

Lienert/Raatz

Testaufbau und Testanalyse

6. Auflage

Landes-L...
des Fürstent...
V a g u:

Psychologie VerlagsUnion

INHALTSVERZEICHNIS

VORWORT	V
EINLEITUNG	VII
Kapitel 1: Grundsätzliches über den Test.	1
1.1 Wesen und Bedeutung des Tests.	1
1.1.1 Zur Begriffsbestimmung des Wortes „Test“	1
1.1.2 Die geschichtlichen Stadien der Testentwicklung	2
1.1.3 Die Bestandteile eines Tests	5
1.1.4 Die Aufgaben des Tests	6
1.2 Die Gütekriterien eines Tests.	7
1.2.1 Die Objektivität eines Tests	7
1.2.2 Die Reliabilität eines Tests	9
1.2.3 Die Validität eines Tests	10
1.2.4 Die Normierung eines Tests	11
1.2.5 Die Vergleichbarkeit eines Tests	12
1.2.6 Die Ökonomie eines Tests	12
1.2.7 Die Nützlichkeit eines Tests	13
1.2.8 Die Wechselbeziehungen zwischen den Gütekriterien	13
1.3 Die Klassifizierungsmöglichkeiten von Tests.	14
1.4 Die Testaufgabentypen	18
1.4.1 Die gebundene Aufgabenbeantwortung	18
1.4.2 Die freie Aufgabenbeantwortung	20
1.4.3 Atypische Aufgabenbeantwortung	22
1.4.4 Der Komplexitätsgrad der Aufgaben	23
Kapitel 2: Planungsprobleme.	24
2.1 Wahl des Aufgabentypus	24
2.1.1 Vor- und Nachteile der RF (Richtig-Falsch)-Aufgabe	24
2.1.2 Vor- und Nachteile der EG (Ergänzungs)-Aufgabe	25
2.1.3 Vor- und Nachteile der MW (Mehrfach-Wahl)-Aufgabe	26
2.1.4 Vor- und Nachteile der ZO (Zuordnungs)-Aufgabe	27

XII Inhaltsverzeichnis

2.1.5	Vor- und Nachteile der UO (Umordnungs)-Aufgabe	27
2.1.6	Vor- und Nachteile der KA (Kurzaufsatz)-Aufgabe	28
2.1.7	Zusammenfassung	28
2.2	Die Kriterien einer guten Testaufgabe	29
2.2.1	Die Objektivität der Aufgabe	29
2.2.2	Die Reliabilität der Aufgabe	29
2.2.3	Die Validität einer Aufgabe	30
2.2.4	Die Schwierigkeit einer Aufgabe	30
2.2.5	Die Beziehung zwischen den einzelnen Aufgabenkriterien	31
2.2.6	Die inhaltlichen Kriterien einer guten Testaufgabe	32
2.3	Testlänge und Testzeit	33
2.3.1	Die Testlänge	33
2.3.2	Die Testzeit	34
2.3.3	Schnelligkeit und Niveau	34
2.4	Homogenität und Validität des Aufgabeninhaltes	36
2.4.1	Homogenität und Heterogenität	36
2.4.2	Statistische und psychologische Validität	36
2.5	Die Arbeitsteilung im Rahmen der Testentwicklung	37
Kapitel 3:	Der Testentwurf	40
3.1	Die Abgrenzung von Validitäts- und Geltungsbereich	40
3.1.1	Der Validitätsbereich	40
3.1.2	Der Geltungsbereich	41
3.1.3	Rationales oder empirisches Testkonzept	42
3.2	Merkmalsanalyse und Teststruktur	43
3.2.1	Die Merkmalsanalyse	43
3.2.2	Der strukturelle Aufbau des Tests	44
3.3	Testanweisung und Aufgabenbewertungsplan	45
3.3.1	Die Testanweisung	45
3.3.2	Die Testvorgabe	48
3.3.3	Die Aufgabenbewertung	48
Kapitel 4:	Die Aufgabenkonstruktion	50
4.1	Das Aufgabenkonzept	50
4.2	Regeln für den sprachlichen Aufbau von Aufgaben	52
4.3	Der provisorische Testaufbau	53

Kapitel 5: Die Aufgabenanalyse	57
5.1 Das Ziel der Aufgabenanalyse	57
5.1.1 Bedeutung der Qualität der Aufgaben	57
5.1.2 Bedeutung der Testpunktverteilung	58
5.2 Planung und Durchführung der Aufgabenanalyse	59
5.2.1 Die Wahl des Analysenkriteriums	59
5.2.2 Die Wahl der Analysenstichprobe	60
5.2.3 Das Erprobungs-Ernstfall-Dilemma	61
5.2.4 Einzel- und Gruppendurchführung bei der Aufgabenanalyse ..	61
5.2.5 Die Darbietungstechniken von Aufgaben	62
5.2.6 Die Kontrolle der Testdurchführung	66
5.3. Die Rohwertermittlung bei der Analysenstichprobe	66
5.3.1 Die einfache Rohwertermittlung	68
5.3.2 Die Rohwertermittlung mit Zufallskorrektur	68
5.3.3 Die Rohwertermittlung mit empirischer Fehlerkorrektur	69
5.3.4 Die Testzeit als Rohwert	71
5.3.5 Die Rohwertermittlung bei „gewogener“ Aufgabenbewertung	72
Kapitel 6: Berechnung von Schwierigkeitsindex, Trennschärfenkoeffizient und Aufgaben-Interkorrelation	73
6.1 Die Schwierigkeitsanalyse	73
6.1.1 Die Bedeutung des Schwierigkeitsindex	73
6.1.2 Die Bestimmung des Schwierigkeitsindex	74
6.2 Die Trennschärfenanalyse	78
6.2.1 Die Bedeutung des Trennschärfenkoeffizienten	78
6.2.2 Die Bestimmung des Trennschärfenkoeffizienten bei einem quantitativen Analysenkriterium	79
6.2.3 Die Trennschärfebestimmung bei einem qualitativen Analysenkriterium	84
6.2.4 Die Bestimmung der Trennschärfe bei einem echt alternativen Analysenkriterium	91
6.2.5 Andere Trennschärfe-Beurteilungsmethoden	94
6.3 Die Trennschärfe beeinflussende Faktoren	96
6.3.1 Scheinkorrelation gegenüber dem Testrohwert	96
6.3.2 Stellung der Aufgabe	97
6.3.3 Reliabilität des Analysenkriteriums	98
6.4 Die Bedeutung der Aufgaben-Interkorrelationen	98

6.5 Analyse und Revision der Alternativ-Antworten bei Mehrfach-Wahl-Aufgaben	101
6.5.1 MW-Aufgaben mit Best-Antwort	101
6.5.2 MW-Aufgaben ohne Best-Antwort	104
6.6 Die Beziehungen zwischen Schwierigkeit, Trennschärfe und Aufgaben-Interkorrelation	104
6.6.1 Schwierigkeit und Trennschärfe	104
6.6.2 Schwierigkeit und Homogenität	108
6.6.3 Die Verteilung der Schwierigkeitsindizes	109
6.6.4 Die Streuung der Aufgabenpunktwerte	109
6.7 Spezielle Probleme der Aufgabenanalyse	112
6.7.1 Zeitbegrenzung und Aufgabenindizes	112
6.7.2 Zufallseinfluß und Aufgabenindizes	112
6.7.3 Die Faktorenanalyse der Aufgaben	113
Kapitel 7: Die Verwertung der Analysedaten	114
7.1 Aufgabenselektion und -revision aufgrund der Analysedaten	114
7.1.1 Die Aufgabenselektion	114
7.1.2 Die Aufgabenrevision	123
7.1.3 Die Konsequenzen aus Selektion und Revision	125
Demonstration der Aufgabenauswahl an einem Beispiel	126
7.2.1 Die rationale Selektion	128
7.2.2 Die Kennwert-Selektion	129
7.2.3 Die GuixiKSEN-Technik	131
7.2.4 Die Validitätsvorhersage	134
7.2.5 Die Reliabilitätsvorhersage	135
7.3 Aufgabenbewertung und Aufgabenreihung aufgrund der Analysedaten	137
7.3.1 Die Aufgabenbewertung	137
7.3.2 Die Aufgabenreihung	140
7.4 Paralleltestaufbau aufgrund der Analysedaten	141
7.4.1 Aufgabenpaarung nach der Aufgabenanalyse	141
7.4.2 Aufgabenpaarung vor der Aufgabenanalyse	142
7.4.3 Aufbau von Pseudo- und Quasi-Parallelform	143
7.5 Die Unterlassung der Aufgabenanalyse und ihre Durchführung in einem späteren Testentwicklungsstadium	144
7.5.1 Die Unterlassung der Aufgabenanalyse	144
7.5.2 Die Durchführung der Aufgabenanalyse in einem späteren Stadium der Testentwicklung	145

Kapitel 8: Die Analyse der Rohwertverteilung	147
8.1 Die Überprüfung der Normalverteilungshypothese	147
8.1.1 Die Verteilung der Rohwerte	147
8.1.2 Die Überprüfung der Verteilung auf Normalität	147
8.2 Ursachen für anomale Rohwertverteilungen	151
8.2.1 Anomale Verteilung infolge Stichprobenheterogenität	152
8.2.2 Anomale Verteilung infolge mangelhafter Testkonstruktion.	155
8.2.3 Anomale Verteilung infolge eines nichtnormalverteilten Persönlichkeitsmerkmals.	158
8.3 Die praktische Bedeutung anomaler Verteilungsformen	159
8.3.1 Unterschiedliche Differenzierung in verschiedenen Skalenbereichen.	159
8.3.2 Die Testrevision aufgrund der Verteilungsanalyse.	160
Kapitel 9: Die Entwicklung der Testendform	162
9.1 Der Aufbau der Testendform	162
9.1.1 Herstellung des Testmaterials.	162
9.1.2 Anordnung der Aufgaben	163
9.2 Die Darbietung der Testendform	163
9.2.1 Die Festlegung der Testzeit für die Testendform.	164
9.2.2 Die Testinstruktion.	164
9.2.3 Motivation und Anspannung.	167
9.2.4 Die äußeren Bedingungen.	167
9.3 Die Auswertung der Testendform	168
9.3.1 Die Korrektur des Rohwertes.	168
9.3.2 Die Auswertungstechniken.	170
Kapitel 10: Die Kontrolle der Reliabilität eines Testes	173
10.1 Die experimentelle Planung der Reliabilitätskontrolle	173
10.1.1 Der Zeitpunkt der Reliabilitätskontrolle.	173
10.1.2 Wahl und Größe der Kontrollstichprobe.	174
10.2 Die Aspekte der Reliabilität	175
10.3 Die Methoden der Reliabilitätsbestimmung	180
10.3.1 Die Testwiederholungsmethode.	180
10.3.2 Die Paralleltestmethode.	182

XVI Inhaltsverzeichnis

10.3.3 Die Testhalbierungsmethode	182
10.3.4 Die Methode der Konsistenzanalyse	191
10.4 Vergleichende Bewertung von Reliabilitätskoeffizienten	201
10.4.1 Die Bedeutung der verschiedenen Reliabilitätskennwerte	201
10.4.2 Faktoren, die die Reliabilität beeinflussen	202
10.5 Die Möglichkeiten der Reliabilitätsverbesserung	209
10.5.1 Die Testverlängerung	209
10.5.2 Weitere Möglichkeiten	213
10.6 Spezielle Probleme der Reliabilitätskontrolle	213
10.6.1 Schwierigkeiten bei der Reliabilitätsbestimmung	213
10.6.2 Die Reliabilität von Sekundär-Rohwerten	214
10.6.3 Schnelligkeit und Niveau als Grundlagen der Reliabilitäts- beurteilung	218
Kapitel 11: Die Kontrolle der Validität eines Tests	220
11.1 Die kriterienbezogene Validität	220
11.2 Die inhaltliche Validität	224
11.3 Die Konstruktvalidität	226
11.4 Das Validitätskriterium	228
11.4.1 Die Einteilung der Validitätskriterien	229
11.4.2 Die Gewinnung der Validitätskriterien	230
11.5 Die Bewertung von Validitätskriterien	234
11.5.1 Einfache objektive Kriterien	234
11.5.2 Kombinierte objektive Kriterien	234
11.5.3 Subjektive Kriterien	237
11.6 Die Methoden der Validitätsbestimmung	241
11.6.1 Die Planung der Validitätskontrolle	241
11.6.2 Statistische Methoden zur Ermittlung eines Validitäts- kennwertes	242
11.7 Faktoren, die die Validität mitbestimmen	255
11.7.1 Die teilweise Unvereinbarkeit von Reliabilität und Validität	255
11.7.2 Testlänge und Validität	255
11.7.3 Die Reliabilität des Validitätskriteriums	257
11.7.4 Die Ermittlung der Reliabilität des Validitätskriteriums	261

11.8 Die Ermittlung des Validitätskoeffizienten aus irrepräsentativen Stichproben	264
11.9 Anforderungen an die Validität eines Tests	269
11.9.1 Die Validität und der Verwendungszweck	269
11.9.2 Allgemeine Richtlinien über die erforderliche Höhe von Validitätskoeffizienten	270
Kapitel 12: Die Testeichung	272
12.1 Probleme, Planung und Durchführung der Eichung	272
12.1.1 Reliabilität und Eichung	272
12.1.2 Die Organisierung einer Eichstichprobe	273
12.1.3 Der Umfang der Eichstichprobe	276
12.1.4 Die Durchführung der Eichung	276
12.2 Die Analyse der Eichstichprobenverteilung	276
12.2.1 Die Häufigkeitsverteilung der Testrohwerter	276
12.2.2 Varianzanalyse und Homogenitätsprüfung bei der Aufstellung von Gruppennormen	277
12.3 Norm-Maßstäbe	281
12.3.1 Äquivalenz— und Variabilitätsnormen	282
12.3.2 Einfach— und Mehrfachnormen	282
12.3.3 Standardnormen und Prozentrangnormen	283
12.3.4 Grob- und Feinnormen	288
12.3.5 Verteilungsfreie und verteilungsgebundene Normen	290
12.4 Die rechnerische Ableitung und Darstellung der gebräuchlichen Normen	291
12.4.1 Die Normierung normal verteilter Eichstichproben-Rohwerter	291
12.4.2 Die Normierung anomal verteilter Eichstichproben-Rohwerter	293
Kapitel 13: Die Entwicklung von Paralleltests und Testserien	297
13.1 Wesen und Bedeutung von Paralleltests und Testserien	297
13.1.1 Paralleltests	297
13.1.2 Testserien	298
13.2 Die Äquivalenzkriterien	299
13.2.1 Die Äquivalenz der Validität	300
13.2.2 Die Äquivalenz der Reliabilität	300
13.2.3 Die Äquivalenz der Verteilungskennwerter	300

XVIII Inhaltsverzeichnis

13.2.4 Die Äquivalenz der Häufigkeitsverteilung	301
13.2.5 Simultanüberprüfung von Mittelwerten, Varianzen und Kovarianzen	301
13.3 Paralleltetaufbau und Äquivalenzüberprüfung	306
13.3.1 Paralleltetaufbau	306
13.3.2 Die Äquivalenzüberprüfung	307
13.4 Methoden zur Angleichung nicht-äquivalenter Parallelförmigkeiten	308
13.4.1 Angleichung durch Gleichsetzung der Mittelwerte	309
13.4.2 Angleichung durch lineare Transformation	309
13.4.3 Angleichung durch Regression	309
13.4.4 Angleichung über die Aquiperzentil-Methode	310
13.5 Der Aufbau von Entwicklungstestserien	315
Kapitel 14: Testprofile und Testbatterien.	318
14.1 Zur Charakteristik von Testprofilen und Testbatterien	318
14.1.1 Das Testprofil	318
14.1.2 Die Testbatterie	319
14.1.3 Die möglichen Kombinationen von Tests und Validitätskriterien	320
14.2 Die Analyse von Testprofilen	321
14.2.1 Die Profildarstellung	322
14.2.2 Die Profilverlässlichkeit	323
14.2.3 Die Profilvalidität	324
14.2.4 Die Profilverkorrelationen	325
14.2.5 Differentielle Eignungstests	325
14.3 Der formale Aufbau von Testbatterien	326
14.3.1 Die Arten der Testbatterien	327
14.3.2 Die Kombination von Einzeltests zu Batterien	327
14.4 Die Reliabilität von Testbatterien	328
14.4.1 Die empirische Ermittlung	328
14.4.2 Die rechnerische Ermittlung	329
14.5 Die differentielle Wägung der Untertests	332
14.5.1 Die implizite Wägung	332
14.5.2 Die explizite Wägung	333
14.5.3 Die Ermittlung der Testkennwerte für eine Batterie aus gewogenen und nicht gewogenen Untertests	336

14.6 Die Validität von Testbatterien	337
14.6.1 Die Validität einer Testbatterie als Maßkorrelation eines Kriteriums mit einer Summe von Untertest-Rohwerten.	338
14.6.2 Multiple Validität und Suppressor-Funktion.	340
14.6.3 Die multiple Validität als Korrelation zweier Tests gegenüber einem Kriterium.	342
14.6.4 Die allgemeine Formel für den multiplen Validitätskoeffizienten für eine Batterie mit mehreren Untertests.	345
14.6.5 Rohwert- und Beta-Gewichte als Koeffizienten in der multiplen Regressionsgleichung.	346
14.6.6 Die Ermittlung der Beta-Gewichte.	347
14.6.7 Die Verwendung der Beta-Gewichte.	350
14.7 Die Untertest-Selektion	353
14.7.1 Die Selektionskennwerte.	353
14.7.2 Die rationale Untertest-Selektion.	354
14.7.3 Die statistische Untertest-Selektion.	356
14.8 Die Eichung von Testbatterien	360
14.8.1 Die Normierung von Testbatterien.	361
14.8.2 Die Normierung von Profilbatterien.	362
 Kapitel 15: Zur Interpretation objektiver Testresultate	364
15.1 Zur Interpretation einfacher Tests	364
15.1.1 Der Standardmeßfehler als Unsicherheitsmaß für individuelle Testpunktwerte.	364
15.1.2 Die praktische Bedeutung des Standardmeßfehlers.	365
15.1.3 Die Beurteilung interindividueller Unterschiede.	368
15.1.4 Die Beurteilung intraindividueller Unterschiede.	370
15.1.5 Beurteilung weiterer Unterschiede.	371
15.2 Zur Interpretation von Testprofilen	373
15.2.1 Die Bedeutung des Profilreliabilitätskoeffizienten.	373
15.2.2 Die speziellen Aussagemöglichkeiten eines Testprofils.	374
15.2.3 Eine allgemeine Formel zur Prüfung von Profilunterschieden.	377
15.2.4 Die Ähnlichkeitsbeurteilung von Profilen.	378
15.3 Die Kriteriumvorhersage bei Tests und Testbatterien	381
15.3.1 Die einfache Regression bei der Interpretation empirisch valider Tests.	382
15.3.2 Die multiple Regression bei der Kriteriumvorhersage.	386

XX	Inhaltsverzeichnis	
15.4	Die Unsicherheit der Kriteriumvorhersage	387
15.4.1	Der Standardschätzfehler als Unsicherheitsmaß der Kriteriumvorhersage	387
15.4.2	Die Unsicherheit der Voraussage eines dichotomischen Kriteriums	389
	Literaturverzeichnis	395
	Anhang	407
	Sachverzeichnis	427