	562
22.5	Erweiterte Anwendungen der LCA	567
22.6	Zusammenfassung	571
22.7	EDV-Hinweise	572
22.8	Kontrollfragen	572
	Literatur	573
23	Exploratorische Faktorenanalyse (EFA)	575
	<i>Holger Brandt</i>	
23.1	Einleitung	577
23.2	Faktormodell (Fundamentaltheorem)	578
23.3	Methoden der Faktorextraktion	585
23.4	Abbruchkriterien der Faktorextraktion	590
23.5	Faktorenrotation	595
23.6	Modellevaluation und Itemauswahl	604
23.7	Neue Verfahren	608
23.8	Abschließende Bemerkungen	610
23.9	Zusammenfassung	611
23.10	EDV-Hinweise	611
23.11	Kontrollfragen	611
	Literatur	612
24	Konfirmatorische Faktorenanalyse (CFA)	615
	<i>Jana C. Gäde, Karin Schermelleh-Engel und Holger Brandt</i>	
24.1	Grundlagen	617
24.2	Spezifikation eines Messmodells	619
24.3	Eindimensionale Modelle: Stufen der Messäquivalenz	629
24.4	Mehrdimensionale Modelle	634
24.5	Parameterschätzung	643
24.6	Modellevaluation	648
24.7	Modifikation der Modellstruktur	651
24.8	Modellvergleiche	652
24.9	Messinvarianztestung	653
24.10	Zusammenfassung	656
24.11	EDV-Hinweise	656
24.12	Kontrollfragen	656
	Literatur	657

25	Multitrait-Multimethod-Analysen (MTMM-Analysen)	661
	<i>Karin Schermelleh-Engel, Christian Geiser und G. Leonard Burns</i>	
25.1	Einleitung	663
25.2	Konvergente und diskriminante Validität	663
25.3	Methodeneffekte	664
25.4	Das MTMM-Design	666
25.5	Korrelationsbasierte Analyse der MTMM-Matrix	669
25.6	Faktorenanalytische Ansätze: Klassische CFA-MTMM-Modelle	672
25.7	Faktorenanalytische Ansätze: Neuere CFA-MTMM-Modelle	678
25.8	Zusammenfassung	683
25.9	EDV-Hinweise	684
25.10	Kontrollfragen	684
	Literatur	684
26	Latent-State-Trait-Theorie (LST-Theorie)	687
	<i>Augustin Kelava, Karin Schermelleh-Engel und Axel Mayer</i>	
26.1	Einleitung	688
26.2	LST-Theorie als Erweiterung der KTT	692
26.3	Modelltypen	697
26.4	Anwendungen der LST-Theorie	703
26.5	Zusammenfassung	708
26.6	EDV-Hinweise	709
26.7	Kontrollfragen	709
	Literatur	709
27	Konvergente und diskriminante Validität über die Zeit: Integration von Multitrait-Multimethod-Modellen (MTMM-Modellen) und der Latent-State-Trait-Theorie (LST-Theorie)	713
	<i>Fridtjof W. Nussbeck, Michael Eid, Christian Geiser, Delphine S. Courvoisier und David A. Cole</i>	
27.1	Einleitung	715
27.2	Längsschnittliche MTMM-Modelle	721
27.3	Multiconstruct-LST- und Multimethod-LST-Modell in der empirischen Anwendung	730
27.4	Praktische Hinweise zur Analyse longitudinaler multimodaler Modelle	735
27.5	Zusammenfassung	736
27.6	EDV-Hinweise	736
27.7	Kontrollfragen	737
	Literatur	737
	Serviceteil	739
	Übersicht der griechischen Buchstaben	740
	Verteilungsfunktion der Standardnormalverteilung (z-Tabelle)	741
	Glossar	744
	Stichwortverzeichnis	761